Edited by

A VICTOR **HOFFBRAND**

DOUGLAS R **HIGGS**

DAVID M **KEELING**

ATUL B **MEHTA**

Seventh Edition

Postgraduate Haematology

WILEY Blackwell

Postgraduate Haematology

Companion website

This book has a companion website:

www.wiley.com/go/hoffbrand/ph7

with:

- Figures and tables from the book for downloading

Postgraduate Haematology

EDITED BY

A Victor Hoffbrand MA, DM, FRCP, FRCPath, FRCP (Edin), DSc, FMedSci
Emeritus Professor of Haematology, University College London, London, UK

Douglas R Higgs MD, FRCP, FRS
Professor of Molecular Haematology & Director, Weatherall Institute of Molecular Medicine, John Radcliffe Hospital, Oxford, UK

David M Keeling BSc, MD, FRCP, FRCPath
Consultant Haematologist, Oxford Haemophilia and Thrombosis Centre, Oxford University Hospitals, Oxford, UK

Atul B Mehta MA, MD, FRCP, FRCPath
Consultant Haematologist, Lysosomal Storage Disorders Unit, Department of Haematology, Royal Free Hospital, London, UK

Seventh Edition

WILEY Blackwell

Registered office: John Wiley & Sons, Ltd, The Atrium, Southern Gate, Chichester, West Sussex, PO19 8SQ, UK

Editorial offices: 9600 Garsington Road, Oxford, OX4 2DQ, UK
The Atrium, Southern Gate, Chichester, West Sussex, PO19 8SQ, UK
111 River Street, Hoboken, NJ 07030-5774, USA

For details of our global editorial offices, for customer services and for information about how to apply for permission to reuse the copyright material in this book please see our website at www.wiley.com/wiley-blackwell

Library of Congress Cataloging-in-Publication Data

Postgraduate haematology / edited by A. Victor Hoffbrand, Douglas R. Higgs, David M. Keeling, Atul B. Mehta. – Seventh edition.
p. ; cm.
Includes bibliographical references and index.
ISBN 978-1-118-85432-7 (cloth)
I. Hoffbrand, A. V., editor. II. Higgs, Douglas R., editor. III. Keeling, David (David Michael), editor. IV. Mehta, Atul B., editor.
[DNLM: 1. Blood Physiological Phenomena. 2. Hematologic Diseases. 3. Leukemia–blood. 4. Thrombosis–blood. WH 100]
RC633
616.1′5–dc23 2015019099

A catalogue record for this book is available from the British Library.

Wiley also publishes its books in a variety of electronic formats. Some content that appears in print may not be available in electronic books.

Cover image: GettyImages-464401418 by Frentusha

Set in 9.5/12pt MinionPro by Aptara Inc., New Delhi, India

1 2016

Contents

Contributor list

Shubha Allard
Barts Health NHS Trust and NHS Blood and Transplant, London, UK

Paul R Ames
Department of Haematology, Haemostasis and Thrombosis, St George's Hospital, London, UK

Trevor Baglin
Cambridge University Hospitals NHS Trust, Addenbrookes Hospital, Cambridge, UK

Barbara J Bain
Professor in Diagnostic Haematology, St Mary's Hospital Campus of Imperial College Faculty of Medicine, London and Honorary Consultant Haematologist, St Mary's Hospital, London, UK

Imelda Bates
Professor of Tropical Haematology, Liverpool School of Tropical Medicine, Liverpool, UK

Paola Bianchi
Fondazione IRCCS Ca' Granda Ospedale Maggiore Policlinico Milano, Oncohaematology Unit, Physiopathology of Anaemias Unit, Milan, Italy

Joan Bladé
Hospital Clinic de Barcelona, Institut d'Investigacions Biomediques August Pi I Ferrer (IDIBAPS), Barcelona, Spain

Piers Blombery
UCL Cancer Institute, School of Life and Medical Sciences, University College London, London, UK

Niccolo Bolli
Division of Hematology, Fondazione IRCCS Istituto Nazionale dei Tumori, University of Milan, Milan, Italy

Alessandro Broccoli
Institute of Haematology 'L. e A. Seràgnoli', University of Bologna, Bologna, Italy

Alan K Burnett
Professor, Institute of Cancer and Genetics, Cardiff University, Cardiff, UK

Clara Camaschella
Vita-Salute University, Milan, Italy

Peter J Campbell
Head, Cancer Genetics and Genomics at the Institute, Wellcome Trust Sanger Institute, Cambridge, UK

Elias Campo
Hospital Clinic, University of Barcelona, Barcelona, Spain

Ronjon Chakraverty
Professor of Haematology and Cellular Immunotherapy, University College London, London, UK

Subarna Chakravorty
Department of Paediatrics, Imperial College Healthcare and Imperial College London, London, UK

Peter W Collins
Professor, Cardiff Institute of Infection & Immunity, School of Medicine, Cardiff University, University Hospital of Wales, Cardiff, UK

Marcela Contreras
Professor of Transfusion Medicine, Royal Free and University Medical Schools, London, UK

Charles Craddock
Professor and Consultant Haematologist, Centre for Clinical Haematology, Queen Elizabeth Hospital, Birmingham, UK

Kate Cwynarski
Consultant Haematologist and Honorary Senior Lecturer (UCL), Department of Haematology, Royal Free Hampstead NHS Trust, London, UK

Geoff Daniels
International Blood Group Reference Laboratory, NHS Blood and Transplant, Bristol, UK

Emma de Pater
Erasmus Stem Cell Institute and Department of Cell Biology, Erasmus Medical Center, Rotterdam, Netherlands

Inderjeet S Dokal
Chair of Child Health and Honorary Consultant in Haematology, Barts and The London School of Medicine and Dentistry, Queen Mary University of London, Barts Health NHS Trust, London, UK

Maria Domenica Cappellini
Foundation IRCCS Ca' Granda Policlinico and DISSCO University of Milan, Milan, Italy

Mark Drayson
Professor Clinical Immunodiagnostics, Director, Clinical Immunology Service, Honorary Consultant University Hospitals Birmingham and Heart of England, Birmingham, UK

Elaine Dzierzak
Professor of Cell Biology, Erasmus Stem Cell Institute, Erasmus Medical Centre, Rotterdam, Netherlands and Centre for Inflammation Research, University of Edinburgh, Edinburgh, UK

Ivy Ekem
Department of Haematology, University of Ghana and Korle Bu Teaching Hospital, Accra, Ghana

Modupe O Elebute
Honorary Consultant, Department of Haematology, King's College Hospital, London, UK

Adele K Fielding
Reader in Haematology, Cancer Institute, University College London, London, UK

Vanya Gant
University College London Hospitals NHS Trust, London, UK

Simon DJ Gibbs
National Amyloidosis Centre, Royal Free and University College London Medical School, London, UK

Eliza Gil
University College London Hospitals NHS Trust, London, UK

Keith Gomez
Haemophilia Centre and Thrombosis Unit, Royal Free London NHS Foundation Trust, London, UK

Anthony R Green
Professor of Haemato-Oncology, University of Cambridge; Departments' of Haematology and Oncology, Cambridge University Hospitals NHS Foundation Trust; Cambridge Institute for Medical Research, Wellcome Trust/MRC Stem Cell Institute, and Department of Haematology, University of Cambridge, Cambridge, UK

David Grimwade
Professor of Molecular Haematology, Department of Medical and Molecular Genetics, King's College London, UK and Honorary Consultant Haematologist, Guy's and St Thomas' NHS Foundation Trust, London, UK

Torsten Haferlach
MLL Munich Leukemia Laboratory, Munich, Germany

Claire Harrison
Professor and Consultant Haematologist, Department of Haematology, Guy's and St Thomas' NHS Foundation Trust, London, UK

Paul Harrison
Senior Lecturer, School of Immunity and Infection, College of Medical and Dental Sciences, University of Birmingham, Birmingham, UK

Philip N Hawkins
National Amyloidosis Centre, Royal Free and University College London Medical School, Centre for Amyloidosis and Acute Phase Proteins, London, UK

Deborah Hay
Weatherall Institute of Molecular Medicine, John Radcliffe Hospital, Oxford, UK

Chaim Hershko MD
Department of Medicine, Shaare Zedek Medical Center; Professor Emeritus, Hebrew U Hadassah Medical School, Jerusalem, Israel

Douglas R Higgs
Professor of Molecular Haematology and Director, Weatherall Institute of Molecular Medicine, John Radcliffe Hospital, Oxford, UK

Peter Hillmen
Professor of Experimental Haematology and Honorary Consultant Haematologist, Leeds Teaching Hospitals NHS Trust, Leeds, UK

A Victor Hoffbrand
Emeritus Professor of Haematology, University College London, London, UK

Ronald Hoffman
Tisch Cancer Institute, Icahn School of Medicine at Mount Sinai, New York, New York, USA

Derralynn A Hughes
Senior Lecturer and Honorary Consultant Haematologist, RFH Lysosomal Storage Disorders Unit, Royal Free Hospital, London, UK

Timothy P Hughes
Haematologist and Head of Department (RAH site), Department of Haematology, SA Pathology, Adelaide; Clinical Professor, Discipline of Medicine, University of Adelaide and South Australian Health and Medical Research Institute, Adelaide, Australia

Beverley J Hunt
Professor of Thrombosis and Haemostasis and Consultant Haematologist, Guy's and St Thomas' Foundation Trust, London, UK

David Keeling
Consultant Haematologist, Oxford Haemophilia and Thrombosis Centre, Oxford University Hospitals, Oxford, UK

Rachel Kesse-Adu
Consultant, Department of Haematology, Guy's and St Thomas' Hospital, London, UK

Panagiotis Kottaridis
Department of Haematology, University College London, London, UK

Marina Kremyanskaya
Tisch Cancer Institute, Icahn School of Medicine at Mount Sinai, New York, New York, USA

Austin G Kulasekararaj
Department of Haematology, King's College Hospital/King's College London, London, UK

Michael A Laffan
Centre for Haematology, Imperial College School of Medicine, Imperial College, Hammersmith Hospital, London, UK

Swee Lay Thein
King's College London, Molecular Haematology, Faculty of Life Sciences and Medicine, and Department of Haematological Medicine, King's College Hospital NHS Foundation Trust, London, UK

David Linch
Professor of Clinical Haematology, UCL Cancer Institute, School of Life and Medical Sciences, University College London, London, UK

Pier Mannuccio Mannucci
A Bianchi Bonomi Hemophilia and Thrombosis Center, IRCCS Cà Granada, Ospedale Maggiore, Milan, Italy

Robert Marcus
Consultant Haematologist, King's College Hospital NHS Foundation Trust, and Department of Clinical Haematology, King's College Hospital, Denmark Hill, London, UK

Maurizio Margaglione
Medical Genetics, Department of Clinical and Experimental Medicine, University of Foggia, Italy

Anne Marsh
Department of Hematology/Oncology, UCSF Benioff Children's Hospital Oakland, Oakland, CA, USA

Judith CW Marsh
Professor, King's College London, and Department of Haematological Medicine, King's College Hospital, London, UK

John Mascarenhas
Tisch Cancer Institute, Icahn School of Medicine at Mount Sinai, New York, New York, USA

John H McVey
Professor of Cardiovascular Biology, School of Biosciences and Medicine, University of Surrey, Guildford, UK

Atul B Mehta

Consultant Haematologist, Lysosomal Storage Disorders Unit, Department of Haematology, Royal Free Hospital, London, UK

Marzia Menegatti

A Bianchi Bonomi Hemophilia and Thrombosis Center, Fondazione IRCCS Ca' Granda Ospedale Maggiore Policlinico, Università degli Studi di Milano, Milan, Italy

Narla Mohandas

Laboratory of Red Cell Physiology, New York Blood Center, New York, NY, USA

Emili Montserrat

Institute of Haematology and Oncology, Hospital Clinic, University of Barcelona, Barcelona, Spain

Neil V Morgan

Lecturer in Cardiovascular Genetics, Cardiovascular and Respiratory Sciences, School of Clinical and Experimental Medicine, College of Medical and Dental Sciences, University of Birmingham, Birmingham, UK

Paul Moss

Professor and Head, School of Cancer Sciences, University of Birmingham, and Queen Elizabeth Hospital, Birmingham, UK

Ghulam J Mufti

Department of Haematological Medicine, King's College Hospital, UK

Adrian C Newland

Professor of Haematology, Academic Haematology Unit, Blizard Institute; Barts and The London School of Medicine and Dentistry, Queen Mary University of London, London, UK

Jessica Okosun

Centre for Haemato-Oncology, Barts Cancer Institute, University of London, London, UK

Roberta Palla

Department of Pathophysiology and Transplantation, Università degli Studi di Milano, A Bianchi Bonomi Hemophilia and Thrombosis Center, Milan, Italy

K John Pasi

Professor and Consultant Haematologist, Barts and The London School of Medicine and Dentistry, Royal London Hospital London, UK

Flora Peyvandi

A Bianchi Bonomi Hemophilia and Thrombosis Center, IRCCS Cà Granada, Ospedale Maggiore, Milan, Italy

Stefano A Pileri

Professor of Pathology, Bologna University School of Medicine, and Director of the Haematopathology Unit, European Institute of Oncology, Bologna, Italy

Drew Provan

Centre for Haematology, Institute of Cell and Molecular Science, Queen Mary University of London, London, UK

Kavita Raj

Professor of Haematological Oncology, King's College London, London, UK

David Rees

Consultant Paediatric Haematologist, King's College London/King's College Hospital, Department of Haematological Medicine, King's College Hospital, Denmark Hill, London, UK

Irene Roberts

Professor of Paediatric Haematology, Department of Paediatrics and Weatherall Institute of Molecular Medicine, University of Oxford, Oxford, UK

Clare J Rowntree

Department of Haematology, University Hospital of Wales, Cardiff, UK

Noémi Roy

Academic Clinical Lecturer, Weatherall Institute of Molecular Medicine, John Radcliffe Hospital, Oxford, UK

Jesús San-Miguel

Clinica Universidad de Navarra, Centro de Investigación Medica Aplicada (CIMA), Pamplona, Spain

Jecko Thachil

Department of Haematology, Manchester Royal Infirmary, Manchester, UK

Cheng-Hock Toh

Professor of Haematology, Roald Dahl Centre, Royal Liverpool University Hospital, Liverpool, UK

William Townsend

University College Hospitals NHS Foundation Trust and King's College London, London, UK

George Vassiliou

Haematological Cancer Genetics, Wellcome Trust Sanger Institute Hinxton, Cambridge, UK

Elliott P Vichinsky

Medical Director, Hematology/Oncology, UCSF Benioff Children's Hospital Oakland, Oakland, CA, and Professor of Pediatrics, University of California San Francisco, CA, USA

Ajay Vora

Consultant Paediatric Haematologist and Honorary Professor of Haematology (University of Sheffield), Department of Haematology, Sheffield Children's Hospital, Sheffield, UK

Stephen P Watson

Professor in Cardiovascular Sciences and Cellular Pharmacology, Centre for Cardiovascular Sciences, Institute for Biomedical Research, College of Medical and Dental Sciences, University of Birmingham, Birmingham, UK

Henry G Watson

Consultant Haematologist, Aberdeen Royal Infirmary, Aberdeen, UK

David TO Yeung

Haematologist, Department of Haematology, SA Pathology, Adelaide; Clinical Associate Lecturer, Discipline of Medicine, University of Adelaide, Adelaide, Australia

Neal S Young

National Heart, Lung and Blood Institute, National Institutes of Health, Bethesda, MD, USA

Pier Luigi Zinzani

Professor, Institute of Haematology 'L. e A. Seràgnoli', University of Bologna, Bologna, Italy

Preface to the seventh edition

Since the sixth edition of *Postgraduate Haematology* was published in 2011, substantial advances have been made in our understanding of the pathogenesis of inherited and acquired haematological diseases. This progress has largely resulted from the application of next generation sequencing of the relevant exomes and genomes to identify the DNA mutations responsible for these diseases. For example, mutation of the myeloid differentiation primary response gene (MYD88) has been found in over 90% of cases of Waldenstrom's macroglobulinaemia; mutation of calreticulin has been found in most of the JAK2 negative cases of essential thrombocythaemia and primary myelofibrosis; and multiple driver mutations have been shown to underlie myelodysplasia and acute myeloid leukaemia revealing the complexity of these diseases and the wide individual variation that is relevant to their treatment and prognosis. It seems likely that understanding the genetic complexity of hamatological malignancies will play an increasingly important role in providing personalised treatment for specific tumours.

These advances have been accompanied by the introduction of new, effective, targeted therapies, based on the knowledge that has been gained of the key signalling pathways on which the malignant cells depend for their proliferation and survival. For example, inhibitors of the B-cell receptor signalling pathway have proved life saving in patients with chronic lymphocytic leukaemia resistant to other therapies, and JAK2 inhibitors are extending survival and quality of life in patients with myelofibrosis. Other advances in therapy include many new monoclonal antibodies used for treating Hodgkin and non-Hodgkin lymphomas, and new immunomodulatory and proteasome inhibitory drugs that are increasing life expectancy in multiple myeloma. The more widespread use of orally active, direct inhibitors of coagulation and of the orally active iron chelating drugs are also having a major impact on patient care.

The seventh edition of *Postgraduate Haematology* reflects these exciting developments in the diagnosis and treatment of blood diseases, with revised text, new scientific diagrams and tables. Douglas Higgs, David Keeling and Atul Mehta have formed an Editorial team with the original Editor, Victor Hoffbrand, and many new and previous authors have contributed superb, up to date, well-illustrated, chapters. We thank most warmly Danny Catovsky, Edward Tuddenham and Tony Green for their major contribution as Editors of previous editions. We also thank Claire Bonnett, Rob Blundell and Tom Bates of Wiley Blackwell who have been responsible for the publishing process throughout the preparation of this edition and have been unstinting in their support, patience and professional expertise. Thanks also to Kathy Syplywczak who project managed this edition, and we are also grateful once again to Jane Fallows for her superb art work and scientific diagrams.

AVH, DRH, DK, ABM
London and Oxford

Preface to the first edition

In this book the authors combine an account of the physiological and biochemical basis of haematological processes with descriptions of the clinical and laboratory features and management of blood disorders. Within this framework, each author has dealt with the individual subjects as he or she thought appropriate. Because this book is intended to provide a foundation for the study of haematology and is not intended to be a reference book, it reflects, to some extent, the views of the individual authors rather than providing comprehensive detail and a full bibliography. For these the reader is referred to the selected reading given at the end of each chapter. It is hoped that the book will prove of particular value to students taking either the Primary or the Final Part of the examination for Membership of the Royal College of Pathologists and the Diplomas of Clinical Pathology. It should also prove useful to physicians wishing to gain special knowledge of haematology and to technicians taking the Advanced Diploma in Haematology of the Institute of Medical Laboratory Technology, or the Higher National Certificate in Medical Laboratory subjects.

We wish to acknowledge kind permission from the editors and publishers of the *British Journal of Haematology*, the *Journal of the Royal College of Physicians of London* and the *Quarterly Journal of Medicine* for permission to reproduce Figures 4.1, 4.5, 4.10, 4.11, 4.12, 9.4 and 9.10, also the publishers of *Progress in Haematology* for Figure 7.2, and many other publishers who, together with the authors, have been acknowledged in the text. We are particularly grateful to Professor JV Dacie for providing material which formed the basis of many of the original illustrations in Chapters 4–8. We are greatly indebted to Mrs T Charalambos, Mrs J Cope and Mrs D Haysome for secretarial assistance and to Mrs P Schilling and the Department of Medical Illustration for photomicrography, art work and general photography.

Finally, we are grateful for the invaluable help and forbearance we have received from Mr R Emery and William Heinemann Medical Books.

London, 1972
AVH
SML

CHAPTER 1

Stem cells and haemopoiesis

1

Emma de Pater and Elaine Dzierzak

Erasmus Stem Cell Institute, Erasmus Medical Centre, Rotterdam, Netherlands and University of Edinburgh, Centre for Inflammation Research, UK

Introduction

Haemopoietic stem cells (HSCs) are the foundation of the adult blood system and sustain the lifelong production of all blood lineages. These rare cells are generally defined by their ability to self-renew through a process of asymmetric cell division, the outcome of which is an HSC and a differentiating cell. In health, HSCs provide homeostatic maintenance of the system through their ability to differentiate and generate the hundreds of millions of erythrocytes and leucocytes needed each day. In trauma and physiological stress, HSCs ensure the replacement of the lost or damaged blood cells. The tight regulation of HSC self-renewal ensures the appropriate balance of blood cell production. Perturbation of this regulation and unchecked growth of HSCs and/or immature blood cells results in leukaemia. Over the last 50 years, great success has been achieved with bone marrow transplantation as a stem cell regenerative therapy. However, insufficient numbers of HSCs are still a major constraint in clinical applications. As the pivotal cells in this essential tissue, HSCs are the focus of intense research to: (1) further our understanding of their normal behaviour and the basis of their dysfunction in haemopoietic disease and leukaemia and (2) provide insights for new strategies for improved and patient-specific stem cell therapies. This chapter provides current and historical information on the organization of the adult haemopoietic cell differentiation hierarchy, the ontogeny of HSCs, the stromal microenvironment supporting these cells, and the molecular mechanisms involved in the regulation of HSCs.

Hierarchical organization and lineage relationships in the adult haemopoietic system

The haemopoietic system is the best-characterized cell lineage differentiation hierarchy and, as such, has set the paradigm for the growth and differentiation of tissue-specific stem cells. HSCs are defined by their high proliferative potential, ability to self-renew and potential to give rise to all haemopoietic lineages. HSCs produce immature progenitors that gradually and progressively, through a series of proliferation and differentiation events, become restricted in lineage differentiation potential. Such restricted progenitors produce the terminally differentiated functional blood cells.

The lineage relationships of the variety of cells within the adult haemopoietic hierarchy (Figure 1.1) are based on results of *in vivo* transplantation assays in irradiated/myeloablated recipient mice and many *in vitro* differentiation assays that became available following the identification of haemopoietic growth factors. These assays facilitated measurement of the maturational progression of stem cells and progenitors, at or near the branch points of lineage commitment. Clonal analyses, in the form of colony-forming unit (CFU) assays or single cell transplantation assays, were developed to define the lineage differentiation potential of the stem cell or progenitor, and to quantitate the number/frequency of such cells in the population as a whole. In general, the rarer a progenitor is and the greater its lineage differentiation potential, the closer it is in the hierarchy to the HSC. *In vitro* clonogenic assays measure the most immature

Postgraduate Haematology, Seventh Edition. Edited by A Victor Hoffbrand, Douglas R Higgs, David M Keeling and Atul B Mehta.

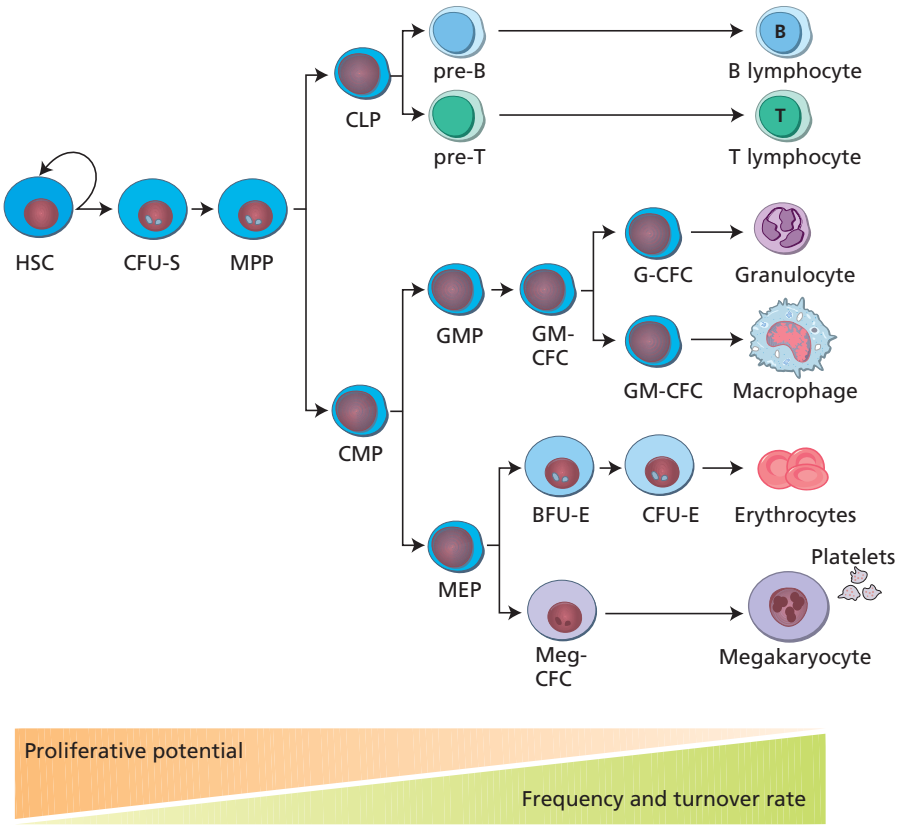


Figure 1.1 The adult haemopoietic hierarchy. Haemopoietic stem cells are at the foundation of the hierarchy. Through a series of progressive proliferation and differentiation steps the mature blood cell lineages are produced. Haemopoietic stem cells have the greatest proliferative and multilineage differentiation potential, while the mature blood cells are not proliferative and are lineage restricted. While large numbers of mature cells are found in the blood and turn over rapidly, the bone marrow contains long-lived quiescent haemopoietic stem cells at a very low frequency.

progenitor CFU-GEMM/Mix (granulocyte, erythroid, macrophage, megakaryocyte), bipotent progenitors CFU-GM (granulocyte, macrophage) and restricted progenitors CFU-M (macrophage), CFU-G (granulocyte), CFU-E (erythroid) and BFU-E (burst-forming unit-erythroid). While such *in vitro* clonogenic assays measure myeloid and erythroid potential, lymphoid potential is revealed only in fetal thymic organ cultures and stromal cell cocultures in which the appropriate microenvironment and growth factors are present. Long-term culture assays (6–8 week duration), such as the cobblestone-area-forming cell (CAFC) and the long-term culture-initiating cell (LTC-IC) assays, reveal the most immature of haemopoietic progenitors. Currently, the major hurdle in studies and clinical applications of HSCs is the fact that HSCs cannot be expanded and are poorly maintained in culture. The only way to detect a *bona fide* HSC is *in vivo*.

In vivo, the heterogeneity of the bone marrow population of immature progenitors and HSCs is reflected in the time periods at which different clones contribute to haemopoiesis. Short-term *in vivo* repopulating haemopoietic progenitor cells such as CFU-S (spleen) give rise to macroscopic erythro-myeloid colonies on the spleen within 14 days of injection. *Bona fide* HSCs give rise to the long-term high-level engraftment of all haemopoietic lineages. Serial transplantations reveal the ability of the long-term repopulating HSCs to self-renew. The clonal nature of engraftment and the multilineage potential of HSCs has been demonstrated through radiation, retroviral and barcode marking of bone marrow cells. Such studies suggest that, at steady state, several HSC clones contribute to the haemopoietic system at any one time. Further analyses of bone marrow HSCs show that this compartment consists of a limited number of distinct HSC subsets, each with predictable behaviours, as described by their repopulation kinetics in myeloablated adult recipients. In general, the bone marrow haemopoietic cell compartment, as measured by *in vitro* clonogenic assays and *in vivo* transplantation assays, shows a progression along the adult differentiation hierarchy from HSCs to progenitors and fully functional blood cells with decreased multipotency and proliferative potential.

The use of flow cytometry to enrich for HSCs and the various progenitors in adult bone marrow has been instrumental in refining precursor–progeny relationships in the adult haemopoietic hierarchy. HSCs are characteristically small 'blast' cells, with a relatively low forward and side light scatter and low metabolic activity. Both mouse and human HSCs are negative for expression of mature haemopoietic lineage cell-surface markers, such as those found on B lymphoid cells (CD19, B220), T lymphoid cells (CD4, CD8, CD3), macrophages (CD15,

Mac-1) and granulocytes (Gr-1). Positive selection for mouse HSCs relies on expression of Sca-1, c-kit, endoglin and CD150 markers and for human HSCs on expression of CD34, c-kit, IL-6R, Thy-1 and CD45RA markers. Similarly, cell types at lineage branch points have been identified, including the CMP (common myeloid progenitor), CLP (common lymphoid progenitor) and GMP (granulocyte macrophage progenitor). Recently, using the Flt3 receptor tyrosine kinase surface marker along with many other well-studied markers, the LMPP (lymphoid primed multipotent progenitor) has been identified within the lineage negative, Sca-1 positive, c-kit positive (LSK) enriched fraction of HSCs. These cells have granulocyte/macrophage, B lymphoid and T lymphoid potential, but little or no megakaryocyte/erythroid potential. This suggests that the first lineage differentiation event is not a strict separation into common lymphoid and myeloid pathways. While these cell-surface marker changes and functional restriction events are represented by discrete cells in the working model of the haemopoietic hierarchy as depicted in textbooks and Figure 1.1, it is most likely that there is a continuum of cells between these landmarks and/or alternative differentiation paths. The currently identified progenitor cells in the hierarchy represent the cells present at stable and detectable frequencies and for which we currently have markers and functional assays. As more cell-surface markers are identified and the sensitivity of detection is increased, additional intermediate cell subsets are likely to be identified. Together with single cell transcriptomic approaches, it may be possible to predict the molecular events needed for the HSC state and the differentiation of the entire haemopoietic system.

Sites of adult haemopoiesis

Bone marrow, spleen, thymus and lymph nodes are the haemopoietic sites in the adult, and each tissue plays a special role in supporting the growth and differentiation of particular haemopoietic cell lineages and subsets. Equally important is the blood itself, which is a mobile haemopoietic tissue, with mature blood cells travelling through the circulation to function in all parts of the body. Not only do the terminally differentiated cells, such as erythrocytes and lymphocytes, move by means of the circulation, but HSCs (at low frequency) also migrate through the circulation from the bone marrow to other haemopoietic tissues. HSCs are mostly concentrated in the bone marrow and are found in the endosteal and vascular niches (Figure 1.2). HSCs can be induced to circulate by administration of granulocyte colony-stimulating factor (G-CSF). Recent improvements in confocal microscopy have allowed the visualization of the migration of circulating HSCs to the bone marrow endosteal niche by time-lapse imaging in the mouse.

The estimated frequency of HSCs is 1 per 10^4–10^5 mouse bone marrow cells and 1 per 20×10^6 human bone marrow cells.

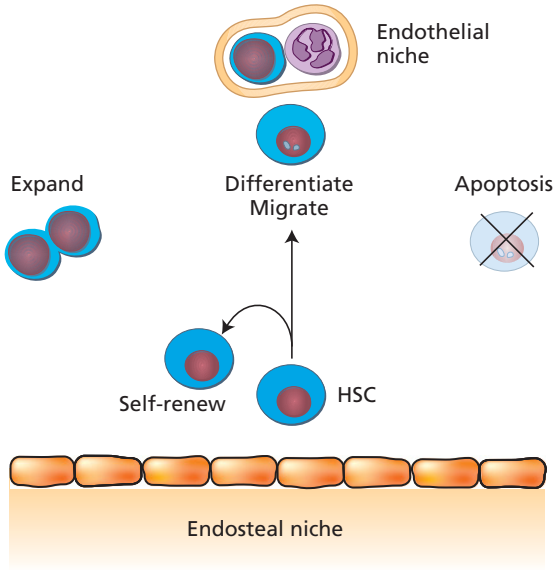


Figure 1.2 The bone marrow haemopoietic niches. Haemopoietic stem cells are found in the endosteal and endothelial niches of the bone marrow. These niches support the maintenance, self-renewal, expansion, differentiation, migration and survival of haemopoietic stem cells through local growth factor production and cell–cell interactions.

HSCs are also found in the mouse spleen at approximately a 10-fold lower frequency and in the circulating blood at a 100-fold lower frequency. The capacity for HSCs to migrate and also be retained in bone marrow supportive niches is of relevance to clinical transplantation therapies. HSCs injected intravenously in such therapies must find their way to the bone marrow for survival and effective haemopoietic engraftment. For example, stromal-derived factor (SDF)-1 and its receptor CXCR4 (expressed on HSCs) are implicated in the movement of HSCs and the retention of HSCs in the bone marrow. Indeed, HSC mobilization can be induced through AMD3100, an antagonist of SDF-1, and by the administration of G-CSF. Mobilization strategies with G-CSF are used routinely to stimulate bone marrow HSCs to enter the circulation, allowing ease of collection in the blood rather than through bone marrow biopsy.

Development of HSCs

Waves of haemopoietic generation in embryonic development

Until the mid-1960s it was thought that blood cells were intrinsically generated in tissues such as the liver, spleen, bone

marrow and thymus. Survival studies in which cells from un-irradiated tissues were injected into lethally irradiated mice showed that it was the bone marrow that contains the potent cells responsible for rescue from haemopoietic failure. Later, through clonal marking studies, it was demonstrated that the bone marrow harbours HSCs during the adult stages of life. But where, when and how are HSCs generated during ontogeny? In the 1970s, examination of mouse embryo tissues suggested that adult haemopoietic cells are generated in the yolk sac, migrate and colonize initially the fetal liver and subsequently the bone marrow, where they reside throughout adult life. However, studies in non-mammalian vertebrate models (avian and amphibian) demonstrated that the aorta region in the body of the embryo generates the long-lived adult blood system, while the yolk sac (or equivalent tissue) produces the transient embryonic haemopoietic system. In agreement with these studies, the aorta–gonad–mesonephros (AGM) region of mammalian embryos was later found to generate the first HSCs of the permanent adult blood system.

The development of the haemopoietic system is complex. As a growing organism, the embryo itself needs rapid haemopoiesis to allow it to thrive before the adult system is generated. Thus, a simple transient haemopoietic system is generated at early embryonic stages to rapidly produce primitive erythroid and myeloid cells. In the yolk sac, both haemopoietic and endothelial cells are simultaneously generated from a common mesodermal precursor cell, the haemangioblast (Figure 1.3). Thereafter, other haemopoietic progenitor and differentiated cell types are generated in both the yolk sac and the intraembryonic AGM region to create an intermediate haemopoietic system. These progenitors and differentiated cell types arise from a specialized population of endothelial cells that have haemogenic potential (haemogenic endothelial cells). At both these early times in ontogeny, the mouse embryo contains no HSCs. Hence, in the absence of HSCs, the embryo generates a haemopoietic system that is short-lived and lacks the important qualitative characteristics (longevity and self-renewability) of the adult haemopoietic system. However, some early yolk sac progenitors provide long-lived tissue resident macrophages, such as the glial cells in the brain. The independent and distinct waves of haemopoiesis that supply the embryo and adult are likely derived from different subsets of mesodermal precursor cells (Figure 1.4).

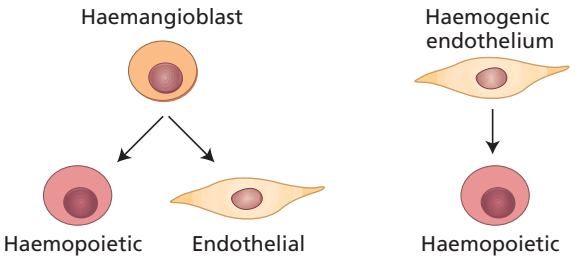


Figure 1.3 Precursors to haemopoietic cells in embryonic stages. The mesodermal precursor to haemopoietic and endothelial lineages at early stages of development is the haemangioblast. Later, haemogenic endothelial cells are the precursors to haemopoietic stem cells and progenitor cells. Both precursors appear to exist during a short window of developmental time.

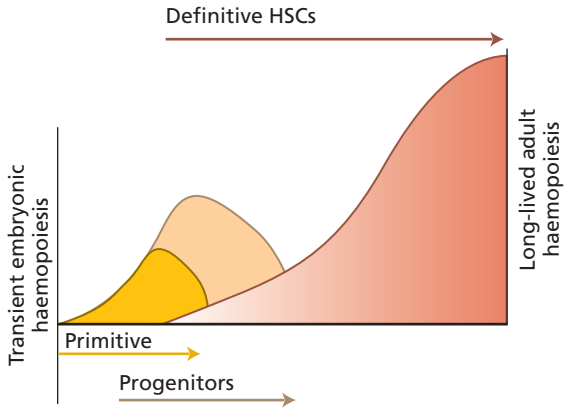


Figure 1.4 Waves of haemopoietic cell emergence during embryonic stages. The earliest haemopoietic cells are produced during the first wave of haemopoietic fate determination. The onset of this wave occurs in the yolk sac blood islands and produces transient primitive erythroid cells. This wave continues with the production of erythroid-myeloid progenitors in the absence of *bona fide* haemopoietic stem cells. True long-lived self-renewing definitive haemopoietic stem cells (adult repopulating stem cells) are generated in the second wave of haemopoietic cell emergence in the AGM region. In this wave, haemogenic endothelial cells bud into the aortic lumen as these cells take on haemopoietic stem cell fate.

The adult system has its foundation in a cohort of initiating HSCs. The first adult HSCs are autonomously generated in the mouse AGM at E10.5 and in the human AGM beginning at week 4 of gestation. Recently, the process of HSC generation has been visualized in real time in the mouse embryo. This remarkable demonstration confirms that HSCs are derived via a transdifferentiation event in which specialized endothelial cells lining the aorta bud into the lumen to form round cells with HSC fate (Figure 1.5), and shows that haemopoietic development is conserved between mammalian and non-mammalian species. The emerging mouse aortic HSCs are characterized by the loss of cell-surface markers for endothelium, such as Flk-1 and VE-cadherin, and the gain of expression of haemopoietic markers CD41 and CD45 and HSC markers Sca1, c-kit and endoglin. The emerging aortic HSCs are as functionally potent as bone marrow HSCs, since these sorted cells can form a complete long-term haemopoietic system and self-renewing HSCs after transplantation into irradiated adult recipient mice.

Lineage tracing experiments in the mouse embryo have indicated that the adult haemopoietic system is generated during a

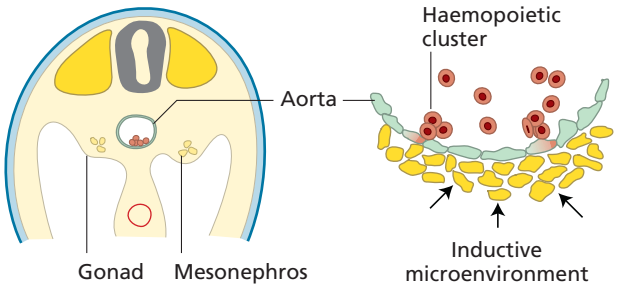


Figure 1.5 Schematic diagram of the aorta–gonad–mesonephros (AGM) region and haemopoietic cell clusters emerging from the dorsal aorta. The haemopoietic stem-cell-inductive microenvironment is localized in the ventral aspect of the aorta. Tissues ventral to the AGM, such as the gut and mesenchyme provide HSC-inducing signals, whereas dorsal tissues such as the notochord and the neural tube suppress HSC induction.

short window of development, spanning E9–E12. Using Cre-lox recombination (temporally and cell-lineage controlled) to mark endothelial cells in the mid-gestation embryo, it was found that almost all the blood cells in the circulation and haemopoietic tissues of the adult mice were derived from VE-cadherin expressing cells. Moreover, these cells require the Runx1 transcription factor, as demonstrated by Runx1 conditional deletion in this mouse model. Other lineage tracing experiments marking the earliest cells expressing the *Runx1* and *SCL* transcription factor genes, showed that the progeny contributed to the bone marrow cells in the adult. Thus, the progeny of haemogenic endothelial cells in the major vasculature of the embryo contribute to a cohort of adult bone marrow HSCs that form the foundation of haemopoiesis throughout adult life.

Embryonic haemopoietic sites and haemopoietic migration

The AGM and yolk sac are not the only sites where haemopoietic cells are found in the early conceptus. The placenta is a highly haemopoietic tissue and much like the early-stage yolk sac, the mouse placenta can produce erythro-myeloid progenitors. Embryos deficient for the Ncx1 gene lack a heartbeat and circulation, and thus were used to study the origins of early haemopoietic progenitors. Ncx1 deficient embryos were shown to contain erythro-myeloid progenitors in the yolk sac and placenta, demonstrating that these haemopoietic progenitors are generated by these tissues. Unfortunately, the embryos die before the onset of HSC generation at mid-gestation, precluding analysis of HSC production in the yolk sac and placenta. In normal embryos where the circulation is established between the embryo body and the extraembryonic tissues at E8.25, HSCs are detected in the placenta and yolk sac only beginning at E11, subsequent to the first HSC generation in the AGM at E10.5. A recent study revealed the presence and generation of HSCs in the E10.5/E11 head vasculature through lineage marking. It remains uncertain whether the placenta (or the yolk sac) can generate HSCs *de novo* since there is no method at present by which cells can be uniquely marked in these developing tissues. Nonetheless, quantitative studies in which HSC numbers in each of these tissues was determined suggest that the AGM cannot generate all the HSCs that are found in the fetal liver (a tissue that harbours haemopoietic cells but does not generate them) and later in the adult bone marrow (Figure 1.6). Since the placenta at mid-gestation contains an abundance of HSCs, it is possible that this highly vascularized tissue generates HSCs from haemogenic endothelium and/or that the placenta is a highly supportive and proliferative microenvironment for AGM-derived HSCs.

The development of the haemopoietic system in the human conceptus closely parallels that in the mouse conceptus. Like the mouse placenta, the developing human placenta contains HSCs. Already at week 6 of gestation HSCs can be detected, as analysed by *in vivo* xenotransplantation into immunodeficient mice; also, haemopoietic progenitors are found at these early stages. Phenotypic characterization shows that HSCs and progenitors are in both the CD34-positive and CD34-negative fractions at week 6 of gestation and are exclusively in the CD34-positive fraction by week 19. These cells are in close association with the placental vasculature. The placenta may be considered a source of haemopoietic progenitors and HSCs in addition to

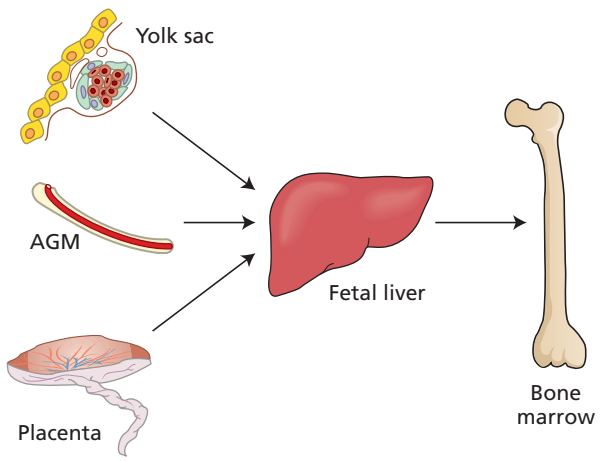


Figure 1.6 Haemopoietic sites during development. The first haemopoietic stem cells arise in the AGM region. Other haemopoietic cells and progenitors are generated in the yolk sac and placenta. It is as yet undetermined whether the yolk sac and placenta can generate haemopoietic stem cells. Haemopoietic cells generated in these three tissues migrate and colonize the fetal liver. Subsequently, the long-lived haemopoietic cells (primarily the haemopoietic stem cells) migrate and colonize the bone marrow, where they reside in the adult stages of life.

umbilical cord blood for preclinical studies and potential clinical therapies.

HSC quiescence, proliferation and ageing

Somatic stem cells undergo lifelong self-renewal and possess the potential to produce the differentiated cells of the tissue. HSCs are considered to be relatively dormant stem cells, dividing rather infrequently. They are enriched in the quiescent fraction of adult bone marrow and are resistant to 5-fluorouracil (which is an antimetabolite drug that results in the death of rapidly dividing cells). Recent studies in mice using a label-retaining method for analysis of cycling versus non-cycling cells show that under homeostatic conditions, dormant HSCs cycle only once every 21 weeks. The adult mouse possesses approximately 600 dormant LSK $CD150^{+}CD48^{-}CD34^{-}$ HSCs. Interestingly, 38% of HSCs in G0, considered to be the dormant HSCs, can be activated by myelo/lymphodepletion during injury, 5-fluorouracil or G-CSF administration, and can return to the dormant state after the re-establishment of homeostasis.

The maintenance of HSC dormancy is thought to be an important strategy for preventing stem cell exhaustion during adult life. Serial transplantations in the mouse demonstrate that HSC self-renewal is limited to about six rounds of transplantation and that there is a progressive decrease in the ability of the transplanted stem cells to repopulate/self-renew. It has been proposed that accumulating DNA mutations and loss of telomere repeats adversely affect HSC function. Studies of chromosome shortening in human HSCs suggest that self-replication is limited to about 50 cell divisions. Recently, it was found that HSC characteristics are changed in aged mice. Comparison of various inbred mouse strains has shown that the rate of haemopoietic cell cycling is inversely correlated with their mean lifespan. The decrease in HSC quality was due to cell-intrinsic genetic or epigenetic factors. Causative genes were identified by transcriptional profiling comparisons between the HSCs of the different strains. Of particular interest are chromatin modifiers involved in prevention of HSC exhaustion through maintenance of a stem-cell-specific transcriptional programme. Changes in chromatin structure associated with high HSC turnover would result in stem cell senescence (which is thought to protect stem cells from malignant transformation by oncogenic events).

Transplantations of single HSCs from both the fetal liver and adult bone marrow have revealed HSC heterogeneity in lineage differentiation output related to developmental stage and aging: some HSCs give a balanced lineage differentiation output of myeloid and lymphoid cells, whereas others yield a predominant lymphoid or myeloid cell lineage output. During fetal stages, HSCs with a balanced lineage output are at a higher frequency than in adult BM. During aging the frequency of BM HSCs with a predominant myeloid output increases as compared to the frequency of HSCs with a balanced lineage output or predominant lymphoid output. HSCs with a predominant myeloid output can also be found in prenatal life. Thus, the myeloid type HSC is not unique to aging – it is the prevalence to maintain these HSCs that is.

It is unclear why such heterogeneity in HSCs exists. HSCs generally do not undergo apoptosis in response to DNA damage and have adopted several mechanisms to preserve stemness rather than self-renewal, to reduce DNA damage and/or to prevent inappropriate differentiation leading to loss of HSCs. Both developmental and stem cell protective mechanisms may assist in providing maximum HSC fitness during reproductive life, providing an evolutionary benefit. Altered gene expression, however, may drive lymphoid differentiation, deplete lymphoid-biased HSCs and thus contribute to the relative predominance of myeloid-biased HSCs.

Haemopoietic-supportive microenvironments

Adult bone marrow microenvironment

Most tissue-specific stem cells are maintained in special microenvironments/niches that support long-term cell growth and self-renewal. To provide the continuous production of human blood over the many decades of adulthood, HSCs are maintained in the specialized haemopoietic-supportive niches of the adult bone marrow (Figure 1.2). The importance of the bone marrow haemopoietic niche and the interactions between supportive cells and HSCs was first demonstrated in mice. In transplantation studies of anaemic mouse strains naturally deficient in the c-kit receptor tyrosine kinase (W mice) or kit-ligand (KL; Steel mice) it was revealed that bone marrow from W mutant mice could not repopulate the haemopoietic system of wild-type irradiated recipient mice, whereas bone marrow from Steel mutant mice could. In contrast, W mutant mice could be repopulated by wild-type donor bone marrow cells, whereas Steel recipients were defective for repopulation by wild-type donor cells. It was proposed that a receptor–ligand interaction was involved to support HSCs within the bone marrow microenvironment. It was subsequently shown that HSCs express c-kit and bone marrow stromal cells express KL. The development of *ex vivo* culture systems to study this complex microenvironment allowed further dissection of the cellular and molecular aspects of the bone marrow microenvironment. These studies were aided by the isolation of mesenchymal stromal cells.

Stromal cell lines have been derived from the adult mouse bone marrow and fetal tissues. These are generally of mesenchymal lineage, as determined by cell-surface marker expression and their osteogenic and adipogenic potentials. Although widely heterogeneous in their ability to support haemopoiesis, some stromal lines (MS5 and AFT024, for example) have been shown

to support the growth and/or maintenance of HSCs in cocultures for long periods. Moreover, they have been instrumental in further characterization of these haemopoietic-supportive niches. Comparative transcriptional profiling and database analysis of HSC supportive and non-supportive stromal cell lines has revealed a complex genetic programme involving a wide variety of known molecules and molecules whose function in haemopoiesis is currently under investigation.

The *in vivo* bone marrow microenvironment is very complex, containing osteoblastic niches and vascular niches localized within the trabecular regions of the long bones. HSCs are maintained in close association with the so-called 'stromal cells' of the niches (osteoblasts and vascular endothelial cells). Along with KL, some of the key molecular regulators within the bone marrow niches include N-cadherin and CD150, and signalling pathway molecules SDF1, Notch, Wnt, Hedgehog, Tie2/angiopoietin, transforming growth factor (TGF), bone morphogenetic protein (BMP) and fibroblast growth factor (FGF). These regulators are implicated in a variety of cellular processes, such as HSC maintenance, differentiation, self-renewal and homing. Indeed, live tracking of haemopoietic progenitor/stem cells in the mouse model has shown the homing ability of these cells to bone marrow niches, and mouse models as well as *in vitro* culture systems are beginning to reveal the specific molecular mechanisms involved.

Microenvironments important for haemopoietic development in the conceptus

Prior to the establishment of an adult haemopoietic-supportive microenvironment, the embryo contains several haemopoietic microenvironments that are supportive and/or inductive. The extraembryonic yolk sac and placenta, and the intraembryonic AGM generate haemopoietic progenitor cells, whereas the AGM region generates the first adult repopulating HSCs (Figure 1.6). Little is known about the differences between the microenvironments of the embryonic haemopoietic tissues. However, the AGM microenvironment is the most-well characterized due to the simplicity of its structure, with the aorta at the midline of the embryo and the laterally located gonads and mesonephroi (Figure 1.5). The avian AGM microenvironment contains different types of mesenchymal cells and a population of aorta-associated stem cells called 'mesoangioblasts' that contribute to cartilage, bone and muscle tissues, and also to blood. In the mouse AGM region, cells more typical of mesenchymal stromal cells have been found. Interestingly, mapping and frequency analysis in the mouse conceptus show that mesenchymal progenitors, with the potential to differentiate into cells of the osteogenic, adipogenic and/or chondrogenic lineages, reside in most of the sites harbouring haemopoietic cells, suggesting that both the HSC and mesenchymal stromal cell microenvironment develop in parallel. Phenotypic characterization of haemopoietic-supportive AGM stromal lines places them in the vascular smooth muscle cell (VSMC) hierarchy, in between a mesenchymal stem cell and a VSMC. Other niche cells include cells of the nervous system and endothelial cells.

Stromal cell lines established from the AGM region, placenta and fetal liver can support immature haemopoietic progenitors and HSCs and are more supportive as compared to adult bone marrow cell lines. Some can also support the haemopoietic differentiation of embryonic stem (ES) cells. Such stromal cell lines in a re-aggregate culture system have been able to support the differentiation of cells with a haemogenic endothelial phenotype ($VE\text{-}Cad^{+}CD45^{-}CD41^{+}$ cells from mouse embryos before the onset of HSC generation) into long-term repopulating HSCs. This highlights that in an *ex vivo* controlled environment, cells with a potential to become HSCs, can be influenced to do so by other cells. However, it is still unknown whether the inductive factors in the stromal/re-aggregate cultures are the same factors produced in the *in vivo* physiologic HSC-inductive microenvironment. It is likely that HSC induction is a complex process requiring a variety of spatial and temporal cues emanating from several cell types in the niches of the embryo.

Within the normal physiology of the embryo, the AGM lies between the ventral tissue that includes mesenchyme and the endoderm-derived gut, and the dorsal tissue including the notochord and the ectoderm-derived neural tube (Figure 1.5). Mouse AGM explant culture experiments have shown that dorsal tissues/signals repress AGM HSC activity and ventral tissues/signals enhance HSC emergence. In both mouse and human AGM regions, cells expressing HSC markers are closely adherent to the vascular endothelium on the ventral aspect of the aorta. In the mouse, at precisely E10.5, single endothelial cells bud into the lumen as they take on HSC identity (Figure 1.5). Importantly, HSC activity, as determined by functional transplantation assays, is localized exclusively to the ventral aspect of the mouse mid-gestation aorta. Thus there is a strong positive ventral positional influence on HSC generation in the AGM, and morphogens and local signals emanating from the ventral endodermal tissues may be responsible for establishing the HSC-inductive microenvironment.

Haemopoietic transcription factors required for HSC generation such as Gata2 and Runx1 are expressed in cells of the ventral aortic clusters and endothelium. Deletion of *Gata2* and *Runx1* genes in mice leads to mid-gestation embryonic lethality, with complete absence of adult haemopoiesis (although embryonic haemopoiesis occurs), thus demonstrating that these two pivotal transcription factors promote the HSC genetic programme. Zebrafish and frog embryos have been useful models for dissecting the cascade of upstream events that lead to HSC induction. Developmental growth factor signalling pathways, such as the BMP, Hedgehog and Notch pathways, converge to activate expression of the two transcription factors in aortic haemopoietic cells and promote the HSC programme. In both the mouse and human embryo, BMP4 is expressed in the mesenchyme underlying the ventral aspect of the aorta at the time

of haemopoietic cluster formation. Culture experiments have demonstrated the positive influence of BMP4 exposure to mouse and human HSC-containing cell populations. BMP4 has been found to act directly on HSCs in the AGM and, in addition, may stimulate the microenvironment to produce HSC effectors. Similarly, Hedgehog signalling regulates HSCs in the AGM region, likely in an indirect way through VEGF. Other ventrally localized HSC regulators include the Notch signalling molecules, as well as Wnt3a and interleukin (IL)-1.

High-throughput chemical screens offer a means of identifying molecules involved in HSC growth, maintenance and expansion. Through such a screen in zebrafish embryos, prostaglandin E2 (PGE2) was recently identified as a regulator of HSC number. When tested in the murine transplantation model, *ex vivo* exposure of bone marrow cells to PGE2 enhanced short-term repopulation by haemopoietic progenitors and increased the frequency of long-term repopulating bone marrow HSCs. PGE2 modifies the Wnt signalling pathway, which in turn is thought to control HSC self-renewal and bone marrow repopulation. Extracellular environmental cues, such as blood flow, also affect HSC generation. A zebrafish chemical screen identified modulators of blood flow such as nitric oxide synthetase (NOS). Inhibition or deficiency of NOS reduces murine bone marrow HSC number/function. Thus, together with general physiological cues, such as the haemopoietic growth factors, KL, IL-3, Flt3 and thrombopoietin, chemical modulators and developmental regulators may be useful for expansion of HSC number and enhancement of HSC function for therapeutic purposes.

Haemopoietic regenerative and replacement therapies

Stem cell transplantation

For over 50 years, HSC transplantation has been the most successful and significant clinical cell regenerative therapy (see Chapter 35). Initially, whole bone marrow was the source of cells used in clinical transplantation, but through experience and much research new and/or improved sources of transplantable HSCs were found. These now include the $CD34^+CD38^-$ fraction of adult bone marrow, mobilized peripheral blood HSCs and the $CD34^+CD38^-$ fraction of umbilical cord blood. The cumulative data from the large number of patients worldwide receiving a bone marrow transplant provide valuable information on the success of autologous versus allogeneic transplantation, the number of human leucocyte antigen (HLA) differences that are tolerated by the recipient, the incidence of graft-versus-host disease (GVHD), and the unexpected and advantageous graft-versus-leukaemia effect.

Interestingly, umbilical cord blood (UCB) appears to offer a beneficial source of HSCs for several reasons: UCB HSCs are young, being harvested at the neonatal stage of development, thus circumventing concerns about the ageing of HSCs; UCB transplantation induces less frequent and less severe GVHD, since UCB contains many fewer activated T cells than adult bone marrow; also, UCB HSCs are highly proliferative. However, only relatively small numbers of cells are harvested (approximately 10-fold lower than those in adult bone marrow) and this limits their use to paediatric patients, unless multiple UCB units are transplanted. Despite increases in the number of UCB units (400 000) stored in cord blood banks (>50) around the world (catalogued and recorded by EUROCORD and other coordinating efforts) and HLA donor-cell selection for rare haplotypes, the supply of HSCs is still limited.

Gene therapy and gene editing for haemopoietic disease

Monogenic disorders of the blood are the first targets of gene therapy approaches. To effect a cure for a haematologic disease in which a single gene or regulatory element is mutated, a viral vector containing a normal copy of the gene is used to introduce and express the gene in HSCs. Gene therapy for β-haemoglobinopathies, such as β-thalassaemia and sickle cell disease, were among the first proposed and tested in mouse models. Primary immunodeficiencies (PID) are also monogenic disorders and result in the absence of (parts of) the innate and adaptive immune system. Patients can be cured with allogeneic HLA matched (related) HSC transplantation. However, donor availability is limited. For patients without an allogeneic donor, gene therapy of their own bone marrow HSCs and subsequent autologous transplantation is the only option for curative treatment. Lentiviral vector infection offers an efficient mode of delivery of a functional copy of the mutated gene into the genome of the patient's own HSCs used for transplantation. In initial gene therapy trials of immunodeficient patients, lenti-viral vector insertions in the genome of some transplanted HSCs resulted in activation of oncogenes, the selective growth of these HSC clones and the onset of leukaemia. More recent trials have incorporated a safety feature in the lentivirus that reduces (but has not completely eliminated) the unwanted activation of oncogenes in the case of viral insertion. In 2010 a new gene therapy clinical trial was initiated for Wiscott–Aldrich (WAS) patients who suffer from thrombocytopenia, eczema, recurrent infections, autoimmune disorders and high susceptibility to develop tumours. To date, all patients are alive and show significant increase in platelets and T cells, although long-term follow-up is required. Similar results have been obtained for adenosine deaminase-deficient severe combined immunodeficiency (ADA SCID) where 40 patients have been treated since 2000 in Italy, the UK and the US without any reports of malignant occurrences. X-linked SCID and chronic granulomatous disease (CGD) patients have also undergone gene therapy treatment,

albeit with less success. The treatment for CGD was impaired since the earliest trials did not make use of myeloablation to enhance chimerism of the gene-manipulated graft. Gene therapy trials for β-thalassaemic patients are ongoing and encouraging, but for successful treatment, higher levels of gene expression and HSC chimerism will be needed.

Clinical trials with gene therapy are promising. However, lentiviral vector insertions that may result in malignant clone outgrowth remain a risk. New gene editing techniques offer new hope for gene correction directly within the gene of concern. Gene editing makes use of endonucleases (zinc finger nucleases, TALENs or CRISPR/Cas9) to target a specific genomic site and repair the mutated gene or insert a functional gene under the control of its own promoter. This method leaves no extra genetic modulation. At present this approach requires prolonged cell culture and a selection step for the corrected HSCs, and thus requires further research developments in HSC growth and expansion before it will be clinically useful. Gene therapy for genetic disorders of coagulation proteins is discussed in Chapter 38.

New sources of HSCs for transplantation

The ability to expand HSCs *ex vivo* is a theoretically practical and attractive means to obtain an accessible and limitless source of HSCs for transplantation therapies. Unfortunately, despite many years of research using different culture systems and combinations of haemopoietic growth factors and proliferation-stimulating agents, *ex vivo* expansion of HSCs has not been achieved. However, HSC developmental studies have begun to provide new insights into the processes directing the generation and growth of HSCs. If cells such as the haemogenic endothelial cells of the embryonic aorta are present in the adult vasculature or could be obtained from ES/iPS cells, they could provide a novel source of inducible HSC precursors, particularly if they can be sustained and expanded to large numbers in culture.

Embryonic stem cells and induced pluripotent stem cells

Pluripotent embryonic stem (ES) and induced pluripotent stem (iPS) cells have been used to generate differentiated cells in many tissue systems, including the haemopoietic system. Developmental studies revealing the temporally and spatially limited production of HSCs in the embryonic vasculature, the components of the specific microenvironment, and the knowledge of the molecular programme of endothelial to haemopoietic cell transition have yielded insight into how HSCs may be induced and/or expanded without undergoing differentiation in such cultures. Haemopoietic-directed differentiation of human iPS cells towards endothelial cells, haemogenic endothelial cells and HSCs would be a potentially attractive alternative to conventional sources of HSCs. Furthermore, such a cell culture system would make possible the use of novel gene editing approaches for monogenetic disorders. These gene correction approaches could be used in combination with patient-derived iPS cells. Studies using mouse and human ES cells have optimized culture conditions to include temporally changing combinations of growth factors (ActivinA, BMP4, VEGF, etc.) and signalling pathway antagonists to control differentiation to the mesodermal, vascular and thereafter the haemopoietic lineage. The ES-cell-derived haemopoietic cells arise from haemangioblasts and/or primitive endothelial-like cells that express PECAM-1, FLK-1 (KDR) and VE-cadherin, and are thought to represent the types of precursors, progenitors and differentiated cells found normally in the yolk sac. These results have strengthened the idea that ES cell differentiation proceeds via a 'haemogenic endothelial' differentiation step, before definitive haemopoietic cells can be produced and require activation of the Wnt-β-catenin pathway (Figure 1.7).

An alternative approach to produce HSCs *ex vivo* has recently emerged. This molecular reprogramming approach aims to reprogramme non-haemopoietic or differentiated haemopoietic cells directly to HSCs, without going through a pluripotent stem cell state. Such induced HSCs, or iHSCs would be generated through reprogramming directed by transcription factors pivotal to HSC generation and/or growth. Several laboratories have been able to generate haemopoietic progenitors or stem cells from more differentiated cells using four to eight different haemopoietic transcription factors previously identified from HSC transcriptome databases. In one case, mouse endothelial-like precursor cells (Sca1+, Prominin 1+ and expressing a human CD34 reporter) have been converted into haemopoietic progenitors using the factors Gata2, Gfi1b, cFos and Etv6. A human myeloid precursor (CD34+CD45+) cell has been converted into a haemopoietic progenitor cell (CD34+CD38−) using HOXA9, SOX4, RORA, and MYB; however, neither study was able to generate long-term repopulating HSCs. A study converting mouse committed B cell progenitors using a mix of eight transcription factors (Run1t1, Hlf, Lmo2, Prdm5, Pbx1, Zfp37, Myc-n and Meis1) has resulted in long-term repopulating HSCs. In this method, B cell progenitors are transduced with the eight factors, and immediately transplanted into irradiated recipients. In this way the native bone marrow niche preserves the new iHSCs and allows them to be maintained and function in the physiologic context of the recipient (Figure 1.7). Whereas this study demonstrates that transcription factor transduction of haemopoietic cells can yield HSCs, this approach is limited in applications for research or therapy. The fact that each of these studies uses a completely different panel of transcription factors to make induced HSCs or HPCs indicates that there may be more than one way to reprogramme cells to the haemopoietic lineage and that a further understanding of HSC biology is required.

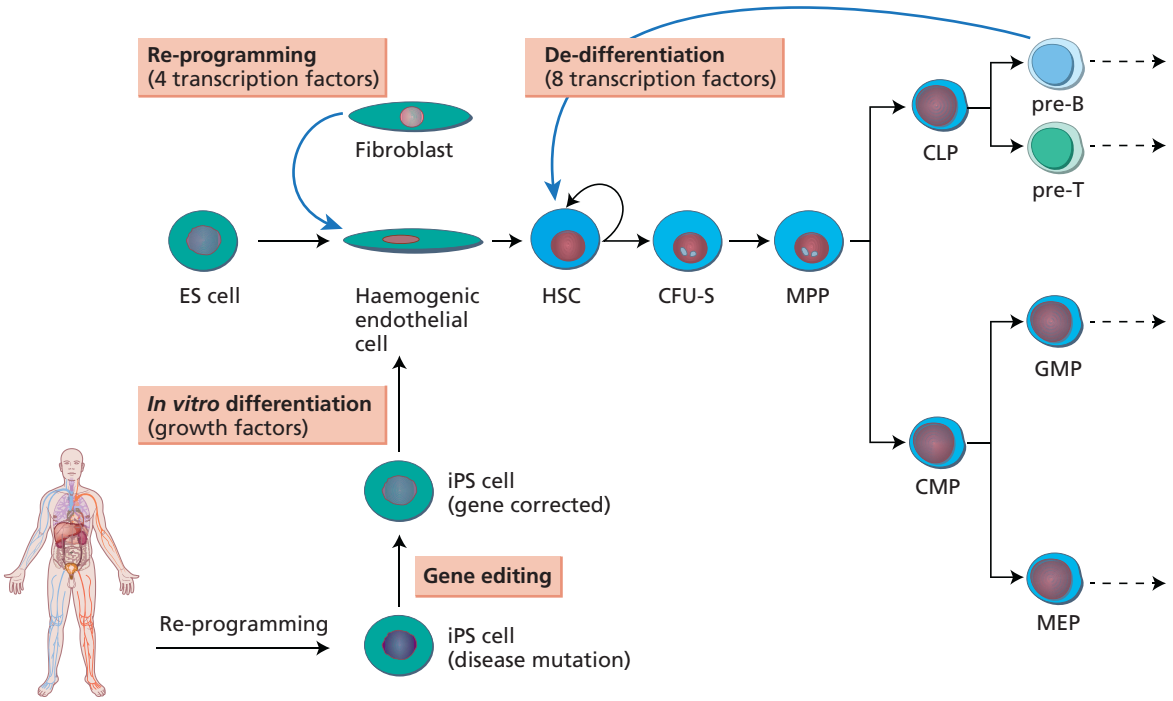


Figure 1.7 Several experimental approaches to generate HSCs *de novo*. De-differentiation of pre-B cells with eight transcription factors and immediate transplantation *in vivo*, allows for the production of multilineage, self-renewing HSCs due to the presence of functional HSC niches that are as yet not attainable in *in vitro* cultures. To date, direct reprogramming of B cell progenitors and immediate transplantation into irradiated mouse recipients has been the only study successful in generating HSCs. Reprogramming with four pivotal haemopoietic transcription factors has yielded haemogenic endothelial cells and haemopoietic progenitors, but not HSCs. *In vitro* haemopoietic differentiation of ES and iPS cells relies on the addition of developmental and haemopoietic growth factors to induce the progressive differentiation of these pluripotent cells to mesoderm, endothelial, haemogenic endothelial and haemopoietic fates. As this culture system improves, it may be possible to make iPS cells from patients with monogenic disease and correct the gene mutation by gene editing. These cells may then be differentiated to HSC fate and used for clinical treatment.

Selected bibliography

Boisset JC, Cappellen G, Andrieu C, Galjart N, Dzierzak E, Robin C (2010) *In vivo* imaging of hematopoietic stem cells emergence. *Nature* **464**: 116–20.

Dzierzak E, Speck NA (2008) Of lineage and legacy: the development of mammalian haematopoietic stem cells. *Nature Immunology* **9**: 129–36.

Gluckman E, Rocha V (2009) Cord blood transplantation: state of the art. *Haematologica* **94**: 451–4.

Morrison SJ, Scadden DT (2014) The bone marrow niche for haematopoietic stem cells. *Nature* **505**: 327–34.

Nienhuis AW (2013) Development of gene therapy for blood disorders: an update. *Blood* **122**: 1556–64.

Pereira C-F, Lemischka I, Moore K (2014) 'From blood to blood': de-differentiation of hematopoietic progenitors to stem cells. *The EMBO Journal* **33**(14): 1511–13.

Snoeck, H-W (2013) Aging of the hematopoietic system. *Current Opinion in Hematology* **20**: 355–61.

CHAPTER 2

Erythropoiesis

2

Douglas R. Higgs, Noémi Roy and Deborah Hay

Weatherall Institute of Molecular Medicine, John Radcliffe Hospital, Oxford, UK

Introduction

Erythropoiesis is the generation of red blood cells carrying the respiratory pigment haemoglobin, for the transport of oxygen to the tissues. This process, from the erythroid commitment of multipotent haemopoietic stem cells (HSCs), through the maturation of erythroblasts, to the terminal differentiation of red blood cells, is governed by complex transcriptional and epigenetic programmes, in response to extracellular signalling. Erythropoiesis normally maintains the steady state of an individual's red cell mass, producing 10^{11}–10^{12} new cells per day to replace those that are lost through senescence or premature destruction. Furthermore, erythropoiesis must be able to respond rapidly to erythroid stress such as haemorrhage and haemolysis. Perhaps unsurprisingly, this system is remarkably sensitive to systemic disease, with anaemia being a common manifestation of a wide range of inherited and acquired clinical disorders. Understanding the basic biology of erythropoiesis provides a logical basis for the diagnosis and treatment of the inherited and acquired anaemias that are so frequently encountered in clinical practice. In this chapter, we outline the normal mechanisms underlying erythroid specification, differentiation and maturation, and highlight some of the ways in which this complex system may fail in erythroid diseases.

The origins of erythroid cells during development

Both primitive (embryonic) and definitive (fetal/adult) HSCs arise in close association with endothelial cells. HSCs and endothelial cells are thought to arise from a common progenitor, the haemangioblast, which has the potential to form both blood and vessels (Chapter 1). Erythropoiesis occurs in waves that emerge from several sites in the developing embryo, and begins at the same time as development of the circulatory system. Primitive erythropoiesis is first evident at around three weeks of gestation, and arises from the blood islands of the extraembryonic yolk sac. A second wave of haemopoietic activity emerges from the yolk sac at approximately 4 weeks' gestation, and marks the onset of definitive erythropoiesis. Erythroid progenitors released into the circulation at this time pass to the liver, which becomes the main site of erythropoiesis in the fetus. A final wave of haemopoietic activity occurs in the aorta–gonad–mesonephros (AGM) region, the placenta and the major vessels at approximately 4–6 weeks of gestation. By 10–12 weeks, haemopoiesis starts to migrate to the bone marrow, where blood formation becomes established during the last three months of gestation (Chapter 1).

Primitive and definitive erythropoietic cells are distinguished by their morphology, cytokine responsiveness, growth kinetics, transcription factor programmes, epigenetic programmes and patterns of gene expression. Importantly, the types of haemoglobin produced are quite distinct in embryonic (Hb Gower I $\zeta_2\varepsilon_2$, Hb Gower II $\alpha_2\varepsilon_2$ and Hb Portland $\zeta_2\gamma_2$), fetal (HbF $\alpha_2\gamma_2$) and adult (HbA $\alpha_2\beta_2$ and HbA_2 $\alpha_2\delta_2$) erythroid cells. These specific patterns of globin expression provide critical markers for identifying the developmental stages of erythropoiesis. It remains unclear whether primitive and definitive haemopoiesis in mammals have entirely separate origins or if they are both derived from common stem cells that arise during early development. Accurately defining the embryological origins of these cells (Chapter 1) is important for understanding

Postgraduate Haematology, Seventh Edition. Edited by A Victor Hoffbrand, Douglas R Higgs, David M Keeling and Atul B Mehta.

the normal mechanisms that establish and maintain HSCs and how these programmes are subverted in common haematological disorders.

Specifying the erythroid lineage

At a cellular level, the precise mechanism by which HSCs differentiate into lineage-committed progenitors remains unknown and is currently under intense investigation. However, it is clear that as HSCs differentiate, they initially form multipotential progenitor cells such as CFU-GEMM – colony-forming units that have the ability to produce granulocytes, erythrocytes, monocytes and megakaryocytes. Such cells retain short-term repopulating ability, but lose long-term repopulating potential. Further differentiation progressively restricts the lineage potential of these cells, and reduces their proliferative capacity, resulting in bipotential progenitors with the ability to form megakaryocytes or erythroid cells (MEPs). MEPs further differentiate either into megakaryocytes and platelets or into fully committed erythroid precursors and red blood cells. These cells are functionally defined by their growth potential and characteristics as assayed by *in vitro* cultures: this explains the names 'burst-forming' erythroid units (BFU-E) and 'colony-forming' erythroid units (CFU-E) (Figures 2.1 and 2.2). They are not morphologically recognizable in the bone marrow, but can be defined as discrete cell populations when assayed by cell-surface markers. CFU-Es defined in these culture systems most closely correspond *in vivo* to pronormoblasts (also known as proerythroblasts), the earliest morphologically recognizable erythroid precursor in the bone marrow.

Expression of critical transcription factors specifies the erythroid lineage

Over the past few years, it has emerged that key haemopoietic transcription factors play a major role in regulating the formation, survival, proliferation and differentiation of multipotent stem cells as they undergo the transition to erythroid cells. These transcription factors may operate on their own or as members of multicomponent complexes involved in the activation and/or repression of gene expression. At present, the key transcription factors known to be involved in the specification and maintenance of HSCs include RUNX1, TAL1, LMO2, TEL, MLL and GATA2 (Figure 2.1). During normal erythroid development, GATA2 probably initiates the erythroid programme and plays an important role in the expansion and maintenance of haemopoietic progenitors. It is replaced during terminal erythroid maturation by GATA1, with the level of GATA2 declining as GATA1 increases. GATA1 is first expressed in MEPs and is essential for the terminal differentiation and maturation of both megakaryocytes and erythroid cells.

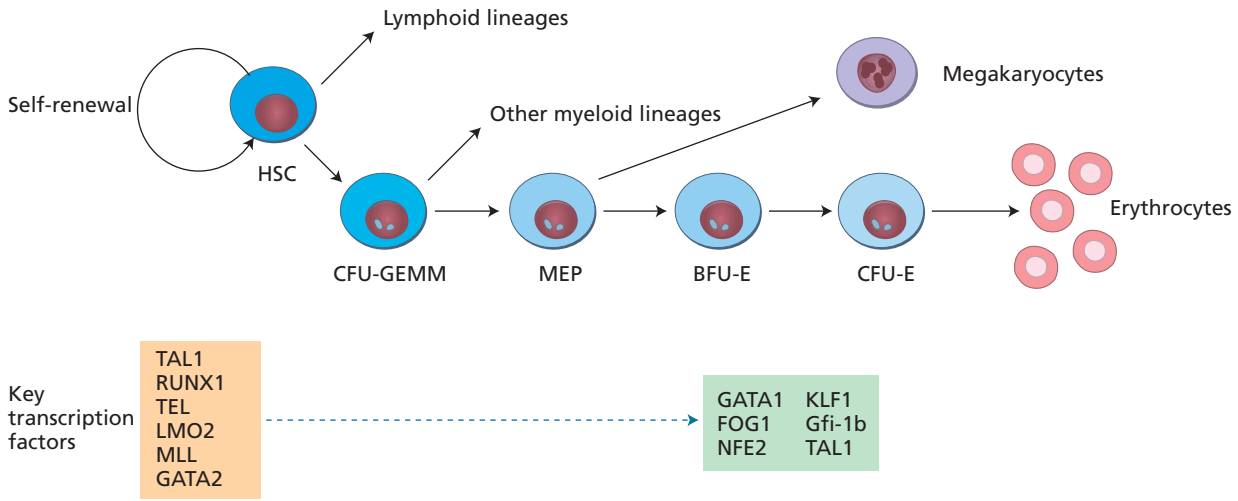


Figure 2.1 Summary of some steps in self-renewal, lineage specification and differentiation of haemopoietic stem cells to red cells. Some of the key transcription factors involved in this process are summarized beneath the diagram. HSC, haemopoietic stem cell; CFU-GEMM, colony-forming unit - granulocyte erythrocyte monocyte megakaryocyte; MEP, megakaryocyte-erythroid progenitor; BFU-E, burst-forming unit - erythroid; CFU-E, colony-forming unit - erythroid; TAL1, T-cell acute lymphoblastic leukaemia 1; TEL, translocation Ets leukaemia; LMO2, LIM domain only 2; MLL, mixed lineage leukaemia; GATA1, GATA-binding factor 1; GATA2, GATA-binding factor 2; FOG, friend of GATA; NFE2, nuclear factor erythroid-derived 2; KLF1, Kruppel-like factor 1 (erythroid); GFi-1b, growth-factor independent 1b.

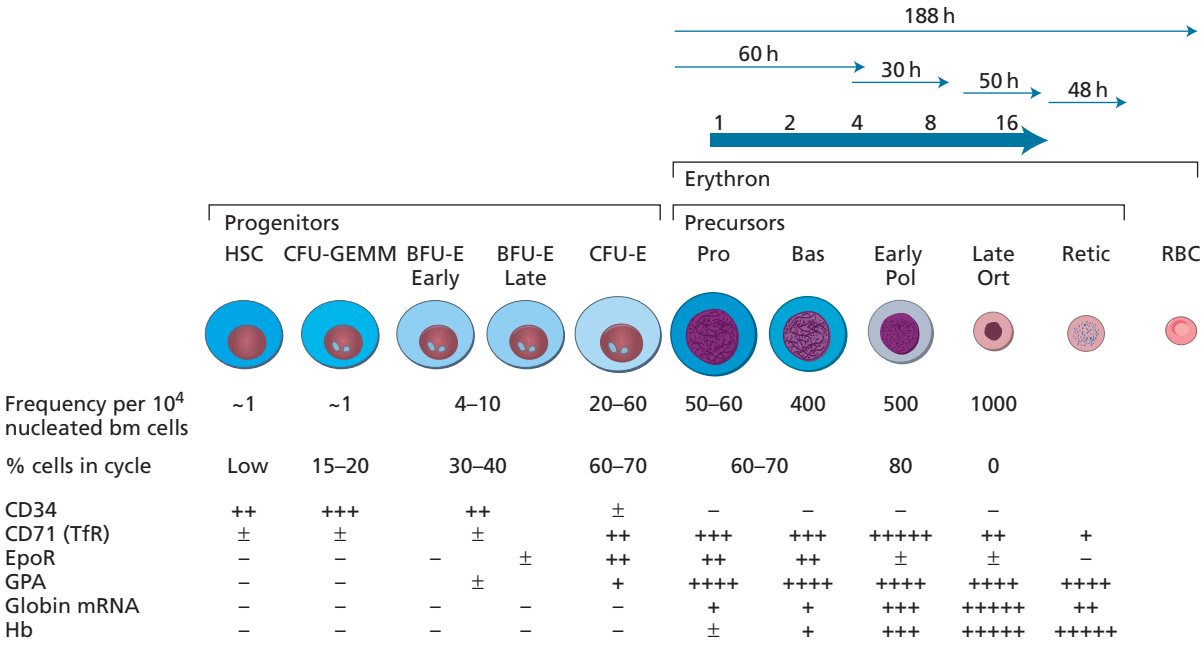


Figure 2.2 The specification and terminal differentiation of erythroid cells from haemopoietic stem cells. At the top, the estimated times for maturation of terminally differentiating cells are shown. The precursors are as follows: pronormoblasts (Pro), basophilic erythroblasts (Bas), polychromatic erythroblasts (Pol), orthochromatic erythroblasts (Ort), reticulocytes (Retic), mature red blood cells (RBCs). The number of divisions from pronormoblasts to orthochromatic normoblasts (1–16) is also shown. Some examples of the expression patterns of key cell-surface markers are shown below. TfR, transferrin receptor; EpoR, erythropoietin receptor; GPA, glycophorin A; bm, bone marrow.

One emerging principle in our understanding of lineage commitment is that factors affiliated with different lineages such as GATA1 (erythroid) and PU.1 (lymphocytes and granulocytes) are both present in uncommitted progenitors, reflecting the potential of these cells to develop along alternative pathways (so-called multilineage priming). It is now known that GATA1 and PU.1 interact and cross-antagonize each other. Therefore, as cells differentiate, reinforcement of the transcriptional programme of one lineage may actively suppress those of the alternative lineages.

Terminal maturation of committed erythroid cells

Once the erythroid programme has been specified, the final phase of erythropoiesis involves the maturation of committed erythroid progenitors to fully differentiated red cells. The earliest recognizable erythroid precursor in the bone marrow is the pronormoblast. Division of pronormoblasts leads to progressively smaller basophilic normoblasts, early polychromatic and finally late polychromatic/orthochromatic normoblasts (Figure 2.3). It has been estimated that, on average, four divisions occur within the morphologically recognizable proliferating precursor pool, so that each newly formed pronormoblast develops into 16 red cells (Figure 2.2). As a small amount of cell death (ineffective erythropoiesis) normally occurs, the average amplification is slightly less than 16-fold. The majority (60–80%) of pronormoblasts, basophilic normoblasts and early polychromatic normoblasts are in cell cycle. By contrast, late polychromatic/orthochromatic erythroblasts are postmitotic, non-dividing cells. In the final stages of terminal maturation, the nucleus condenses further and is eventually extruded. This produces the mature reticulocyte, which has no nucleus, but retains a few mitochondria and ribosomes. The cytoplasm of reticulocytes is predominantly pink on Wright–Giemsa staining because of the high concentration of haemoglobin, but it has a greyish tint due to the presence of ribosomes. When stained supravitally, the ribosomes precipitate into basophilic granules or a reticulum. Reticulocytes continue to synthesize haemoglobin for 24–48 hours after leaving the bone marrow.

Changes in the expression of transcription factors during terminal maturation

Once progenitor cells have committed to become erythroid cells, GATA1 and its cofactor FOG-1 (friend of GATA1) are

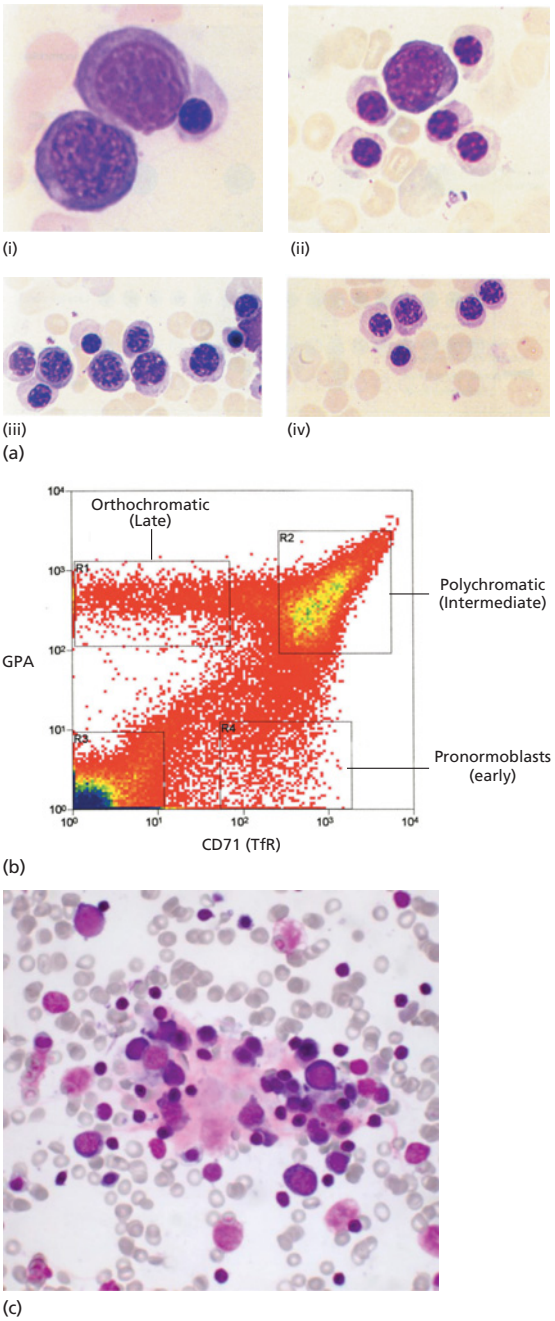


Figure 2.3 (a) Examples of pronormoblasts; (i) basophilic and polychromatic erythroblasts, and (ii) polychromatic and orthochromatic erythroblasts (iii and iv). All these different cell types can also be conveniently viewed at http://hsc.virginia.edu/medicine/clinical/pathology/educ/innes/text/nh/mature.html. (b) An example of early (pronormoblasts), intermediate (polychromatic erythroblasts) and late (orthochromatic erythroblasts) erythroid precursors separated on the basis of their cell-surface markers (CD71 and GPA). (c) An erythroblastic island with its central macrophage surrounded by erythroid progenitors at various stages of differentiation.

among the factors needed for them to proceed through terminal differentiation. There are GATA-binding motifs in the promoters and/or enhancers of virtually all erythroid-specific genes studied to date, including the globin genes, haem biosynthetic enzymes, red cell membrane proteins (e.g. blood group antigens) and erythroid transcription factors, such as TAL1, KLF1, Gfi1b, NFE2 and GATA1 (Figure 2.1). TAL1 is a basic helix–loop–helix transcription factor, which plays an important role in both the stem cell compartment and in developing the erythroid programme. KLF1 is a zinc-finger-like transcription factor expressed only in erythroid cells. Its binding sites are found in the regulatory elements of many erythroid-specific genes, including the β-globin gene. All of these transcription factors play key roles in coordinating erythroid maturation and globin gene regulation. Mutations of GATA1 and KLF1 are rare, but have been described in families with abnormalities of haemoglobin synthesis (Chapter 6), disorders of the red cell membrane (Chapter 8), abnormal haem synthesis (Chapter 3) and other abnormalities of erythropoiesis manifesting, for example, as congenital dyserythropoietic anaemia, Diamond–Blackfan anaemia (Chapter 10) and sideroblastic anaemia (Chapter 3). We anticipate that, with genome-wide sequencing, mutations in the other erythroid transcription factors (FOG1, TAL1, Gfi1b, NFE2) will be found underlying some rare forms of anaemia.

Changes in the expression of erythroid proteins during terminal maturation

As multipotent progenitors enter terminal differentiation, the expression of many genes (~6,000) is downregulated, reflecting the commitment to a single specialized lineage. By contrast ~600 mRNAs encoding proteins that characterize the red cell phenotype, are, in general, upregulated. Examples include blood group antigens, red cell membrane proteins, red cell glycolytic pathway enzymes, carbonic anhydrase and enzymes of the haem synthesis pathway. A full catalogue of these changes in gene expression can be found at https://cellline.molbiol.ox.ac.uk/eryth/index.html (Human Erythroid Maturation database).

Changes in gene expression are reflected in the cell-surface phenotypes of erythroid progenitors and precursors, in turn setting up the different signalling programmes of erythroid cells as they differentiate. Receptors for the key erythroid hormone erythropoietin (Epo, discussed in more detail below) first appear in small numbers on late BFU-Es, increasing in CFU-Es and pronormoblasts and subsequently declining and disappearing in later erythroid precursors (Figure 2.2). Similarly, CD71 (the transferrin receptor, TfR), which allows transferrin-bound iron to be taken into the cell, is present on early haemopoietic cells but is considerably upregulated on cells that are actively synthesizing haemoglobin, reaching a peak of 800,000 molecules per cell on polychromatic normoblasts. CD71 levels diminish in the

late phase of terminal differentiation and the receptor is undetectable on mature erythrocytes, which no longer have a need for iron uptake. In addition, developing erythroid cells express cell-surface adhesion molecules that interact with the extracellular matrix at high levels in early precursors. These are lost as maturation proceeds, freeing erythroid cells from the bone marrow niche (see below) to enter the circulation.

Key among the transcripts upregulated during erythroid maturation are those of α and β globin. The globins are first expressed in pronormoblasts and early basophilic erythroblasts, with the number of transcripts reaching 20,000 molecules per cell in late polychromatic and orthochromatic erythroblasts. During the later stages of erythroid cell maturation, the amount of RNA per cell and the rate of total protein synthesis decline, but the unusual stability of globin mRNA ensures that globin remains the predominant protein made in late erythroblasts and reticulocytes (Figure 2.2). Disorders of α and β globin structure and synthesis are the most frequent causes of inherited anaemia throughout the world and are discussed in detail in Chapters 6 and 7.

The individual components of the haemoglobin synthetic pathway (iron, free porphyrins, haem and monomeric globin chains) are all toxic to the cell, and feedback loops have evolved to ensure that cells are not damaged by these intermediates. In particular, the synthesis of globin is accurately matched with the synthesis of haem, in which some steps occur in the cytoplasm and others in the mitochondria (Figure 2.4). mRNAs encoding many components of the haem biosynthetic pathway (e.g. ALAS and porphobilinogen deaminase) are coordinately upregulated in terminal erythroid differentiation and their genes contain similar *cis*-regulatory elements, which are bound by the transcription factors GATA1, SCL and KLF1. As will be explained in Chapter 3, one of the main mechanisms by which haem and globin syntheses are coordinated is at the level of mRNA and translation via iron-responsive elements (IREs). There is also a need to coordinate the availability of iron to the requirements of erythropoiesis, by controlling iron absorption from the gut and release of iron from its stores. This is achieved by the master regulator of iron, hepcidin. This aspect of haemopoiesis and the diseases arising from abnormalities in iron metabolism are reviewed in Chapters 3 and 4.

The high levels of protein synthesis and rapid cell proliferation that characterize the erythroid compartment, even in a steady state, render these cells exquisitely sensitive to perturbations in levels of the key substrates required for erythropoiesis. Among these are a variety of nutritional factors and cofactors, particularly iron (Chapters 3 and 4), vitamin B_{12} and folate (Chapter 5), but also manganese, cobalt, vitamin C, vitamin E, vitamin B_6 (pyridoxine), thiamine, riboflavin, pantothenic acid and amino acids. Absolute or relative deficiencies of any of these factors can impair normal erythropoiesis and result in anaemia.

Control of erythropoiesis via cell signalling

The process of erythropoiesis must also be sensitive to changes in the circulating capacity for oxygen carriage and varying physiological demands. This precise regulation hinges on sensing

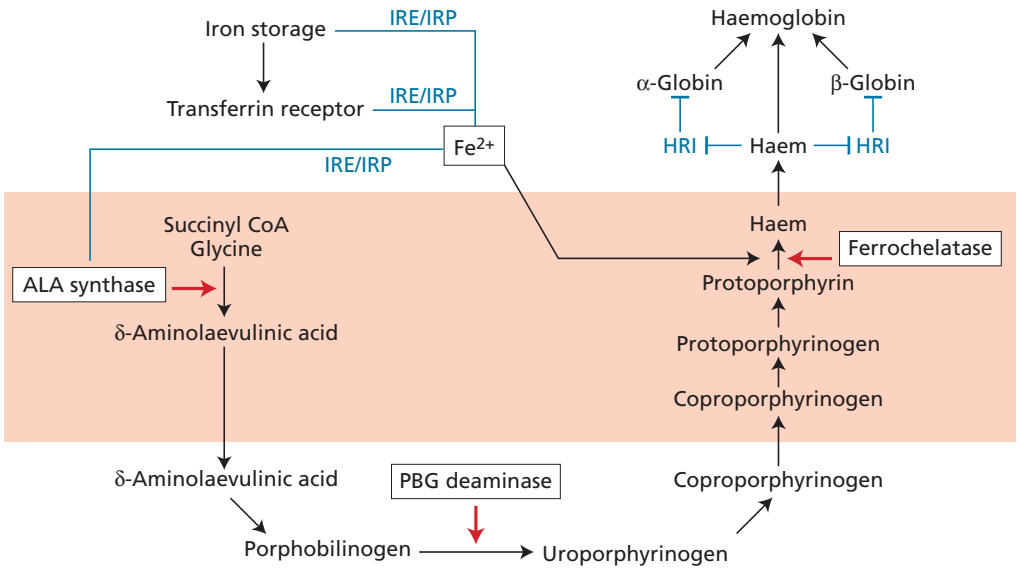


Figure 2.4 Coordination of globin synthesis, haem synthesis and iron regulation. Blue lines indicate some of the known regulatory feedback systems. The red shaded box indicates reactions occurring in the mitochondria. Rate-limiting controls of haem synthesis are shown in black boxes. ALA, δ-aminolaevulinic acid; HRI, haem-regulated eIF2alpha kinase; IRE, iron-responsive element; IRP, iron-responsive binding protein; PBG, porphobilinogen.

hypoxia and a tight control over the supply of erythroid precursors. Over the past 25 years there has been great progress in understanding the mechanisms by which cells sense hypoxia and orchestrate their response. The most important mediator of this response is the transcription factor HIF (hypoxia-inducible factor), which activates genes influencing adaptive responses to hypoxia (Figure 2.5). These include the genes for erythropoietin (Epo) to boost erythropoiesis, glycolytic pathway enzymes to maintain energy availability despite hypoxia, the transferrin receptor to ensure increased iron availability for erythropoiesis, and VEGF to promote angiogenesis. In rare cases of inherited polycythaemia, constitutive mutations in *HIF* or *vHL* (von Hippel Lindau; Figure 2.5) result in deregulated oxygen sensing and an erythropoietic drive in the absence of hypoxia (Chapter 35).

Epo is a 166-amino-acid 34.4-kDa glycoprotein, found in serum at baseline levels of 5–25 iU/L that can be elevated 1000-fold by severe anaemia. It contains about 40% carbohydrate, is rich in sialic acid residues, and has a half-life of 7–8 hours in plasma, whereas non-glycosylated Epo is rapidly cleared from the circulation. The main site of Epo production is the interstitial cells of the kidney. Under normoxic conditions, little or no Epo mRNA is detectable in the kidneys; hypoxia results in the rapid induction of its transcription such that levels may increase up to 200-fold over baseline within 30 minutes.

Epo upregulation is accomplished through a hypoxia-response element (HRE) at the 3′-end of the Epo gene. Under hypoxic conditions, HIF-1α is stabilized and can bind to the HRE of the Epo gene (and other hypoxia-sensitive genes), to increase transcription and therefore increase serum Epo levels. The positive effect of HIF is greatly increased by two cofactors, HNF-4, and the coactivator CBP/p300, also under hypoxic control. Linking Epo production to tissue oxygenation ensures that when there is reduced ambient oxygen tension, blood loss or shortened red cell survival, the level of Epo rises, stimulating red cell production and ultimately providing a greater source of local oxygen delivery.

The effects of Epo on red cell production are mediated by both increasing proliferation and reducing apoptosis of erythroid precursors. To respond to erythroid stress, the marrow must continuously produce an excess of erythroid precursors that can be called upon to differentiate into mature red cells whenever an immediate increase in erythroid output is needed. The low numbers (20–50) of Epo receptors (EpoR) on BFU-E explain the relative Epo unresponsiveness of these cells; much higher levels (~1000) are found in CFU-E, pronormoblasts and basophilic erythroblasts (Figure 2.2). Late BFU-E, CFU-E and pronormoblasts may all require continuous signalling via the EpoR to prevent their apoptosis, with signals from EpoR, via the JAK2–STAT5 pathway, inducing or stabilizing expression of the antiapoptotic protein Bcl-X_L (Figures 2.6 and 2.7). The overall effect is that, secondary to a hypoxia-driven rise in Epo, a population of erythroid progenitors normally destined for apoptosis at steady-state receive antiapoptotic signals, survive and expand the erythroid precursor pool. While the most important effect of

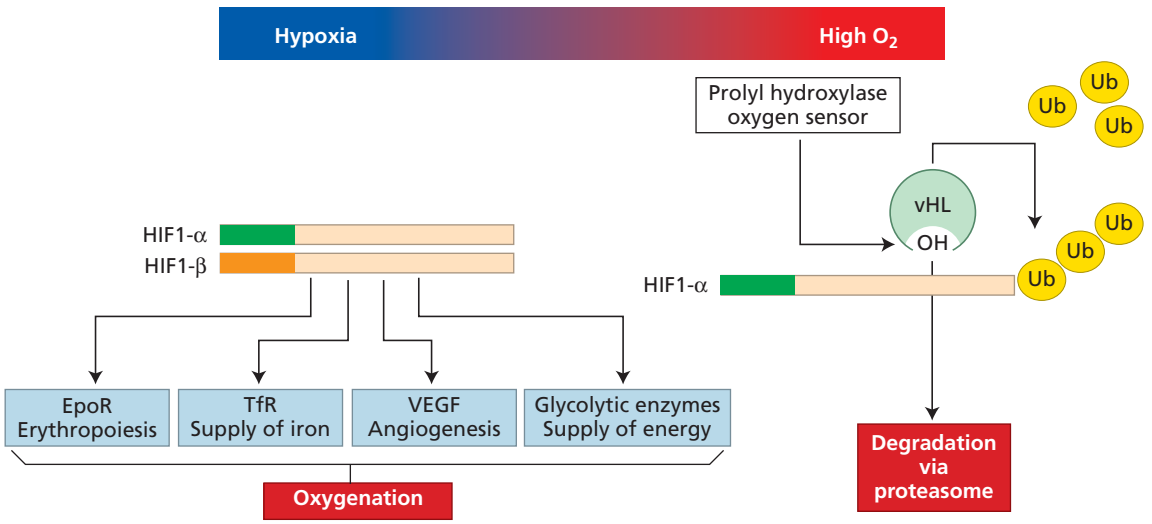


Figure 2.5 The oxygen-sensing system. Ub, ubiquitination; vHL, von Hippel–Lindau protein; HIF, hypoxia inducible factor. vHL is an E3 ubiquitin ligase. In the context of normal oxygenation, HIF-1α is hydroxylated, providing a binding site for vHL, which ubiquitinates it, thereby targeting it for degradation by the proteasome. At low oxygen tension, hydroxylation cannot occur and vHL cannot bind and ubiquitinate HIF-1α, the half-life of which is therefore greatly increased. As a result, HIF-1 is able to carry out its function as a transcription factor, upregulating the expression of its target genes, such as EpoR, TfR, VEGF, and glycolytic enzymes in response to hypoxia.

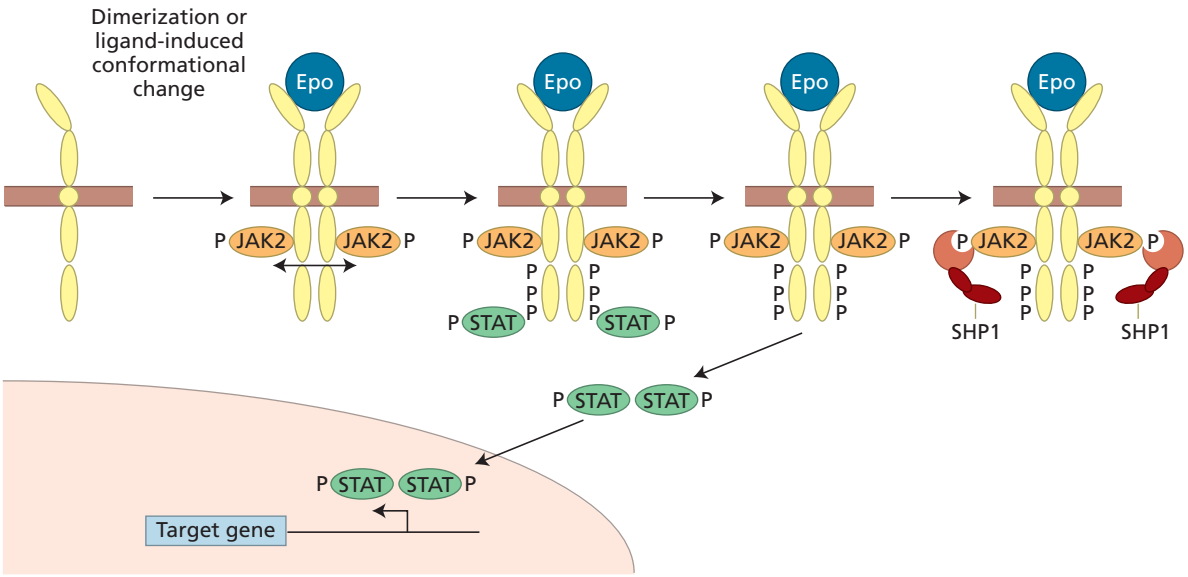


Figure 2.6 A summary of signalling via the erythropoietin (Epo) receptor as described in the text. P denotes regions of phosphorylation. The diagram shows Epo-induced dimerization or conformational change with *trans*-phosphorylation of JAK2, followed by phosphorylation of the Epo receptor. This is followed by binding and phosphorylation of STAT5. Binding of SHP1 (far right) to the Epo receptor activates its phosphatase activity, which can then dephosphorylate JAK2 and terminate signalling.

Epo is to increase the number of progenitor cells, which rapidly respond by proliferating and differentiating into viable pronormoblasts, it has also been suggested that Epo is able to speed up the rate of terminal differentiation by shortening the cell cycle and maturation times of erythroblasts.

Recent work in the mouse has shown that Epo stimulation through the JAK-STAT pathway causes upregulation of the protein erythroferrone, which mediates suppression of hepcidin synthesis. In this way, the signals that result in higher numbers of erythroid progenitors are matched with those resulting in increased iron absorption (see also Chapter 3). Current work is investigating whether this control loop also exists in humans.

Erythropoiesis is also influenced by pathways other than Epo–EpoR. Erythroid progenitors express receptors for SCF (stem cell factor), insulin-like growth factor (IGF-1) and insulin. After Epo, the second most important signalling system for erythropoiesis involves SCF (stem cell factor also called c-Kit). SCF was originally identified by its ability to stimulate proliferation of multipotent haemopoietic progenitors, but it is also effective in supporting growth of committed progenitors, including erythroid progenitors, acting synergistically with Epo.

In addition to SCF and Epo, stimulation of the nuclear hormone receptors for dexamethasone (glucocorticoid receptor) and oestrogen (oestrogen receptor) produces sustained proliferation of erythroid progenitors. Furthermore, the nuclear hormone receptors for thyroid hormone (c-ErbA/thyroid hormone receptor), all-*trans* retinoic acid (retinoic acid receptor) and 9-*cis*-retinoic acid can promote erythroid differentiation. Such observations are consistent with reports showing that many endocrine disorders (hypothyroidism, hypopituitarism, Addison's disease and male hypogonadism) can be associated with normochromic normocytic anaemia.

The erythroid niche

The control of erythropoiesis ultimately depends on integration of information delivered by cell–cell contacts, local growth factors and systemic hormonal signalling. Although erythropoiesis can be largely recapitulated using liquid cultures, it is significantly less efficient in this context than *in vivo*, particularly in terms of proliferation and terminal enucleation, suggesting that erythropoiesis in its natural setting is modified by additional factors.

Inspection of the bone marrow architecture shows that maturing erythroblasts are not randomly distributed; rather, they are arranged in characteristic 'erythroid islands' made up of a central macrophage surrounded by up to 30 erythroid cells at various stages of differentiation (Figure 2.3). This cellular structure is referred to as the erythroid niche. Erythroid cells cultured *in vitro* with macrophages proliferate threefold more than those cultured alone, and macrophages secrete soluble factors such as IGF-1 and burst-promoting activity that may influence

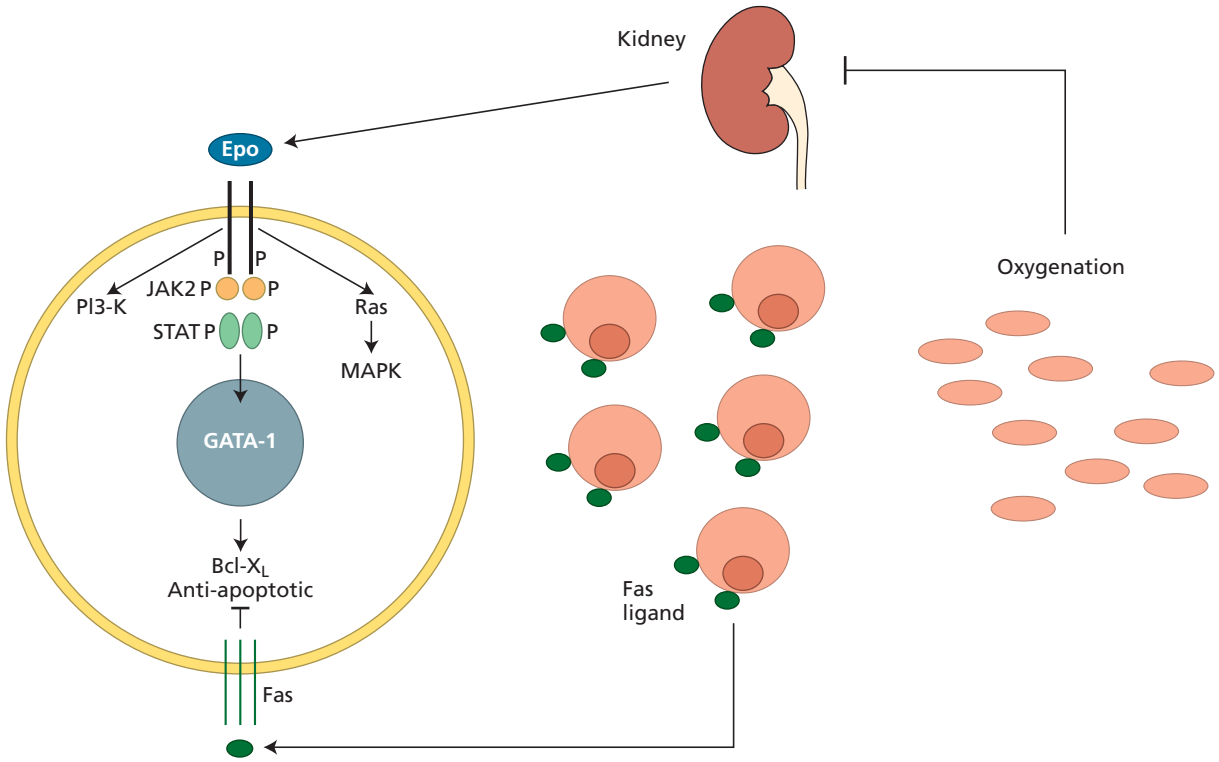


Figure 2.7 A summary of the apoptotic pathways (Epo and Fas) in erythroid progenitors. These cells (BFU-E and CFU-E) undergo apoptosis in the absence of Epo signalling or in the presence of Fas signalling. Bcl-X_L may be the key pathway through which these effects are mediated.

erythroid proliferation. Furthermore, direct contacts between the macrophages and surrounding erythroblasts (e.g. by cell adhesion molecules) appear to be critical for red cell maturation, stress erythropoiesis and red cell clearance. The central macrophage also phagocytoses and degrades the extruded red cell nuclei, and may play a role in the delivery of iron to the maturing erythroblast.

The juxtaposition of mature and maturing erythroblasts in erythroid islands also permits a further mechanism for the regulation of erythroid expansion. In addition to the Epo-Bcl-X_L pathway, apoptosis of erythroid precursors can also occur via activation of the Fas receptor (FasR), which is present on both early and late erythroid precursors. Crucially, its activating ligand (FasL) appears only on late erythroblasts. Binding of FasL to FasR activates proteolytic caspases that cleave intracellular proteins, including GATA1. The absence of GATA1 leads to apoptosis of erythroid precursors, and may also prevent induction of the antiapoptotic protein Bcl-X_L. Thus, FasL-activated GATA1 cleavage results in loss of Bcl-X_L and increased apoptosis of erythroid precursors – but only in the context of adequate numbers of their more mature counterparts. The result is a negative feedback loop to control erythroid proliferation (Figure 2.7).

Red cell senescence and clearance

Mature red cells live for approximately 120 days in the circulation under normal conditions, suggesting that mechanisms exist to monitor their senescence and control their removal from the circulation. Since mature red cells have no nucleus, they lack the capacity to synthesize new cellular components. Their ability to maintain cellular integrity becomes compromised with age, and characteristic features of the ageing red cell include increased glycation of haemoglobin and deamination of cytoskeletal components such as protein 4.1. Microvesiculation, which represents an effective means of removing damaged or ineffective red cell components, results in the continual loss of small fragments of the red cell, producing cells that are more dense and less deformable in the microcirculation than their younger counterparts.

As well as these changes, however, specific cues for clearance of the aged red cell from the circulation are thought to exist. Phosphatidylserine and phosphatidylethanolamine are key constituents of the red cell membrane that are normally confined to the inner aspect of the lipid bilayer. This asymmetry is maintained by an ATP-dependent aminophospholipid translocase. In senescent cells, phosphatidylserine is found in

the outer leaflet, where it is able to bind to macrophages in the liver and spleen and prompt erythrophagocytosis. Other senescence-related neo-antigens on the red cell surface may be generated by the clustering of membrane proteins such as band 3, thought to occur in response to oxidative change. Following erythrophagocytosis, the red cell components, including the iron from its haem groups, are recycled for subsequent red cell synthesis.

It is possible to compensate for a small decrease in the lifespan of mature red cells through increased Epo production and a reduction in apoptosis. Even when compensatory mechanisms are able to drive increased red cell synthesis, more significant reductions in red cell survival will lead to haemolytic anaemia, as discussed in Chapters 6–9. Systemic illness may also limit red cell survival, with the short red cell lifespan in uraemia and the anaemia of inflammation being well described, if not well understood.

Assessing erythropoiesis

Erythropoiesis is disturbed in a wide range of primary haematological conditions and, to some extent, in almost all multisystem diseases. Defects may arise in the production of committed erythroid progenitors; in the response of the oxygen sensing system, mediated via Epo and EpoR, and its effect on terminal erythroid differentiation and maturation; or in the red cell output achieved, which in turn exerts a major influence on the production of Epo, thus completing the regulatory loop (Figure 2.8).

Simple tools are available to test the circuit in a logical manner. Erythroid activity may be estimated by the ratio of myeloid precursors to erythroid precursors in the marrow (normally about 4:1, but with a very broad normal range). In the past, total erythropoiesis was measured accurately using radioactive (^{59}Fe) ferrokinetic assays. The plasma iron turnover measures the total (i.e. effective and ineffective) amount of erythropoiesis, whereas the red cell iron utilization assay measures the degree of effective erythropoiesis. To a large extent, these two parameters can now be assessed much more simply by measuring the levels of soluble TfR and the reticulocyte count. Soluble TfR is a truncated form of the receptor that circulates in a complex with transferrin. Erythroblasts rather than the reticulocytes are the main source of soluble TfR (Figure 2.2) and, when iron stores are adequate and available, measuring the level of soluble TfR (normal range 5.0 ± 1.0 mg/mL) is a good guide to the total level of erythropoiesis. Soluble TfR levels are increased when erythropoiesis is stimulated and decreased when it is diminished. The interpretation of soluble TfR levels is complicated in iron deficiency as this condition independently raises the level of soluble TfR. The reticulocyte count (0.5–2.0% or 25–75 × 10^9/L) is raised in proportion to the degree of anaemia when erythropoiesis is effective (e.g. uncomplicated response to bleeding), but is relatively low when erythropoiesis is ineffective (e.g. β thalassaemia [Chapter 6], myelodysplasia [Chapter 25] and CDA [Chapter 10]) or when an extrinsic abnormality prevents a normal response (e.g. nutritional deficiency; Figure 2.8).

The output of the system, the red cell mass, can be accurately measured by radioactive dilution techniques using ^{51}Cr, but can often be reliably estimated from the haematocrit or concentration of haemoglobin. Changes in red cell size, shape and haemoglobin content, often reflected in the red cell morphology, may provide important guides to specific abnormalities in red cell maturation (e.g. haemoglobinopathies, thalassaemia, nutritional deficiencies). If the red cell mass is appropriate to meet the demands for oxygenation, then Epo production will be suppressed and the serum level will be in the normal range. If there is inadequate oxygenation, the level of Epo will generally be raised in proportion to the degree of anaemia, unless there is some impediment to Epo production (e.g. chronic renal failure, anaemia of chronic disease). For any given degree of anaemia the level of Epo in the blood may vary, depending on the underlying pathology. For example, levels tend to be very high in aplastic anaemia and less than anticipated in thalassaemia. This may reflect the different numbers of precursors in the marrow that are able to bind available Epo molecules, thus altering the number of free Epo molecules that are measured.

In clinical practice, erythropoiesis can be assessed by examination of the full blood count and red cell indices, the reticulocyte count and examination of the peripheral blood film and bone marrow morphology. The usefulness of these and other

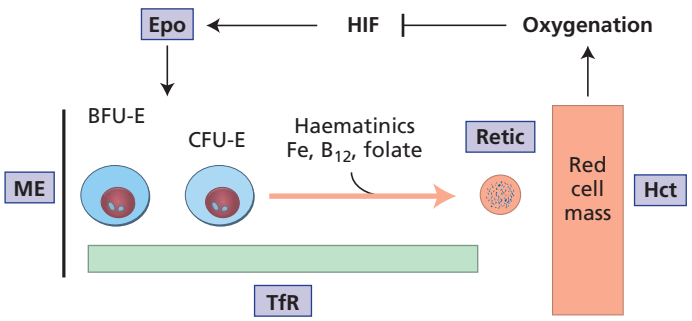


Figure 2.8 Summary of the regulation of erythropoiesis with the key points for assessment boxed in blue. ME denotes assessment of the myeloid/erythroid ratio in the bone marrow. Hct, haematocrit.

specialized tests in the investigation of anaemia will depend on the initial assessment of the patient by a detailed history and physical examination. The specific tests used to diagnose the inherited or acquired disorders that may perturb each phase of erythropoiesis are highlighted in the relevant subsequent chapters as each anaemia is discussed in detail.

Conclusions

The process of erythropoiesis needs to be tightly controlled to ensure adequate oxygen delivery to the tissues across a range of physiological and pathological conditions. An appreciation of the mechanisms of erythroid specification, maturation and feedback has already led to an improved understanding of the pathophysiology of many inherited and acquired anaemias, and provides a framework for the logical investigation and treatment of anaemia in the clinical setting.

Selected bibliography

Palis J. (2014) Primitive and definitive erythropoiesis in mammals. *Frontiers in Physiology* **28**(5): 3.

Kerenyi MA, Orkin SH (2010) Networking erythropoiesis. *Journal of Experimental Medicine* **207**(12): 2537–41.

Crispino JD, Weiss MJ. (2014) Erythro-megakaryocytic transcription factors associated with hereditary anemia. *Blood* **123**(20): 3080–8.

Haase VH. (2013) Regulation of erythropoiesis by hypoxia-inducible factors. *Blood Reviews* **27**(1): 41–53.

Dzierzak E, Speck NA (2008) Of lineage and legacy: the development of mammalian haematopoietic stem cells. *Nature Immunology* **9**: 129–36.

Orkin SH, Zon LI (2008) Hematopoiesis: an evolving paradigm for stem cell biology. *Cell* **132**: 631–44.

Cantor AB, Orkin SH (2002) Transcriptional regulation of erythropoiesis: an affair involving multiple partners. *Oncogene* **21**: 3368–76.

Unger FE, Thompson AM, Blank MJ *et al.* (2010) Erythropoiesis-simulating agents: time for a reevaluation. *The New England Journal of Medicine* **362**: 189–192.

CHAPTER 3

Iron metabolism, iron deficiency and disorders of haem synthesis

3

Clara Camaschella[1], A Victor Hoffbrand[2] and Chaim Hershko[3]

[1]Vita-Salute University, Milan, Italy
[2]University College London, London, UK
[3]Shaare Zedek Medical Center, Jerusalem, Israel

Introduction

Iron is essential for many metabolic processes. It shares with other transition metals two properties of particular importance in biology: the ability to exist in more than one relatively stable oxidation state and the ability to form many complexes. Its ability to exist in both ferric and ferrous states underlies its role in critical enzyme reactions concerned with oxygen and electron transport and the cellular production of energy. As well as physiologically active iron compounds, many of which are haem proteins, there are also specialized proteins of iron absorption, transport and storage. The latter are necessary to enable iron to remain in solution at neutral pH, at which ferric iron is insoluble, and to limit the potential toxicity of this reactive metal. The insolubility of ferric iron also means that although the Earth's crust contains approximately 4% iron and iron may be plentiful in the diet, much of this is not bioavailable. As a result, the body is limited in the adjustments it can make to excessive loss of iron, which frequently occurs due to haemorrhage, and iron deficiency is the most common cause of anaemia throughout the world, affecting in the order of 2 billion people. The general need to conserve the metal is reflected in the absence of any physiological mechanism for excretion of iron, control of iron balance being at the level of iron absorption. This is important in the rarer, but potentially fatal disorders of iron overload (see Chapter 4).

Distribution of body iron

The amount of iron in the adult human body is normally about 50 mg/kg in males and 40 mg/kg in females. The largest component is circulating haemoglobin, with 450 mL (1 unit) of whole blood containing about 200 mg of iron (Figure 3.1). Much of the remainder is contained in the storage proteins ferritin and haemosiderin. These are found mainly in the macrophages of the liver, spleen and bone marrow (which gain iron from breaking down red cells), and in parenchymal liver cells (which normally obtain most of their iron from the plasma iron-transporting protein transferrin).

Proteins important in iron metabolism

Haem proteins and iron-containing enzymes

Haemoglobin contains four haem groups linked to four globin chains, and can bind four molecules of oxygen. Myoglobin accounts for 4–5% of body iron and has a single haem group attached to its one polypeptide chain. It has a higher affinity for oxygen than haemoglobin and behaves as an oxygen reserve in muscles. The mitochondria contain a series of haem and non-haem iron proteins (including the cytochromes *a*, *b* and *c*, succinate dehydrogenase and cytochrome oxidase) that form an

Postgraduate Haematology, Seventh Edition. Edited by A Victor Hoffbrand, Douglas R Higgs, David M Keeling and Atul B Mehta.

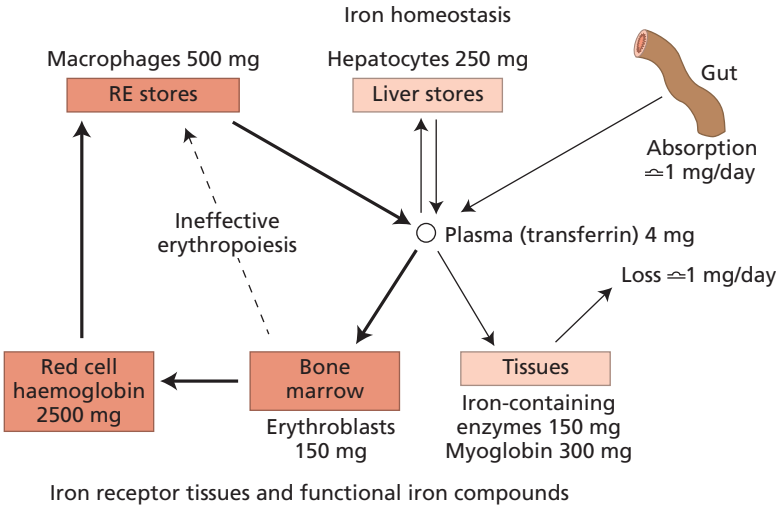


Figure 3.1 Iron homeostasis. The major compartments of iron in a 70-kg man. Iron supply for erythropoiesis and release of iron from senescent red cells dominate internal iron exchange. RE, reticuloendothelial. The dotted line indicates the low proportion of iron that derives from ineffective erythropoiesis in normal subjects.

electron transport pathway responsible for the oxidation of intracellular substrates and the simultaneous production of adenosine triphosphate (ATP). Haem is an essential component of microsomal and mitochondrial cytochrome P450, which is concerned with hydroxylation reactions (including drug detoxification by the liver), and of cyclooxygenase, involved in prostaglandin synthesis. Other haem proteins include the enzymes catalase and lactoperoxidase, which are concerned with peroxide breakdown, and tryptophan pyrrolase, involved in the oxidation of tryptophan to formylkynurenine. There is a smaller group of iron sulfur proteins (e.g. xanthine oxidase, reduced nicotinamide adenine dinucleotide dehydrogenase and aconitase). Iron is also necessary for the function of ribonucleotide reductase, a key enzyme in DNA synthesis.

Ferritin and haemosiderin

Ferritin is the primary iron storage protein and provides a reserve of iron. It consists of an approximately spherical apoprotein shell enclosing a core of ferric hydroxyphosphate (up to 4000 iron atoms). Human ferritin is made up from 24 subunits of two immunologically distinct types: H and L. There are multiple gene copies, which are mostly pseudogenes. An intronless gene codes for mitochondrial ferritin, an H-type ferritin. The internal cavity of the ferritin molecule communicates with the exterior via six channels, through which ferrous iron may enter (to interact with a ferroxidase centre on the ferritin H subunit) or leave (after reduction, e.g. by dihydroflavins or ascorbic acid).

The way in which ferritin iron is mobilized is poorly understood, and a process in which the entire ferritin molecule is degraded within lysosomes prior to iron release has been suggested. Variation in the proportion of H to L subunits explains the heterogeneity of ferritin from different tissues on isoelectric focusing: L-rich ferritins (from spleen and liver) are more basic than H-rich ferritins (from heart and red cells). The small amount of ferritin normally present in serum contains little iron and consists almost exclusively of L subunits. It is also heterogeneous, owing to glycosylation. This glycosylation and the direct relationship of serum concentration to storage iron in macrophages suggest that serum ferritin is secreted by macrophages in response to changing iron levels.

Haemosiderin, unlike ferritin, is a water-insoluble, crystalline, protein–iron complex that is visible by light microscopy when stained by the Prussian blue (Perls') reaction. It has an amorphous structure, with a higher iron/protein ratio than ferritin, and is probably formed by the partial digestion of ferritin aggregates by lysosomal enzymes. In normal subjects, the majority of storage iron is present as ferritin, and haemosiderin is predominantly found in macrophages rather than hepatocytes. In iron overload, the proportion present as haemosiderin increases considerably in both cell types.

Transferrin and transferrin receptors

Transferrin is a single-chain polypeptide present in plasma (1.8–2.6 g/L) and extravascular fluid (Table 3.1). It has a plasma half-life of 8–11 days. The protein is synthesized predominantly by the liver, synthesis being inversely related to iron stores. Two atoms of ferric iron can bind to each molecule. Although transferrin contains only about 4 mg of body iron at any time, it is vital to iron transport, with over 30 mg iron passing through this compartment each day (Figure 3.1). The uptake of iron from transferrin requires that the protein is attached to specific receptors on the cell surface. The transferrin receptor gene (*TFRC*) codes for TFR1, a transmembrane protein (identified as CD71), each molecule of two subunits binding one transferrin molecule. A second receptor, TFR2, also binds transferrin (Table 3.1). Through their binding with HFE, TFR1 and TFR2 are involved in regulating hepcidin synthesis (see Figure 3.2).

Table 3.1 Iron transport proteins, oxidoreductases, storage proteins and regulators.

Protein (gene)	Chromosome location	Tissue expression	Structure	Function	Regulation	Mutations and disease
Duodenal cytochrome b_1 (*CYBRD1*)	2q31	Duodenal enterocytes	TMP, 6TMD	Ferric reductase	Fe (hepcidin)	Not known
DMT1 (*SLC11A2*)	12q13	Widespread	TMP, 568 aa	Fe uptake	Fe (3′-IRE)	Mk mouse, Belgrade rat Human microcytic anaemia
Hemojuvelin (*HFE2*)	1q21.2	Liver, heart, muscle	Membrane-bound receptor or secreted protein	Regulator of hepcidin synthesis	?	Juvenile HC
Frataxin (*FXN*)	9q21.11	Heart, spinal cord, cerebellum	Mitochondrial protein,	Mitochondrial iron donor	?	Friedreich ataxia
FLVCR (*FLVCR1*)	1q32.3	Erythroid	Major facilitator family	Receptor for feline leukaemia virus C; haem export	?	Not known
Ferroportin (*SLC11A3*)	2q32	Liver, spleen, enterocyte	TMP, 571 aa, 9TMD	Fe export	Fe (5′-IRE)	Human HC, autosomal dominant
Hepcidin (*HAMP*)	19q13.1	Plasma (liver)	Mature peptide of 25 aa	Regulator of iron homeostasis	Fe (BMP6), IL6	Juvenile HC (digenic HC)
Hephaestin (*HEPH*)	Xq11–q12	Enterocyte	TMP, 1TMD, copper protein with homology to caeruloplasmin	Fe^{2+} oxidase	–	Sla mouse: iron deficiency anaemia
Haemochromatosis (*HFE*)	6p21.3	Widespread	HLA class I heavy chain	Regulates *TFRC*, iron uptake and hepcidin expression	?	Human HC, autosomal recessive
Mitoferrin 1 (*SLC25A28*)	8p21	Erythroid, liver, skeletal muscle, heart	Mitochondrial inner membrane	Mitochondrial iron importer	?	Not known
STEAP3	2q14.2	Erythroid, placenta (with TFR1)	Six-transmembrane epithelial antigen of the prostate-3	Ferric reductase (also reduces copper)	?	Human microcytic anaemia
Transferrin receptor (*TFR*)	3q26.2–qter	Widespread: highest number in erythroblasts	TMP dimeric polypeptide	Binds transferrin	Fe (3′-IRE)	Lethal in knockout mouse
Transferrin receptor 2 (*TFR2*)	7q22	Liver, erythroid cells	60% similarity in extracellular domain to TFR	Binds transferrin, iron homeostasis, regulator of hepcidin synthesis	No IRE	Human HC, autosomal recessive
Transferrin (*TF*)	3q21	Plasma, extravascular space	Single-chain polypeptide, glycoprotein	Iron transport	Iron stores	Atransferrinaemia, autosomal reeessive

(*Continued*)

Table 3.1 (*Continued*)

Protein (gene)	Chromosome location	Tissue expression	Structure	Function	Regulation	Mutations and disease
Ferritin heavy chain (*FTH1*)	11q13	Widespread, cytosolic	Subunit of ferritin	Iron storage (catalytic subunit for iron incorporation)	Fe (5′-IRE)	Autosomal dominant Fe overload (very rare)
Ferritin light chain (*FTL*)	19q13.3–q13.4	Widespread, cytosolic	Subunit of ferritin	Iron storage	Fe (5′-IRE)	Hyperferritinaemia and cataract syndrome Neuroferritinopathy*
IRP1 (*ACO1*)	9p21.1	Widespread	Cytoplasmic, with 4Fe–4S cluster	Regulation of synthesis of FTH, FTL, TFRC, DMT1, ferroportin, ALAS2	Cell iron	Not known
IRP2 (*IREB2*)	15	Widespread	Cytoplasmic, no 4Fe–4S cluster	As IRP1	Cell iron	Not known
Matriptase-2 (*TMPRSS6*)	22q13.1	Mainly liver	Type II transmembrane serine protease	Cleaves hemojuvelin	Unknown	Human IRIDA, autosomal recessive
Erythroferrone	2q37.3	Erythroid, muscle	Secreted protein	Hepcidin inhibitor	Erythropoietin	Not known

*In hyperferritinaemia/cataract syndrome, mutations affect the 5′-UTR IRE. In neuroferritinopathy, mutations affect *FTL* coding sequences.
aa, amino acid; BMP, bone morhphogenetic protein 6; HC, haemochromatosis; IL6, interleukin 6; IRIDA, iron refractory iron deficiency anaemia; IRE, iron response element; IRP, iron regulatory protein; TMD, transmembrane domain; TMP, transmembrane protein.

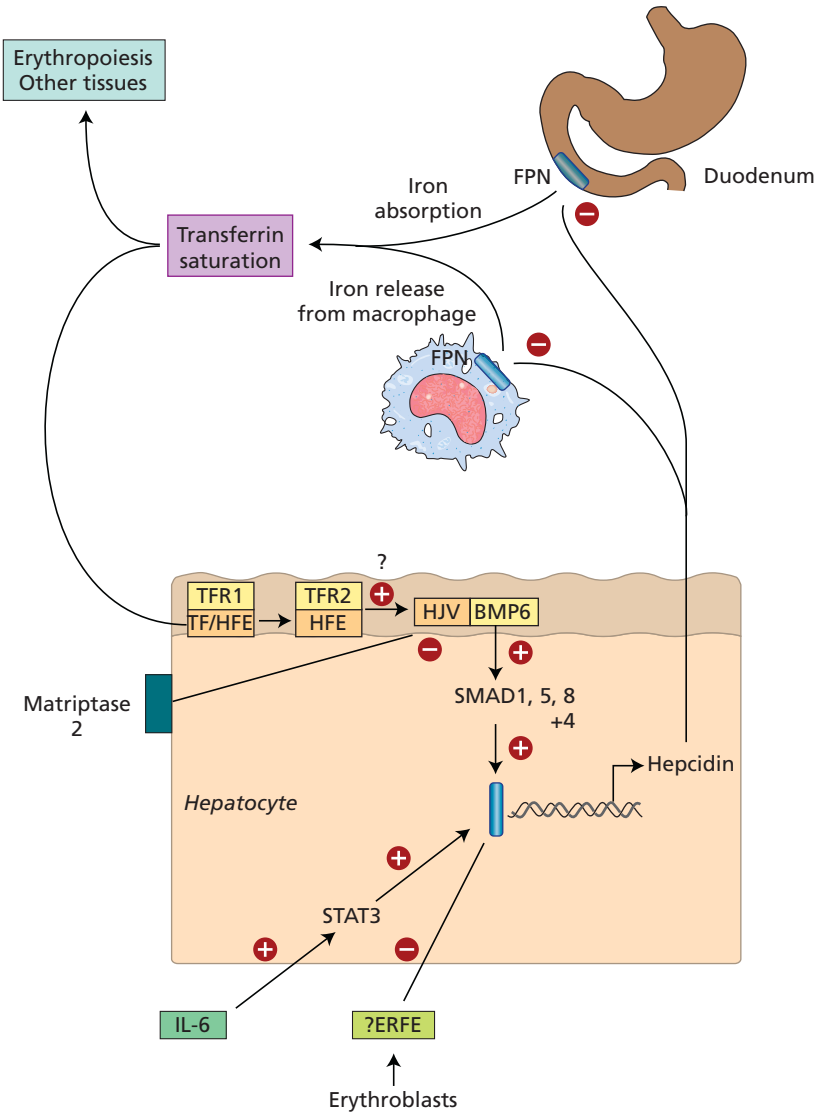


Figure 3.2 Stimulatory and inhibitory signals of hepcidin regulation. Hepcidin, as well as hemojuvelin (HJV), transferrin receptor 2 (TFR2) and HFE, are all produced in the hepatocyte. High plasma iron and inflammation stimulate hepcidin synthesis. This is mediated by SMADs and STAT3, respectively. Conversely, low plasma iron, increased rates of erythropoiesis (including ineffective erythropoiesis) and hypoxia inhibit hepcidin production. This is mediated by matriptase and ERFE. Hepcidin binds ferroportin (FPN), causing its destruction and so inhibits iron absorption and iron release from macrophages into plasma and from intracellular compartments. BMP, bone morphogenetic protein; ERFE, erythroferrone; The ? indicates uncertainty of the ERFE function in humans; GDF-15 may be the human equivalent of ERFE.

Lactoferrin is a glycoprotein that is structurally related to transferrin. It is found in milk and other secretions and in neutrophils. It has a bacteriostatic action at secreting surfaces by depriving microorganisms of the iron needed for their growth.

Divalent metal transporter 1

Divalent metal transporter (DMT)1 is an electrogenic pump that requires proton cotransport in order to transfer Fe^{2+} across cell membranes. This occurs at the apical membrane and subapical endosomes of the duodenal enterocyte and the transferrin-cycle endosome, both of which have a low pH. The intestinal DMT is produced by different mRNA splicing from that which produces endosomal DMT1. DMT1 expression is upregulated in iron deficiency (see later) and may be involved in absorption of other divalent metal cations, including Mn^{2+}, Co^{2+}, Zn^{2+}, Cu^{2+} and Pb^{2+}.

Ferroportin

This transmembrane domain protein is the basolateral transporter of iron, essential for iron release from macrophages, the intestinal absorptive enterocyte, hepatocytes and placental syncytiotrophoblasts. It is also present in intracellular compartments. Caeruloplasmin is required for the cell-surface localization of ferroportin, whose concentration is controlled by hepcidin, which triggers its internalization and degradation in lysosomes.

Other proteins

The roles in iron metabolism of hemojuvelin (HJV), bone morphogenetic protein-6 (BMP-6), *small mothers against*

*decapentaplegic (*SMADs), ferrioxidative and reduction enzymes, and caeruloplasmin are discussed under the headings of hepcidin regulation, iron absorption, iron uptake by erythroid cells and haem synthesis.

Hepcidin

Hepcidin has a central role in the regulation of iron metabolism and absorption (Figure 3.2). A product of the *HAMP* gene (Table 3.1), it is a small peptide (25 amino acids) released from a large prepropeptide of 84 amino acids. It is predominantly expressed in the liver. It regulates iron homeostasis by binding to cell-surface ferroportin, causing its degradation in lysosomes. It therefore acts to inhibit iron absorption, iron release from macrophages and iron transport across the placenta. It is bound in plasma to α_2-macroglobulin and the major route of clearance is the kidney. Hepcidin can be measured in serum or urine by ELISA or mass spectrometry-based techniques.

Regulation of hepcidin expression

The regulation of hepcidin expression is transcriptional. Hepcidin expression is increased in response to raised serum iron, iron overload and inflammation, and is suppressed by iron deficiency, hypoxia and increased erythropoietic activity. Under basal conditions, expression depends on signalling through the BMP/SMAD pathway (Figure 3.2). HJV is a member of the repulsive guidance molecules (RGM) family that is highly expressed in liver, skeletal muscles and the heart. It is either associated with cell membranes through a glycosylphosphatidylinositol anchor or released as a soluble form. Membrane-bound HJV participates in the pathway regulating hepcidin expression as a BMP coreceptor, whereas soluble HJV antagonizes BMP-6. BMP-6 is the master hepcidin activator *in vivo* in murine models.

HFE and TFR2 are also involved in hepcidin expression (Figure 3.2). HFE is able to bind TFR1 and TFR2. During low or basal serum iron conditions, HFE and TFR1 exist as a complex at the plasma membrane, TFR1 serving to sequester HFE to silence its activity. Diferric serum transferrin (Fe^{2+}-TF) competes with HFE for binding to TFR1. Increased serum transferrin saturation therefore results in dissociation of HFE from TFR1. Acting as an iron sensor, HFE then binds to TFR2 and conveys the Fe^{2+}-TF status to the signal transduction effector complex. HJV binds to BMP, then phosphorylates SMADs to form a SMAD-1/-5/-8–SMAD-4 complex, which translocates to the nucleus and stimulates hepcidin production by activating its promoter. In keeping with this model, genetic mutations of HFE, TFR2, HJV and hepcidin all result in haemochromatosis with low serum hepcidin levels (see Chapter 4). Iron levels also seem to control BMP-6 production.

A second type of transcriptional hepcidin regulation occurs in inflammation. Interleukin (IL)-6 and IL-1β induce transcription of the hepcidin gene by activating STAT3 (signal transducer and activator of transcription 3) and its binding to a regulatory element in the hepcidin promoter. It may converge on a final shared SMAD-4-dependent pathway.

The hepcidin response is remarkably rapid. In humans, iron ingestion results in a sharp increase in urinary hepcidin excretion within 12–24 hours of starting treatment. Likewise, infusion of recombinant IL-6 results in significant increase in urinary hepcidin and decreased serum iron and transferrin saturation within 2 hours of infusion. These observations imply that hepcidin expression is directly controlled by serum iron (probably by transferrin saturation) and IL-6 and not by long-term gradual accumulation of iron in tissues.

Response to anaemia and hypoxia

Hepcidin levels are reduced or undetectable in iron deficiency anaemia. Iron absorption is accelerated in iron deficiency, ineffective erythropoiesis and hypoxia (Figure 3.2) likely through common mechanisms mediated by the 'erythroid regulator'. Erythroid precursors secrete growth differentiation factor-15 (GDF-15) and TWSG1, two cytokines of the TGF-β family, which have been proposed to inhibit hepcidin production in the liver. Serum concentrations of GDF-15 are greatly increased in thalassaemia major and other conditions associated with ineffective erythropoiesis. However, the role of these cytokines *in vivo* is uncertain. Recently erythroferrone, a novel protein, has been identified that plays the role of physiological erythroid regulator in mice. It is released by erythroid precursors and suppresses hepcidin after bleeding and erythropoietin treatment in order to increase intestinal iron absorption and macrophage iron release according to the erythropoietic needs. It remains to be seen if this hormone plays a similar role in human erythropoiesis.

Hypoxia directly increases iron absorption through increased duodenal iron importer DMT1 and indirectly through hepcidin suppression by increased erythropoiesis. The same mechanism may contribute to the harmful accumulation of iron in response to chronic anaemia associated with ineffective erythropoiesis in thalassaemia and other dyserythropoietic anaemias.

Matriptase-2

This is a type 2 member of the transmembrane serine protease family mainly expressed in the liver. Membrane-bound matriptase-2 regulates hepcidin expression by cleaving membrane-bound HJV, releasing soluble HJV fragments. The factors that regulate matriptase-2 expression need to be elucidated. Matriptase-2 activity over-rides all known activating stimuli of hepcidin synthesis. Homozygous or compound heterozygous *TMPRSS6* mutations in humans and homozygous inactivation of the gene in mice result in marked upregulation

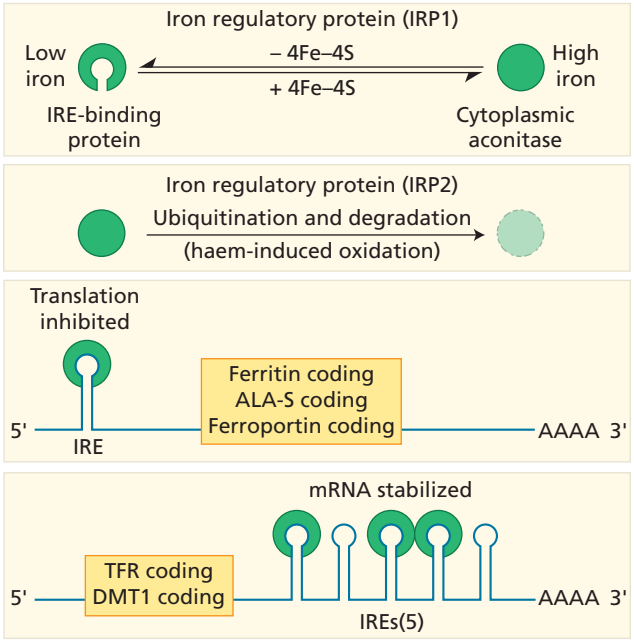


Figure 3.3 Coordinate regulation of expression of ferritin and transferrin receptor: the role of the iron response element (IRE)–iron regulatory protein (IRP) mechanism. When cellular iron levels are low, IRP binds to the IRE stem and loop structures of mRNA to inhibit translation of ferritin and ALA-S, but increases translation of transferrin receptors (TFR) and DMT1 by preventing degradation of the mRNA. When iron levels are high, the IRP functions as a cytoplasmic aconitase and no longer binds to the IREs. Ferritin synthesis can thus proceed, while TFR synthesis is reduced. IRP2 binds to IREs when iron levels are low, but is degraded after ubiquitination (initiated by haem-induced oxidation) when iron levels are high. The IRPs therefore provide two ways of sensing iron requirements, either involving Fe–S proteins or haem proteins.

of hepcidin and blockade of intestinal and macrophage iron transport into plasma, leading to a refractory hypochromic microcytic anaemia (see below).

Intracellular iron homeostasis

Synthesis of several of the proteins involved in iron metabolism is regulated at the level of RNA translation by two cytoplasmic iron-dependent proteins, IRP1 and IRP2 (Table 3.1). These are capable of binding to mRNAs that contain a sequence forming a stem-and-loop structure called an *iron-responsive element* or IRE (Figure 3.3). The 3′-untranslated region (3′-UTR) of TFR1 contains five IREs, whereas the 5′-UTR region of ferritin mRNA contains a single IRE. Binding of IRP when there are low levels of intracellular iron protects TFR1 mRNA from cytoplasmic degradation, but inhibits translation of ferritin mRNA, by interfering with the binding of initiation factors. In contrast, when intracellular iron is increased, the opposite effects occur. Thus, coordinated regulation of TFR1 and ferritin acts to maintain a constant intracellular iron content over the short term by balancing cellular iron uptake and storage.

Erythroid δ-aminolaevulinic acid synthase (ALAS2) mRNA also has an IRE in its 5′-UTR region, whereas 'housekeeping' ALAS1 mRNA does not. The IRP–IRE system is therefore involved in matching iron supply to haem synthesis, with repression of protoporphyrin synthesis in iron-deficient erythroblasts (see p. 29). DMT1 like TFR1 has a 3′-IRE and is upregulated in iron deficiency. Ferroportin has a 5′-IRE (Table 3.1) but, as occurs for DMT1, ferroportin may be translated in a non-IRE isoform that escapes the IRP control in some tissues, for example in duodenal cells.

Normal iron balance

The amount of iron in the body at birth depends on the blood volume and haemoglobin concentration, birth weight (which determines blood volume) being particularly important. Delay in clamping the cord leads to an increased red cell mass by placental transfusion. The level of maternal iron stores has little effect on fetal iron. The newborn contains about 80 mg/kg at full term. Neonatal iron reserves are utilized for growth, and from 6 months to 2 years virtually no iron stores are present. Thereafter, iron stores gradually accumulate during childhood to around 5 mg/kg. In men, there is a further increase between 15 and 30 years to about 10–12 mg/kg (total up to approximately 1 g), whereas iron stores remain lower in women (average 300 mg) until the menopause.

Requirements are higher in menstruating women and during periods of rapid growth in infancy and adolescence (Table 3.2). Menstrual blood loss has a median value of 30 mL, but the 95th centile value is 118 mL per month (equivalent to 1.9 mg iron per day), which has been found to be significantly associated with iron deficiency. Requirements are highest of all in pregnancy.

Iron absorption

Iron absorption depends not only on the amount of iron in the diet but also, and more importantly, on the bioavailability of that iron, as well as the body's needs for iron. A normal Western diet provides approximately 15 mg of iron daily. Of that iron, digestion within the gut lumen releases about half in a soluble form, from which about 3 mg may be taken up by mucosal cells and only about 1 mg (or 5–10% of dietary iron) transferred to the portal blood in a healthy man. Iron absorption can thus be influenced at several different stages.

Table 3.2 Daily iron losses and requirements.

Group (age, years)	Daily loss (mg)		Requirement for growth (mg)	Total (mg)
	Urine, skin, faeces, etc.	Menses		
Children				
0.5–1	0.17	–	0.55	0.72
1–3	0.19	–	0.27	0.46
4–6	0.27	–	0.23	0.50
7–10	0.39	–	0.32	0.71
Males				
11–14	0.62		0.55	1.17
15–17	0.90	–	0.60	1.50
18+	1.05	–	–	1.05
Females				
11–14*	0.65	–	0.55	1.20
11–14	0.65	0.48†	0.55	1.68
15–17	0.79	0.48†	0.35	1.62
18+	0.87	0.48†	–	1.35
Post menopause	0.87	–	–	0.87
Lactating‡	1.15	–	–	1.15

*Non-menstruating.
†Median loss.
‡Average dietary requirement during pregnancy is 3–4 mg.
Source: WHO (2001) *Iron Deficiency Anaemia. Assessment, Prevention and Control. A Guide for Programme Managers*. World Health Organization, Geneva. Reproduced with permission of WHO.

Dietary and luminal factors

Much of dietary iron is non-haem iron derived from cereals (commonly fortified with additional iron in the UK), with a lesser component of haem iron from meat and fish. Even in iron deficiency, the maximum iron absorption from a mixed Western diet is no more than 3–4 mg daily. This figure is much less with the predominantly vegetarian, cereal-based diets of most of the world's population since iron is better absorbed from animal than vegetable sources.

Iron is released from protein complexes by acid and proteolytic enzymes in the stomach and small intestine, and haem is liberated from haemoglobin and myoglobin. Iron is maximally absorbed from the duodenum and less well from the jejunum, probably because the increasingly alkaline environment leads to the formation of insoluble ferric hydroxide complexes. Acid pH, vitamin C and some low-molecular-weight chelates (e.g. sugars, amino acids) enhance absorption. Therapeutic ferrous iron salts are well absorbed on an empty stomach, but when taken with a meal, absorption is reduced as a result of the same ligand-binding processes that affect dietary non-haem iron; phytates, tannates in tea and bran inhibit absorption.

Mucosal factors: molecular aspects of iron absorption and its regulation

A variety of mechanisms for the binding of non-haem iron to the mucosal membrane have been described. Specific, saturable and receptor-mediated mechanisms, and passive diffusion at higher doses, may occur. Iron absorption is regulated both at the stage of mucosal uptake and at the stage of transfer to the blood. The process is illustrated in Figure 3.4. Non-haem iron is released from food as Fe^{3+} and reduced by duodenal cytochrome b_1 (DCytb) to Fe^{2+}. This is transported across the brush border membrane by DMT1, which is upregulated in iron deficiency and by hypoxia inducible factor 2 (HIF2-α) activation in the local hypoxic environment of duodenum. It is assumed that iron enters the labile pool and some may be incorporated into ferritin and lost when the cells are exfoliated. Iron destined for retention by the body is transported across the serosal membrane by ferroportin before uptake by transferrin as Fe^{3+}. The regulation of duodenal cell iron release by hepcidin through its action on ferroportin is rapid. Hephaestin is a copper-containing ferroxidase expressed predominantly in villous cells of the small intestine that converts Fe^{2+} to Fe^{3+} in the basolateral transfer step of iron absorption.

Haem iron is initially bound by haem receptors at the brush border membrane and released intracellularly by haem oxygenase before entering the labile iron pool and following a common pathway with iron of non-haem origin.

Iron uptake by erythroid cells

About 85% of transferrin iron normally enters developing red cells for incorporation into haemoglobin. This tissue distribution of transferrin-bound iron reflects the expression of transferrin receptors, which are present in high concentration on cells with a high iron requirement. The latter includes any rapidly dividing cells, but is normally dominated by the cells of the erythron. A soluble truncated form of the transferrin receptor derived from these cell surfaces is detectable in serum.

Transferrin receptors have the highest affinity for diferric transferrin. The transferrin–receptor complex is taken up by a process of receptor-mediated endocytosis (Figure 3.5). The iron is released at the low pH of the endosome, reduced from Fe^{3+} to Fe^{2+} by STEAP3, a ferrireductase, before the apotransferrin and receptor are recycled to the plasma and the cell membrane, respectively. Iron release from the endosome is via DMT1 (Figure 3.5) and the iron is transported into mitochondria by mitoferrin or enters ferritin. A mitochondrial version of ferritin, which has an unknown physiological role, is elevated in congenital or acquired sideroblastic anaemias. Direct transfer of storage iron from macrophages to erythroblasts (rhopheocytosis) may also occur.

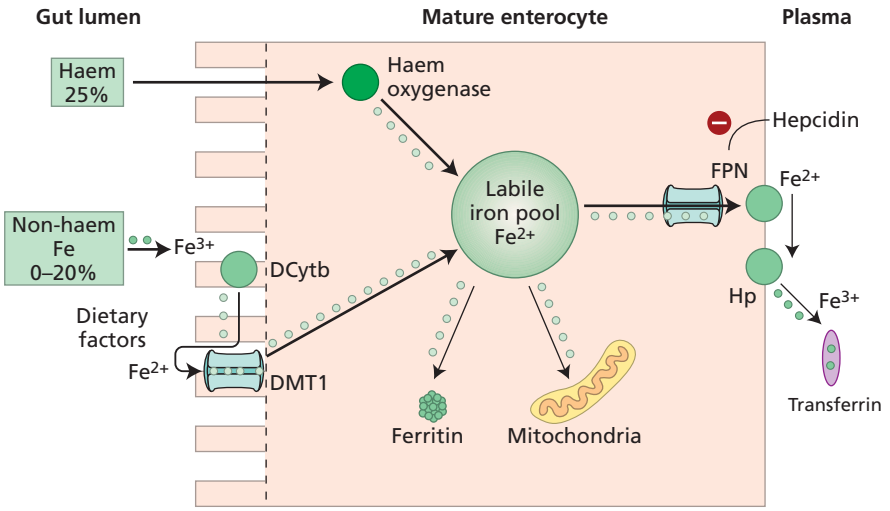


Figure 3.4 Molecular pathways of iron absorption. The area enclosed in the dotted box refers to the uptake of iron from the plasma in the developing enterocyte in the intestinal crypt. Otherwise, the diagram refers to iron absorption by the villous epithelial cell. DMT1, divalent metal transporter 1; FPN, ferroportin. Hp, hephaestin. For further details see text and Table 3.1.

Some 80–90% of iron taken into developing erythroblasts is converted to haem within 1 hour. Any iron taken up in excess of the requirement for haem synthesis is incorporated in ferritin (Figure 3.5). The red cell ferritin content is therefore increased when haemoglobin synthesis is impaired, as in thalassaemia or sideroblastic anaemia. Excess iron may be seen in the cytoplasm of mature red cells as one or more siderotic granules. These are composed of haemosiderin, and stain blue with Perls' reaction and purplish blue with Romanowsky stains, when they are called Pappenheimer bodies. The spleen removes these granules by its pitting action.

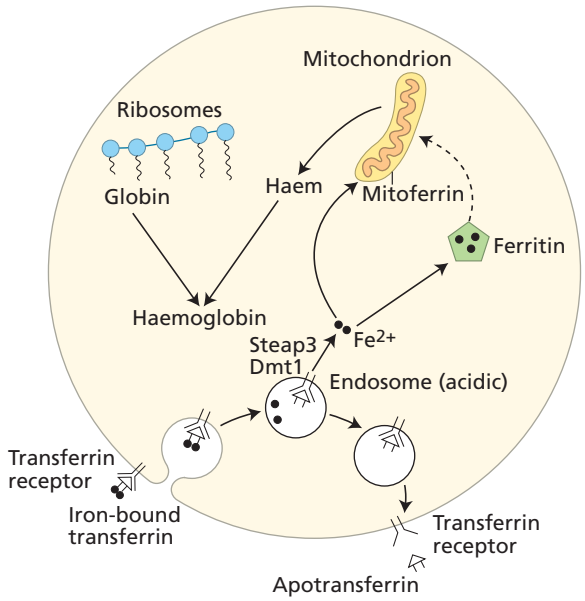


Figure 3.5 Incorporation of iron from plasma transferrin into haemoglobin in developing red cells. Uptake of transferrin iron is by receptor-mediated endocytosis.

Haem synthesis and mitochondrial iron metabolism

Haem consists of a protoporphyrin ring with an iron atom at its centre. Haem is synthesized from the precursors succinyl-CoA and glycine, which condense to form δ-aminolaevulinic acid (ALA) under the action of ALAS, with pyridoxal phosphate as a coenzyme (Chapter 2). Housekeeping ALAS is ALAS1; in erythroid cells erythroid-specific ALAS2 predominates. ALA can be utilized for the formation of both purines and haem. Four molecules of porphobilinogen condense under the influence of porphobilinogen deaminase and uroporphyrinogen cosynthase to form the tetrapyrrole ring compound uroporphyrinogen III. The latter is converted to protoporphyrin IX. Finally, iron in the ferrous form is incorporated under the influence of the enzyme ferrochelatase. Iron in haem has six coordinating valencies: four link the iron to nitrogen atoms in each pyrrole ring, whereas the remaining two link haem to histidine residues in the globin chain, the distal bond being unstable and easily replaced by oxygen to form oxyhaemoglobin. A haem exporter, group C feline leukemia virus receptor (FLVCR) (Table 3.1), is present in erythroid cells and rids them of any haem made in excess.

The mitochondria play a major role in haem synthesis as they contain ALAS, coproporphyrinogen oxidase and ferrochelatase, the enzyme sequence from ALA to coproporphyrinogen being situated in the cytoplasm. Besides haem, mitochondria utilize iron in the synthesis of iron–sulfur clusters, prosthetic groups essential for the function of several mitochondrial (respiratory chain complexes, ferrochelatase) and cytosolic (aconitase) proteins. A number of porphyrins are formed by side reactions during the synthesis of protoporphyrin. In the porphyrias (see p. 35), many of these compounds accumulate in the major sites of haem synthesis, the liver and the red cells.

The assembly of iron–sulfur clusters is a complex, incompletely understood process that requires multiple proteins. The

Table 3.3 Differential diagnosis of hypochromic anaemia.

	Iron deficiency	Chronic disease	Thalassaemia trait (α or β)	Sideroblastic anaemia	IRIDA
MCV/MCH	↓	↓ or N	↓	↓ (congenital) ↑N (acquired)	↓
Serum iron	↓	↓	N	↑	↓
TIBC	↑	↓ or N	N	N	
Transferrin saturation	↓	↓	N	↑	↓
Serum ferritin	↓	N or ↑	N	↑	N
Serum TFR	↑	N	N	N or ↑	↑
Serum hepcidin	↓	↑	N	↓	N or ↑
Bone marrow iron stores	↓	N or ↑	N	N or ↑	↑
Erythroblast iron	↓	↓	N	Ring forms	

IRIDA, iron refractory iron deficiency anaemia; MCH, mean corpuscular haemoglobin; MCV, mean corpuscular volume; N, normal; TFR, transferrin receptor; TIBC, total iron-binding capacity.

best known is frataxin, a likely iron donor in this pathway. The gene encoding frataxin is mutated in Friedreich ataxia. An enzyme active in the pathway, whose function is still unclear, is glutaredoxin-5, which has been found mutated in a rare recessive form of sideroblastic anaemia. ABCB7, a member of a family of transmembrane proteins characterized by the ABC domain that binds and hydrolyses ATP, transfers iron–sulfur clusters from mitochondria to the cytosol. Mutations of ABCB7 cause an X-linked form of sideroblastic anaemia with ataxia (see below).

The mitochondria are also the site of the citric acid cycle, which supplies succinate. The mature red cell, which lacks mitochondria, is therefore unable to synthesize haem.

Intracellular transit iron and plasma non-transferrin-bound iron

It has been suggested, but not proved, that there is a transit pool of 'metabolically active' or 'labile' iron within cells, which receives iron from degraded haem or ferritin, exchanges with transferrin and is incorporated into newly synthesized iron-containing proteins. This iron is considered to be sensed by iron regulatory proteins and available for chelation. Within cells, low-molecular-weight chelates (e.g. with citrate) may be present. In iron overload, non-transferrin-bound iron may also exist as oligomeric iron oxide, either free or bound to albumin, and is particularly toxic to various organs (Chapter 4).

Breakdown of haemoglobin

After phagocytosis by macrophages, haem from senescent red cells is broken down by haem oxygenase (HMOX1) to release iron (Chapter 8). As ferrous iron, it can then either enter ferritin (where it is oxidized to ferric iron by the ferritin protein) or be released into plasma (via ferroportin), where its binding to transferrin (also as the ferric form) may be facilitated by a plasma ferrous oxidase (e.g. caeruloplasmin). The release of macrophage iron is controlled by hepcidin, with high levels, as in inflammation or iron overload, reducing iron release. Changes in hepcidin concentration may account for the diurnal rhythm of serum iron concentration, which is highest in the morning and lowest in the evening. A diurnal increase in serum hepcidin at noon and 8 p.m. is observed in healthy volunteers.

Diagnostic methods for investigating iron metabolism

The large amount of iron present as haemoglobin means that the degree of any anaemia must always be considered in assessing iron status. Reduced amounts of haemoglobin accompany an overall reduction in body iron in iron deficiency anaemia or after acute blood loss. In other anaemias, including the anaemia of chronic disease and most haemolytic and megaloblastic anaemias, iron is redistributed from the red cells to macrophage iron stores, with a corresponding increase in marrow-stainable iron and serum ferritin. No single measurement is ideal for all clinical circumstances, as all are affected by confounding factors and changes may develop sequentially (as in progressive negative iron balance) or may affect particular body iron compartments. Table 3.3 summarizes the changes in measures of iron status accompanying various types of hypochromic anaemia. The assessment of *iron overload* is discussed in Chapter 4.

Storage iron

Serum ferritin

In healthy subjects, the serum ferritin concentration correlates with iron stores. Normal concentrations of serum ferritin range from about 15 to 300 μg/L, and are higher in men (median about 90 μg/L) than in premenopausal women (median 30 μg/L). In neonates, the concentration in cord blood (median approximately 100 μg/L) rises further over the first 2 months of life as fetal haemoglobin is broken down, and thereafter falls to low levels (median 20–30 μg/L) throughout childhood and adolescence.

Serum ferritin concentrations below 15 μg/L are virtually specific for storage iron depletion, but normal values do not exclude this and values above 300 μg/L do not necessarily, or even usually, indicate iron overload. This is because ferritin synthesis is influenced by factors other than iron (in particular, it behaves as an acute-phase reactant in many inflammatory diseases). For this reason, serum ferritin concentrations below 100 μg/L may be associated with a lack of storage iron in patients with the anaemia of chronic disease (ACD).

Bone marrow aspiration

Staining the bone marrow for iron gives an indication of stores as well as erythroblast iron (Figure 3.6a). In iron deficiency anaemia, macrophage and erythroblast iron are absent (Figure 3.6b). In the anaemia of chronic diseases macrophage iron is present, but erythroblast iron is absent.

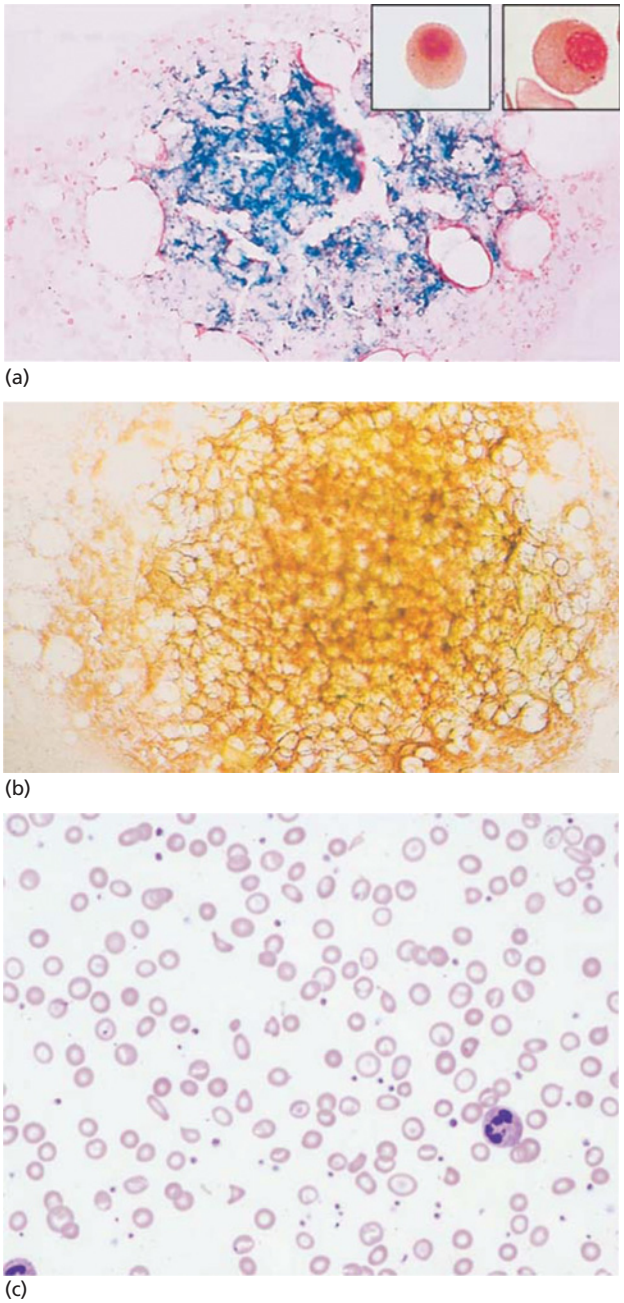


Figure 3.6 (a) Normal bone marrow showing plentiful iron in macrophages (Perls' stain) with iron granules in erythroblasts (insets). (b) Iron deficiency: bone marrow showing absence of stainable iron (Perls' stain). (c) Iron deficiency: peripheral blood film showing hypochromic microcytic red cells.

Iron supply to the tissues

Serum iron and iron-binding capacity

The serum iron and, more particularly, the saturation of the total iron-binding capacity of transferrin (TIBC) give a measure of the iron supply to the tissues. A serum transferrin saturation less than 15% is insufficient to support normal erythropoiesis. A rise in TIBC is characteristic of iron deficiency. A reduced serum iron concentration with a normal or reduced TIBC is a characteristic response to inflammation (see below).

Serum transferrin receptors

Plasma concentrations reflect both the number of erythroid precursors and iron supply to the bone marrow, TRs being raised in erythroid hyperplasia and in iron deficiency. In ACD, the assay has been proposed as a valuable indicator of deficiency of body iron stores. Serum TRs only increase in this situation in the absence of storage iron. The test, however is not validated for clinical use.

Red cell protoporphyrin

When iron supply to the erythron is limited, iron incorporation into haem is restricted, leading to accumulation of the immediate precursor, protoporphyrin IX bound to zinc. This is lost only slowly from circulating red cells; concentrations greater than the normal upper limit of 80 μmol/mol haemoglobin therefore indicate that a reduction in iron supply has been present over the previous few weeks. Protoporphyrin levels may also increase in

patients with sideroblastic anaemias and lead poisoning. The test is not uniformly available, but analysers that are portable and easy to operate are useful in large-scale field studies of iron deficiency anaemia as an initial screening test.

Percentage of hypochromic red cells

As iron supply to the erythron diminishes, the new red cells produced are increasingly hypochromic. Assessment of the haemoglobin content of individual red cells, which is possible using some automated cell counters, allows measurement of the percentage of hypochromic cells. Values rising to above 6% may help in the early identification of impaired iron supply in patients with chronic renal failure who are receiving treatment with recombinant erythropoietin, when associated inflammatory disease means that other measures of iron status can be misleading. Measurement of the reticulocyte content of haemoglobin is now also used as a test of functional iron status and is discussed next.

Reticulocyte haemoglobin content

Measurement of the reticulocyte content of haemoglobin is now used as a test of functional iron status. This parameter (CHr or Re-He according to the type of automated counter used) is useful in screening iron status, particularly for the assessment of iron-deficient erythropoiesis, as in dialysis patients treated with erythropoietin. Moreover, CHr may serve as a predictor of the response of anaemia to iron treatment. When response to treatment is favourable, an increase in CHr may be discerned within a few days of starting treatment before overall mean corpuscular haemoglobin (MCH) has changed.

Iron deficiency anaemia

An attempt to evaluate the global anaemia burden in the last 20 years showed that, notwithstanding some improvement in the most recent years, iron deficiency remains the top-ranking cause of anaemia worldwide, especially in developing countries, due to a combination of inadequate iron intake and blood loss from hookworm, schistosomiasis infestation and other causes of haemorrhage.

Sequence of events

Depletion of iron stores

When the body is in a state of negative iron balance, the first event is depletion of body stores, which are mobilized for haemoglobin production. Iron absorption is increased when stores are reduced, before anaemia develops and even when the serum iron level is still normal, although the serum ferritin will have already fallen.

Iron-deficient erythropoiesis

With further iron depletion, manifested by a serum ferritin below 15 μg/L and a fall in serum transferrin saturation to less than 15%, iron-deficient erythropoiesis develops with increasing concentrations of serum transferrin receptor and red cell protoporphyrin. At this stage, the haemoglobin, mean corpuscular volume (MCV) and MCH may still be within the reference range, although they may rise significantly when iron therapy is given.

Iron deficiency anaemia

If the negative balance continues, frank iron deficiency anaemia develops. The red cells become obviously microcytic and hypochromic (Figure 3.6c), and poikilocytosis becomes more marked. The MCV and MCH are reduced, and target cells may be present. The reticulocyte count is low for the degree of anaemia. The serum TIBC rises and the serum iron falls, so that the percentage saturation of TIBC is usually less than 10%.

The erythroblasts, devoid of siderotic granules, have a ragged vacuolated cytoplasm and relatively pyknotic nuclei. The bone marrow macrophages show a total absence of iron, except where very rapid blood loss outstrips the ability to mobilize the storage iron. Platelets are frequently increased.

Tissue effects of iron deficiency

When iron deficiency is severe and chronic, widespread tissue changes may be present, including koilonychia (ridged or spooned nails, breaking easily), hair thinning, angular stomatitis, glossitis and pharyngeal webs (Paterson–Kelly syndrome). Partial villous atrophy, with minor degrees of malabsorption of xylose and fat, reversible by iron therapy, has been described in infants suffering from iron deficiency, but not in adults. Pica is sometimes present; in some who eat clay or chalk, this may be the cause rather than the result of iron deficiency.

Iron-dependent enzymes in the tissues are usually better preserved than other iron-containing compounds. In severe iron deficiency, however, these enzymes are not inviolate and their levels may fall. This may be partly responsible for the general tissue changes, poor lymphocyte transformation and diminished cell-mediated immunity, and impaired intracellular killing of bacteria by neutrophils.

A particular concern has been the finding that infants with iron deficiency anaemia may have impaired mental development and function, and that this deficit may not be completely restored by iron therapy. It remains controversial whether impaired work performance seen in adults results from the anaemia or from depletion of mitochondrial iron-containing enzymes. It is also unclear to what extent some of the other tissue effects of iron deficiency can occur even in the absence of anaemia. One study has found that treatment with intravenous iron in premenopausal women with fatigue and serum ferritin <15 μg/L, but non-anaemic, significantly improved symptoms compared to a placebo-treated group. Routine prophylaxis with

Table 3.4 Causes of iron deficiency.

Dietary
Especially vegetarian diet
Malabsorption
Gluten-induced enteropathy (child or adult), gastrectomy, duodenal bypass, atrophic gastritis, IBD, chronic inflammation, heart failure, clay eating, etc.
Blood loss
Uterine: menorrhagia, postmenopausal bleeding, parturition
Gastrointestinal: oesophageal varices, hiatus hernia, *Helicobacter pylori*, peptic ulcer, aspirin ingestion, hookworm, schistosomiasis, hereditary telangiectasia, carcinoma of the stomach, caecum or colon, ulcerative colitis, angiodysplasia, Meckel diverticulum, diverticulosis, haemorrhoids, etc.
Renal tract: haematuria (e.g. renal or bladder lesion), haemoglobinuria (e.g. paroxysmal nocturnal haemoglobinuria, mechanical haemolyis)
Pulmonary tract: overt haemoptysis, idiopathic pulmonary haemosiderosis
Widespread bleeding disorders
Frequent blood donors
Self-inflicted (Munchausen syndrome)

ESA (erythropoiesis-stimulating agents) may cause functional iron deficiency.

iron of African children carries the risk of exacerbating malaria and is not now advised.

Causes of iron deficiency (Table 3.4)

Diet

Defective intake of iron is rarely the sole or major cause of iron deficiency in adults in Western communities. It takes more than 4 years for an adult male to develop iron deficiency anaemia on a diet completely devoid of iron. The diet may contain insufficient or poorly available iron as a result of poverty, religious tenets or food faddism. Iron deficiency is more likely to develop in subjects taking a largely vegetarian diet – the majority of the world's population – and who also have increased physiological demands for iron.

Increased physiological iron requirements

Iron deficiency is common in infancy, when demands for growth may be greater than dietary supplies. It is aggravated by prematurity, infections and delay in mixed feeding. Young children (<5 years) in low and middle income areas are the age group most affected, with little variation in prevalence of anaemia over the last 20 years, according to a recent large study. Iron deficiency is also frequent in adolescence and in females (Table 3.2). The gender gap in iron deficiency anaemia has widened in recent years on a worldwide basis. Pregnancy also favours iron deficiency. The fetus acquires about 280 mg of iron and a further 400–500 mg is required for the temporary expansion of maternal red cell mass. Another 200 mg of iron is lost with the placenta and with bleeding at delivery. Although iron absorption increases throughout pregnancy and increased requirements are partly offset by amenorrhoea, this may not be sufficient to meet the resultant net maternal outlay of over 600 mg iron. Routine iron prophylaxis is not recommended in pregnancy in the UK, but is given for anaemia (haemoglobin <110 g/L in first trimester, <105g/L in second or third trimesters and <100 g/L post partum).

Blood loss

Blood loss is the most common cause of iron deficiency in adults. A loss of more than about 6–8 mL of blood (3–4 mg iron) daily becomes of importance, as this equals the maximum amount of iron that can be absorbed from a normal diet. The loss is usually from the genital tract in women or from the gastrointestinal tract in either sex. The most common cause on a worldwide basis is infestation with hookworm, in which anaemia is related to the degree of infestation. Another important cause in developing countries that can affect any age is schistosomiasis, infection by *Schistosoma (S.) mansoni* causing bleeding from the gut and *S. haematobium* from the bladder, although in both cases anaemia of inflammation may coexist with iron deficiency. In the UK, menorrhagia, haemorrhoids and peptic ulceration are common, as well as gastric bleeding because of salicylates or other non-steroidal anti-inflammatory drugs, hiatus hernia, colonic diverticulosis and cancer (Table 3.4). Haemodialysis results in loss of 4–8 mg iron daily. Investigation often involves, particularly in older subjects, occult blood tests, upper and lower gastrointestinal endoscopy or computed tomography pneumocolon and if these are negative capsule endoscopy. Screening for colon cancer with serum tumour markers CEA and Ca19-9 is useful, but normal values do not exclude the possibility of colon cancer being present. Some unusual causes of blood loss deserve mention. Cow's milk intolerance in infants may lead to gastrointestinal haemorrhage. Self-induced haemorrhage may occur as an unusual form of Munchausen syndrome. Chronic intravascular haemolysis, such as that in paroxysmal nocturnal haemoglobinuria or mechanical haemolytic anaemia, may be a serious source of urinary iron loss.

Malabsorption

Malabsorption may be the primary cause of iron deficiency or it may prevent the body adjusting to iron deficiency from other causes. Dietary iron is poorly absorbed in gluten-induced enteropathy, in both children and adults. Gluten-induced enteropathy is encountered in about 5% of patients presenting with unexplained iron deficiency anaemia and, conversely, about 50% of patients with newly diagnosed coeliac disease have coexistent iron deficiency anaemia. Patients with this

Table 3.5 Human porphyrias.

Form	Inheritance	Enzyme defect	Clinical features*
Hepatic			
Acute intermittent porphyria	Autosomal dominant	Porphobilinogen deaminase	A
Hereditary coproporphyria	Autosomal dominant	Coproporphyrinogen oxidase	A + P
Porphyria variegate	Autosomal dominant	Protoporphyrinogen oxidase	A + P
Porphyria cutanea tarda	Acquired or (rare) autosomal dominant	Uroporphyrinogen decarboxylase	P
Erythropoietic			
Congenital erythropoietic porphyria	Autosomal recessive	Uroporphyrinogen cosynthase	P
Erythropoietic protoporphyria	Autosomal dominant	Ferrochelatase	P

*Acute attacks (A) of the gastrointestinal and/or nervous system are related to the accumulation of porphyrin precursors (δ-aminolaevulinic acid and porphobilinogen). Photosensitive skin lesions (P) are seen when the level of the enzyme defect in the haem synthetic pathway leads to the accumulation of formed porphyrins.

disease often show decreased or no response to oral therapy with inorganic iron.

Helicobacter pylori gastritis appears to be a common cause of iron deficiency, responding favourably to eradication with triple therapy. *H. pylori* gastritis inhibits gastric hydrochloric acid secretion, interfering with the solubilization and absorption of inorganic food iron, but it is also possible that gastrointestinal blood loss plays a significant role in the causation of the iron deficiency. Achlorhydria associated with autoimmune gastritis, an entity that may progress to pernicious anaemia, is an important cause of iron malabsorption due to impaired food iron solubilization. It is encountered in about 20% of patients with unexplained or refractory iron deficiency anaemia, mostly women of fertile age in whom achlorhydria aggravates the consequences of menstrual blood loss. Inflammatory bowel diseases are conditions in which iron deficiency may coexist with anaemia of inflammation.

Iron deficiency occurs in congestive heart failure due to malabsorption and iron loss and these patients may also have the features of the anaemia of chronic disorders.

Management of iron deficiency

Management entails: (1) identification and treatment of the underlying cause and (2) correction of the deficiency by therapy with inorganic iron. Iron deficiency is commonly due to blood loss and, wherever possible, the site of this must be identified and the lesion treated.

Oral therapy

In most patients, body stores of iron can be restored by oral iron therapy. Iron is equally well absorbed from several simple ferrous iron salts, and as ferrous sulfate is the cheapest, this is the drug of first choice – 200 mg of ferrous sulfate contains 67 mg of iron. A maximum of one tablet is given three times daily with 6 hours between doses. Where smaller doses are required to avoid side-effects, 300 mg of ferrous gluconate provides 36 mg of iron. It is usual to give up to 200 mg of elemental iron each day to adults and about 3 mg/kg per day as a liquid iron preparation to infants and children. The side-effects of oral iron, such as nausea, epigastric pain, diarrhoea and constipation, are related to the amount of iron taken. If iron causes gastrointestinal symptoms, these can usually be ameliorated by reducing the dose or taking the iron with food, but this also reduces the amount absorbed. Enteric-coated and sustained-release preparations should not be used, as much of the iron is carried past the duodenum to sites of poor absorption. Iron reduces absorption of tetracyclines (and vice versa) and of ciprofloxacin.

The minimum rate of response should be a 20 g/L rise in haemoglobin every three weeks, and the usual rate is 1.5–2.0 g/L daily. This will be slower when the dose tolerated is less than 100 mg/day, but this is seldom of clinical importance. It is usually necessary to give iron for 3–6 months to correct the deficit of iron in circulating haemoglobin and in stores (shown by a rise in serum ferritin to normal).

Failure to respond to oral iron is most commonly due to the patient not taking it, although there may be continued haemorrhage or malabsorption. In non-responding patients it is important to reassess the diagnosis to exclude other causes of microcytic anaemia such as iron-loading anaemias. For instance, many patients with thalassaemia trait, sideroblastic anaemia or other anaemias have been treated with iron before haemoglobin studies, bone marrow examination or other tests have revealed the correct diagnosis. A poor response may also be obtained if the patient has an infection, renal or hepatic failure, an underlying malignant disease or anaemia of inflammation and any

other cause of anaemia in addition to iron deficiency. Iron supplementation in children in areas endemic for malaria may worsen the infection, especially cerebral malaria.

Parenteral iron therapy

This is indicated in subjects who genuinely cannot tolerate oral iron, particularly if gastrointestinal disease, such as inflammatory bowel disease, is present. It is also occasionally necessary in gluten-induced enteropathy and when it is essential to replenish body stores rapidly (e.g. where severe iron deficiency anaemia is first diagnosed in late pregnancy) or when oral iron cannot keep pace with continuing haemorrhage (e.g. in patients with hereditary haemorrhagic telangiectasia). Patients with chronic renal failure who are being treated with recombinant erythropoietin also require parenteral iron therapy. In this situation, the demand for iron by the expanded erythron may outstrip the ability to mobilize iron from stores, leading to a 'functional' iron deficiency. Increased red cell loss at dialysis contributes to iron needs and oral iron therapy is usually inadequate to prevent an impaired response to erythropoietin. Serum ferritin in this setting is an unreliable indicator of iron deficiency. High hepcidin levels because of reduced clearance and inflammation contribute to erythropoietin resistance in renal failure. This occurs in other chronic inflammatory conditions: e.g. anaemia of cancer when treated with erythropoietin, requires parenteral iron as well. Intravenous iron has been shown to benefit functional capacity and quality of life in patients with congestive heart failure and iron deficiency, even in the absence of anaemia.

From all parenteral preparations, the iron complex is taken up by macrophages, from which iron is released to circulating transferrin, which then transports it to the marrow. High-molecular-weight iron dextran has been largely abandoned because of potentially severe side-effects, including anaphylaxis. A low-molecular-weight iron dextran, available in the UK as Cosmofer, is relatively safe and can provide all iron necessary (up to 1000 mg) in one infusion. An iron–sucrose complex, Venofer, is given by intravenous infusion to a maximum of 200 mg in 1–2 hours and not repeated in less than 48 hours. The deficit in body iron should be calculated from the degree of anaemia; it is usually 1–2 g. In patients receiving erythropoietin treatment in chronic renal failure, smaller intravenous doses of Venofer (25–150 mg/week) may be used, with regular monitoring of serum ferritin to avoid iron overload. Ferrinject is a macromolecular iron(III)-hydroxide carbohydrate complex (molecular weight approximately 150 kDa). It can be administered as an intravenous bolus (maximum single dose 200 mg) or slow infusion (maximum single dose 1000 mg over 15 minutes). Iron isomaltoside 1000 is given by intravenous infusion or slow intravenous injection.

Headache, flushing, nausea, skin rashes, urticaria, shivering, general aches and pains, dyspnoea and syncope are possible immediate adverse effects of intravenous iron. Delayed reactions include arthralgia, fever and lymphadenopathy and can persist for several days.

Iron refractory iron deficiency anaemia

Homozygous or doubly heterozygous mutations of matriptase-2 are a cause of iron refractory iron deficiency anaemia. The mutations may affect different conserved domains of the protein, including the trypsin-like serine protease domain. The patients show a microcytic hypochromic anaemia due to inappropriately high hepcidin levels and low serum iron and percentage saturation of iron-binding capacity. The patients absorb iron poorly and are refractory to oral iron therapy, but are partially responsive to parenteral iron.

A microcytic hypochromic anaemia with liver iron overload has also been described in a few patients with homozygous or doubly heterozygous mutations of DMT1. Liver iron stores are increased, but erythroid iron utilization is impaired and serum hepcidin levels are low for the degree of iron overload. These patients may respond to erythropoietin injections.

Deficiency of serum transferrin due to mutations of the transferrin gene causes a hypochromic microcytic anaemia with tissue iron overload caused by increased plasma non-transferrin-bound iron and low hepcidin levels. Treatment has been with infusions of fresh frozen plasma or apotransferrin.

Deficiency of caeruloplasmin also causes a mild hypochromic microcytic anaemia with iron overload in the liver and progressive neurodegeneration. There is failure of ferroxidase activity, which impairs iron mobilization from stores.

Pathological alterations in haem synthesis

Porphyrias

These are a group of inherited or acquired diseases, each characterized by a partial defect in one of the enzymes of haem synthesis (Chapter 2). Increased amounts of the intermediates of haem synthesis accumulate, the disorders being classified by whether the effects are predominantly in the liver or the erythron (Table 3.5). A full discussion of these disorders is beyond the scope of this chapter, but those with a particular haematological overlap are mentioned briefly.

Congenital erythropoietic porphyria

This is a very rare autosomal recessive disorder that is due to reduced uroporphyrinogen III synthase activity. Most patients are compound heterozygotes for mutations in the uroporphyrinogen III synthase gene. A single case with a germline mutation of the X-linked erythroid-specific transcription factor GATA1 has been described. Large amounts of porphyrinogens accumulate, and their conversion by spontaneous oxidation to photoactive porphyrins leads to severe, and disfiguring, cutaneous photosensitivity and dermatitis, as well as a

haemolytic anaemia with splenomegaly. Increased amounts of uroporphyrin and coproporphyrin, mainly type I, are found in bone marrow, red cells, plasma, urine and faeces. The age of onset and clinical severity of the disease are highly variable, ranging from non-immune hydrops fetalis to a later onset in which there are only cutaneous lesions. Treatment, including avoidance of sunlight and splenectomy to improve red cell survival, is only partially effective. High-level blood transfusions to suppress erythropoiesis (combined with iron chelation therapy) have been used to reduce porphyrin production sufficiently to abolish the clinical symptoms. Allogeneic bone marrow transplantation has been successful.

Erythropoietic protoporphyria

This is the most common erythropoietic porphyria and is usually caused by an autosomal dominant inherited deficiency of ferrochelatase, which results in increased free (not Zn linked) protoporphyrin concentrations in bone marrow, red cells, plasma and bile. Bone marrow reticulocytes are the primary source of the excess protoporphyrin. This leaks from cells and is excreted in the bile and faeces. Molecular analysis of the ferrochelatase gene has revealed a variety of missense, nonsense and splicing mutations as well as deletions and insertions. The onset of the disease is usually in childhood.

Expression of the gene is variable, and photosensitivity and dermatitis range from mild or absent to moderate in degree. There is little haemolysis, but a mild hypochromic anaemia may occur, and accumulation of protoporphyrins can occasionally lead to severe liver disease. Treatment is by the avoidance of sunlight; β-carotene may also diminish photosensitivity. Iron deficiency should be avoided as this may increase the amount of free protoporphyrin.

Rare patients with a variant of erythropoietic protoporphyria have been discovered to have *gain-of-function* mutation of ALAS2.

Porphyria cutanea tarda

This is the most common of the hepatic porphyrias and occurs worldwide. The incidence in the UK has been estimated at 2–5 per million. Type I or 'sporadic' porphyria cutanea tarda (PCT) accounts for 80% of cases of PCT. The underlying metabolic abnormality is decreased activity of uroporphyrinogen decarboxylase (UROD) in the liver. Type II disease is an autosomal dominant disorder caused by mutations in the *UROD* gene. Type III disease is a rare familial form and appears to result from unknown inherited defects that affect hepatic UROD activity. There is a marked increase in porphyrins in liver, plasma, urine and faeces. The disease is characterized by photosensitivity and dermatitis. It is precipitated in middle or later life, more often in men than women, by factors such as liver disease, alcohol excess or oestrogen therapy. A modest increase in liver iron is a common feature.

Either the homozygous or heterozygous presence of the C282Y and H63D mutations in the *HFE* gene (Chapter 4) may predispose to the development of PCT. Prevalence of the C282Y mutation is increased in both sporadic (type I) and familial (type II) PCT. In the UK, only homozygosity for C282Y (found in about 25% of patients) is significantly more common than in the general population (0.7%). In southern Europe, where C282Y is much less common, the H63D mutation is associated with PCT. Iron is known to inhibit UROD. Removal of the iron by repeated phlebotomy is standard treatment, usually leading to remission.

Lead poisoning

Chronic ingestion of lead in humans causes an anaemia that is usually normochromic or slightly hypochromic. Red cell lifespan is shortened and there is a mild rise in reticulocytes, but jaundice is rare. Basophilic stippling is a characteristic, though not universal, finding and it is thought to be due to precipitation of RNA, resulting from inhibition of the enzyme pyrimidine 5′-nucleotidase. Siderotic granules, and occasionally Cabot rings, are found in circulating red cells. The bone marrow shows increased sideroblasts, in some patients with ring sideroblasts. Red cell protoporphyrin and coproporphyrin are raised, as is urinary excretion of ALA, coproporphyrin III and uroporphyrin I.

The cause of the anaemia appears to be multifactorial. Haemolysis, probably due to the blocking of sulfydryl groups with consequent denaturation of structural proteins, and damage to mitochondria, with defective haemoglobin production due to inhibition of the enzymes of haem synthesis, are the major factors.

Sideroblastic anaemia

The sideroblastic anaemias comprise a group of refractory anaemias (Table 3.6) in which there are variable numbers of hypochromic cells in the peripheral blood, with an excess of iron in the bone marrow; at least 15% of the developing erythroblasts in the primary forms of anaemia contain iron granules arranged in a ring around the nucleus (Figure 3.7). These ring sideroblasts (more than four perinuclear granules per cell and covering one-third or more of the nuclear circumference) are the diagnostic feature of the anaemia. Ring sideroblasts may comprise a small (< 15%) population of the erythroblasts in a wide variety of clinical disorders.

Inherited sideroblastic anaemias

These are rare disorders manifesting mainly in males, usually in childhood or adolescence, but occasionally presenting late in life, when they need to be distinguished from the more common acquired form termed 'refractory anaemia with ring sideroblasts'. Inherited sideroblastic anaemias may show exclusively

Table 3.6 Sideroblastic anaemias.

Hereditary
X-linked
Erythroid-specific ALAS2 mutations
Associated with spinocerebellar ataxia: mitochondrial ATP-binding cassette (*ABCB7*) mutations
Autosomal recessive
Thiamine responsive, THTR-1 (*SLC19A2*) mutations: DIDMOAD syndrome
Glutaredoxin-5 (*GLRX5*) mutations
Associated mitochondrial myopathy: pseudouridine synthase-1 mutations, YARS mutations
SLC25A38 mitochondrial transporter mutations
Mitochondrial
Pearson (marrow–pancreas) syndrome
Kearns–Sayre syndrome
Acquired
Primary
Refractory anaemia with ring sideroblasts (RARS)
Secondary
Drugs: isoniazid, pyrazinamide, cicloserine, chloramphenicol, penicillamine
Mitochondrial toxicity: alcohol, lead poisoning
Copper deficiency in systemic disease: carcinoma, rheumatoid arthritis

DIDMOAD, diabetes insipidus, diabetes mellitus, optic atrophy and deafness

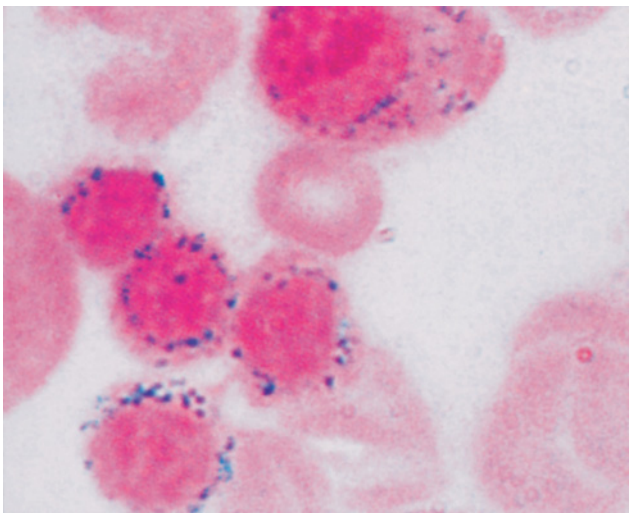


Figure 3.7 Sideroblastic anaemia. Erythroblasts showing perinuclear rings of iron (Perls' stain).

anaemia as the main symptom or present within a syndrome that includes anaemia in the context of other non-haematologic manifestations.

X-linked sideroblastic anaemia

ALAS2 mutations

In most reported families, inheritance has followed an X-linked pattern. More than 25 different mutations of the gene for erythroid-specific ALAS2 have been identified. All have been single-base substitutions. Most lead to changes in protein structure, causing instability or loss of function. They are found scattered over the exons encoding the catalytically active domain of the protein. Mutations affecting the promoter have also been shown to cause disease. Function may be rescued to a variable degree by administration of pyridoxal phosphate (the essential cofactor for ALAS2), the best responses occurring when the mutation affects the pyridoxal phosphate-binding domain of the enzyme. The response is better if iron overload is reduced by phlebotomy or chelation.

Female carriers of X-linked sideroblastic anaemia may show partial haematological expression, usually with only mild or no anaemia, although rarely a severe dimorphic anaemia occurs. This may depend on variation in the severity of the defect, as well as the degree of lyonization of the affected X chromosome. Late onset in some patients suggests that the degree of lyonization may change with age. Iron loading may also aggravate the defect in haem synthesis in both males and females with sideroblastic anaemia.

Patients with X-linked sideroblastic anaemia show a hypochromic, often microcytic, anaemia. There may be a few circulating siderocytes, normoblasts and cells with punctate basophilia, but these features become pronounced only if the spleen has been removed. The bone marrow shows erythroid hyperplasia and the erythroblasts tend to be microcytic with a vacuolated cytoplasm. There are more than 15% ringed sideroblasts. The ineffective erythropoiesis is not usually accompanied by bone deformities, but some bossing of the skull and enlargement of the facial bones may result from the erythroid expansion. The spleen may be enlarged. Patients may present with severe iron overload even when the anaemia is relatively mild, but the rate of iron loading is accelerated if red cell transfusions are needed. Iron loading, however, aggravates the anaemia.

ABCB7 mutations

A rare form of X-linked sideroblastic anaemia is caused by abnormalities in the ATP-binding cassette transporter gene (*ABCB7*). This form is associated with early-onset, non-progressive cerebellar ataxia. A useful diagnostic distinction is the presence within the red cells of increased zinc protoporphyrin, despite adequate iron stores, rather than the low/normal levels found in patients with abnormalities in ALAS2. The anaemia is mild to moderately severe. The enzyme deficiency

induces disruption in the maturation of cytosolic iron–sulfur clusters. The ataxia may be due to iron damage to mitochondria in neural cells.

Autosomal sideroblastic anaemia

Mitochondrial myopathy and sideroblastic anaemia

Mitochondrial myopathy, lactic acidosis and sideroblastic anaemia (MLASA) results from a homozygous mutation in the nuclear-encoded gene for pseudouridine synthase. As in Pearson syndrome, there are defects of the mitochondrial electron transport chain affecting reduced access to ferrochelatase. Mutations affecting the pseudouridylate synthase 1 (PUS1) gene cause deficient pseudouridylation of mitochondrial tRNAs. Mutations of the YARS2 gene product, mitochondrial tyrosyl-tRNA synthetase may cause the same MLASA phenotype through respiratory chain dysfunction.

Abnormalities of a high-affinity transporter of thiamine

Abnormalities of *SLC19A2* gene encoding a putative thiamine transporter (THTR-1) are responsible for thiamine-responsive megaloblastic anaemia (Roger syndrome). This syndrome, inherited in an autosomal recessive manner, may also combine diabetes insipidus, diabetes mellitus, optic atrophy and deafness (DIDMOAD), which respond in varying degrees to pharmacological doses of thiamine (vitamin B_1). Ring sideroblasts in varying numbers are typically present and onset is usually in childhood, although some symptoms may be present in infancy. A direct link between the presence of the mutation and mitochondrial iron loading has yet to be demonstrated.

Mutations of glutaredoxin-5

This enzyme participates in iron–sulfur cluster formation. Two families have been described with autosomal recessive hypochromic microcytic anaemia with ring sideroblasts with inherited mutations of the enzyme.

Mutations of Solute Carrier Family 25, Member 38 (SLC25A38)

This mitochondrial carrier is a transporter that may be involved in glycine transfer to the mitochondria, an essential step in synthesis of ALA. Mutation in congenital sideroblastic anaemia is described in several families from different ethnic groups. The degree of anaemia, which is microcytic and hypochromic, is usually severe since infancy and transfusion-dependent. Some patients have been cured by allogeneic bone marrow transplantation.

Mitochondrial DNA mutations

Deletions of mitochondrial DNA, sometimes associated with duplications, are known to be the cause of the Pearson marrow–pancreas syndrome, typically consisting of sideroblastic anaemia, pancreatic exocrine dysfunction and lactic acidosis. This is a severe disorder of early onset, presenting usually with failure to thrive, persistent diarrhoea and lactic acidosis. All haemopoietic cell lineages can be affected, and the anaemia is typically macrocytic with prominent vacuoles in cells of both myeloid and erythroid lineages.

Mitochondrial DNA has its own genetic code and encodes mitochondrial tRNA and ribosomal RNA as well as several mitochondrial proteins. In Pearson syndrome, deletions may encompass tRNA as well as mitochondrial genes and therefore have an effect on the function of all mitochondrion-encoded proteins, causing considerable loss of mitochondrial function. The presence of many different mitochondria within nucleated cells enable the coexistence of normal and abnormal species, the proportion of which is likely to vary within different tissues, a phenomenon known as *heteroplasmy*. The extent to which different tissues are affected depends to some extent on this proportion and detection often requires the study of different tissues. Inheritance is difficult to determine for the same reason and most cases are described as of 'sporadic' occurrence.

Different types of mitochondrial DNA deletion characterize also the Kearns–Sayre syndrome, which presents with ophthalmoplegia and cardiomyopathy and other developmental abnormalities.

Recently a new syndromic form of recessive sideroblastic anaemia was described in the context of immunodeficiency, periodic fevers, and developmental delay (SIFD), caused by mutations in transfer RNAs (tRNA)-nucleotidyltransferase 1 (*TRNT1*), the gene encoding the enzyme adding the nucleotide sequence CCA to the 3′ end of tRNA, essential for maturation of both nuclear and mitochondrial tRNAs.

There are still a number of cases of inherited sideroblastic anaemia in which the exact underlying genetic defect remains obscure, that may or may not show sex-linked inheritance and often show a macrocytic or dimorphic picture.

Acquired sideroblastic anaemia (see also Chapter 25)

Refractory anaemia with ring sideroblasts

This is a form of myelodysplasia and arises as a clonal disorder of haemopoiesis. The anaemia is often macrocytic with raised red cell protoporphyrin concentrations, in contrast to X-linked sideroblastic anaemia. Marked erythroid hyperplasia may be present, together with increased iron stores. In these patients, abnormalities in the white cell or platelet precursors are usually absent and the risk of transformation to acute myeloid leukaemia appears less than in other myelodysplastic disorders. Somatic mutations of the spliceosome SF3B1 gene have been found in more than 90% of patients with refractory anaemia with ring sideroblasts (RARS) and allow diagnosis by molecular test (Chapter 25). Smaller numbers (<15%) of ring sideroblasts may be present in patients with any of the other myelodysplastic and myeloproliferative diseases. Acquired defects of mitochondrial DNA may underlie iron transport abnormalities in refractory anaemia with ring sideroblasts.

Secondary sideroblastic anaemias

Sideroblastic anaemia may be found as a complication of antituberculous chemotherapy, particularly with isoniazid and cicloserine (pyridoxine antagonists). Sideroblastic anaemia occurs in alcoholism if there is associated malnutrition and folate deficiency. Suggested mechanisms include interference with haem formation and pyridoxine metabolism. The anaemia rapidly reverses with abstinence from alcohol, a normal diet and pyridoxine therapy. Chloramphenicol inhibits mitochondrial protein synthesis and in some patients causes ring sideroblast formation, presumably as a result of impaired haem formation in the mitochondria. Lead inhibits several enzymes involved in haem synthesis and may damage structural mitochondrial proteins. In some cases, ring sideroblasts are visible in the marrow. Ring sideroblasts may also occur in erythropoietic porphyrias.

Treatment

Pyridoxine

Some patients with X-linked sideroblastic anaemia respond to pyridoxine, given initially in doses of 100–200 mg daily. The response is usually partial. Patients may require only small doses (less than 10 mg daily) to maintain a higher haemoglobin concentration.

Pyridoxine therapy is almost always ineffective in RARS. However, some secondary sideroblastic anaemias may be completely reversed by pyridoxine therapy. This has been described in alcoholism, haemolytic anaemia and gluten-induced enteropathy, as well as in patients receiving antituberculous chemotherapy, in whom the drugs have been stopped and pyridoxine administered.

Other forms of treatment

Folic acid may benefit patients with secondary anaemia. For refractory patients, the anaemia may remain stable and, if the patient is transfusion independent, no treatment is needed. Patients requiring regular red cell transfusions require iron chelation therapy (see Chapter 4). Iron loading may aggravate the anaemia and, in some patients, improvement in the anaemia has followed iron removal by phlebotomy or iron chelation therapy. Splenectomy should usually be avoided, as it does not benefit the anaemia and leads to persistently high platelet counts postoperatively, with a high incidence of thromboembolism. Bone marrow transplantation may be proposed to severe non-syndromic forms of sideroblastic anaemia.

Selected bibliography

Bottomley SS, Fleming MD (2014) Sideroblastic anemia: diagnosis and management. *Hematology/Oncology Clinics of North America* **28**(4): 653–70.

Camaschella C. (2015) Iron deficiency anemia. *The New England Journal of Medicine* **372**(19): 1832–43.

Camaschella C (2013) Iron and hepcidin: a story of recycling and balance. *Hematology ASH Education Program* **2013**: 1–8.

Hershko C, Camaschella C (2014) How I treat unexplained refractory iron deficiency anemia. *Blood* **123**(3): 326–33.

Hoffbrand AV (ed.) (2009) Recent advances in the understanding of iron metabolism and iron-related diseases. *Acta Haematologica* **122**: 75–184, 12 articles covering all aspects of iron metabolism and related diseases.

Kassenbaum NJ, Jasrasaria R, Naghavi M *et al.* (2014) A systematic analysis of global anemia burden from 1990 to 2010. *Blood* **123**: 615–624.

Mastrogiannaki M, Matak P, Peyssonnaux C (2013) The gut in iron homeostasis: role of HIF-2 under normal and pathological conditions. *Blood* **122**(6): 885–92.

Meynard D, Babitt JL, Lin HY (2014) The liver: conductor of systemic iron balance. *Blood* **123**: 168–76.

Pavord S, Myers B, Robinson S *et al.* (2012) UK guidelines on the management of iron deficiency in pregnancy. *British Journal of Haematology* **156**: 588–600.

Thomas DW, Hinchcliffe RF, Briggs C *et al.* (2013) Guideline for the laboratory diagnosis of functional iron deficiency. *British Journal of Haematology* **161**: 639–48.

Weiss G, Schett G (2013) Anaemia in inflammatory rheumatic diseases. *Nature Reviews Rheumatology* **9**: 205–15.

CHAPTER 4

Iron overload

4

Clara Camaschella[1], A Victor Hoffbrand[2] and Maria Domenica Cappellini[3]

[1]Vita-Salute University and IRCCS San Raffaele Scientific Institute, Milan, Italy
[2]University College London, London, UK
[3]Foundation IRCCS Ca' Granda Policlinico and DISSCO University of Milan, Milan, Italy

Introduction

Excessive iron accumulation may lead to tissue damage. Iron overload of the parenchymal cells of the liver, endocrine organs and the heart commonly arises when there is excessive iron absorption, whereas iron administered parenterally (e.g. as multiple transfusions) is first taken up by macrophages, which destroy senescent red cells. However, there is no absolute distinction between the two sources of iron loading, as iron in macrophages is slowly released to transferrin, from which it is taken up by parenchymal cells. Causes of iron overload are shown in Table 4.1. Severe iron overload, arbitrarily defined as an excess of storage iron of more than 5 g, is confined to the hereditary forms of haemochromatosis, and the iron-loading anaemias.

Hereditary haemochromatosis

Hereditary haemochromatosis (HH) is now classified according to the genetic defect causing iron overload (Table 4.2). The vast majority of cases are of Type 1, involving the *HFE* gene (Chapter 3). In populations of northern European origin, most patients with HH are homozygous for the *HFE* pCys282TyrC mutation (pC282Y). All the other types, often indicated as non-HFE HH, are very rare. Types 1, 2 and 3 HH are autosomal recessive diseases and share common features due to hepcidin deficiency, including high transferrin saturation and hepatocyte iron accumulation. Type 4 HH, inherited as a dominant condition, is a heterogeneous disease with variable clinical phenotype. Digenic inheritance has been rarely described in patients who are heterozygous for *HFE* pC282Y and have a mutation in the one of the other genes.

HFE haemochromatosis

HFE HH is one of the most common genetic conditions in populations of northern European origin, associated with mutations in the *HFE* gene.

The *HFE* gene is an atypical MHC class I-like gene located at 6p21. The association of HH with HLA-A3 suggested a founder mutation in a chromosome carrying the A3 haplotype. Homozygosity for a G→A substitution at nucleotide 845 of the *HFE* gene results in a cysteine to tyrosine substitution at amino acid 282 (pC282Y). A second variant (187C→G) results in a histidine to aspartic acid substitution at amino acid 63 (pHis63Asp or pH63D).

HFE genotypes have been reported from large population studies throughout the world. In a study of 100,000 multiethnic participants from primary care practices, the estimated prevalence of the pC282Y homozygous genotype was higher in white (1 in 227, 0.44%) than in black individuals and was extremely rare in Asian people. The frequency is higher in northern Europe than in Italy and Greece. In the UK, about 1 in 8 people are carriers of the pC282Y mutation, and about 1 in 200 are homozygous for this mutation. Homozygosity is strongly associated with HH, with about 90% of patients with genetic iron overload having this genotype in northern Europe and about 60% in southern Europe. The HH gene may have increased in frequency because

Postgraduate Haematology, Seventh Edition. Edited by A Victor Hoffbrand, Douglas R Higgs, David M Keeling and Atul B Mehta.

Table 4.1 Causes of iron overload.

Severe iron overload (>5 g excess)
Excess iron absorption
Hereditary haemochromatosis
Massive ineffective erythropoiesis (e.g. β-thalassaemia intermedia, sideroblastic anaemia, congenital dyserythropoietic anaemia)
Increased iron intake
Sub-Saharan dietary iron overload (in combination with a genetic determinant of increased absorption)
Excess parenteral iron therapy
Repeated red cell transfusions
Congenital anaemias (e.g. β-thalassaemia major, sickle cell anaemia, red cell aplasia)
Acquired refractory anaemias (e.g. myelodysplasia, aplastic anaemia)

of a selective advantage for heterozygotes, protection against iron deficiency anaemia. Homozygotes would be unlikely to suffer the effects of iron overload before reproducing.

The pH63D mutation is found throughout the world with a mean allele frequency of 15%. A small proportion of HH patients are compound heterozygotes for the pC282Y and pH63D mutations. Other *HFE* gene mutations associated with iron accumulation have been described, mostly in individual families.

That *HFE* was the HH gene was confirmed by the demonstration that *HFE* knockout mice and mice homozygous for the pC282Y mutation develop iron overload. In addition, tissue-specific inactivation of HFE protein in hepatocytes, but not in duodenal cells or macrophages, leads to iron overload, indicating a specific function of HFE in the liver. The more recent model of the pathophysiology of the disease suggests that HFE competes with diferric transferrin for TFR1 binding and that when diferric transferrin binds TFR1, HFE binds TFR2, thus forming a complex that activates hepcidin synthesis (Chapter 3). Low serum hepcidin explains the findings in the early stages of HH: increased iron absorption, a raised serum iron and transferrin saturation, iron accumulation in hepatocytes with a periportal distribution and a paucity of iron in macrophages.

Iron parameters and clinical status of patients with HFE mutations

Transferrin saturation and serum ferritin are significantly higher in pC282Y/pC282Y subjects than in other genotypes. Increased levels of iron parameters are present in 75–100% of males and 40–60% of females in different studies. In population surveys slightly, but significantly, higher values for serum iron and transferrin saturation have been found in heterozygotes for either pC282Y or pH63D compared with subjects lacking these mutations. The differences in serum ferritin levels are small and not significant, except in rare cases where mutations in other genes involved in iron metabolism may be present. In compound heterozygotes, and those homozygous for pH63D, significant iron accumulation is rare and liver fibrosis exceptional. The frequency of homozygosity or heterozygosity for pC282Y or pH63D mutations is not increased in patients with arthritis, diabetes and heart disease.

In pC282Y homozygotes, there is a gradual accumulation of iron, leading to tissue damage, which may present as cirrhosis of the liver, diabetes, hypogonadism, cardiomyopathy, arthritis and a slate-grey skin (melanin) pigmentation. Only the most severe forms develop cardiac involvement that in the early stages is characterized by diastolic dysfunction and arrhythmias, in the later stages by dilated cardiomyopathy. Hepatocellular carcinoma may develop up to 25% of established cases with cirrhosis. Most patients present between the ages of 40 and 60 years, but the clinical penetrance is low. Menstrual losses and pregnancies account for a generally delayed onset in women. Full phenotypic expression of the disorder is dependent on other factors, including dietary iron intake, blood donations or blood loss, and other genetic factors modifying the genotype. Alcohol is a definite risk

Table 4.2 Classification of hereditary haemochromatosis.

Type	Gene	Inheritance and phenotype	Severity	Incidence
1	*HFE*	AR, parenchymal iron overload	Highly variable	Common
2 (juvenile)	Hemojuvelin (*HFE2*)	AR, parenchymal iron overload	Severe	Rare
	Hepcidin (*HAMP*)	AR, parenchymal iron overload	Severe	Rare
3	*TFR2*	AR, parenchymal iron overload	Variable	Rare
4a	Ferroportin (*SLC11A3*)	AD, RE iron	Variable	Rare
4b	*SLC11A3* (mutations in binding site for hepcidin)	AD, parenchymal iron overload	Severe	Rare

AD, autosomal dominant; AR, autosomal recessive; RE, reticuloendothelial.
For gene symbols see Table 3.1.

factor for the development of cirrhosis in patients homozygous for pC282Y.

Subjects homozygous for *HFE* pC282Y compared with control subjects with normal *HFE* genotype have shown the same frequencies of haemochromatosis symptoms and complications such as lethargy, arthralgia and diabetes. There is, however, a small but significant increase in subjects with either raised serum transaminases or fibrosis/cirrhosis in the pC282Y homozygous group. Homozygosity for pC282Y is more frequent in patients with cirrhosis and hepatoma than in the general population. The relative risk for hepatocellular carcinoma in cirrhotic patients with HH is approximately 20, with an annual incidence of 3–4%. Monitoring the serum α-fetoprotein is needed in both untreated and treated cases.

There is evidence that ferritin concentration above 1000 μg/L is a risk factor for liver fibrosis with a prevalence of 20–45%. A serum ferritin level >1000 μg/L has 100% sensitivity and 70% specificity for identification of cirrhosis. These patients should undergo liver biopsy or magnetic resonance imaging (MRI) and Fibroscan evaluation (see below) and complete cardiac and endocrinological evaluation. The ferritin cut-off might be lower in the presence of cofactors for liver disease such as alcohol. Only about 50% of homozygous women have raised transferrin saturation, and progression through iron accumulation and tissue damage is usually, but not always, slower.

Population surveys have shown that less than 5% of subjects homozygous for pC282Y ever receive a diagnosis of HH. Most men who are homozygous for pC282Y will have a raised transferrin saturation before the age of 30 years; a proportion will have an elevated serum ferritin concentration, but only a minority will eventually develop fibrosis and cirrhosis of the liver. Studies in Australia and Norway found that 5% of homozygote males, but no females, showed cirrhosis. Time-dependent ferritin increase is slow.

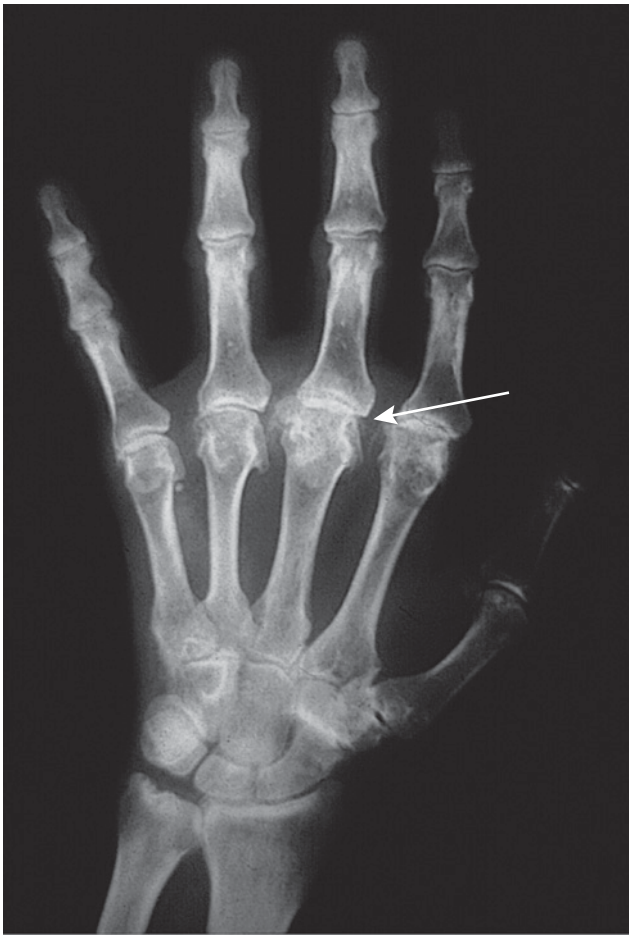


Figure 4.1 Radiograph of hand: patient with haemochromatosis showing loss of joint space and erosion of cartilage at the metacarpophalangeal joints (arrow).

Clinical and laboratory diagnosis

Although the biochemical penetrance of pC282Y homozygosity is substantial (about 70%), the disease penetrance is significantly lower (about 25%) among men and (1%) among women.

The variety of clinical presentations and their lack of specificity for HH mean that a high degree of clinical suspicion is needed. Fatigue, abdominal pain, diabetes mellitus, signs and symptoms of gonadal failure and arthritis may be present for several years before the diagnosis is made. Arthritis particularly affects the second and third metacarpophalangeal joints (Figure 4.1), and destructive arthropathy of hip and knee joints occurs in 10% of patients. There is chondrocalcinosis with pyrophosphate deposition in the joints. Abdominal pain may result from hepatic enlargement or hepatocellular carcinoma. Grey skin pigmentation results from excess melanin deposition.

In asymptomatic subjects, iron accumulation is indicated by a raised transferrin saturation (>45%). Most men and about 50% of women who are homozygous for *HFE* pC282Y will have a raised transferrin saturation. As iron accumulates, the serum ferritin concentration rises, and values in excess of 200 μg/L (women) and 300 μg/L (men) suggest iron overload. Serum ferritin concentrations largely reflect iron turnover in phagocytic cells and do not provide an early indication of iron accumulation in liver parenchymal cells. Thus, measurement of transferrin saturation is essential for early detection of iron loading. However, in patients with infection, inflammation or malignancy or in those undergoing surgery, transferrin saturation may be depressed and the serum ferritin concentration elevated. In most cases, *HFE* genotyping will confirm the diagnosis of HH.

Testing adult first-degree relatives of a pC282Y homozygous patient may identify subjects homozygous for *HFE* pC282Y with normal serum ferritin concentration and without evidence of liver disease: they should have transferrin saturation and serum ferritin measured at yearly intervals. Compound heterozygotes are at lesser risk of iron overload, but should also be tested by

measuring iron parameters. Widespread population screening by iron status or genetic testing is considered unwarranted as the level of risk for a pC282Y homozygote developing iron overload appears to be low. Screening in high-risk populations needs further studies.

Associations with other conditions

The pC282Y mutation is relatively common (see above); heterozygosity, and even homozygosity, may occur with other haematological conditions, including inherited sideroblastic anaemia, β-thalassaemia trait or other haemolytic anaemias and aggravate a tendency to iron load. The mutation may worsen liver disease in conjunction with hepatitis C infection, alcohol and non-alcoholic fatty liver disease. Porphyria cutanea tarda, which may be associated with iron overload, an important trigger of disease manifestations, is discussed in Chapter 3.

The *HFE* pC282Y mutation was identified in multiple genome-wide association studies for quantitative loci as erythrocyte and iron traits, blood pressure, HbA1C levels in diabetics and others, likely through indirect effects or due to the association of the gene with the HLA complex.

Body iron quantitation

Liver iron concentration is a good estimate of total body iron. Liver iron can be measured on a dried fragment of liver biopsy. The normal ranges are 0.17–1.8 mg/g dry weight. Values in excess of 3.0 mg/g dry weight indicate iron overload. Since the advent of genetic testing, liver biopsy is not usually performed. Iron quantitation in liver, heart and other organs is performed by non-invasive methods, mainly MRI techniques (see p. 47).

In patients homozygous for pC282Y with evidence of liver disease, hepatomegaly, increased transaminases and serum ferritin concentration >1000 μg/L, liver biopsy may be advisable to assess tissue damage. The biopsy reveals the hepatocyte iron accumulation with typical periportal predominance (Figure 4.2a). Fibroscan, a non-invasive technique valuable to assess and monitor liver fibrosis, is increasingly used instead of liver biopsy. In patients with an unexplained raised transferrin saturation and serum ferritin, who are not homozygous for pC282Y, a liver biopsy may be required to identify the different pattern of iron overload and to determine subsequent diagnostic tests.

The amount of iron removed to reach iron depletion can be calculated (see Treatment) and provides a good estimate of total body iron. The amount of iron removed at each venesection is calculated by weighing the blood bag before and after venesection (density of blood is 1.05 g/mL) and assuming that 450 mL of blood (haemoglobin concentration 13.5 g/dL) contains approximately 200 mg of iron. Iron absorption should be allowed for at the rate of 3 mg daily (20 mg/week). With these assumptions, 25 weekly venesections will remove 4.5 g of iron. The amount of storage iron measured by the technique in normal adults has been shown to be about 750 mg in men and 250 mg in women.

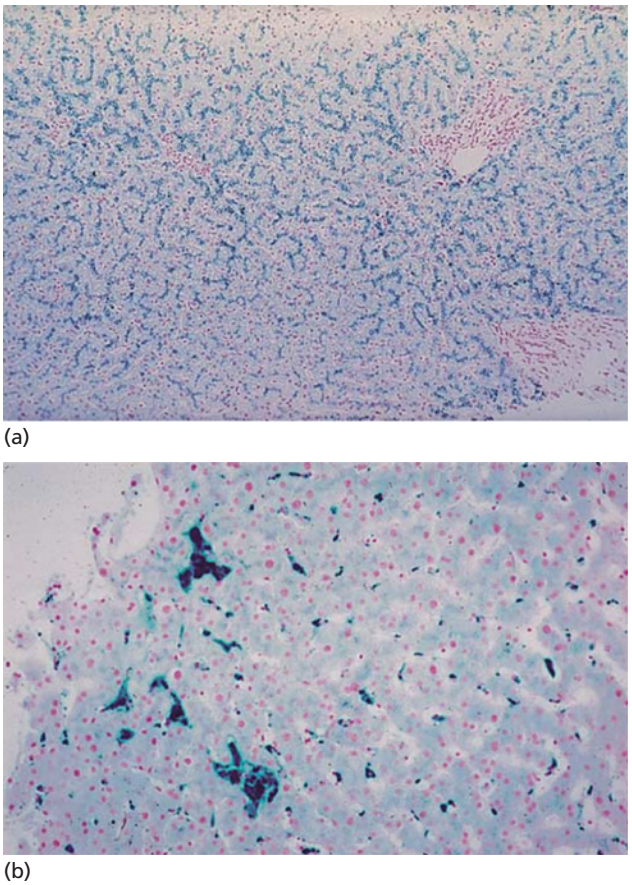


Figure 4.2 Liver histology (Perls' stain). (a) Liver biopsy from a patient with Type 1 haemochromatosis, showing staining predominantly in parenchymal cells. (b) Liver biopsy from a patient with Type 4a haemochromatosis, showing iron staining predominantly in Kupffer cells.

Treatment

Although no randomized controlled trial is available to compare phlebotomy versus no phlebotomy in the treatment of HH, retrospective studies show that removal of excess iron by regular phlebotomy greatly reduces the morbidity and mortality from cardiac and hepatic failure. However, hepatocellular carcinoma accounts for a substantial proportion of deaths in those with established cirrhosis. Early diagnosis is therefore a priority, as patients identified and treated before the onset of cirrhosis and/or diabetes have a normal life expectancy.

Since no test is available to identify those patients who will progress to fibrosis and those who will not, the present recommendation is to phlebotomize all patients with evidence of iron overload. No evidence-based guidelines for HH treatment by

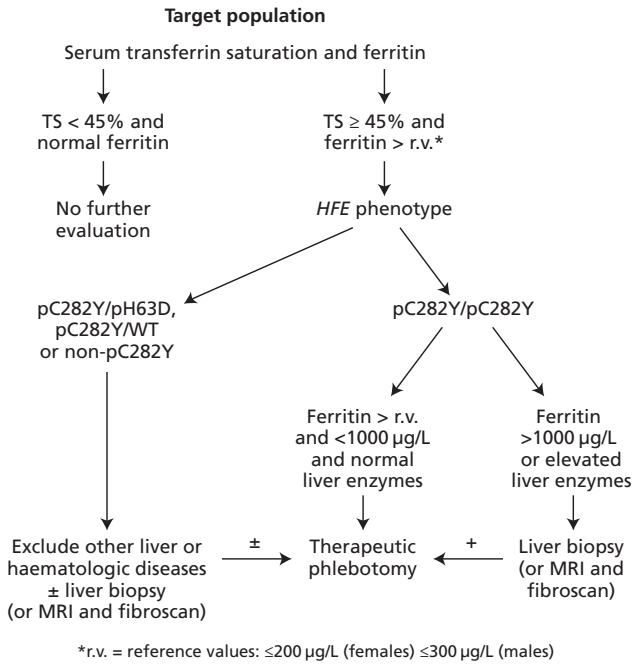


Figure 4.3 Algorithm for haemochromatosis diagnosis and treatment. The scheme is modified from the recommendations of the American Association for the Study of Liver Diseases (AASLD). TS = transferrin saturation. MRI = magnetic resonance imaging.

phlebotomy are available. Treatment is conventionally initiated when serum ferritin levels exceed the normal range. An algorithm for diagnosis and treatment of the disease based on the recommendations of the American Association for the Study of Liver Diseases is shown in Figure 4.3.

Phlebotomy should be at a rate of 400–450 mL of blood each week and is performed until iron depletion is reached (a target serum ferritin <50–100 µg/L is conventionally used). Haemoglobin levels should be measured weekly and the rate of venesection reduced if anaemia develops. Serum ferritin should be monitored monthly. Weekly phlebotomy will need to be continued to remove total iron excess, which is usually greater than 5 g in established symptomatic disease, but may be >20 g. When iron stores are exhausted, the frequency of phlebotomy should be reduced to two to four units each year, to continue indefinitely. The aim is to prevent re-accumulation of iron by maintaining a serum ferritin in the normal range (50–100 µg/L). Fatigue, abdominal pain, transaminase elevation, mild fibrosis and skin pigmentation usually reverse on venesection. In some patients, diabetes mellitus, hypogonadism and arthralgia improve, but cirrhosis and destructive arthritis are irreversible. Early cardiac disease may respond to phlebotomy, while severe cardiomyopathy requires iron chelation (see p. 48). Patients with end-stage liver disease and HCC may undergo orthotopic liver transplantation, after iron depletion. Post-transplantation survival rates in untreated patients are reduced compared with those of non-iron-loaded patients due to iron toxicity and infectious complications.

No diet recommendation is needed except limiting alcohol intake and avoiding raw shellfish for possible infection with *Vibrio vulnificus*, which is a particular risk in iron-loaded patients.

Iron chelation with subcutaneous desferrioxamine (DFO) (see below) may be used in patients who do not tolerate phlebotomy or have concomitant anaemia. DFO given as a continuous intravenous infusion with or without an oral iron chelator (see below) may have a role in the short-term management of patients with life-threatening cardiac failure. Oral iron chelators such as deferasirox at low dose are likely to be used for selected patients in the future. Targeting the hepcidin pathway to increase hepcidin production ameliorates iron overload in preclinical models of the disease.

Non-*HFE* haemochromatosis

Non *HFE* HH is a clinical definition applied to all the other types of HH (Table 4.2). These forms are all rare, accounting for <5% of all cases. In juvenile HH, clinical symptoms appear in the second and third decades of life. The disease can be due either to hemojuvelin mutations, that strongly decrease hepcidin production (see Chapter 3), or, more rarely, to mutations in the hepcidin (*HAMP*) gene. In both cases iron absorption is greater than in *HFE* HH and iron deposition occurs not only in the hepatocytes, but also in cardiac myocytes, pancreas and pituitary; symptoms of cardiomyopathy, diabetes and hypogonadism are prominent (Table 4.2). HH Type 3 is phenotypically similar to *HFE* HH, although it may present at an earlier age. It is due to mutations in the gene for transferrin receptor 2 (*TFR2*) (Table 4.2). All types of HH except Type 4 (Table 4.2) are recessive and characterized by low hepcidin secretion, the defect being the most severe in juvenile HH. Mutations in all of the genes of non-*HFE* HH (hemojuvelin, hepcidin, transferrin receptor 2, and ferroportin) have been found in rare patients from the Asia-Pacific region.

HH Type 4 is inherited as an autosomal dominant trait due to heterozygous missense mutations in the gene for the iron exporter ferroportin. All patients have increased serum ferritin levels, but some have normal transferrin saturation and a mixed pattern of iron accumulation, both in hepatocytes and macrophages. In the typical form (Type 4a, see Table 4.2), so-called ferroportin disease, due to *loss-of-function* ferroportins that are unable to target correctly the cell surface and so export iron from macrophages, the phenotype differs from HH and is similar to that found in the anaemia of chronic disease, with iron accumulation predominantly in the macrophages (Figure 4.2b) and normal or low transferrin saturation. In the rare Type 4b with mutations occurring in the binding site of ferroportin for hepcidin, *gain-of-function* ferroportins reach the cell surface,

but are resistant to the effect of hepcidin, causing increased iron export to plasma, saturation of transferrin and iron deposition in hepatocytes similar to *HFE* HH.

Genetic testing is required for diagnosis of non-*HFE* HH. Phlebotomy is indicated as in pC282Y homozygotes (Figure 4.3). Tolerance to phlebotomy should be carefully monitored and intensive regimens are not indicated in ferroportin disease (Type 4a in Table 4.2). Patients with juvenile forms of HH presenting with life-threatening cardiac failure should be treated with combination of iron chelators (see below).

Other causes of iron overload

Neonatal haemochromatosis

This is a condition that is recognized at birth, but may occur *in utero*. It is characterized by heavy parenchymal iron deposition in several organs and irreversible liver failure. No mutations in the known HH genes have been reported and a genetic cause has been more recently doubted in favour of an immunological pathogenesis, but heterogeneous causes cannot be ruled out. It has been proposed that the disease is due to an alloantibody, but the target antigen is unknown. Infusions of γ-globulin in pregnancy appear to reduce the severity of the condition. The other therapeutic option is liver transplantation.

African iron overload

African iron overload, a condition that occurs in sub-Saharan Africa (Bantu siderosis), results from the combination of a dietary component (a traditional beer that contains iron) and an unknown susceptibility gene unrelated to *HFE*. Iron deposition is of mixed type and occurs in both hepatocytes and macrophages and may lead to hepatic fibrosis and cirrhosis. Serum ferritin is usually elevated, but transferrin saturation may be normal. The iron overload is associated with a poor outcome in patients who also have tuberculosis, an infection that is highly prevalent in sub-Saharan Africa. A pGln248His mutation in ferroportin, a common variant that may be associated with a susceptibility to iron loading and mild anaemia has been reported in people of African origin.

Atransferrinaemia

This is a rare recessive genetic disorder associated with a severe hypochromic anaemia with, in some cases, excessive deposition of non-transferrin-bound iron (NTBI) in the parenchymal cells. In all cases tested, some iron-transferrin has been detected by iron-binding ability or immunologically, defining a type of hypotransferrinaemia. Complete absence of transferrin would presumably lead to fetal death. Treatment is based on plasma or transferrin infusions.

Chronic liver diseases

Moderate iron overload limited to the liver can be found in various chronic viral or metabolic liver diseases in the context of alcohol abuse, necro-inflammatory processes and in the so-called metabolic syndrome. Often these conditions are characterized by increased serum ferritin, but normal transferrin saturation and slightly increased hepcidin in the context of an inflammatory process.

Hereditary hyperferritinaemia-cataract syndrome

This condition is characterized by isolated hyperferritinaemia not associated with iron overload, early-onset bilateral cataracts and normal or low serum iron and transferrin saturation. It is usually due to heterozygous point mutations in the L-ferritin iron-responsive element so that a monoclonal ferritin is synthesized due to impaired negative feedback of ferritin synthesis. The mutation is in the gene promoter and constitutively increases gene transcription. High ferritin accumulates in the lens, causing cataracts.

Iron and neurodegeneration: Acaeruloplasminaemia

This is a rare recessive disorder in which there is a deficiency of ferroxidase activity as a consequence of mutations in the caeruloplasmin gene (Chapter 3). Clinically, the condition presents in middle age, with progressive degeneration of the retina and basal ganglia and with diabetes mellitus. Iron accumulates in the liver, pancreas and brain, with smaller amounts in the heart, kidneys, thyroid, spleen and retina. The serum iron is low and mild anaemia may be present. The total iron-binding capacity of transferrin (TIBC) is normal and ferritin is normal or raised. Unfortunately, no effective treatment is available to reduce neurodegeneration.

Acaeruloplasminaemia belongs to the group of rare genetic disorders that are classified as neurodegeneration with brain iron accumulation (NIBIA). In these conditions, as well as in Friederich ataxia and Parkinson disease, iron accumulation may be found in specific areas of the brain, irrespective of systemic iron. The therapeutic potential of low-dose iron chelation with deferiprone is being explored in prospective double-blind pilot clinical trials.

Iron-loading anaemias

In different forms of anaemia, iron overload occurs when iron intake is increased over a sustained period of time. This may be due either to red blood cell transfusions, as in thalassaemia major or transfusion-dependent thalassaemia (TDT), or to increased iron absorption through the gastrointestinal (GI) tract, because of inhibition of hepcidin synthesis by proteins released from erythroblasts (Chapter 3), as in non-transfusion-dependent thalassaemia (NTDT). Secondary iron overload may occur in patients transfused for other forms of inherited and acquired anaemias. It may be present in patients after successful allogeneic bone marrow or HSC transplantation

Table 4.3 Characteristics of desferrioxamine, deferiprone and deferasirox.

	Desferrioxamine (DFO)	Deferiprone (DFP)	Deferasirox (DFX)
Structure	Hexadentate	Bidentate	Tridentate
Molecular weight (Da)	560	139	373
Iron–chelator complex	1 : 1	1 : 3	1 : 2
Plasma clearance ($t_{1/2}$)	20 min	1–3 hours	1–16 hours
Absorption	Negligible	Peak 45 min	Peak 1–2.9 hours
Iron excretion	Urine + faecal	Urine	Faecal
Therapeutic daily dose	40 mg/kg	75–100 mg/kg	20–30 mg/kg
Route	Parenteral	Oral	Oral
Clinical experience	>40 years	>20 years	>10 years
Side-effects	Ototoxicity, retinal toxicity, growth defects, cartilage and bone abnormalities	Agranulocytosis, arthropathy, gastrointestinal disturbance, transient transaminitis, zinc deficiency	Skin rashes, gastrointestinal disturbance, rising serum creatinine

for various haematological disorders, who have been heavily transfused.

Iron chelation therapy is essential in patients with transfusion-dependent anaemia, to prevent death from iron overload, usually caused by cardiac failure or arrhythmia. Red blood cell requirements are about 160 mL/kg annually in non-splenectomized and 120 mL/kg annually in splenectomized TDT patients. The iron content of each transfusion is determined by volume (mL) × haematocrit × 1.16 mg. Intake of iron ranges from 0.32 to 0.64 mg/kg body weight daily. Patients with anaemias associated with increased iron absorption who are too anaemic to be venesected to remove iron, may also require iron chelation therapy.

The available iron-chelating drugs are: desferrioxamine (DFO), which is given by slow subcutaneous or intravenous infusion, deferiprone (DFP) and deferasirox (DFX), which are orally active. Their iron binding properties, pharmacokinetics and routes of elimination differ (Table 4.3).

Iron chelation therapy is monitored by:

1 tests of body iron burden;

2 tests of damage to the organs sensitive to iron overload (Table 4.4);

3 tests to detect potential side-effects of the particular chelating drug being used.

Tests of body iron burden

Serum ferritin

Serum ferritin generally relates to body iron stores and it is relatively inexpensive and easily measured repeatedly. Serum ferritin is useful in monitoring changes in body iron, although measures do not always predict body iron or trends in body iron accurately. This is partly because inflammation increases serum ferritin, while other factors affect it, such as vitamin C status, and partly because the distribution of liver iron between macrophages (Kupffer cells) and hepatocytes has a major impact on plasma ferritin. A sudden increase in serum ferritin should prompt a search for hepatitis, other infections, or inflammatory

Table 4.4 Monitoring for iron-induced organ damage.

Cardiac function
ECG ± exercise
24-hour monitoring
Echocardiography, MUGA ± stress test
Doppler echography, MRI

Liver structure and function
Liver function tests
Liver histology
Liver ultrasound
Fibroscan

Bone
Osteoporosis: bone density (Dexa scan)

Endocrine system
Diabetes: urine glucose, HbA_{1c}, glucose tolerance test, IGF-1
Growth and sexual development: sitting and standing height, Tanner staging, radiography for bone age, testosterone, estradiol, LH, FSH, SHBG, pulsatile GnRH release, sperm tests
Thyroid: T_4, TSH
Parathyroid: calcium, phosphate, PTH

ECG, electrocardiogram; FSH, follicle-stimulating hormone; GnRH, gonadotrophin-releasing hormone; HbA_{1c}, glycated haemoglobin; IGF, insulin-like growth factor; LH, luteinizing hormone; MRI, magnetic resonance imaging; MUGA, multigated acquisition scan; PTH, parathyroid hormone; SHBG, sex hormone-binding globulin; TSH, thyroid-stimulating hormone.

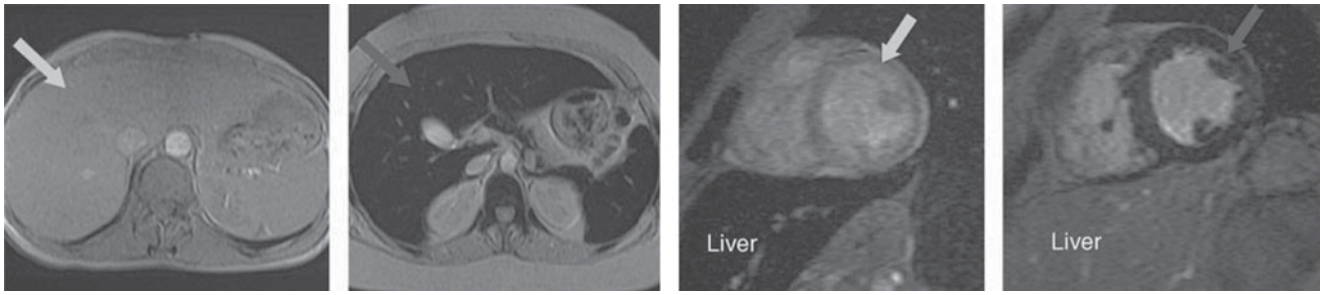


Figure 4.4 Magnetic resonance imaging T2* technique. Tissue appearances of liver and spleen: (a) normal; (b) tissue iron overload; (c) severe liver iron overload with normal cardiac iron; (d) severe cardiac iron deposition with minimal liver iron deposition. White arrows point to normal tissue, grey arrows to iron loaded tissue. Source (parts (c) and (d)): Anderson *et al*., 2001 [*Eur Heart J* 2001; **22**: 2171–9]. Reproduced with permission of Oxford University Press.

conditions. The Thalassaemia International Federation (TIF) guidelines recommend maintaining the level below 1000 µg/L in TDT.

Liver iron

Liver iron may be measured chemically after liver biopsy or by MRI. Chemical estimation was the gold standard, but can be inaccurate if fibrosis is present. Levels >7 mg/g tissue dry weight, and in non-transfusion-dependent anaemia >5 mg/g dry weight are almost always associated with organ damage, while levels <3 mg/g are not associated with clinical complication. Although no clear guidelines are available with liver iron between 3–7 mg/g dry weight, the target of chelation therapy should be to achieve a LIC <3 mg/g dry weight.

MRI techniques are being increasingly used as indirect measures of liver and cardiac iron (Figure 4.4). They have the advantage of being non-invasive and are now widely available. MRI is also the only practical method of performing sequential studies of iron in the liver and heart. Pituitary or other endocrine organs may also be evaluated by MRI, but validation of the software is still required. Different MRI techniques have been used. They all rely on a shortening of relaxation time and thus reduction in signal intensity with iron overload. The R2 or Ferriscan technique is convenient and reproducible for liver iron over the range of clinical interest. The normal ranges of liver iron content by Ferriscan are 0.17–1.8 mg/g dry weight. Gradient-echo imaging with the calculation of the T2* is also applicable, it has a short total imaging time, reducing movement artefacts. It is also more sensitive and reproducible and the best technique for the heart.

Cardiac iron

As cardiac failure or arrhythmia is the usual cause of death in transfusional iron overload, it is essential to monitor cardiac iron. Iron is deposited in myocytes and interstitial fibrosis develops. Direct measurement of cardiac iron by endomyocardial biopsy is inappropriate and inaccurate.

T2* cardiovascular magnetic resonance offers a reproducible (around 5% coefficient of variation between different observers or between two studies of the same patient), sensitive, albeit indirect measure: the lower the T2* value, the greater the cardiac iron (Figure 4.5). The majority of patients with T2* >20 ms have normal left ventricular (LV) function.

A T2* <20 ms correlates with the presence of cardiac dysfunction detected by echocardiography (left and right ventricles) or by 24-hour rhythm monitoring or the need for cardiac therapy (Figure 4.5). Most patients who develop cardiac failure have T2* <10 ms. The risk of developing heart failure increases with T2* values <10 ms, which is associated with 160-fold increased risk of heart failure in the next 12 months without adequate chelation. This risk may be substantially less in patients taking regular chelation (see below).

Poor correlation has been found between myocardial iron and liver iron (MRI derived) or serum ferritin in patients receiving iron chelation; thus liver iron and serum ferritin cannot be used as surrogate measures of cardiac iron in these patients.

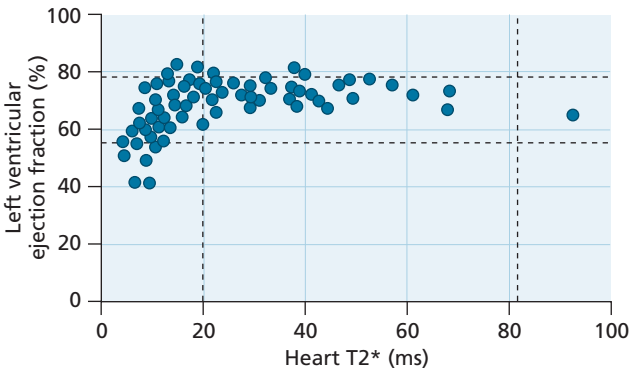


Figure 4.5 Relationships of myocardial T2* and left ventricular ejection fraction in patients with thalassaemia major and iron overload. Source: Anderson *et al*., 2001 [*Eur Heart J* 2001; **22**: 2171–9]. Reproduced with permission of Oxford University Press.

Urine iron excretion

Iron excretion after a single infusion of a standard dose of DFO or an oral dose of DFP is related to body iron. Urine iron is derived from the labile iron pool chelated mainly extracellularly with DFO and probably intracellularly with DFP. About half of total iron is excreted in urine by DFO or over 80% by DFP. With DFO (but not DFP), urine iron excretion is increased by ascorbate and is proportionately higher if the haemoglobin is lower. The test is useful when commencing therapy with DFO or DFP, with which iron excretion is highly dose related, and for monitoring therapy, in some studies correlating closely with liver and cardiac iron. However, several estimations must be performed at any given dose in view of the variability found.

Non-transferrin-bound iron (NTBI)

This is present in plasma in patients with gross iron overload and high transferrin saturation. It is highly toxic, promoting the formation of free radicals that cause peroxidation of membrane lipids. Part of the early improvement in liver and cardiac function with chelation therapy may be due to removal of this fraction, even before iron burden is substantially lowered. Its clearance by DFO is short-lived as it reappears in plasma within hours of stopping an infusion. DFX, with its long clearance time, provides 24-hour removal of NTBI after a single oral dose. DFP removes NTBI for about 3 hours after a single dose. NTBI is absent from plasma of well-chelated patients.

Tests for organ damage

The tests that are usually needed are listed in Table 4.4. Heart function is best tested by measurement of left ventricular ejection fraction (LVEF) by echocardiography or by MRI and by tests for rhythm disturbance. Liver function assessment requires routine liver function tests, as well as liver biopsy to assess liver structure, particularly in the presence of chronic viral hepatitis C. When MRI is available to measure the liver iron concentration, abdominal ultrasound and Fibroscan may be used to avoid liver biopsy. α-Fetoprotein should be tested regularly in patients with cirrhosis. The endocrine system is also damaged by iron and appropriate tests are listed in Table 4.4. The anterior pituitary is particularly sensitive, with damage resulting in reduced growth and impaired sexual maturation. Direct damage to the ovaries or testes may also occur, but is usually less important. Hypogonadic hypogonadism, defects of growth hormone secretion and its receptor, and deficiency of insulin-like growth factor mainly account for growth failure; DFO may also cause this. Diabetes mellitus and prediabetes, due to iron deposition in the pancreatic islets, are frequent, especially in patients with genetic susceptibility. Hypothyroidism and hypoparathyroidism are also common in poorly chelated patients. Osteoporosis is well recognized in iron-overloaded thalassaemia patients; it is due to multiple factors and is detected by bone density studies.

Iron chelation therapy

The available iron chelating drugs are DFO, DFP and DFX. Other drugs are currently under development.

Desferrioxamine (DFO)

Pharmacokinetics

DFO is not absorbed orally, and after parenteral injection is rapidly cleared from the plasma, being excreted in the urine, taken up by hepatocytes or metabolized in the tissues (Table 4.3). This accounts for the much greater mobilization of iron by continuous intravenous or slow subcutaneous infusions, which allow more prolonged exposure of the drug to the chelatable iron than with intramuscular injection. It is a trihydroxamic acid (hexadendate), one molecule binding covalently to all six oxygen sites on one ferric ion to form the red chelate, ferrioxamine which is excreted in urine and bile. Faecal iron is derived from hepatocytes. Urine iron also derives, at least partly, from hepatocytes, although other body sources, especially iron released from macrophages, contribute. Urinary iron excretion tends to level off at higher doses, but this does not occur with bile excretion so bile iron may predominate at high doses, and this is also the major route of excretion when total body iron has been reduced to relatively low levels. Increased erythropoiesis, as in haemolytic anaemias, is associated with an increase in urine iron excretion in relation to body iron stores.

Clinical studies

DFO was the first chelator used and shown to reduce complications and improve survival in TDT. Its main disadvantage is that it must be administered parenterally, which may often result in non-compliance. The increased toxicity of DFO at low levels of body iron means that guidelines have been conservative, generally recommending not starting therapy until serum ferritin levels reach 1000 ng/mL after approximately 12 units of blood and with care to avoid over-chelation, especially in children below this serum ferritin value.

Most studies have involved TDT, but patients with other inherited anaemias (e.g. Diamond–Blackfan syndrome, Fanconi anaemia, sickle cell anaemia, sideroblastic anaemia) or acquired disorders (especially myelodysplasia, myelofibrosis, red cell aplasia or aplastic anaemia) may require iron chelation therapy. In these conditions, as well as in elderly patients with acquired, transfusion-dependent, refractory anaemias and otherwise good prognosis, DFO was sporadically used if iron overload was likely to cause significant morbidity or mortality. In children, tissue damage from iron may be present from very early life; regular iron chelation should begin in TDT after transfusion of about 12 units of blood or when serum ferritin exceeds 1000 μg/L, but it should be started at 20 mg/kg to prevent tissue damage due to iron without causing toxicity due to excess DFO. A local anaesthetic cream (e.g. EMLA) reduces pain from the needle insertion. Oral chelation in preference to

DFO is often chosen by patients or their parents. The standard adult dose of DFO is 40 mg/kg administered subcutaneously given as an 8–12 hour infusion on at least 5 days each week. It is given intravenously at the time of blood transfusion only in patients not complying with any type of chelation and grossly iron overloaded. Vitamin C increases iron excretion by increasing the availability of chelatable iron, but if given in excess may increase the toxicity of iron. It is recommended not to give more than 2–3 mg/kg/day taken at the time of DFO infusion. The response to DFO chelation depends on the amount of blood and consequently of iron being transfused. Regimes of up to 50 mg/kg daily and 7 days a week may be needed.

For those with iron-induced cardiomyopathy, continuous intravenous DFO may be given via an indwelling catheter (e.g. Hickman) or Port-a-Cath chamber. Removal of liver iron is more rapid than removal of cardiac iron with this intensive chelation regimen. Combined therapy with DFP and DFX may also be indicated if there is iron-induced cardiac disease.

Growth and pubertal development are improved in many, but not all, patients; diabetes and other endocrine abnormalities still occur frequently and MRI studies have shown that about a third of TDT patients chelated with DFO have cardiac T2* values <10 ms. Through lack of compliance with DFO, premature deaths, usually from iron-induced cardiac damage may occur. Combination therapy or orally active drugs are improving the survival in these DFO failures.

Side-effects

These include rare generalized sensitivity reactions, local soreness related to the site of injection (usually due to the needle being inserted too superficially) and exacerbation of some infections, notably of the urinary tract and precipitation of *Yersinia* enterocolitis. Auditory (high-tone sensorineural hearing loss) and visual neurotoxicity (night blindness, visual field loss, retinal pigmentation) are relatively frequent. Growth and bone defects may also occur. The spine may be affected, with sitting height reduced; rickets-like bone lesions, *genu valgum* and metaphyseal changes are described, especially in children.

Auditory, visual and growth side-effects of DFO occur mainly if the body iron burden is low and doses of DFO high, particularly in children. A therapeutic index can be calculated as follows: mean daily dose (mg/kg)/current serum ferritin (μg/L). If this is below 0.025 at all times, these side-effects of DFO do not occur.

Deferiprone (DFP)

Pharmacokinetics

Deferiprone (1,2-dimethyl-3-hydroxypyrid-4-one) is an orally rapidly absorbed iron chelator, appearing in plasma within 15 min of ingestion (Table 4.3). The chelator–iron complex is excreted with the free drug and glucuronide derivative in urine. Its iron chelation site is inactivated by glucuronidation, the speed of which varies from patient to patient. This explains much of the individual variation in response. It is available both as tablets (500 mg) and as a liquid formulation containing 100 mg/mL, which is particularly useful in children. DFP mobilizes iron from parenchymal and reticuloendothelial pools and from transferrin, ferritin and haemosiderin. Its enhanced ability to cross cell membranes may underlie its superior ability compared with DFO to protect the heart from iron and also the 'shuttle' effect for iron when the two drugs are given simultaneously (Figure 4.6). It crosses the blood–brain barrier and has been used in trials to treat neurological conditions with iron loading in the brain.

Clinical studies

The usual dose used has been 75–100 mg/kg daily given as three divided doses. There are numerous non-randomized cohort studies demonstrating a lowering of serum ferritin at doses of 75 mg/kg/day administered in three doses. MRI studies suggest that liver iron may generally be higher in patients treated with DFP (75 mg/kg 7 days per week) compared with those treated with DFO (40 mg/kg 5 days per week) or DFX (30 mg/kg daily). In the liver, DFO has the advantage of facilitated transport into cells by an active mechanism. DFP, on the other hand, has greater penetration of myocardial cells. Retrospective and prospective

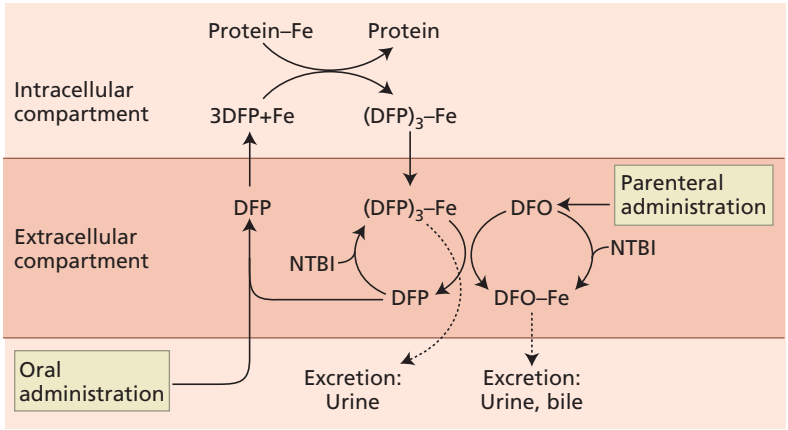


Figure 4.6 The concept of combination therapy: DFO, desferrioxamine; DFP, deferiprone; NTBI, non-transferrin-bound iron.

studies have shown, on the basis of T2* MRI measurement of cardiac iron, echocardiography, clinical incidence of cardiac disease, need for cardiac therapy and survival, that DFP 75 mg/kg is more effective than DFO at protecting the heart from iron-induced cardiomyopathy in routine clinical practice.

Side-effects

The most severe is agranulocytosis (neutrophils $<0.5 \times 10^9$/L on two consecutive counts), with an estimated incidence of around 0.5–1.0% (0.2–0.3 episodes per 100 patient-years). It is most frequent in the first year of therapy. Lesser degrees of neutropenia ($0.5–1.5 \times 10^9$/L) are more frequent, around 3.5%, occurring more frequently in non-splenectomized patients. Agranulocytosis and neutropenia spontaneously recover when the drug is immediately discontinued, usually within 4–28 days (median duration 9 days), but occasionally are more prolonged. Granulocyte colony-stimulating factor may produce a faster recovery. The mechanism appears to be idiosyncratic, more common in females, with no definite evidence of an immune mechanism established. Patients should be monitored by blood counts every week for at least the first 12 weeks of therapy and every 2 weeks thereafter for 2 years. Agranulocytosis may be more frequent in patients with Diamond–Blackfan syndrome. Deferasirox is safer for this disease and preferable for other conditions such as aplastic anaemia and myelodysplasia with pre-existing neutropenia.

Painful joints, especially the knees, occur in around 5–10% in most large series. The incidence has been highest in Indian patients. Some but not all studies show that this complication is most frequent in the most iron-loaded patients and with higher doses of DFP. It usually, but not invariably, resolves with withdrawal of the drug and it is often possible to reintroduce the drug, commencing with lower doses. About 2.0% of patients permanently discontinue therapy because of joint symptoms.

GI side-effects (e.g. nausea and abdominal pain) occur in about 30% of patients in the first year but decrease to 3% in subsequent years. In most, the drug can be reintroduced long term, initially at a lower dose. The liquid preparation appears to produce fewer GI symptoms. Zinc deficiency has been described in diabetic and prediabetic patients. Rarely, it can lead to clinical features, such as skin rashes and hair loss. It is easily treated by oral zinc therapy. Transient increases in liver enzymes have been associated with DFP therapy in about 7% of patients and about 1% of patients have been withdrawn from therapy because of a persistent rise in liver enzyme levels. There have been no reports of renal, cardiac or neurological side-effects. Embryo toxicity and teratogenicity have been reported in non-iron-loaded animals treated with deferiprone. Women of childbearing age should be counselled to avoid the drug or use contraception, but a few uneventful pregnancies with healthy newborns have been reported.

DFO and DFP combination therapy

Urine iron excretion when DFO and DFP are given simultaneously is equivalent to or greater than the sum of the excretion when the drugs are given on separate days. There is evidence for a 'shuttle' effect in which DFP enters cells, chelates iron and then returns to plasma, where the iron is transferred to DFO for excretion in urine or bile and DFP may re-enter cells (Figure 4.6). All studies of combination therapy, for example DFP on 7 days a week and DFO on 2 days, have shown a significant fall in serum ferritin and improvement in cardiac and liver iron over 6–18 months. It has been associated with improved survival and reversal of endocrinological complications including diabetes, hypothyroidism and hypogonadism. The combination has also proved successful in reversing severe myocardial siderosis. Alternating therapy has also been practised, for example 4 or 5 days of DFP and 2 or 3 days of DFO each week, with improved compliance and improved iron status in previously poorly compliant (with DFO) children or adults.

Deferasirox (DFX)

Pharmacokinetics

Deferasirox (DFX), 4-[3,5-bis(2-hydroxyphenyl)-1,2,4-triazol-1-yl]benzoic acid, is an oral tridentate chelator forming a 2:1 chelator–iron complex and increases predominantly faecal iron excretion. After a single oral dose, only 6% of iron excretion occurs in the urine (see Table 4.3). It is highly selective for iron. Peak plasma concentration after a single oral dose occurs at about 2 hours, and the drug is still detectable in plasma in almost all patients at 24 hours, with a mean elimination half-life of between 11 and 19 hours after multiple-dose administration. The single daily dose ranges from 20 to 40 mg/kg.

Clinical studies

DFX has been shown to be effective at eliminating NTBI in plasma and reducing serum ferritin and liver iron in heavily iron-loaded patients. The effect is dependent on dose and on transfusion requirements of the patient. The drug has been shown to be safe and effective in children as young as 2 years. The starting dose in children is 20 mg/kg daily with subsequent dose adjustments. In adults 30–40 mg/kg daily may be required, according to iron stores. Trials lasting for up to 5 years on children and adults have not shown any progressive renal, hepatic or bone marrow dysfunction and there are no reports of DFX having negative impact on growth or sexual development.

Emerging data suggest that DFX is effective at removing cardiac iron and preventing cardiac siderosis in TDT. It has also been shown to maintain or reduce iron overload in transfusion-dependent patients with myelodysplasia, Diamond–Blackfan anaemia and aplastic anaemia and in iron-loaded sickle cell anaemia patients. Adverse effects in these groups appear similar to those with TDT. The use of DFX is also being explored in

HH, chronic hepatitis C infection, porphyria cutanea tarda and mucormycosis.

Side-effects

The most common adverse effects have been abdominal pain, nausea, diarrhoea, vomiting and skin rashes. These decrease in frequency annually. They usually respond to dose adjustments, taking the drug in the evening or adding products such as Lactaid to the diet. Non-progressive increases in serum creatinine (defined as a rise above the mean pretreatment measurement by more than 33% on two consecutive tests) occur in about one-third of patients. These increases are dose-dependent and resolve spontaneously. Serum creatinine should be measured in duplicate before therapy and then monthly, with significantly increased levels managed by dose reduction or interruptions.

DFX and DFO combination therapy

Experience with this drug combination is relatively limited compared with that of DFO and DFP. Recent prospective studies of combined DFO and DFX therapy in patients with cardiac iron loading show it is effective at lowering cardiac iron and improving cardiac function as well as LIC in a follow-up of 2 years. No patients with severe myocardial iron loading (T2* values <10 ms) developed heart failure over a 2-year period while taking DFO and DFX combination.

Non-transfusion-dependent thalassaemia (NTDT)

Non-transfusion-dependent thalassaemia (NTDT) includes anaemic thalassaemic patients (Hb levels between 70 and 100 g/L) who do not require regular transfusion for survival, as the various genetic forms of β-thalassaemia intermedia, haemoglobin H disease and mild/moderate forms of HbE/β-thalassemia (Chapter 6). For these patients, who are not transfusion-dependent or only need a few transfusions each year, iron loading occurs mainly through increased iron absorption due to hepcidin suppression. Since venesection is not applicable in most cases because of the severity of the anaemia and danger of increasing bone marrow expansion and bone deformities, DFO has been occasionally used in the past. A recent randomized clinical trial using the oral iron chelator DFX has been shown to be effective in 'de-ironing' such patients, potentially reducing serum ferritin and liver iron to normal. Liver iron concentration should be targeted to below 5 mg/g dry weight to avoid iron toxicity related complications. It has been shown that in NTDT a serum ferritin level of 300 μg/L corresponds to approximately 3.0 mg LIC, while a serum ferritin level of 800 μg/L corresponds with a level of 5 mg/dry weight.

A rise in haemoglobin level may occur in iron-chelated patients with NTDT. This may be due to removal of iron from the renal oxygen sensor, augmenting the effect of hypoxia and increasing erythropoietin secretion from the kidney. The rise in haemoglobin may also be a result of reducing ineffective erythropoiesis and haemolysis. Improved haemopoiesis has also been described in low-risk myelodysplastic syndrome after intensive iron chelation (Chapter 25).

Acute iron poisoning

Acute oral iron poisoning produces a severe necrotizing gastritis and enteritis, followed by metabolic acidosis and, after a day or two, cardiovascular collapse and evidence of liver damage. DFO should be given parenterally. The instillation of 5 g into the stomach after a 1% sodium bicarbonate gastric lavage (to reduce further absorption) and an injection of 1–2 g intramuscularly may be tried. If a large number of tablets have been taken, an intravenous DFO infusion up to a maximum dose of 80 mg/kg in 24 hours should be used. DFP and DFX have not yet been used in this setting.

Selected bibliography

Angelucci E, Santini V, Di Tucci AA *et al.* (2014) Deferasirox for transfusion-dependent patients with myelodysplastic syndromes: safety, efficacy, and beyond (GIMEMA MDS0306 Trial). *European Journal of Haematology* **92**(6): 527–36.

Aydinok Y, Kattamis A, Viprakasit V (2014) Current approach to iron chelation in children. *British Journal of Haematology* **165**: 745–55.

Bacon BR, Adams PC, Kowdley KV, Powell LW, Tavill AS, American Association for the Study of Liver Diseases (2011) Diagnosis and management of hemochromatosis: 2011 practice guideline by the American Association for the Study of Liver Diseases. *Hepatology* **54**(1): 328–43.

Barton JC, Acton RT, So S *et al.* (2012) Increased risk of death from iron overload among 422 treated probands with HFE hemochromatosis and serum ferritin greater than 1000 mg/L at diagnosis. *Clinics in Gastroenterology and Hepatology* **10**: 412–16.

Cappellini MD, Porter JB, El-Beshlawy A *et al.* (2010) Tailoring iron chelation by iron intake and serum ferritin: prospective EPIC study of deferasirox in 1744 patients with transfusion-dependent anemias. *Haematologica* **95**(4): 557–66.

Ekanayake D, Roddick C, Powell LW (2015) Recent advances inhemochromatosis: a 2015 update: a summary of proceedings of the 2014 conference held under the auspices of Hemochromatosis Australia. *Hepatology International* **9**(2): 174–82.

Hoffbrand AV, Taher A, Cappelini MD (2012) How I treat transfusional iron overload. *Blood* **120**: 3657–69.

Pennell DJ, Udelson JE, Arai AE *et al.,* American Heart Association Committee on Heart Failure and Transplantation of the Council on Clinical Cardiology and Council on Cardiovascular Radiology and Imaging. (2013) Cardiovascular function and treatment in β-thalassemia major: a consensus statement from the American Heart Association. *Circulation* **128**(3): 281–308.

Shenoy N, Vallumsetla N, Rachmilewitz E, Verma A, Ginzburg Y (2014) Impact of iron overload and potential benefit from iron chelation in low-risk myelodysplastic syndrome. *Blood* **124**(6): 873–81.

Taher A, Vichinsky E, Musallam K, Cappellini MD, Viprakasit V (2013) Guidelines for the Management of Non Transfusion Dependent Thalassaemia (NTDT) [Internet]; Weatherall D, editor. Nicosia, Cyprus: Thalassaemia International Federation.

Taher AT, Temraz S, Cappellini MD (2013) Deferasirox for the treatment of iron overload in non-transfusion-dependent thalassemia. *Expert Reviews in Hematology* **5**: 495–509.

Aydinok Y, Kattamis A, Cappellini MD *et al.* (2015) HYPERION Investigators (2015). Effects of deferasirox-deferoxamine on myocardial and liver iron in patients with severe transfusional iron overload. *Blood* **125**(25): 3868–77.

Ware HM, Kwiatkowski JL (2013) Evaluation and treatment of transfusional iron overload in children. *Pediatric Clinics of North America* **60**(6): 1393–406.

CHAPTER 5

Megaloblastic anaemia

5

A Victor Hoffbrand

University College London, London, UK

Introduction

The megaloblastic anaemias are a group of disorders characterized by the distinctive morphological appearances of developing erythroid cells in the bone marrow. The cause is usually deficiency of either cobalamin (vitamin B_{12}) or folate, but megaloblastic anaemia may arise because of inherited or acquired abnormalities affecting the metabolism of these vitamins or because of defects in DNA synthesis not related to cobalamin or folate (Table 5.1).

Underlying basic science

Biochemical basis of megaloblastic anaemia

The common feature of all megaloblastic anaemias is a defect in DNA synthesis that affects rapidly dividing cells in the bone marrow and other tissues. All conditions that give rise to megaloblastic changes share in common a reduced rate of synthesis or of polymerization of the four immediate precursors of DNA: the deoxyribonucleoside triphosphates (Figure 5.1). In deficiencies of either folate or cobalamin there is a failure to convert deoxyuridine monophosphate (dUMP) to deoxythymidine monophosphate (dTMP). The coenzyme 5,10-methylene tetrahydrofolate polyglutamate is needed for this reaction and the availability of this co-enzyme is reduced in either deficiency (see further on).

The reduced supply of deoxythymidine triphosphate (dTTP) in megaloblastic anaemia due to folate or cobalamin deficiency slows elongation and ligation of newly originated replicating segments of DNA at multiple sites of origin during cell mitosis. Thus, small DNA fragments accumulate, single-stranded areas become points of weakness where mechanical or enzymatic breakage may occur, and the failure to form bulk DNA impairs contraction of newly replicated lengths of DNA, leaving the chromosomes elongated, despirillated and with random breaks. Late-replicating DNA is particularly affected and some cells become arrested at this stage and die by apoptosis. An alternative hypothesis for megaloblastic anaemia in cobalamin or folate deficiency is the misincorporation of uracil into DNA because of a build-up of deoxyuridine triphosphate (dUTP) at the replication fork as a consequence of the block in conversion of dUMP to dTMP (Figure 5.1). Data on this hypothesis are conflicting. It does not explain megaloblastic anaemia due to defects of DNA synthesis at sites other than at thymidylate synthesis, for example with drugs such as hydroxycarbamide (hydroxyurea), cytosine arabinoside or 6-mercaptopurine, or with enzyme deficiencies such as orotic aciduria or thiamine-responsive megaloblastic anaemia (see further on).

Cobalamin–folate relationship

Folate is required for many other reactions in mammalian tissues, including two in purine synthesis (Table 5.2), but impairment of these is far less important clinically. Only two reactions in the body are known to require cobalamin (Figure 5.2). Methylmalonyl-CoA isomerization, which requires deoxyadenosyl(ado)-cobalamin, is discussed later.

The methylation of homocysteine to methionine requires both 5-methyltetrahydrofolate (methyl-THF) as methyl donor and methylcobalamin as co-enzyme (Figures 5.1, 5.3). This reaction, which is almost completely irreversible, is the first step in the pathway by which methyl-THF, which enters bone

Postgraduate Haematology, Seventh Edition. Edited by A Victor Hoffbrand, Douglas R Higgs, David M Keeling and Atul B Mehta.

Table 5.1 Causes of megaloblastic anaemia.

Cobalamin deficiency or abnormalities of cobalamin metabolism (Table 5.4)
Folate deficiency or abnormalities of folate metabolism (Table 5.6)
Therapy with antifolate drugs (e.g. methotrexate)
Independent of either cobalamin or folate deficiency and refractory to cobalamin and folate therapy:
Some cases of acute myeloid leukaemia, myelodysplasia*
Therapy with drugs interfering with DNA synthesis (e.g. cytarabine, hydroxycarbamide, 6-mercaptopurine, azidothymidine)
Orotic aciduria (responds to uridine)
Lesch–Nyhan syndrome (? responds to adenine)

*Folate deficiency also occurs frequently in these diseases.

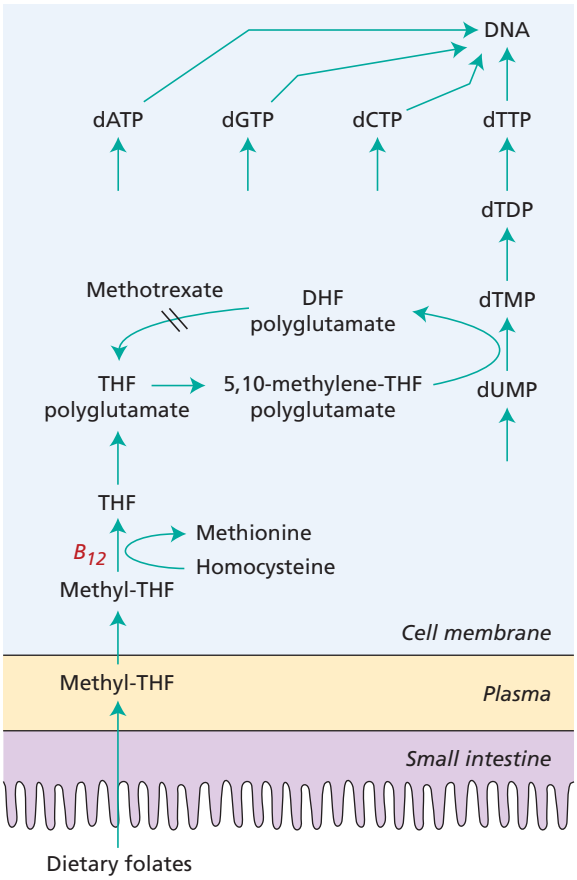


Figure 5.1 Role of folate (as 5,10-methylene-THF polyglutamate) and methylcobalamin in DNA synthesis. THF, tetrahydrofolate; MP, monophosphate; TP, triphosphate; d, deoxyribose; A, adenine; T, thymine; C, cytosine; G, guanine; B12, vitamin B12, cobalamin.

marrow and other cells from plasma, is converted into all the intracellular folate co-enzymes (Figure 5.1). The co-enzymes are all polyglutamated (the larger size aiding retention in the cell), but the enzyme folate polyglutamate synthase requires THF and not methyl-THF as substrate. In cobalamin deficiency, methyl-THF accumulates in the plasma, while intracellular folate concentrations fall due to failure of formation of intracellular folate polyglutamates including 5,10-methylene-THF because of 'THF starvation'.

This theory explains the abnormalities of folate metabolism that occur in cobalamin deficiency (high serum folate, low cell folate, reduced thymidylate synthesis, raised purine precursor AICAR excretion; Table 5.2) and also why the anaemia that occurs in cobalamin deficiency will respond to folic acid in large doses. The explanation of why serum cobalamin falls in folate deficiency may also be related to impairment of the homocysteine–methionine reaction, with reduced formation of methylcobalamin, the main form of cobalamin in plasma, but other mechanisms may be responsible.

Clinical features

Many symptomless patients are detected through the finding of a raised mean corpuscular volume (MCV) on a routine blood count. The main clinical features in more severe cases are those of anaemia. Anorexia is usually marked and there may be weight loss, diarrhoea or constipation. Other particular features include glossitis, angular cheilosis, a mild fever in the more severely anaemic patients, jaundice (unconjugated) and reversible melanin skin hyperpigmentation. Thrombocytopaenia sometimes leads to bruising (and this may be aggravated by vitamin C deficiency in malnourished patients. The (anaemia and) low leucocyte count may predispose to infections, particularly of the respiratory or urinary tracts. Cobalamin deficiency has also been associated with impaired bactericidal function of phagocytes.

General tissue effects of cobalamin and folate deficiencies

Epithelial surfaces

These deficiencies, when severe, affect all rapidly growing (DNA-synthesizing) tissues. After the marrow, the next most affected tissues are the epithelial cell surfaces of the mouth, stomach, small intestine and respiratory, urinary and female genital tracts. The cells show macrocytosis, with increased numbers of multinucleate and dying cells. The deficiencies may cause cervical smear abnormalities.

Complications of pregnancy

The gonads are also affected and infertility is common in both men and women with either deficiency if severe. Maternal folate

Table 5.2 Biochemical reactions of folate co-enzymes.

Reaction	Co-enzyme form of folate involved	Single-carbon unit transferred	Importance
Formate activation	THF	–CHO	Generation of 10-formyl-THF
Purine synthesis			
Formation of glycinamide ribonucleotide	5,10-Methenyl-THF	–CHO	Formation of purines needed for DNA, RNA synthesis, but reactions probably not rate limiting
Formylation of amino-imidazole-carboxamide-ribotide (AICAR)	10-Formyl-THF	–CHO	
Pyrimidine synthesis			
Methylation of deoxyuridine monophosphate (dUMP) to thymidine monophosphate (dTMP)	5,10-Methenyl-THF	$-CH_3$	Rate limiting in DNA synthesis Oxidizes THF to DHF Some breakdown of folate at the C-9–N-10 bond
Amino acid interconversion			
Serine–glycine interconversion	THF	$=CH_2$	Entry of single-carbon units into active pool
Homocysteine to methionine	5-Methyl-THF	$-CH_3$	Demethylation of 5-methyl-THF to THF; also requires cobalamin, flavine adenine dinucleotide, ATP and adenosylmethionine
Forminoglutamic acid to glutamic acid in histidine catabolism	THF	–HN–CH=	Basis of the Figlu test (now obsolete)

DHF, dihydrofolate; THF, tetrahydrofolate.

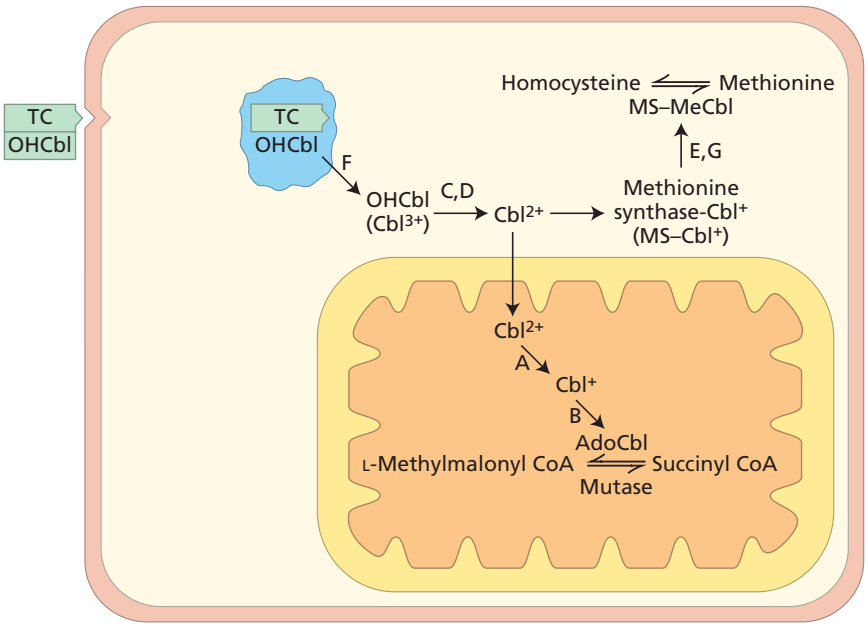


Figure 5.2 Intracellular cobalamin metabolism. Cbl^{1+}, Cbl^{2+} and Cbl^{3+} refer to the oxidation state of the central cobalt atom of cobalamin Cbl^{3+} being the most oxidized. A–G refer to the sites of blocks that have been identified by complementation analysis in infants with metabolic defects. AdoCbl, adenosylcobalamin; MeCbl, methylcobalamin; TC, transcobalamin. The mitochondrial, lysosomal and cytoplasmic compartments are indicated. Source: Lilleyman JS, Hann IM, and Blanchette VS (eds) (1999) *Paediatric Haematology*, 2nd edn. Churchill Livingstone, Edinburgh. Reproduced with permission from Elsevier.

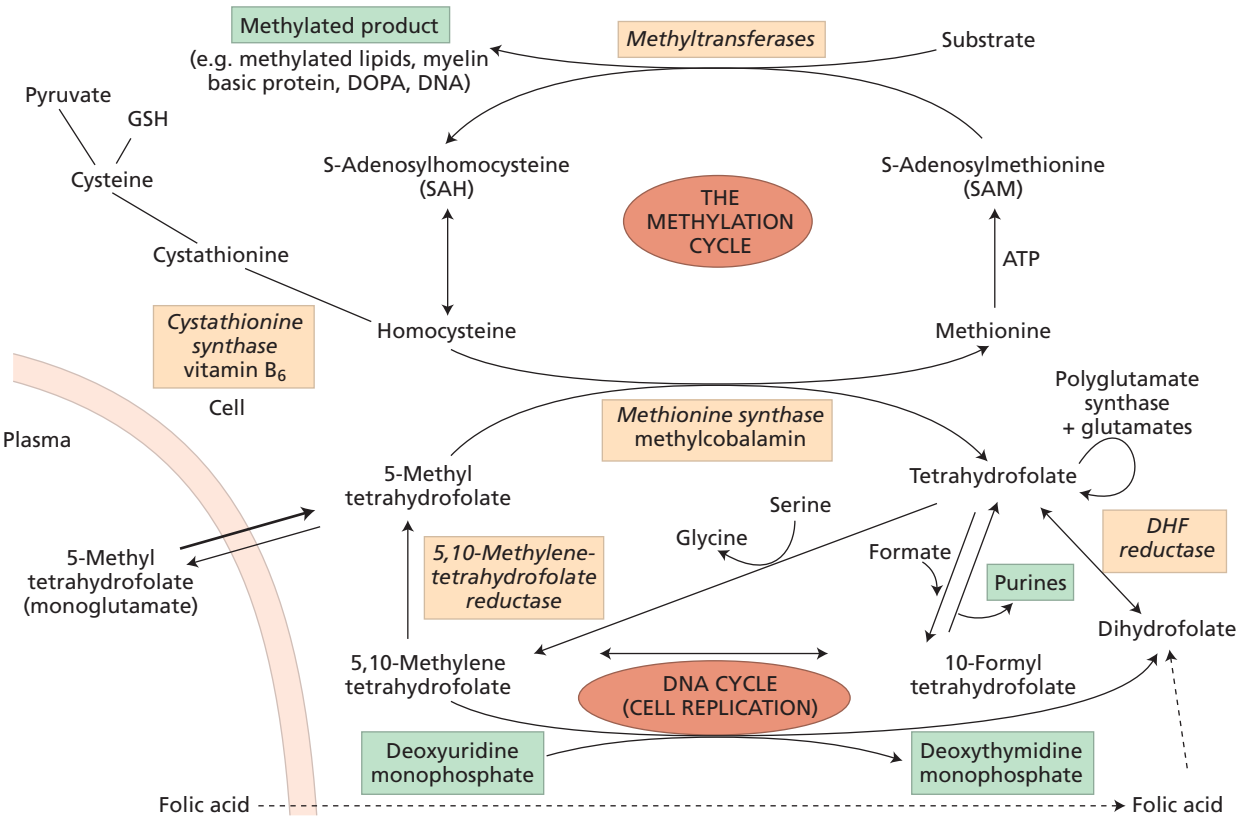


Figure 5.3 The role of folates in DNA synthesis and in formation of *S*-adenosylmethionine (SAM), which is involved in numerous methylation reactions. Enzymes are shown in pink boxes. (Figure prepared in conjunction with Professor John Scott.)

deficiency has been implicated as a cause of prematurity and both folate and cobalamin deficiency have been implicated in recurrent fetal loss.

Neural tube defects

Folic acid supplements at the time of conception and in the first 12 weeks of pregnancy reduce by about 70% the incidence of neural tube defects (NTDs) – anencephaly, meningomyelocele, encephalocele and spina bifida – in the fetus. Most of this protective effect can be achieved by taking folic acid 0.4 mg daily. The incidence of cleft palate and harelip can also be reduced by prophylactic folic acid. There is no clear simple relationship between maternal folate status and these fetal abnormalities, although the lower the maternal folate (and vitamin B_{12}) plasma levels, the greater the risk of NTD or, in diabetic mothers of other birth defects, in infants. NTDs can also be caused by antifolate and antiepileptic drugs.

An underlying maternal folate metabolic abnormality has also been postulated. One abnormality has been identified: reduced activity of the enzyme 5,10-methylenetetrahydrofolate reductase (MTHFR) (Figure 5.3) caused by a common 677C→T polymorphism in the *MTHFR* gene. In one study, the prevalence of this polymorphism was found to be higher in the parents of NTD fetuses and in the fetuses themselves: homozygosity for the TT mutation was found in 13% compared with 5% in control subjects. The polymorphism codes for a thermolabile form of MTHFR. The homozygous state results in lower mean serum and red cell folate compared with control subjects, as well as significantly higher serum homocysteine levels. Tests for mutations in other enzymes possibly associated with NTDs, have been negative.

Cardiovascular disease

Children with severe homocystinuria (blood levels of 100 µmol/L or more) due to deficiency of one of three enzymes, methionine synthase, MTHFR or cystathionine synthase (Figure 5.3; Chapter 44), suffer from vascular disease (e.g. ischaemic heart disease, cerebrovascular disease or pulmonary embolus) as teenagers or in young adulthood. Meta-analysis shows a significant association between lesser degrees of raised serum homocysteine (normal range 5–15 µmol/L) and of homozygosity for mutated *MTHFR* with ischaemic heart disease, stroke, deep vein thrombosis and pulmonary embolism.

It remains possible that homocysteine levels may be high as a consequence of vascular damage or may merely be a marker for some other underlying factor that is responsible for both the

vascular damage and the raised homocysteine. Folate deficiency, for example, may be such a factor. Folate levels have been found in various studies to be lower in patients with myocardial infarct and carotid artery disease than control subjects. Meta-analysis of data from large multicentre prospective trials of folic acid in prevention of coronary arterial disease, however, do not show a positive effect, whereas a reduction of the risk of stroke by 18% has been reported, but even for this there are conflicting data. For coronary disease aspirin may over-ride any effect of folic acid. Extremely high levels of homocysteine (>50 μmol/L) are toxic to endothelia. When homocysteine levels are only mildly or moderately elevated, other mechanisms have been proposed, including oxidant damage interaction of homocysteine with cysteine residues on coagulation factors, platelets, adhesion molecules or endothelial cells or promotion of vascular wall inflammation.

Malignancy

Prophylactic folic acid in pregnancy has been found in some, but not all, studies to reduce the subsequent incidence of acute lymphoblastic leukaemia (ALL) in childhood. A significant negative association has been found with the *MTHFR* 677C→T and 1298A→C polymorphisms and the incidence of both paediatric and adult ALL. There are various positive and negative associations between polymorphisms in other folate-dependent enzymes and the incidence of paediatric and adult ALL. Other tumours that have been associated with folate polymorphisms or status include follicular lymphoma, breast cancer and gastric cancer.

Meta-analysis of ten randomized trials of folic acid at doses 2 mg or more for an average of 5.2 years, as well as of three trials for prevention of colonic adenomas (total 49,621 individuals) has not shown a significant difference for cancer generally or for any specific cancer between folic acid and placebo. This is reassuring, both for those taking folic acid as a vitamin supplement and for countries where food fortification with folic acid is carried out.

Other tissues

Folate deficiency causes reduced regeneration of cirrhotic liver. Patients with gluten-induced enteropathy and those with sickle cell anaemia have also been reported to show stunted growth, which has been improved coincidentally with commencement of folic acid therapy, but it is not certain how much the growth improvement in these children was due to folic acid and how much to other, simultaneously administered vitamins. There are experimental data in animals and some clinical observations that long-term vitamin B_{12} deficiency can affect bone formation. In the fragile X syndrome, sister chromatid exchange and DNA breaks are increased *in vitro* in a folate-deficient medium, apparently at the Xq28 site. No *in vivo* abnormality of folate metabolism can be detected.

Neurological manifestations

Cobalamin deficiency may cause bilateral peripheral neuropathy or degeneration (demyelination) of the posterior and pyramidal tracts of the spinal cord and, less frequently, optic atrophy or cerebral symptoms. The patient classically presents with paraesthesiae in the legs, muscle weakness or difficulty in walking and sometimes dementia, psychotic disturbances or visual impairment. Long-term nutritional cobalamin deficiency in infancy leads to poor brain development and impaired intellectual development. Folate deficiency may cause mental changes such as depression and slowness, and has been suggested to cause organic nervous disease, but this is uncertain. Methotrexate injected into the cerebrospinal fluid may cause brain or spinal cord damage. Neural tube defects in the fetus are discussed above.

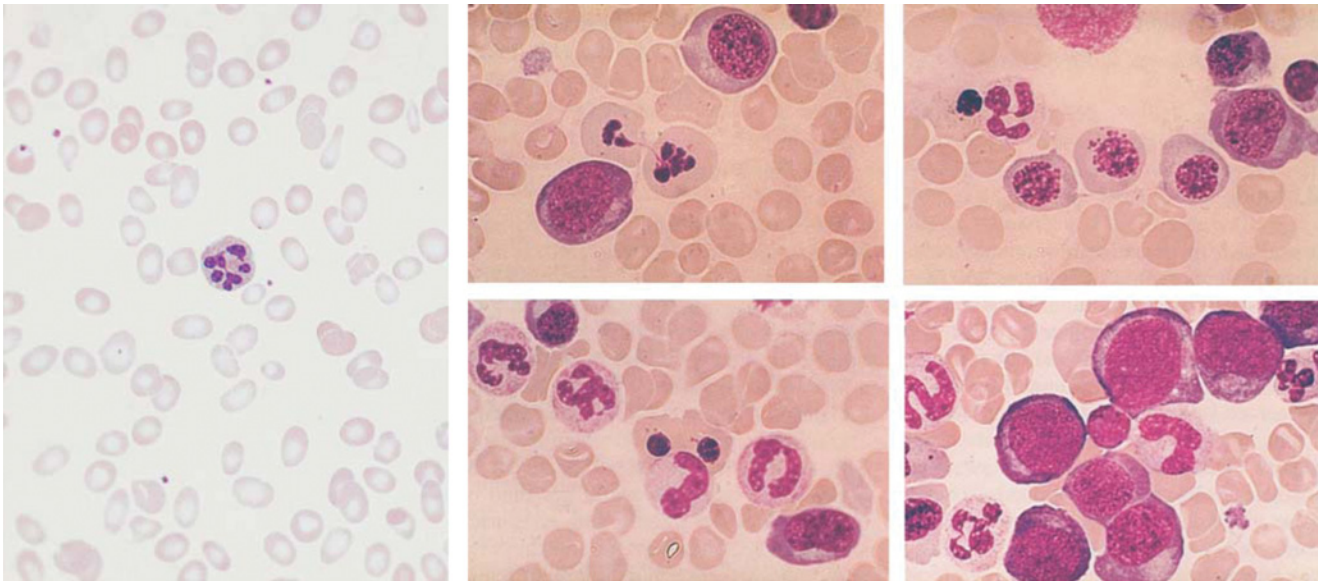
The biochemical basis for cobalamin neuropathy remains obscure. Its occurrence in the absence of methylmalonic aciduria in transcobalamin deficiency, and in monkeys given nitrous oxide (N_2O), suggests that the neuropathy is related to the defect in conversion of homocysteine to methionine. Accumulation of *S*-adenosylhomocysteine (SAH) in the brain, resulting in inhibition of transmethylation reactions due to an altered *S*-adenosylmethionine (SAM) to SAH ratio, has been suggested. SAM is needed in methylation of biogenic amines (e.g. dopamine), as well as of proteins, phospholipids and neurotransmitters in the brain (Figure 5.3).

Studies showing an association between lower serum levels of folate or cobalamin and higher homocysteine levels and Alzheimer disease, loss of cognitive function or brain volume loss have been reported. However, trials of supplementation with folic acid, vitamin B_{12} and vitamin B_6 have not consistently shown a benefit in preventing progression of the dementia compared with a control group, or in improving cognitive function.

Haematological findings

Peripheral blood

Oval macrocytes, usually with considerable anisocytosis and poikilocytosis, are the main feature (Figure 5.4a). The MCV is usually more than 100 fL unless a cause of microcytosis (e.g. iron deficiency or thalassaemia trait) is present, when there is a raised red cell distribution width (RDW) and the film is dimorphic. The MCV may also be normal owing to excess fragmentation of red cells. Some of the neutrophils are hypersegmented (more than five nuclear lobes). Both macrocytosis and hypersegmented neutrophils may also occur in other situations (Table 5.3). Together, however, they strongly suggest megaloblastic haemopoiesis. There may be leucopoenia due to a reduction in granulocytes and lymphocytes; the platelet count may be moderately reduced, rarely to less than 40×10^9/L.



(a) (b)

Figure 5.4 Severe megaloblastic anaemia: (a) peripheral blood; (b) bone marrow.

Occasionally, a leuco-erythroblastic blood picture is seen. In the non-anaemic patient, the presence of a few macrocytes and hypersegmented neutrophils in the peripheral blood may be the only abnormalities.

Table 5.3 Conditions in which macrocytosis or hypersegmented neutrophils may occur in the absence of megaloblastic anaemia.

Macrocytosis
Alcohol
Liver disease (especially alcoholic)
Reticulocytosis (haemolysis or haemorrhage)
Aplastic anaemia or red cell aplasia
Hypothyroidism
Myelodysplasia
Myeloma and macroglobulinaemia
Leucoerythroblastic anaemia
Myeloproliferative disease
Pregnancy
Newborn
Congenital dyserythropoietic anaemia (type II)
? Chronic respiratory failure
Hypersegmented neutrophils
Renal failure
Congenital (familial)
? Iron deficiency

Note: Falsely high MCV recorded when cold agglutinins, paraproteins or marked leucocytosis are present.

Bone marrow

In severely anaemic patients, the marrow is hypercellular with accumulation of primitive cells due to selective death of more mature forms. There is dissociation between nuclear and cytoplasmic development in the erythroblasts, with the nucleus maintaining a primitive appearance despite maturation and haemoglobinization of the cytoplasm; fully haemoglobinized (orthochromatic) erythroblasts may be seen. The nucleus of the megaloblast has an open, fine, lacy appearance; the cells are larger than normoblasts and an increased number of cells with eccentric lobulated nuclei or nuclear fragments may be present (Figure 5.4b). Mitotic and dying cells are seen more frequently than normal. Giant and abnormally shaped metamyelocytes, and enlarged hyperpolyploid megakaryocytes are characteristic. Severe florid megaloblastic changes may be confused with acute erythroid leukaemia. Rarely, the marrow may be hypocellular or red cell precursors are lost almost completely from the marrow and a mistaken diagnosis of myeloid leukaemia may be made. Iron staining shows increase in both reticuloendothelial stores and in the developing megaloblasts.

In less anaemic patients, the changes in the marrow may be difficult to recognize. The terms 'intermediate', 'mild' and 'early' have been used. The changes may be mild and difficult to recognize, even in a severely anaemic patient, if the anaemia is largely due to other factors (e.g. iron deficiency, infection, malignant disease, haemolysis) and the megaloblastosis is an incidental phenomenon. The term 'megaloblastoid' has several different connotations, including the dysplastic changes seen in the myelodysplastic syndromes, and is best avoided.

Chromosomes

Bone marrow cells, transformed lymphocytes and other proliferating cells in the body show a variety of changes including random chromosomal breaks, reduced contraction of chromatin, spreading of the centromere, and exaggeration of secondary chromosomal constrictions and prominent satellites. Similar abnormalities may be produced by antimetabolite drugs (e.g. cytarabine, hydroxycarbamide and methotrexate) that interfere with either DNA replication or folate metabolism and which also cause megaloblastic appearances.

Ineffective haemopoiesis

There is accumulation of unconjugated bilirubin in plasma due to the death of nucleated red cells in the marrow (ineffective erythropoiesis). Other evidence for this includes raised urine urobilinogen, reduced haptoglobins and positive urine haemosiderin, raised serum lactate dehydrogenase and raised serum iron, non-transferrin-bound iron and ferritin levels. Carbon monoxide production is also increased. Serum lysozyme may also be raised, suggesting ineffective granulopoiesis.

In rare patients, ineffective haematopoiesis is associated with features of disseminated intravascular coagulation, with raised serum fibrin degradation products. Thrombocytopenia, when it occurs, is usually caused by ineffective megakaryopoiesis. A weakly positive direct antiglobulin test due to complement can lead to a false diagnosis of autoimmune haemolytic anaemia.

Cobalamin

Cobalamin (vitamin B_{12}) exists in a number of different chemical forms. The molecule consists of two halves: a planar group and a nucleotide set at right angles to it (Figure 5.5). The planar group is a corrin ring and the nucleotide consists of a base, 5,6-dimethylbenzimidazole, and a phosphorylated sugar, ribose-5-phosphate. In nature, the vitamin is mainly in the 5′-deoxyadenosyl (ado) form. This is the main form in human tissues and is located in the mitochondria. It serves as the co-factor for methylmalonyl-CoA mutase. The other major natural cobalamin is methylcobalamin, the main form in human plasma and cell cytoplasm, which serves as the co-factor for methionine synthase. There are also minor amounts of hydroxocobalamin, the form to which methyl- and ado-cobalamin are rapidly converted by exposure to light, hydroxocobalamin having its cobalt atom in the fully oxidized Cbl^{3+} state, whereas the cobalt exists as reduced Cbl^{1+} in the methyl- and ado-cobalamin forms (see Figure 5.2).

Dietary sources and requirements

Cobalamin is synthesized solely by microorganisms. Ruminants obtain cobalamin from the foregut, but the only source for humans is food of animal origin. The highest amounts are found in liver and kidney (up to 100 μg per 100 g), but it is also present in shellfish, organ and muscle meats, fish, chicken and dairy products (eggs, cheese and milk) in small amounts (6 μg/L). Vegetables, fruits and all other foods of non-animal origin are free from cobalamin unless they are contaminated by bacteria. Cooking does not usually destroy cobalamin.

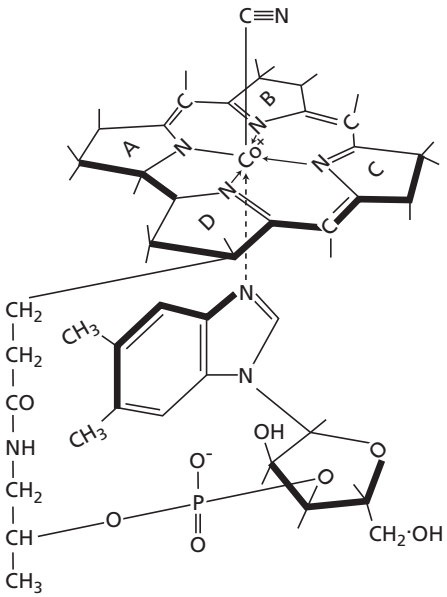


Figure 5.5 The structure of vitamin B_{12} (cyanocobalamin). There is a corrin ring with a cobalt atom at its centre. Above is attached the cyano, hydroxo, methyl or deoxyadenosyl ligand. Below is the nucleotide 5,6-dimethylbenzimidazole.

A normal Western diet contains 5–30 μg of cobalamin daily. Adult daily losses (mainly in the urine and faeces) are about 1–2 μg and because the body does not have the ability to degrade cobalamin, daily requirements are also about 1 μg. Body stores are of the order of 2–3 mg and are sufficient for 3–4 years if supplies are completely cut off.

Absorption

Two mechanisms exist for cobalamin absorption. One is passive, occurring equally through the duodenum and the ileum; it is rapid but inefficient as less than 1% of an oral dose can be absorbed by this process. Passive absorption of cobalamin can also occur through other mucous membranes such as the sublingual and nasal mucosae. The other mechanism is active; it occurs through the ileum in humans and is efficient for small (a few micrograms) oral doses of cobalamin. This is the normal mechanism by which the body acquires cobalamin and is mediated by gastric intrinsic factor (IF).

Dietary cobalamin is released from protein complexes by enzymes in the stomach, duodenum and jejunum; it combines

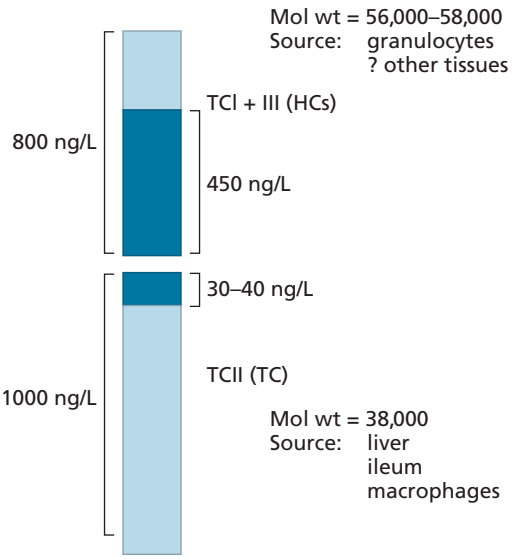


Figure 5.6 The serum cobalamin-binding proteins (TCs, transcobalamins). Dark blue rectangles indicate endogenous serum cobalamin; light blue rectangles indicate unsaturated cobalamin-binding protein; HCs, haptocorrins.

rapidly with a salivary glycoprotein (R binder) related to plasma transcobalamin I (TCI). These proteins are isoforms of cobalamin-binding proteins known as haptocorrins (HCs), which differ from each other only in their patterns of glycosylation. They are products of a single gene (*TCN1*), and they occur in saliva, gastric juice, bile, milk and other body fluids. Subsequently, HC is digested by pancreatic trypsin and the cobalamin transferred to IF. Binding of cobalamin to IF is favoured by an alkaline pH; it binds one molecule for one molecule. All forms of cobalamin are absorbed by the same IF mechanism (Figure 5.6). Intrinsic factor is a glycoprotein (molecular weight 45,000) encoded by a gene on chromosome 11q13. It is produced in gastric parietal cells in the fundus and body of the stomach. The IF–cobalamin complex, in contrast with free IF, is resistant to enzyme digestion, having a more closed structure. It passes to the ileum, where IF attaches to a specific receptor (cubilin) on the brush border surface of the ileal absorptive cells. Cubilin is also present in the yolk sac and renal proximal tubular epithelium. The attachment of the IF–cobalamin complex requires calcium ions and a pH around neutral. It is probably a physical process, not requiring energy. Cubilin traffics by means of the protein called amnionless (AMN). AMN binds tightly to cubilin and directs sublocalization and endocytosis of cubilin with its ligand (IF–cobalamin complex). Defects in cubilin and AMN are implicated in autosomal recessive megaloblastic anaemia, characterized by intestinal malabsorption of cobalamin (see p. 63). A third protein, megalin (LRP2), has been suggested to play a role in stabilizing the cubilin–AMN complex.

Cobalamin then enters the ileal cell. IF is digested and after a delay of about 6 hours absorbed cobalamin appears in portal blood attached to transcobalamin (TCII), which is synthesized in the ileum, either by mucosal cells or by venous endothelial cells in the submucosa.

The ileum has a restricted capacity to absorb cobalamin because of limited receptor sites. Although 50% or more of a single dose of 1 μg of cobalamin may be absorbed, with doses above 2 μg the proportion absorbed falls rapidly. Moreover, after one dose of IF–cobalamin complex has been presented, the ileal cells become refractory to further doses for about 6 hours.

Enterohepatic circulation

Between 0.5 and 5.0 μg of cobalamin enter the bile each day. This binds to IF and a portion of biliary cobalamin is reabsorbed together with cobalamin derived from sloughed intestinal cells. Bile may enhance cobalamin absorption. Cobalamin deficiency develops more rapidly in individuals who malabsorb cobalamin than it does in vegans, who ingest no cobalamin but in whom reabsorption of biliary cobalamin is intact.

Transport

Two main cobalamin-binding proteins exist in human plasma; they both bind cobalamin one molecule for one molecule (Figure 5.6). One HC (also known as TCI) is a glycoprotein. TCIII is a minor isoprotein of TCI in plasma. These HCs are derived primarily from the specific granules in neutrophils and are normally about two-thirds saturated with cobalamin, which they bind tightly. They do not enhance cobalamin entry into tissues. Glycoprotein receptors on liver cells are concerned in the removal of HCs from plasma, and HC may have a role in the transport of cobalamin analogues to the liver for excretion in bile. The gene, *TCN1* on chromosome 11q11–q12.3, has nine exons and codes a protein of 433 amino acids. Heterozygous, homozygous or compound heterozygous mutations in the gene may cause mild or severe reductions in serum vitamin B_{12}, respectively, with no known clinical consequences.

The other major cobalamin transport protein in plasma is transcobalamin (TC, also known as TCII). This is synthesized by liver, and by other tissues, including macrophages, endothelial and possibly ileal cells. It normally carries only 20–60 ng of cobalamin per litre of plasma and readily gives up cobalamin to marrow, placenta and other tissues, which it enters by receptor-mediated endocytosis via clathrin-coated pits. TC is not reutilized. The gene is on chromosome 22q11–q13.1 and, as for IF and HC, there are nine exons. The three proteins are likely to have a common ancestral origin. TC has 20% amino acid homology and more than 50% nucleotide homology with human HC and with rat IF. The regions of homology of HC, TC and IF are involved in cobalamin binding. Five different inherited isoproteins of TC have been described; all are functionally active. TC is

also present in cerebrospinal fluid and binds cobalamin (approximately 10 ng/L) there. Alterations may occur in TC and HC levels in a variety of disease states. In general, an increase in HC is found particularly in myeloproliferative diseases and hepatoma and causes an increase in serum cobalamin, whereas an increase in TC does not.

The TC receptor belongs to the low-density lipoprotein receptor family. It is more heavily expressed in dividing than quiescent cells, and is recycled to the cell surface. Megalin (LRP-2) is also involved in the endocytosis of TC–cobalamin.

Cobalamin analogues

Cobalamin analogues are corrinoids, which exist as cobamides (containing substitutions in the place of ribose, e.g. adenosyl) or as cobinamides (which have no nucleotide whatever). They do not bind to IF and are not absorbed. HC may carry analogues to the liver for excretion in the bile. It is unclear whether they are inert or inhibit cobalamin-dependent reactions. The proportion of analogues derived from diet, gut bacteria or endogenous breakdown of cobalamins is unknown. They are present in fetal blood and tissues.

Causes of cobalamin deficiency

Cobalamin deficiency is usually due to malabsorption. The only other cause is inadequate dietary intake. Cobalamin deficiency due to excess degradation does not occur. Inactivation may occur as a result of exposure to the anaesthetic gas N_2O. N_2O which causes irreversible oxidation of the active Cbl^{1+} during catalytic shunting of labile methyl groups in the methionine synthase reaction (see Figure 5.2).

Inadequate dietary intake

Adults

Dietary cobalamin deficiency arises in vegans who omit meat, fish, eggs, cheese and other animal products from their diet. The largest group in the world consists of Hindus, and many millions of Indians are at risk of deficiency on a nutritional basis. However, not all vegans develop cobalamin deficiency of sufficient severity to cause anaemia or neuropathy, even though subnormal cobalamin levels have been found in up to 50% of randomly selected, young, adult Indian vegans. Dietary cobalamin deficiency may also arise rarely in non-vegetarian subjects who exist on grossly inadequate diets because of poverty or psychiatric disturbance. Nutritional cobalamin deficiency may not progress to megaloblastic anaemia because the diet is not totally lacking cobalamin, the enterohepatic circulation is intact and body losses will reduce as stores diminish.

Infants

Cobalamin deficiency occurs in infants born to severely cobalamin-deficient mothers. These infants develop megaloblastic anaemia at about 3–6 months of age since they are born with low stores of cobalamin and are then fed breast milk of low cobalamin content. This occurs most commonly in Indian vegans, but also in unrecognized maternal pernicious anaemia and in strict practitioners of veganism living in Western countries whose offspring have shown growth retardation, impaired psychomotor development and other neurological sequelae.

Gastric causes of cobalamin malabsorption (Tables 5.4 and 5.5)

Pernicious anaemia

Pernicious anaemia (PA) may be defined as a severe lack of IF due to gastric atrophy. It is a common disease in northern Europeans but occurs in all countries and ethnic groups. The overall incidence is about 120 per 100,000 population in the UK, but there is wide variation between one area and the next. The prevalence rate in Western countries may be as high as 2–3%. The ratio of incidence in men and women is approximately 1:1.6 and the peak age of onset is 60 years, with only 10% of patients being less than 40 years of age. In some ethnic groups, notably black people and Latin Americans, the age of onset of PA is generally lower. The disease occurs more commonly than by chance in close relatives, in subjects or families with other organ-specific autoimmune diseases of the thyroid, adrenal or parathyroid, in those with premature greying, blue eyes and vitiligo, and in persons of blood group A. An association with human leucocyte antigen (HLA)-3 has been reported in some but not all series and in those with endocrine disease, with HLA-B8, -B12 and -BW15. The life expectancy has been estimated as normal in women once regular treatment has begun. Men have been reported to

Table 5.4 Causes of cobalamin deficiency causing megaloblastic anaemia.

Nutritional
Vegans
Malabsorption
Pernicious anaemia
Gastric causes
Congenital intrinsic factor deficiency or functional abnormality
Total or partial gastrectomy
Intestinal causes
Intestinal stagnant loop syndrome: jejunal diverticulosis, ileocolic fistula, anatomical blind loop, intestinal stricture, etc.
Ileal resection and Crohn's disease
Selective malabsorption with proteinuria
Tropical sprue
Transcobalamin deficiency
Fish tapeworm

Table 5.5 Malabsorption of cobalamin may occur in the following conditions but is not usually sufficiently severe and prolonged to cause megaloblastic anaemia.

Gastric causes
Simple atrophic gastritis (food cobalamin malabsorption)
Zollinger–Ellison syndrome
Gastric bypass surgery
Use of proton pump inhibitors
Intestinal causes
Gluten-induced enteropathy
Severe pancreatitis
HIV infection
Radiotherapy
Graft-versus-host disease
Deficiency states
Cobalamin, folate, protein, ?riboflavin, ?nicotinic acid
Drug therapy
Colchicine, *p*-aminosalicylate, neomycin, slow-release potassium chloride, anticonvulsant drugs, ?metformin, ?phenformin, cytotoxic drugs
Alcohol

have a slightly subnormal life expectancy as a result of a higher incidence of carcinoma of the stomach than in control subjects.

Diagnosis

This is usually suspected from the clinical picture and the findings of megaloblastic anaemia due to cobalamin deficiency. Tests for circulating gastric autoantibodies and serum gastrin levels are important. The serum gastrin level is usually raised in PA (>200 μg/L), the hormone coming from endocrine cells in the gastric fundus. Raised serum gastrin also occurs in simple atrophic gastritis. Serum pepsinogen I levels are low (<30 μg/L) in over 90% of those affected and a low ratio of serum pepsinogen I to pepsinogen II correlates with the presence of chronic atrophic gastritis.

Gastric biopsy

This shows atrophy of all layers of the body and also fundal atrophy, with loss of glandular elements, an absence of parietal and chief cells and replacement by mucous cells, a mixed inflammatory cell infiltrate and sometimes intestinal metaplasia. The infiltrate of plasma cells and lymphocytes contains an excess of CD4 cells. The antral mucosa is usually well preserved. *Helicobacter (H.) pylori* infection is infrequent in PA, but it has been suggested that *H. pylori* gastritis may represent an early phase of atrophic gastritis, often associated with iron deficiency, which is gradually replaced, in some individuals, by an immune process with disappearance of *H. pylori* infection. Corticosteroid therapy may restore acid and IF secretion temporarily.

Antibodies to gastric antigens:

1 *IF antibodies.* Two types of IF antibody may be found in the sera of patients with PA, both being IgG. One, the 'blocking' or 'type I' antibody, prevents the combination of IF and cobalamin, whereas the other, the 'binding', 'type 2' antibody, which attaches to IF whether joined to cobalamin or not, prevents attachment of IF to ileal mucosa. Type 1 antibody occurs in the serum of about 55% of patients, type 2 antibody in 35%. IF antibodies cross the placenta and cause temporary IF deficiency in the newborn infant. Patients with PA also show cell-mediated immunity to IF. IF antibodies are rarely found in conditions other than PA. Type I antibody has been detected rarely in the sera of patients without PA but with thyrotoxicosis, myxoedema, Hashimoto disease or diabetes mellitus, and in relatives of PA patients. IF antibodies have also been detected in gastric juice in about 80% of patients with PA. These antibodies may reduce absorption of dietary cobalamin by combining with small amounts of remaining IF in the gastric juice. Achlorhydria favours the formation of this antigen–antibody complex.

2 *Parietal cell and gastrin receptor antibodies.* Parietal cell antibody is present in the sera of almost 90% of adult patients with PA, but it is frequently present in other subjects. Thus, it occurs in as many as 16% of randomly selected female subjects aged over 60 years and in a smaller proportion of younger control subjects; it is found more frequently than in control subjects in relatives of PA patients. These antibodies are also found more frequently in patients with simple atrophic gastritis, chronic active hepatitis and thyroid disorders and their relatives, as well as in Addison disease, rheumatoid arthritis and other conditions. The parietal cell antibody is directed against the α- and β-subunits of the gastric proton pump (H^+/K^+-ATPase). The sera of PA patients may also contain an autoantibody to the gastrin receptor, although this test is not used clinically.

Hypogammaglobulinaemia

PA is found more often than by chance in patients with a deficiency of IgA or with complete hypogammaglobulinaemia. These subjects resemble others with PA, except that they often present relatively early (before the age of 40 years), they have a lower incidence of serum IF and parietal cell antibodies, and they may show intestinal malabsorption. They may also have a history of recurrent infections. The gastric lesion is similar to that in other cases, except that plasma cells are absent from the inflammatory cell infiltrate and the antrum is involved. Serum gastrin levels are normal.

Juvenile pernicious anaemia

This usually occurs in older children and resembles PA of adults. Gastric atrophy, achlorhydria and serum IF antibodies are all

present, although parietal cell antibodies are usually absent. About half of these patients show an associated endocrinopathy such as autoimmune thyroiditis, Addison disease or hypoparathyroidism; in some, mucocutaneous candidiasis occurs.

Congenital intrinsic factor deficiency or functional abnormality

The affected child usually shows no demonstrable IF, but has a normal gastric mucosa and normal secretion of acid. The inheritance is autosomal recessive. These patients usually present with megaloblastic anaemia in the first, second or third year of life when stores of cobalamin accumulated from the mother *in utero* are used up; a few have presented as late as the second decade. Parietal cell and IF antibodies are absent. Variants have been described in which the child is born with IF that can be detected immunologically, but which is unstable or functionally inactive, being unable either to bind cobalamin or to facilitate its uptake by the ileum.

Gastrectomy

Following total gastrectomy, cobalamin deficiency is inevitable and prophylactic cobalamin therapy should be commenced immediately following the operation. After partial gastrectomy, 10–15% of patients also develop this deficiency. The exact incidence and time of onset are most influenced by the size of the resection and the pre-existing size of the cobalamin body store. Gastric plication for obesity may also cause malabsorption of cobalamin.

Simple atrophic gastritis (food cobalamin malabsorption)

The normal IF-mediated mechanism of cobalamin absorption requires adequate gastric output of acid and pepsin to ensure the release of food cobalamin. Failure of this mechanism is believed to be responsible for a condition more common in the elderly known as *food cobalamin malabsorption*. The syndrome has also been described in association with *H. pylori* infection, long-term use of histamine H_2-receptor antagonists and proton-pump inhibitors, chronic alcoholism, pancreatic exocrine failure, Sjögren syndrome and systemic sclerosis. The syndrome is associated with low serum cobalamin levels, with or without evidence of cobalamin deficiency, such as raised serum levels of methylmalonic acid and homocysteine. A minority of patients with food cobalamin malabsorption may go on to develop clinically significant cobalamin deficiency, but the frequency of occurrence and reasons for this progression are not clear.

Intestinal causes of cobalamin malabsorption

Malabsorption of cobalamin occurs in a variety of intestinal lesions in which there is colonization of the upper small intestine by faecal organisms. This may occur in patients with jejunal diverticulosis, entero-anastomosis, intestinal stricture or fistula, or with an anatomical blood loop due to Crohn's disease, tuberculosis or an operative procedure. Bacterial overgrowth in the small intestine may also cause spurious elevation of serum methylmalonate (see further on). Removal of 1.2 m or more of terminal ileum causes malabsorption of cobalamin. In some patients, following ileal resection, particularly if the ileocaecal valve is incompetent, colonic bacteria may contribute further to cobalamin deficiency.

Nearly all patients with acute and subacute tropical sprue show malabsorption of cobalamin; this may persist as the principal abnormality in the chronic form of the disease, when the patient may present with megaloblastic anaemia or neuropathy due to cobalamin deficiency. Absorption of cobalamin usually improves after antibiotic therapy and, in the early stages, after folic acid therapy. Malabsorption of cobalamin occurs in about 30% of untreated patients with gluten-induced enteropathy and correlates with the degree of steatorrhoea. Cobalamin deficiency is not usually severe in these patients and is probably never the cause of megaloblastic anaemia unless another lesion causing malabsorption of cobalamin (e.g. stagnant loop syndrome) is present. The absorption improves when these patients are treated with a gluten-free diet.

Selective malabsorption of cobalamin with proteinuria (also known as Imerslünd syndrome, Imerslünd–Grasbeck syndrome, megaloblastic anaemia Type 1 (MGA1) is an autosomal recessive disease and is the most common cause of megaloblastic anaemia due to cobalamin deficiency in infancy in Western countries. There are familial clusters in Finland, Norway, the Middle East and North Africa. The patients usually present with anaemia between the ages of 1 and 5 years, and secrete normal amounts of IF and gastric acid. In some cases, such as in Finland, impaired synthesis, processing or ligand binding of cubilin due to inherited mutations, for example c.391C→T (named FM1) and a mutation at an intron causing a truncated protein (FM2), have been implicated. In others, for example in Norway, mutation of the gene for AMN has been reported. Other tests of intestinal absorption are normal. Over 90% of these patients show non-specific proteinuria but renal function is otherwise normal and renal biopsy has not shown any consistent renal defect. A few of these patients have shown aminoaciduria and congenital renal abnormalities, such as duplication of the renal pelvis.

The fish tapeworm (*Diphyllobothrium latum*) lives in the small intestine of humans and accumulates cobalamin from food, rendering this unavailable for absorption. People acquire the worm by eating raw or partly cooked fish. Infestation is common around the lakes of Scandinavia, Germany, Japan, North America and Russia. Megaloblastic anaemia or cobalamin neuropathy occurs only in those with a heavy infestation, with the worm high in the small intestine. Many carriers have no cobalamin deficiency.

In severe chronic pancreatitis, lack of trypsin is thought to be why dietary cobalamin attached to gastric non-IF (R) binder is unavailable for absorption. It has also been proposed that in

pancreatitis, the concentration of calcium ions in the ileum falls below the level needed to maintain normal cobalamin absorption. Serum cobalamin levels tend to fall in patients with HIV infection and are subnormal in 10–35% of those with AIDS. Malabsorption of cobalamin not corrected by IF has been shown in some, but not all, patients with subnormal serum cobalamin levels. Malabsorption of cobalamin has been reported in Zollinger–Ellison syndrome. It is thought that there is a failure to release cobalamin from R binding protein due to inactivation of pancreatic trypsin by high acidity, as well as interference with IF binding of cobalamin.

Both total body irradiation and local radiotherapy to the ileum (e.g. as a complication of radiotherapy for carcinoma of the cervix) may cause malabsorption of cobalamin. Graft-versus-host disease commonly affects the small intestine: malabsorption of cobalamin due to abnormal gut flora, as well as damage to ileal mucosa, is frequent. Some drugs cause malabsorption of cobalamin, but rarely megalololoblastic anaemia (Table 5.5). Metformin may lower serum cobalamin by lowering TCI. The use of histamine H_2-blockers for treatment of peptic ulcer disease causes decreases in cobalamin absorption, and continued use may lower serum cobalamin level. Other causes of cobalamin malabsorption are listed in Table 5.5.

Abnormalities of cobalamin metabolism

Congenital transcobalamin deficiency or abnormality

Infants with TC deficiency usually present with megaloblastic anaemia within a few weeks of birth. Serum cobalamin and folate levels are normal, but the anaemia responds to massive (e.g. 1 mg three times weekly) injections of cobalamin, which cause free cobalamin to enter marrow cells by passive diffusion in the absence of functional TC. Some cases show neurological complications. In some cases, the protein is present in normal amounts, but is unable to bind cobalamin or to attach to the cell surface and so is functionally inert. Genetic abnormalities found include mutations of an intraexonic cryptic splice site, extensive or single nucleotide deletion, nonsense mutation and an RNA editing defect. These infants do not show methylmalonic aciduria, but malabsorption of cobalamin occurs in all cases and reduced immunoglobulins in some. Less severe cases present later in childhood. Failure to institute adequate cobalamin therapy or treatment with folic acid may lead to neurological damage.

Congenital methylmalonic acidaemia and aciduria

Infants with this abnormality are ill from birth, with vomiting, failure to thrive, severe metabolic acidosis, ketosis and mental retardation. Anaemia, if present, is normocytic and normoblastic. The condition may arise as a result of a functional defect in either the mitochondrial methylmalonyl-CoA mutase or its co-factor ado-cobalamin (Figure 5.2). Mutations in methylmalonyl-CoA mutase are not responsive, or only poorly responsive, to treatment with cobalamin. Two disorders result in cobalamin-responsive methylmalonic acidaemia. In cobalamin (Cbl)A disease, there is failure of reduction of cobalamin III (Cbl^{3+}) or cobalamin II (Cbl^{2+}) to cobalamin I (Cbl^{1+}) in mitochondria; in CblB disease, there is a defect of an adenosyltransferase required for synthesis of ado-cobalamin (Figure 5.2). A proportion of infants with CblA and CblB disease respond to cobalamin in large doses, whereas others are unresponsive. In those who do not respond to cobalamin, the enzyme methylmalonyl-CoA mutase is lacking (mut^0) or defective (mut^-). Some children have combined methylmalonic aciduria and homocystinuria due to defective formation of both cobalamin co-enzymes. The defects are in the transfer of cobalamin from the endocytic compartment of lysosomes to the cytoplasm (CblF disease) or in the reduction of cobalamin 3^+ to cobalamin 2^+ after transfer to the cytoplasm (CblC and CblD diseases). Over 100 cases of CblC disease have been described. It usually presents in the first year of life with feeding difficulties, developmental delay, microcephaly, seizures, hypotonia and megaloblastic anaemia.

Some patients present with homocystinuria and megaloblastic anaemia, often with neurological defects, but without methylmalonic aciduria. There is a selective deficiency of methylcobalamin. These conditions have been termed CblE and CblG disease and are due to lack of association of methylcobalamin with methionine synthase.

Acquired abnormality of cobalamin metabolism: nitrous oxide inhalation

N_2O irreversibly oxidizes methylcobalamin from its active, fully reduced Cbl^{1+} state to an inactive Cbl^{2+} precursor. This has been shown to inactivate methylcobalamin and methionine synthase. This occurs in both humans and experimental animals and was of importance in the megaloblastic anaemia that occurred in patients undergoing prolonged N_2O anaesthesia (e.g. in intensive care units). A neuropathy resembling cobalamin neuropathy has been described in dentists and anaesthetists who are repeatedly exposed to N_2O and in monkeys exposed to the gas for many months. Recovery from N_2O exposure requires regeneration of methionine synthase, as this protein is damaged by active oxygen derived from the N_2O–cobalamin reaction.

Diagnosis of cobalamin deficiency

The diagnosis of cobalamin or folate deficiency has traditionally depended on the recognition of the relevant abnormalities in the peripheral blood and/or bone marrow and subsequent analysis of the blood levels of the vitamins. Other causes of macrocytosis and hypersegmented neutrophils are listed in Table 5.3. However, assays of serum methylmalonic acid and homocysteine (see further on) have shown these to be raised in some subjects without haematological abnormalities, including a proportion with normal levels of serum cobalamin and folate in whom, nevertheless, the levels of the metabolites fall to normal with cobalamin and/or folate therapy. The significance of these biochemical changes remains controversial. They may imply functional

cobalamin or folate deficiency, not reflected by subnormal levels of the vitamins or by disturbed haemopoiesis. If so, it would imply that the accepted normal serum and red cell levels of the vitamins reflect body stores that are sufficiently high to prevent haematological changes, but which in some subjects may not be optimal for prevention of other complications of the deficiencies, including NTDs in the fetus.

Measurement of serum cobalamin

Serum cobalamin is usually measured by one of a number of enzyme-linked immunosorbent assays. These are frequently automated. Normal serum cobalamin levels range from 160–200 ng/L to about 1000 ng/L (ng × 0.738 = pmol, so 200 ng/L = 148 pmol/L). In patients with megaloblastic anaemia due to cobalamin deficiency, the level is usually less than 100 ng/L. In general, the more severe the deficiency, the lower the serum cobalamin level. In patients with spinal cord damage due to the deficiency, levels are very low even in the absence of anaemia. False normal levels have been described with some of the assays in patients with IF antibodies in serum. Values of between 100 and 200 ng/L are regarded as borderline. They may occur, for instance, in pregnancy, in patients with megaloblastic anaemia due to folate deficiency, and in patients with mutations of the *TCN1* gene that codes for HC (TCI). They are also found in vegans and in patients with other causes of mild cobalamin deficiency (Table 5.5). The relative concentrations of HC and TC also influence the total serum cobalamin level. Raised serum cobalamin levels (if not due to recent therapy) are usually due to a rise in HC or to liver or renal disease with increased saturation of HC and TC.

Serum holotranscobalamin (holoTCII, holoTC)

Since TC is the plasma cobalamin transport protein that is responsible for cellular uptake and delivery of cobalamin, the notion was put forward that measurement of circulating cobalamin that was bound to TC (holoTC) would provide a more meaningful measure of cobalamin status than total serum cobalamin. However, measurement of holoTC is not available or used diagnostically, except in research studies.

Serum methylmalonate and homocysteine levels

In patients with cobalamin deficiency, the serum methylmalonate (MMA) and homocysteine levels are raised. Sensitive methods for measuring MMA and homocysteine in serum have been introduced and recommended for the early diagnosis of cobalamin deficiency, even in the absence of haematological abnormalities or subnormal levels of serum cobalamin or folate. However, serum MMA fluctuates in patients with renal failure. Mildly elevated serum MMA and/or homocysteine levels occur in up to 30% of apparently healthy volunteers, with serum cobalamin levels up to 350 ng/L and normal serum folate levels and in 15% of elderly subjects, even with cobalamin levels above 350 ng/L. These findings bring into question the exact cut-off points for normal MMA and homocysteine levels. It is also unclear at present whether these mildly raised metabolite levels have clinical consequences and how many of the subjects will progress to clinically overt cobalamin deficiency. When cobalamin supplies to the cell are suboptimal, there may be preferential use of methylcobalamin for methionine synthesis compared with adocobalamin for MMA metabolism. Urinary MMA excretion may also be used to screen for cobalamin deficiency, but this is also increased in aminoaciduria (e.g. Fanconi syndrome).

Homocysteine exists in plasma as single molecules, as two molecules linked together (homocystine) and as mixed homocysteine–cysteine disulfides. Serum homocysteine levels are raised in both early cobalamin and folate deficiency, but they may be raised in other conditions, for example chronic renal disease, alcoholism, smoking, pyridoxine deficiency, hypothyroidism, therapy with steroids, ciclosporine and other drugs. Levels are also higher in serum than in plasma, in men than in premenopausal women, in women taking hormone replacement therapy or oral contraceptive users and in elderly subjects and patients with several inborn errors of metabolism affecting enzymes in trans-sulfuration pathways of homocysteine metabolism. Thus, homocysteine levels are not widely used for diagnosis of cobalamin or folate deficiency. However, homocysteine levels are used in assessing for cardiovascular risk (Chapter 44).

Tests for the cause of cobalamin deficiency

Studies of cobalamin absorption were used, but because of the unavailability of radioactive cobalamin have become obsolete. Serum tests for gastrin and antibodies to parietal cells and intrinsic factor aid in the diagnosis of PA. Upper GI endoscopy, including gastric biopsy helps to confirm the diagnosis and exclude gastric neoplasms.

Folate

Dietary folate

Folic acid (pteroylglutamic acid) is a yellow, crystalline, water-soluble substance (molecular weight 441). It is the parent compound of a large family of folate compounds. Pteroylglutamic acid consists of three parts: pteridine, *p*-aminobenzoate and l-glutamic acid (Figure 5.7). It is only a minor component of normal food folates (probably less than 1%), which differ from it in three respects (Figure 5.7): (i) they are partly or completely reduced at positions 4, 5, 7 and 8 in the pteridine portion to dihydrofolate or THF derivatives; (ii) they usually contain a single carbon unit of varying degrees of reduction, such as a methyl group at N-5 or N-10 and (iii) 70–90% of natural folates contain a chain of three or more glutamate residues linked to each other by the unusual γ-peptide bond and are called pteroyl- or

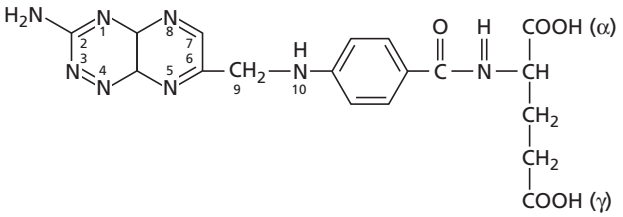


Figure 5.7 The structure of folic acid (pteroylglutamic acid).

folate-polyglutamates. In human cells, four, five and six glutamate residues are usual.

Most foods contain some folate. The highest concentrations are found in liver and yeast (>200 μg per 100 g), spinach, other greens and nuts (>100 μg per 100 g). The total folate content of an average Western diet is about 250 μg daily, but the amount varies widely according to the type of food eaten and the method of cooking. Folate is easily destroyed by heating, particularly in large volumes of water; over 90% may be lost.

Body stores and requirements

Total body folate in the adult is about 10 mg, the liver containing the largest store. Daily adult requirements are about 100–200 μg. Up to 13 μg of folate is lost as such in the urine each day, but breakdown products of folate are also lost in urine. Losses of folate also occur in sweat and skin; faecal folate is largely derived from colonic bacteria. Stores are only sufficient for about 4 months in normal adults, so severe folate deficiency may develop rapidly.

Absorption

The principal site of folate absorption is the upper small intestine, and there is a steep fall-off in absorptive capacity in the lower jejunum and ileum. The absorption of all forms tested is rapid, a rise in blood level occurring within 15–20 min of ingestion.

The small intestine has a tremendous capacity to absorb folate monoglutamates: about 90% of a single dose is absorbed, regardless of whether this is a small (100 μg) or large (15 mg) amount. A proton-coupled high-affinity folate transporter with a low pH optimum, termed PCFT/HCP1, is located at the apical brush border of the duodenal, and to a lesser extent jejunal mucosa and in other cells, including the blood–brain barrier. It accounts for the bulk of folate absorption, including of folic acid itself, and loss of function in hereditary folate malabsorption is not compensated by other folate transporters expressed on intestinal cells.

The absorption of folate polyglutamates with higher numbers of glutamate residues is less. This may be due to the limited capacity of the small intestine to hydrolyse these compounds or to their limited transfer in the mucosal cell. On average, about 50% of food folates is absorbed.

Polyglutamate forms are hydrolysed to monoglutamate derivatives, either in the lumen of the intestine or within the mucosa; they do not enter portal blood intact. Monoglutamate or polyglutamate forms of dietary folate, which are already partly or completely reduced, are converted to 5-methyl-THF within the small intestinal mucosa before entering the portal plasma. The monoglutamates are actively transported across the enterocyte by a carrier-mediated mechanism. Pteroylglutamic acid at doses greater than 400 μg is absorbed largely unchanged and converted to natural folates in the liver. Doses <400 μg are converted to 5-methyl-THF during absorption through the intestine.

Enterohepatic circulation

About 60–90 μg of folate enters the bile each day and is excreted into the small intestine. Loss of this folate, together with the folate of sloughed intestinal cells, accelerates the speed with which folate deficiency develops in malabsorption conditions.

Transport

Folate is transported in plasma, about one-third loosely bound to albumin and two-thirds unbound. In all body fluids (plasma, cerebrospinal fluid, milk, bile) folate is largely, if not entirely, 5-methyl-THF in the monoglutamate form. A carrier-mediated active process is involved in the entry of folate into cells, the rate of uptake being linked to the rate of folate polyglutamate synthesis in the cell, which in replicating cells is related to the rate of DNA synthesis. In most cells, folates are retained with tight binding to folate-binding proteins, three of which are enzymes involved in methyl group metabolism (sarcosine dehydrogenase, dimethylglycine dehydrogenase and glycine *N*-methyltransferase), until the cell dies. Intact liver cells can release folate. Two types of folate-binding protein are involved in entry of methyl-THF into cells. The reduced folate carrier SLC19A1 is a facilitative transporter with a pH optimum of 7.4 and the characteristics of an anion exchanger. Two glycosylphosphatidylinositol (GPI)-linked folate receptors mediate cellular folate uptake by an endocytic mechanism, with internalization in a vesicle (caveola), which is then acidified, releasing folate into the vesicle lumen. Folate is then carried by the membrane folate transporter PCFT/HCPI into the cytoplasm; the caveola recycles to the cell surface, where its high-affinity receptors are reutilized. The GPI-linked transporters may be involved in transport of oxidized folates and folate breakdown products to the liver for excretion in bile.

Biochemical functions

Folates (as the intracellular polyglutamate derivatives) act as co-enzymes in the transfer of single-carbon units from one

compound to another (see Figure 5.3 and Table 5.2). Two of these reactions are involved in purine and one in pyrimidine synthesis necessary for DNA and RNA replication. Folate is co-enzyme for methionine synthesis, in which cobalamin is also involved and THF is generated. THF, the substrate for polyglutamate synthesis, is also the acceptor of single-carbon units newly entering the active pool via conversion of serine to glycine. Methionine, the other product of the methionine synthase reaction, is the precursor for SAM, the universal methyl donor involved in over 100 methyltransferase reactions.

During thymidylate synthesis, 5,10-methylene-THF is converted to dihydrofolate (Figure 5.3). The enzyme dihydrofolate reductase converts this to THF. The drugs methotrexate, pyrimethamine and, mainly in bacteria, trimethoprim inhibit dihydrofolate reductase, and this prevents formation of the active folate co-enzymes from dihydrofolate. A small fraction of the folate co-enzyme is not recycled during thymidylate synthesis but is degraded at the C-9–N-10 bond.

Causes of folate deficiency (Table 5.6)

Nutritional

Dietary folate deficiency is common. Indeed, in most patients with folate deficiency a nutritional element is present. Certain individuals are particularly likely to have diets containing inadequate amounts of folate, including the old, edentulous, poor, alcoholic and psychiatrically disturbed, and patients after gastric operations. With total cessation of intake or absorption, depletion of stores will occur in 3–6 months. In the USA and other countries where fortification of the diet with folic acid has been adopted to reduce the incidence of NTDs, the prevalence of folate deficiency has dropped dramatically and is now almost restricted to high-risk groups with increased folate needs. Nutritional folate deficiency occurs in kwashiorkor and scurvy, and in infants with repeated infections or who are fed solely on goats' milk, which has a low folate content (6 μg/L) compared with human or cows' milk (50 μg/L).

Malabsorption

Malabsorption of dietary folate occurs in tropical sprue and in gluten-induced enteropathy in children and in adults. In the rare recessive congenital syndrome of selective malabsorption of folate, there is an associated defect of folate transport into the cerebrospinal fluid, and these patients show megaloblastic anaemia from the age of a few months, responding to physiological doses of folic acid given parenterally but not orally, or large oral doses of 5-formyl-THF. They show mental retardation, convulsions and other central nervous system abnormalities. Loss-of-function mutations, usually homozygous in the gene coding for the low pH transporter PCFT/HCPI, underlie the disease. Minor degrees of malabsorption may also occur following jejunal resection or partial gastrectomy, in Crohn's disease and in systemic infections but, in these conditions, if severe deficiency occurs, it is usually largely due to poor nutrition.

Malabsorption of folate has been described in patients receiving sulfasalazine, cholestyramine and triamterene. It has also been associated with anticonvulsant drug therapy and alcohol abuse. In the intestinal stagnant loop syndrome, the predominant effect of the small intestinal bacteria is to cause a rise in serum, red cell and urinary folate by synthesizing folate, which is then absorbed.

Table 5.6 Causes of folate deficiency.

Dietary
Particularly in old age, infancy, poverty, alcoholism, chronic invalids and the psychiatrically disturbed; may be associated with scurvy or kwashiorkor

Malabsorption
Major causes of deficiency
Tropical sprue, gluten-induced enteropathy in children and adults, and in association with dermatitis herpetiformis, specific malabsorption of folate, intestinal megaloblastosis caused by severe cobalamin or folate deficiency

Minor causes of deficiency
Extensive jejunal resection, Crohn's disease, partial gastrectomy, congestive heart failure, Whipple disease, scleroderma, amyloid, diabetic enteropathy, systemic bacterial infection, lymphoma, sulfasalazine

Excess utilization or loss
Physiological
Pregnancy and lactation, prematurity

Pathological
Haematological diseases: chronic haemolytic anaemias, sickle cell anaemia, thalassaemia major, myelofibrosis
Malignant diseases: carcinoma, lymphoma, leukaemia, myeloma
Inflammatory diseases: tuberculosis, Crohn's disease, psoriasis, exfoliative dermatitis, malaria
Metabolic disease: homocystinuria
Excess urinary loss: congestive heart failure, active liver disease
Haemodialysis, peritoneal dialysis

Antifolate drugs
Anticonvulsant drugs (phenytoin, primidone, barbiturates), sulfasalazine
Nitrofurantoin, tetracycline, anti-tuberculosis (less well documented)

Mixed causes
Liver diseases, alcoholism, intensive care units

Note: In severely folate-deficient patients with causes other than those listed under Dietary, poor dietary intake is often present.

Excess utilization or loss

Pregnancy

Folate requirements are increased by 200–300 μg to about 400 μg daily in a normal pregnancy, partly because of transfer of the vitamin to the fetus, but mainly because of increased folate catabolism due to cleavage of folate co-enzymes in rapidly proliferating tissues. Megaloblastic anaemia due to this deficiency is now largely prevented by prophylactic folic acid therapy. It occurred in 0.5% of pregnancies in the UK and other Western countries, but the incidence is much higher in countries where the general nutritional status is poor. The deficiency is more common in pregnant women who, because of an inadequate diet, are also iron deficient than in those with normal iron stores. Occasionally, when there is an associated infection, acute arrest of haemopoiesis with pancytopenia may occur; this resembles aplastic anaemia, except that the marrow shows obvious megaloblastic changes.

A number of consequences of folate deficiency in pregnancy have been described, including antenatal and postpartum haemorrhages, prematurity and congenital malabsorption in the fetus. These have not been fully established, but several studies have shown that prophylactic folic acid therapy reduces the incidence of NTDs (see p. 70).

Prematurity

The newborn infant has higher serum and red cell folate concentrations than the adult, but the newborn infant's demand for folate has been estimated to be up to 10 times that of adults on a weight basis and the neonatal folate level falls rapidly to the lowest values at about 6 weeks of age. The falls are steepest and liable to reach subnormal levels in premature babies, a number of whom develop megaloblastic anaemia responsive to folic acid at about 4–6 weeks of age. This occurs particularly in the smallest babies (<1500 g birth weight) and in those who have feeding difficulties or infections, or who have undergone multiple exchange transfusions. In these babies, prophylactic folic acid should be given.

Haematological disorders

Folate deficiency frequently occurs in chronic haemolytic anaemia and other conditions of increased cell turnover because it is not completely reutilized after performing co-enzyme functions, and it is partly lost as pteridines. Patients with primary myelofibrosis may develop folate deficiency at some stage of the illness. There is also a high incidence of mild folate deficiency in patients with leukaemia, lymphoma, myeloma or carcinoma, although it is unusual for this to progress to megaloblastic anaemia. Treatment with folic acid should be avoided (as it may 'feed' the tumour) unless severe megaloblastic anaemia due to folate deficiency is clinically important.

Inflammatory conditions

Chronic inflammatory diseases, such as tuberculosis, rheumatoid arthritis, Crohn's disease, psoriasis, exfoliative dermatitis, bacterial endocarditis and chronic bacterial infections, cause deficiency by reducing the appetite and by increasing the demand for folate. Systemic infections may also cause malabsorption of folate. Severe deficiency is virtually confined to those patients with the most active disease and the poorest diet. Fever *per se* has also been suggested to interfere with folate metabolism by inhibiting temperature-dependent folate enzymes. In patients with subclinical folate deficiency from causes other than infections, intercurrent infections may precipitate severe megaloblastic anaemia.

Homocystinuria

This is a rare metabolic defect in the conversion of homocysteine to cystathionine. Folate deficiency occurring in most of these patients may be due to excessive utilization because of compensatory increased conversion of homocysteine to methionine.

Long-term dialysis

As folate is only loosely bound to plasma proteins, it is easily removed from plasma by haemodialysis or peritoneal dialysis (in contrast, cobalamin is not removed from plasma by dialysis as it is firmly protein bound). The amount of body folate that can be removed in this way is relatively small. Nevertheless, in patients with anorexia, vomiting, infections and haemolysis, folate stores may become depleted and megaloblastic anaemia can supervene. Routine folate prophylaxis is now given.

Congestive heart failure, liver disease

Excess urinary folate losses of more than 100 μg per day may occur in some of these patients. The explanation appears to be release of folate from damaged liver cells.

Antifolate drugs

A large number of epileptics who are receiving long-term therapy with phenytoin (Dilantin) or primidone (Mysoline), with or without barbiturates, develop low serum and red cell folate levels. In some of these patients, megaloblastic anaemia supervenes. A number of mechanisms have been suggested: inhibition of folate absorption, inhibition of the action or synthesis of folate-dependent enzymes, displacement of folate from its plasma transport protein and induction of folate-utilizing enzymes. A dietary element is present in the patients with the severest deficiencies.

Alcohol may also be a folate antagonist, as patients who are drinking spirits may develop megaloblastic anaemia that will respond to normal quantities of dietary folate or to physiological doses of folic acid only if the alcohol is withdrawn. Chronic alcohol intake is associated with macrocytosis even when folate levels are normal. Inadequate folate intake is the major factor in the development of deficiency in spirit-drinking alcoholics. Beer is relatively folate-rich in some countries, depending on the technique used for brewing.

The drugs that inhibit dihydrofolate reductase include methotrexate, pyrimethamine and trimethoprim. Methotrexate

has the most powerful action against the human enzyme, whereas trimethoprim is most active against the bacterial enzyme and is only likely to cause megaloblastic anaemia when used in conjunction with sulfamethoxazole in patients with pre-existing folate or cobalamin deficiency. The activity of pyrimethamine is intermediate. The antidote to these drugs is folinic acid (5-formyl-THF).

Congenital abnormalities of folate metabolism

A number of infants have been described with congenital defects of folate enzymes (e.g. cyclohydrolase or methionine synthase). Some had megaloblastic anaemia.

Diagnosis of folate deficiency

Serum folate

This is measured by an enzyme-linked immunosorbent assay (ELISA). It is low in all folate-deficient patients. In most laboratories, the normal range is from 3.0 μg/L (16.5 nmol/L) to about 15 μg/L. The serum folate is markedly affected by recent diet; inadequate intake for as little as 1 week may cause the level to become subnormal.

The serum folate level rises in severe cobalamin deficiency because of blockage in conversion of methyl-THF, the major circulating form, to THF; raised levels have also been reported in the intestinal stagnant loop syndrome, acute renal failure and active liver damage. (High levels are also obtained when the patient is receiving folic acid therapy, or, if a sample is haemolysed, because of the high concentration of folate in red cells.)

Red cell folate

The red cell folate assay is a valuable test of body folate stores. It is less affected by recent diet and traces of haemolysis than is the serum assay. In normal adults, concentrations range from 160 to 640 μg/L of packed red cells. Subnormal levels occur in patients with megaloblastic anaemia due to folate deficiency, but also occur in nearly two-thirds of patients with megaloblastic anaemia due to cobalamin deficiency. If cobalamin deficiency is excluded, however, a low red cell folate can be used as an indication that severe folate deficiency is present and warrants full investigation and treatment. False normal results may occur if the folate-deficient patient has received a recent blood transfusion or if the patient has a raised reticulocyte count (e.g. due to haemorrhage or haemolytic anaemia).

Serum homocysteine (see previously)

General management of megaloblastic anaemia

It is usually possible to establish which of the two deficiencies, folate or cobalamin, is the cause of the anaemia and to treat only with the appropriate vitamin. In patients who enter hospital severely ill, however, it may be necessary to treat with both vitamins in large doses once blood samples have been taken for cobalamin and folate assay and a bone marrow has been performed (if deemed necessary). Transfusion is usually unnecessary and inadvisable. If it is essential, packed red cells should be given slowly and one or two units will be ample. Exchange transfusion, as well as the usual treatment for heart failure, should be considered in patients with extreme anaemia and congestive heart failure. Platelet concentrates are of value in reducing spontaneous bleeding in rare patients with severe thrombocytopenia.

Treatment of cobalamin deficiency

It is usually necessary to treat patients who have developed cobalamin deficiency with lifelong regular cobalamin therapy. In the UK, the form used is hydroxocobalamin; in the USA cyanocobalamin is used. In a few instances, the underlying cause of cobalamin deficiency can be permanently corrected, for instance the fish tapeworm, tropical sprue or an intestinal stagnant loop that is amenable to surgery.

The indications for starting cobalamin therapy are well-documented megaloblastic anaemia or neuropathy due to the deficiency. It is also necessary to treat any patients with haematological abnormalities due to cobalamin deficiency, even in the absence of anaemia (e.g. hypersegmented neutrophils or megaloblastic erythropoiesis). Patients with borderline serum cobalamin levels, but no haematological or other abnormality should, if practicable, be followed, for example at yearly intervals, to ensure that the cobalamin deficiency does not progress. Alternatively hydroxocobalamin 1 mg every 6 months can be started. Cobalamin should be given routinely to all patients who have had a total gastrectomy or ileal resection. Patients who have undergone gastric reduction for control of obesity or who are receiving long-term treatment with proton pump inhibitors should be screened and given cobalamin replacement as necessary.

Replenishment of body stores should be complete with six 1 mg intramuscular injections of hydroxocobalamin given at 3–7 day intervals. More frequent doses are usually used in patients with cobalamin neuropathy, but there is no evidence that these produce a better response. For maintenance therapy, hydroxocobalamin 1 mg intramuscularly once every 3 months is satisfactory. In the USA, hydroxocobalamin has not yet been approved and marketed for purposes of routine cobalamin replacement. Because of the poorer retention of cyanocobalamin, maintenance treatment protocols generally use higher and more frequent doses (1 mg i.m. monthly). Toxic reactions to cobalamin therapy are extremely rare and are usually due to contamination in its preparation rather than to cobalamin itself. Even when there is complete failure of the physiological IF-dependent mechanism, large daily oral doses (1 mg) of cyanocobalamin can be used for replacement and maintenance of normal cobalamin

status, for example in those who cannot have injections. If this approach is used, it is important to monitor compliance, particularly with elderly forgetful patients. Sublingual and nasal routes have also been proposed, but no long-term follow-up data are available.

Cobalamin therapy has been used in a wide range of conditions such as the chronic fatigue syndrome, multiple sclerosis and other neurological diseases in the absence of cobalamin deficiency. There are no controlled trials showing that any benefit is more than a placebo effect.

Treatment of folate deficiency

There is probably never any need to give folic acid parenterally, except in patients receiving parenteral nutrition who cannot swallow tablets. Oral doses of 5–15 mg folic acid daily are satisfactory, as sufficient folate is absorbed from these extremely large doses even in patients with severe malabsorption. The length of time therapy must be continued depends on the underlying disease. It is customary to continue therapy for about 4 months, when all folate-deficient red cells will have been eliminated and replaced by new folate-replete populations.

Before large doses of folic acid are given, cobalamin deficiency must be excluded and, if present, corrected, otherwise cobalamin neuropathy may develop, despite response of the anaemia of cobalamin deficiency to folate therapy. Long-term folic acid therapy is required when the underlying cause of the deficiency cannot be corrected and the deficiency is likely to recur, for instance in chronic haemolytic anaemias such as thalassaemia major and sickle cell anaemia, and in primary myelofibrosis. It may also be necessary in gluten-induced enteropathy if this does not respond to a gluten-free diet. Where mild but chronic folate deficiency occurs, it is preferable to encourage improvement in the diet after correcting the deficiency with a short course of folic acid. In any patient receiving long-term folic acid therapy, it is important to measure the serum cobalamin level at regular (e.g. once yearly) intervals to exclude the coincidental development of cobalamin deficiency.

Folinic acid (5-formyl-THF)

This is a stable form of fully reduced folate. It is given orally or parenterally to overcome the toxic effects of methotrexate or other dihydrofolate reductase inhibitors (as present in co-trimoxazole).

Prophylactic folic acid

In over 70 countries, food is fortified with folic acid (in grain or flour) to reduce the incidence of NTDs. In Europe this has not occurred, partly due to concern that fortification would delay the diagnosis of pernicious anaemia and allow a neuropathy to develop. The levels of fortification recommended, however, are below those which would elicit a haematological response in cobalamin deficiency; particularly since small amounts of folic acid (<400 μg) in any meal would be converted by the gut to methyl-THF, which cannot be used by cobalamin deficient cells. Studies have shown no exacerbation by high blood folate levels of metabolic abnormalities in subjects with low vitamin B_{12} status. Moreover there has been no increase in the proportion of subjects with low serum cobalamin levels and no anaemia, since fortification of the diet in the USA since 1998 and no reports in any country of an increased incidence of cobalamin neuropathy. Concerns about the risk of causing cancer of any type have been allayed by the meta-analysis involving nearly 50,000 subjects discussed above.

Pregnancy

Folic acid 400 μg daily should be given as a supplement throughout pregnancy. In women who have had a previous fetus with an NTD, 5 mg daily is recommended when pregnancy is contemplated and throughout the subsequent pregnancy. In women of childbearing age, a supplementary intake of folic acid 400 μg daily is recommended, so that this extra intake will be present from conception. Recent studies, however, show poor uptake of folic acid supplements by women of child-bearing age likely to become pregnant, especially in certain ethnic groups.

Prematurity

The incidence of folate deficiency is so high in the smallest premature babies during the first 6 weeks of life that folic acid (e.g. 1 mg daily) should be given routinely to babies weighing less than 1500 g at birth and to larger premature babies who require exchange transfusions or develop feeding difficulties, infections or vomiting and diarrhoea.

Haemolytic anaemia and dialysis

Prophylactic folic acid is usually also given to patients with chronic haemolytic anaemia, for example sickle cell anaemia, or those who are undergoing long-term haemodialysis. An annual serum cobalamin assay is advisable to diagnose any patient developing the deficiency.

Megaloblastic anaemia not due to cobalamin or folate deficiency or altered metabolism

This may occur with many antimetabolic drugs (e.g. hydroxycarbamide, cytosine arabinoside, 6-mercaptopurine) that inhibit DNA replication at a particular point in the supply of precursors or inhibit DNA polymerase. In the rare disease orotic aciduria, two consecutive enzymes in purine synthesis are defective. The condition responds to therapy with uridine, which bypasses the block. In thiamine-responsive megaloblastic

anaemia, there is a genetic defect in the high-affinity thiamine transport (*SLC19A2*) gene. This causes defective RNA ribose synthesis through impaired activity of transketolase, a thiamine-dependent enzyme in the pentose cycle. This leads to reduced nucleic acid production and consequent induction of cell cycle arrest or apoptosis. It may be associated with diabetes mellitus and deafness and the presence of many ringed sideroblasts in the marrow (see p. 38). The biochemical explanation is unclear for megaloblastic changes in the marrow in patients with acute myeloid leukaemia, and other leukaemias and myelodysplasia in the absence of cobalamin or folate deficiency.

Other nutritional anaemias

Protein deficiency

Anaemia is usual in children and adults with severe protein deficiency (kwashiorkor). The anaemia, which may be partly masked by haemoconcentration, is usually normoblastic, but megaloblastic changes have been described in 10–60% of patients in different series. Hypoplasia, or even aplasia, of the marrow has also been reported. The mechanism by which protein deficiency causes anaemia is not completely understood. Lack of protein does not seem to reduce haemoglobin synthesis directly. Studies in experimental animals suggest that the major factor is diminution of erythropoietin secretion. This is probably due to a reduction in general tissue metabolism and therefore oxygen consumption, with a consequent reduced stimulus for erythropoietin secretion. In most patients, other factors contribute to the anaemia. These include infections, deficiencies of folate and iron, and also, possibly, deficiencies of vitamins C, E and B_{12}, and other trace substances. Riboflavin deficiency may also contribute to the anaemia and become apparent only during the response to protein.

Scurvy

There is usually a moderate or severe normocytic, normochromic anaemia in scurvy because of external haemorrhage and haemorrhage into tissues, and from impaired erythropoiesis. In some patients, the anaemia is megaloblastic, which appears to be partly due to associated nutritional folate deficiency and partly due to impairment of folate metabolism caused by vitamin C deficiency. However, vitamin C is not established as playing a role in normal folate metabolism.

Other deficiencies

Deficiencies of nicotinic acid and pantothenic acid cause anaemia in experimental animals, but have not been shown to do so in humans. However, riboflavin deficiency may cause anaemia in humans, resembling the anaemia of protein deficiency. Copper is essential for haematopoiesis and normal iron metabolism, and deficiency of copper causes an anaemia resembling that of iron deficiency in experimental animals. However, anaemia due to copper deficiency has never been documented in humans. Copper excess (as in Wilson disease) causes a haemolytic anaemia.

Selected bibliography

Bunn HF (2014) Vitamin B12 and pernicious anemia: the dawn of molecular medicine. *New England Journal of Medicine* **370**: 773–6.

Carmel R (2012) Subclinical cobalamin deficiency. *Current opnion in Gestroenterology* **28**: 151–58.

Correa A, Gilboa SM, Botto LD *et al.* (2012) Lack of periconceptional vitamins or supplements that contain folic acid and diabetes-associated birth defects. *American Journal of Obstetrics and Gynecology* **206**: 218.e1–13.

Devalia V, Hamilton MS, Molloy AM (2014) Guidelines for the diagnosis and treatment of cobalamin and folate disorders. *British Journal of Haematology* **166**: 496–513.

Ford A, Almeida OP (2012) Effect of homocysteine lowering treatment on cognitive function: a systematic review and meta-analysis of randomized controlled trials. *Journal of Alzheimer's Disease* **29**: 133–49.

Nielsen MJ, Rasmussen MR, Andersen CBF *et al.* (2012) Vitamin B12 transport from food to the body's cells: a sophisticated, multistep pathway. *Nature Reviews Gastroenterology and Hepatology* **9**: 345–354.

Roman-Garcia P, Quiros-Gonzalez I, Mottram L *et al.* (2014) Vitamin B12-dependent taurine synthesis regulates growth and bone mass. *Journal of Clinical Investigation* **124**: 2988–3002.

Shipton MJ, Thachil J *et al.* (2015) Vitamin B12 deficiency – A 21st century perspective. *Clinical Medicine* **15**: 145–50.

Toh B-H, Chan J, Kyaw T, Alderuccio F (2012) Cutting edge issue in autoimmune gastritis. *Clinical Reviews in Allergy and Immunology* **42**: 269–78.

Vollset SE, Clark R, Lewington S *et al.* (2013) Effects of folic acid supplementation on overall and site-specific cancer incidence during randomised trials: meta-analysis of data on 50,000 individuals. *Lancet*, **381**: 1029–36.

Wald D S, Morris J K, Wald N J (2011) Reconciling the evidence on serum homocysteine and ischaemic heart disease; meta-analysis. *PloS One* **6**: e1643.

Yang H-T, Lee M, Hong KS, Ovbiagele B, Saver JL (2012) Efficacy of folic acid supplementation in cardiovascular disease prevention. An updated meta-analysis of randomized controlled trials. *European Journal of Internal Medicine* **23**: 745–54.

CHAPTER 6

Haemoglobin and the inherited disorders of globin synthesis

6

Swee Lay Thein[1,2] and David Rees[1,2]

[1]King's College London, Molecular Haematology, Faculty of Life Sciences & Medicine, London, UK
[2]Department of Haematological Medicine, King's College Hospital NHS Foundation Trust, London, UK

Introduction

Inherited disorders of haemoglobin are the commonest single-gene disorders, with an estimated carrier rate of 7% among the world population. They occur at particularly high frequencies in populations of the tropical and subtropical belt, and consist mainly of the α- and β- thalassaemias, and the haemoglobin variants S, C and E. In many low-income countries, as economic conditions improve and infant death rates from infection and malnutrition fall, the genetic disorders of haemoglobin start to become a major burden on the health services, a phenomenon that has already been observed in many parts of the world. As a result of migrations of populations, these conditions are being seen with increasing frequency in many countries in which they had not been recognized previously.

The structure, genetic control and synthesis of haemoglobin

Different haemoglobins are synthesized in the embryo, fetus and adult, each adapted to their particular oxygen requirements (Figure 6.1). They all have a tetrameric structure made up of two different pairs (one α-like and one β-like) of globin chains, each attached to one haem molecule, the moiety responsible for the reversible binding and transfer of oxygen.

The embryonic haemoglobins include Hb Portland ($\zeta_2\gamma_2$), Hb Gower 1 ($\zeta_2\varepsilon_2$) and Hb Gower 2 ($\alpha_2\varepsilon_2$). In the fetus, HbF ($\alpha_2\gamma_2$) predominates; in adults, HbA ($\alpha_2\beta_2$) comprises over 95% of the total haemoglobin, with a minor component of HbA_2 ($\alpha_2\delta_2$) in the red blood cells. There are two kinds of HbF composed of γ-chains that differ in their amino acid composition at position 136, where they have either glycine or alanine; those with glycine are called $^{G}\gamma$-chains and those with alanine $^{A}\gamma$-chains. The $^{G}\gamma$ and $^{A}\gamma$ chains are the products of separate globin gene loci ($^{G}\gamma$ and $^{A}\gamma$). These different types of haemoglobin are adapted to the changes in physiological requirements that occur during development. Fetal haemoglobin (HbF) exhibits a higher oxygen affinity than adult haemoglobins *in vivo*; the higher oxygen affinity of HbF relative to adult haemoglobin facilitates the transfer of oxygen across the placenta from the maternal to the fetal circulation.

The sigmoid shape of the oxygen dissociation curve, which reflects the allosteric properties of haemoglobin, ensures that oxygen is rapidly taken up at the high oxygen tensions found in the lungs and is released readily at the low tensions encountered in the tissues. It is quite different to myoglobin, a molecule that consists of a single globin chain with haem attached to it and which has a hyperbolic dissociation curve. The transition from a hyperbolic to a sigmoid curve reflects cooperativity between the four haem molecules. When one haem takes on oxygen, the affinity for oxygen of the remaining haem molecules of the tetramer increases markedly. This is because haemoglobin can exist in two configurations, deoxy(T) and oxy(R) (T and R stand for tight and relaxed states, respectively). The T form has a lower affinity than the R form for ligands such as oxygen. At some point during the sequential addition of oxygen to the four haems, transition from the T to R configuration occurs and the oxygen affinity of the partially liganded molecule increases

Postgraduate Haematology, Seventh Edition. Edited by A Victor Hoffbrand, Douglas R Higgs, David M Keeling and Atul B Mehta.

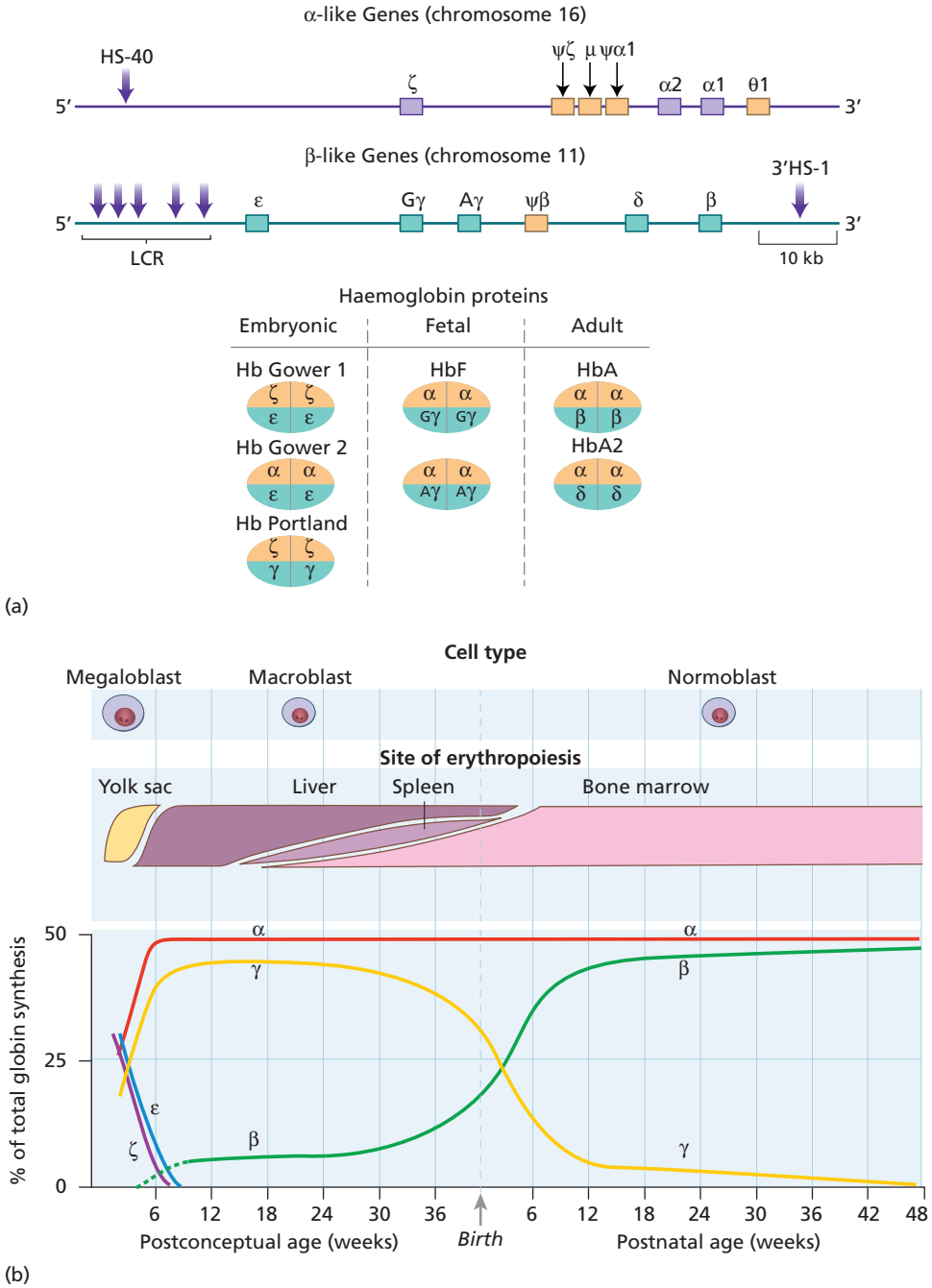


Figure 6.1 (a) The genomic structure of the clusters of α-like and β-like globin genes, on chromosomes 16 and 11, in human beings. (b) The timeline of the expression of the human globin genes from early stages of fetal development to the changes that occur at birth and in the first year of life. (Source: Schechter, 2008 [*Blood* 2008; **112**: 3927–3938]. Reproduced with permission of the American Society of Hematology.)

dramatically. The oxygen dissociation curve, which reflects these changes, can be modified in several ways. First, oxygen affinity is decreased with increasing CO_2 concentrations – the Bohr effect. This facilitates oxygen loading to the tissues, where a drop in pH due to CO_2 influx lowers oxygen affinity. In contrast, in the lungs, efflux of CO_2 and an increase in intracellular pH increases oxygen affinity and hence uptake. Oxygen affinity is also modified by the level of 2,3-diphosphoglycerate

(2,3-DPG) in the red cell. Increasing concentrations shift the oxygen dissociation curve to the right (i.e. reduce oxygen affinity), whereas diminishing concentrations have the opposite effect.

Genetic control, regulation and synthesis

Human haemoglobin production is characterized by two 'switches'. The switch from embryonic to fetal haemoglobin production begins as early as week 5 of gestation and is completed by week 10 (Figure 6.1). Expression of β-globin starts as early as week 8, but synthesis remains low, increasing to approximately 10% at weeks 30–35 of gestation, with a dramatic upregulation of β-globin synthesis just before birth, coinciding with a decrease in γ-globin expression that constitutes the fetal to adult haemoglobin switch. At birth, HbF ($\alpha_2\gamma_2$) comprises 60–80% of the total haemoglobin, falling to about 5% at 6 months of age and eventually reaching the adult level of 0.5–1.0% at 2 years. The relative synthesis of $^{G}\gamma$ and $^{A}\gamma$ chains also changes with the switch, from a $^{G}\gamma/^{A}\gamma$ ratio of 3:1 in fetal life to a ratio of 2:3 in adults. The switch from fetal to adult haemoglobin production is not total, production of variable levels of HbF persisting throughout adult life. These residual amounts are unevenly distributed; erythrocytes that contain measurable amounts of HbF are termed F cells.

Each of the α-like and β-like globin chains is encoded by genetically distinct loci, the α-like cluster on the tip of chromosome 16p and the β-like cluster on chromosome 11p15.5 (Figure 6.1). In both clusters, the genes are arranged along the chromosome in the order in which they are expressed during development: 5′-ε-$^{G}\gamma$-$^{A}\gamma$-ψβ-δ-β-3′ and 5′-ζ-ψζ-ψα2-ψα1-α2-α1–3′. The ψβ, ψζ and ψα-genes are pseudogenes, that is they have sequences that resemble the β, ζ or α-genes, but contain inactivating mutations that prevent them from being expressed. They may be 'burnt out' remnants of genes that were functional at an earlier stage of evolution. Like most mammalian genes, the globin genes have one or more non-coding inserts, called intervening sequences or introns, interrupting the coding sequences or exons.

The β-like globin genes have three exons (coding regions) interrupted by two intervening sequences or introns of 122–130 and 850–900 bp, respectively. The β-genomic sequence codes for 146 amino acids; intron 1 interrupts the sequence between codons 30 and 31, and intron 2 between codons 104 and 105. The α-globin genes code for 141 amino acids and contain similar, but smaller, introns between codons 30 and 31 and between codons 99 and 100. Within each α- and β-globin complex, in addition to the primary *cis* determinants of individual globin gene expression, which are found in the immediate vicinity and within each gene, there are other local regulatory elements known as enhancers, which are located at variable distances from the individual genes.

The local *cis*-acting sequences controlling globin gene expression include the promoter region, splicing donor and acceptors, and poly-A addition sites. The promoter, in the 5′ flanking region, includes blocks of nucleotide homology that are found in analogous positions in many species (Figure 6.2). The three positive *cis*-acting elements include the TATA box (position −28 to −31, i.e. between 28 and 31 bases upstream from the mRNA 'cap' site), a CCAAT box (position −72 to −76), and a CACCC motif which may be inverted or duplicated (position −80 to −140). These promoter elements are recognized by transcription factors and are involved in the initiation of transcription. It is interesting that while the CCAAT and TATA elements are found in many eukaryotic promoters, the CACCC sequence is found predominantly in erythroid cell-specific promoters.

The 5′ untranslated region (UTR) occupies a region of about 50 nucleotides between the 5′ terminus or 'cap' site of globin mRNA and the initiation (ATG) codon. The cap appears to be important for maintaining the stability of the nuclear precursor mRNA. Within the 5′-UTR of the various globin genes there are conserved sequences that are important in the regulation of gene expression. The 3′-UTR constitutes the region between the termination codon and the poly-A tail. It consists of about 130 nucleotides with one conserved sequence, AATAAA, located 20 nucleotides upstream of the poly-A tail. The conserved hexanucleotide AATAAA acts as a signal for cleavage of the 3′ end of the primary transcript and addition of the poly-A tail, which confers stability on the processed mRNA and enhances translation. The importance of all these *cis* elements for normal globin gene expression has been validated by the discovery of thalassaemias caused by several mutations affecting these regions, as well as by deletion experiments.

Throughout development, the appropriate genes of the α- and β-globin gene clusters are coordinately expressed, maintaining a balance in the production of α- and β-like globins needed for the synthesis of normal haemoglobin. The regulation of globin gene expression is mediated at several levels; although most occurs at the transcriptional level, there is some fine-tuning during and after translation. Most DNA that is not involved in gene transcription is tightly packaged into a compact, chemically modified form that is inaccessible to transcription factors and polymerases and which is heavily methylated. Activity is associated with a change in the structure of the chromatin surrounding a gene, which can be identified by enhanced sensitivity to nucleases. Erythroid lineage-specific nuclease hypersensitivity sites are found at several locations in both the β-globin and α-globin gene clusters. A set of five DNase I hypersensitivity sites (designated HSs1–5), distributed 5–25 kb 5′ of the ε-globin gene, constitute the β-locus control region (β-LCR) (see Figure 6.1). This region was originally implicated in the control of the β-globin complex by the discovery of natural mutants that removed sequences upstream, but left the downstream globin genes intact and yet resulted in no output from the cluster. The human β-LCR was the first LCR to be identified and was

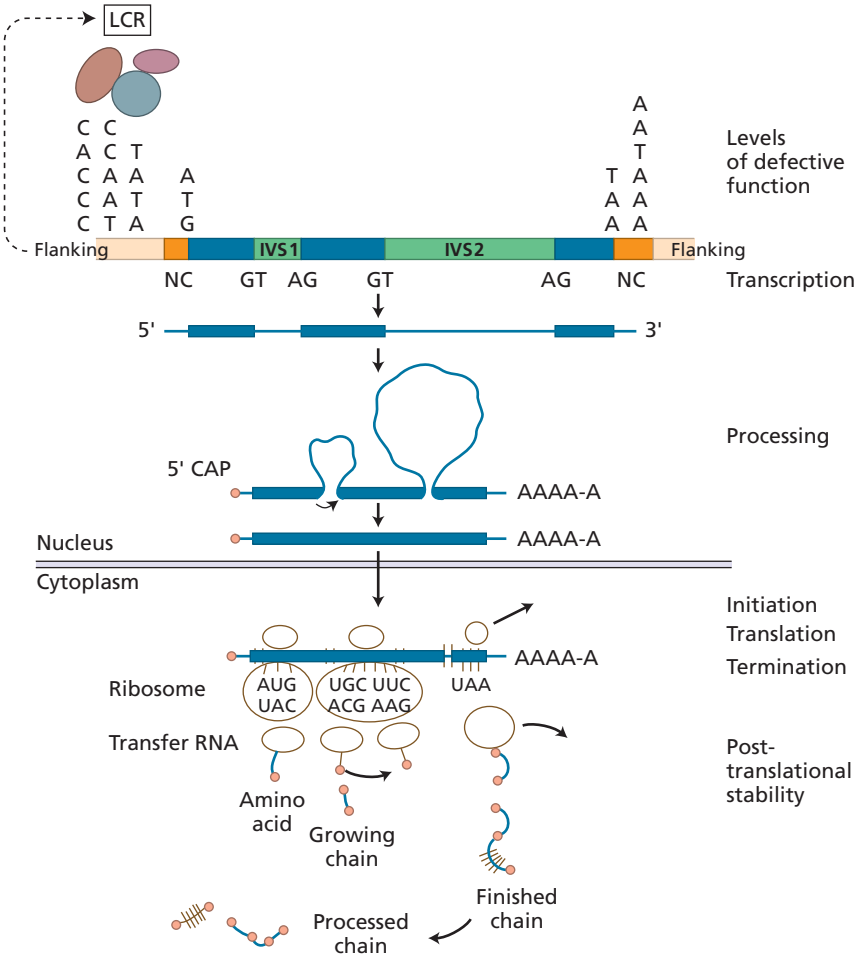


Figure 6.2 A prototype globin gene and the genetic control of globin chain synthesis. Levels of action of mutations are indicated on the right.

functionally defined as a DNA element that provides high levels of tissue-specific expression to a *cis*-linked gene in a copy-number-dependent manner, and which is independent of host-genome integration site. The corresponding region in the α-globin cluster consists of four multispecies conserved sequence (MCS) regions lying 30–70 kb upstream of the α-globin genes called MCS-R1 to MCS-R4. Of these elements, only MCS-R2, which consists of a single DNase hypersensitive site, has been shown to be essential for α-globin expression. MCS-R2, which lies 40 kb upstream of the cluster, is also known as HS-40. The β-LCR establishes a transcriptionally active chromatin domain that encompasses the whole β-globin cluster and acts as a unique enhancer, whereas the α-globin MCS-R2 is most similar to HS2 of the β-LCR and acts as an enhancer. In both clusters, expression of the respective genes is critically dependent on the presence of the upstream regulatory elements.

Several other enhancer sequences have also been identified in both globin gene clusters. All these regulatory regions bind a number of key erythroid-specific transcription factors, notably GATA-1, GATA-2, NF-E2, KLF1 (also known as EKLF) and SCL, as well as various cofactors (e.g. FOG, p300) and factors that are more ubiquitous in their tissue distribution, such as Sp1. Tissue-specific expression may be explained by the presence of binding sites for the erythroid-specific transcription factors. The binding of haemopoietic-specific factors activates the LCR, which renders the entire β-globin gene cluster transcriptionally active. Transcription factors also bind to enhancer and local promoter sequences within each gene, which work in tandem to regulate the expression of the individual genes in the clusters. Some of the transcription factors are developmental-stage specific and may be involved in the (still poorly understood) differential expression of embryonic, fetal and adult globin genes.

The mechanisms by which developmental regulation is controlled are less clear. A dual mechanism has been proposed: autonomous gene silencing and gene competition for direct interaction with the upstream LCR. It appears that the ε and ζ genes are switched on in embryonic cells and autonomously switched off in definitive cells (liver and bone marrow) in which they cannot be substantially reactivated. The second switch, from γ- to β-gene expression, is more complex and involves both autonomous silencing of the γ-genes and competition between the γ- and β-genes for the β-LCR. The

transcriptional factor environment is critical in determining the balance between γ- and β-gene expression and is thought to be mediated by changes in the repertoire and/or abundance of various nuclear factors favouring particular promoter–LCR interactions. So far, the best-defined example of a developmental-stage-specific regulatory factor is the erythroid Krüppel-like factor (EKLF, also known as KLF1)) without which the β-genes cannot be fully activated in the definitive cells. Not only is KLF1 expression restricted mainly to erythroid cells, but it is also a highly promoter-specific activator, binding with high affinity to the β-globin CACCC box. Its greater affinity for the β-globin than the γ-globin promoter accelerates the shutdown of γ in transgenic mice over-expressing klf1. A genetic network regulating the switch from γ- to β-globin expression has emerged, involving the interaction of KLF1, BCL11A and MYB with each other, and other transcription factors (e.g. GATA-1) and corepressor complexes that involve chromatin modelling and epigenetic modifiers. BCL11A, previously known as an oncogene involved in leukaemogenesis, was 'discovered' as an important genetic locus regulating HbF through genome-wide association studies (GWAS). Downstream functional studies in cell lines, primary human erythroid cells and transgenic mice, have shown that BCL11A is a repressor of γ-globin expression. KLF1 is a direct activator of *BCL11A*. KLF1 is key in the switch from γ-globin to β-globin expression; it not only activates the β-globin gene directly, providing a competitive edge, but also silences the γ-globin genes indirectly via activation of *BCL11A*. KLF1 may also play a role in the silencing of the embryonic globin genes. KLF1 has emerged as a major erythroid transcription factor with pleiotropic roles underlying many of the previously uncharacterized anaemias (e.g. congenital dyserythropoietic anaemia type IV). KLF1 variants have also been associated with variable increases in HbF and HbA_2 levels.

Transcription and processing of mRNA

Genetic information in the DNA of the globin genes is transcribed into an RNA copy which is then translated into a specific globin chain (Figure 6.2). The TATA box acts as the initial DNA target for the progressive assembly of an initial transcription complex, which involves the interaction of transcription factors, TATA-binding protein and other proteins with the β-LCR or α-MCS, mediated by RNA polymerase II.

The primary transcript is a large mRNA precursor (pre-mRNA) that contains both introns and exons. While in the nucleus, it undergoes a number of modifications before it can be translated into protein (Figure 6.2). This includes the removal of introns by a complex series of reactions involving several different proteins that constitute the spliceosome, and splicing of the exons. Consensus sequences are universally found encompassing the 5′ (donor) and 3′ (acceptor) ends. Each intron invariantly starts with the dinucleotide GT (5′) and finishes with AG (3′). Mutations that alter the normal consensus sequences or mutations that create similar consensus sequences at new sites in globin genes cause aberrant splicing and constitute the molecular basis of many types of thalassaemia. Other modifications of the nascent mRNAs include the addition of a 'cap' structure at the 5′ end and the addition of a string of adenylic acid residues (poly-A) at the 3′ end. Proper cleavage of the primary RNA transcript and polyadenylation of the 3′ ends of mRNA is guided by a consensus hexanucleotide (AATAAA) sequence about 20 nucleotides upstream of the poly-A tail. The processed mRNA now moves into the cytoplasm to act as a template for globin chain production on defined organelles known as ribosomes.

Translation

Amino acids are transported to the mRNA template on carriers called transfer RNAs; there are specific transfer RNAs for each amino acid. The order of amino acids in a globin chain is determined by the order of nucleotides (reading frame); three bases (codon) code for a particular amino acid. The transfer RNAs also contain three bases, the anticodon, which are complementary to mRNA codons for particular amino acids. The transfer RNAs carry amino acids to the template, where they find the right position by codon–anticodon base-pairing. The mRNA is translated from the 5′ to the 3′ end (left to right) starting with a specific initiation codon (AUG) and ending with a termination codon (UAA, UAG, UGA). When the ribosome reaches the termination codon, translation ceases, the completed globin chain is released, and the ribosomal subunits fall apart and are recycled. Individual globin chains combine with haem, which is synthesized through a separate pathway, and with themselves to form definitive haemoglobin molecules.

Termination codons are also called nonsense codons because they do not usually encode any amino acid. Approximately 50% of the mutations causing β-thalassaemia are caused by termination codons that are premature. Premature termination codons (PTCs) can result from different types of mutations. Single nucleotide substitutions can convert a sense codon to a nonsense codon and are often referred to as nonsense mutations. Frameshift mutations are insertions or deletions of a few bases that are not multiples of three that shift the reading frame, resulting in a nonsense codon that is premature (i.e. PTC). mRNAs with PTCs lead to production of encoded truncated proteins that are potentially harmful. They are kept in check by a cellular surveillance mechanism referred to as nonsense-mediated mRNA decay (NMD). NMD is usually triggered when translation stops prematurely at PTCs, resulting in termination of the mutant mRNA transcript and thus absence of abnormal protein. However, PTCs situated less than 50–55 nucleotides upstream of the 3′-most exon–exon junction or downstream of this junction generally fail to trigger NMD, and result in production of abnormal mRNA species. In the case of the β-globin gene, PTCs within the last exon (exon 3) and 3′ half of exon 2 result in the production of highly unstable β-globin variant chains and a dominantly inherited form of β-thalassaemia. However, there are exceptions to the 50–55 rule; PTCs within β-globin exon 1 and a PTC within

exon 2 have been reported to fail to elicit NMD efficiently despite residing more than 55 nucleotides upstream of the exon 2–exon 3 junction.

The multistep process in the conversion of DNA into protein offers numerous opportunities for mishaps to occur that result in downregulation of gene expression, clearly illustrated by the different mutations causing thalassaemias and haemoglobinopathies.

Classification of the disorders of haemoglobin

Mutations in the globin genes can cause a quantitative reduction in output from that gene, alter the amino acid sequence of the protein produced or a combination of the two (Table 6.1). Quantitative defects cause thalassaemia syndromes, whereas qualitative changes, referred to as haemoglobin variants, cause a wide range of problems, including sickle cell disease (Chapter 7), unstable haemoglobins, decreased oxygen affinity, increased oxygen affinity, and methaemoglobinaemia. However, the majority of qualitative mutations cause no significant change in haemoglobin properties or clinical problems. Some mutations combine both features, resulting in a haemoglobin variant that is made in reduced amounts; HbE (β26 Glu→Lys) is the most common example of this. The substitution at codon 26 (GAG→AAG), which causes HbE, also causes alternative splicing of the β-globin mRNA, leading to a reduction in the normally spliced β-message encoding the variant, and a thalassaemia phenotype. Other haemoglobin variants result in a thalassaemia phenotype caused by extreme instability and functional deficiency of the globin chain variant; for example, Hb Geneva, a dominantly inherited β-thalassaemia. Mutations in the globin gene complex might also alter the switch from fetal to adult haemoglobin synthesis, resulting in persistence of γ-globin expression, and hereditary persistence of fetal haemoglobin (HPFH).

Table 6.1 The thalassaemias and related disorders.

β-Thalassaemia
β^0
Deletion
Non-deletion
β^+
'Silent'
Normal HbA_2
Dominant
α-Thalassaemia
α^0
α^+
Deletion (/−α)
Non-deletion (/$\alpha^T\alpha$)
δβ-Thalassaemia
$^G\gamma^A\gamma\ (\delta\beta)^0$
$^G\gamma\ (^A\gamma\delta\beta)^0$
$(\delta\beta)^+$
γ-Thalassaemia
δ-Thalassaemia
εγδβ-Thalassaemia
Hereditary persistence of fetal haemoglobin
Deletion
Non-deletion
$^A\gamma$
$^G\gamma$

The thalassaemias and related disorders

The thalassaemias are the commonest single-gene disorders. Thalassaemia was first recognized by Cooley and Lee in 1925 as a form of severe anaemia associated with splenomegaly and bone changes in children. The term 'thalassaemia' is derived from the Greek θαλασσα-(meaning 'the sea') since many of the early cases came from the Mediterranean region. However, it is now clear that the disorder is not just limited to the Mediterranean region, but occurs throughout the world, prevalent in the tropical and subtropical regions, including the Middle East, parts of Africa, Indian subcontinent and Southeast Asia. It appears that heterozygotes for thalassaemia are protected from the severe effects of malaria and natural selection has increased and maintained their gene frequencies in these malarious regions.

Definition and classification

The thalassaemias are classified into α-, β-, δβ-, γδβ-, δ-, γ- and εγδβ-thalassaemias, according to the type of globin chain(s) that is produced in reduced amounts (Table 6.1). The two major categories are the α- and β-thalassaemias, while the rare forms include the γ-, δ- and εγδβ- thalassaemias.

Functionally, some thalassaemia mutations cause a complete absence of globin chain synthesis, and these are called α^0- or β^0-thalassaemias; in others, the globin chain is produced at a reduced rate and these are designated α^+- or β^+-thalassaemias. The δβ-thalassaemias are subdivided in the same way. HPFH syndromes refer to the group of disorders in which the switch from fetal to adult haemoglobin production is incomplete and variable levels of fetal haemoglobin levels persist in otherwise normal individuals. Because of their concomitant increased HbF levels, the δβ- and γδβ-thalassaemias are often considered with the HPFH syndromes.

Because thalassaemia occurs in populations in which structural haemoglobin variants are common, it is not unusual to

inherit a thalassaemia gene from one parent and a gene for a structural haemoglobin variant from the other. Furthermore, both α- and β-thalassaemia occur commonly in some countries, and individuals may coinherit genes for both types. These different interactions produce a clinically diverse family of genetic disorders that range in severity from death *in utero* to extremely mild, symptomless, hypochromic anaemias.

Most thalassaemias are inherited in a Mendelian recessive fashion. Heterozygotes are mostly symptomless, although usually they can be recognized by simple haematological analysis. More severely affected patients are either homozygotes for α- or β-thalassaemia or compound heterozygotes for different molecular forms of α- or β-thalassaemia or for one or other form of thalassaemia and a gene for a haemoglobin variant. Clinically, the thalassaemias are classified according to their severity into major, intermediate and minor forms. Thalassaemia major is a severe and transfusion-dependent disorder. Thalassaemia minor is the symptomless trait or carrier state. Thalassaemia intermedia is characterized by anaemia (with or without splenomegaly), though not of such severity as to require regular transfusion. In practice, thalassaemia intermedia encompasses a wide spectrum of clinical severities intermediate between the two extremes of thalassaemia major and trait. Thalassaemia intermedia, also referred to as non-transfusion-dependent thalassaemia (NTDT) remains a clinical definition and includes β-thalassaemia intermedia, HbH disease and the HbE/βthalassaemias.

The β-thalassaemias

The β-thalassaemias pose by far the most important public health problems because they are common and usually produce severe anaemia in their homozygous and compound heterozygous states.

Distribution

The β-thalassaemias occur widely in a broad belt, ranging from the Mediterranean and parts of North and West Africa through the Middle East and Indian subcontinent to Southeast Asia. The disease is particularly common in Southeast Asia, where it occurs in a line starting in southern China and stretching down through Thailand and the Malay Peninsula and Indonesia to some of the Pacific island populations. In this region, and in some of the Mediterranean island and mainland countries, gene frequencies range between 2 and 30%. It should be remembered that β-thalassaemia is not confined entirely to these high-incidence regions and it occurs sporadically in every racial group.

Genetic basis of disease: molecular pathology

The β-thalassaemias are considered to be autosomal recessive disorders since individuals who have inherited one abnormal β-gene (carrier) are asymptomatic and the inheritance of two abnormal β-globin genes is required to produce a clinically detectable phenotype. Molecular analysis of the β-thalassaemia genes has demonstrated a striking heterogeneity. Although almost 300 β-thalassaemia alleles (including deletions) have been characterized, population studies indicate that probably only 20 β-thalassaemia alleles account for more than 80% of the β-thalassaemia mutations in the whole world. This is because in each of the high-frequency areas, only a few (four to six) mutations are common, reflecting local selection due to malaria, with a varying number of rare ones. Each of these populations thus has its own unique group of mutations.

The vast majority (approximately 250) of β-thalassaemia mutations are point mutations (i.e. single-base substitutions) and small insertions or deletions of one to two bases. These may involve any step in globin chain production: transcription, translation or post-translational stability of the globin gene product (Figure 6.2). Approximately half of these mutations completely inactivate the β-gene with no β-globin production resulting in β^0-thalassaemia. Mutations that allow the production of some β-globin cause β^+- or β^{++}-thalassaemia, depending on whether there is a marked or mild reduction in the output of β-chains, respectively.

Transcription

The mutations that interfere with transcription include deletions and point mutations involving the globin gene promoter regions. With the exception of a deletion of about 600 bases at the 3′ end of the β-globin gene, which is restricted to certain Indian populations, major deletions are uncommon. A large number of point mutations involve the promoters or adjacent regions, most of which downregulate the β-globin gene to a variable degree and cause relatively mild forms of β-thalassaemia.

A couple of β-thalassaemia mutations in this class are 'silent' i.e. carriers do not have any evident haematological phenotypes, with red cell indices and HbA_2 levels within the normal range, the only abnormality being imbalanced globin chain synthesis. These β-thalassaemia mutations have usually been 'discovered' in individuals with thalassaemia intermedia resulting from compound heterozygosity for one of those 'silent' mutations in combination with a typical β-thalassaemia mutation. In this case, one parent has typical β-thalassaemia trait and the other is apparently normal. Overall, the 'silent' β-thalassaemia alleles are uncommon, except for the −101 C→T mutation that has been observed fairly frequently in the Mediterranean region, where it interacts with a variety of more severe β-thalassaemia mutations to produce milder forms of β-thalassaemia intermedia.

Processing

A wide variety of mutations interfere with processing of the primary mRNA transcript. Those involving the invariant GT or AG sequences at intron–exon junctions prevent splicing altogether and cause β^0-thalassaemia. Mutations involving the consensus sequences adjacent to the GT or AG dinucleotides in the introns

allows some normal splicing and causes β^+-thalassaemia. Several β-thalassaemia mutations involve other parts of the introns; alternative splicing sites are produced leading to variable degrees of both normal and abnormal mRNA synthesis. An incorrectly spliced mRNA is not functional because it contains intron sequences, generating a frameshift and a PTC. Sequences that resemble the consensus sequences at intron–exon junctions are also present in exons. Mutations may activate these 'cryptic' sites, again leading to abnormal splicing.

Translation

About half of the β-thalassaemia alleles completely inactivate the gene, mostly by generating PTCs, either by single-base substitution to a nonsense codon or through a frameshift mutation. As part of the surveillance mechanism that is active in quality control of the processed mRNA, mRNA harbouring a PTC is destroyed and not transported to the cytoplasm (a phenomenon called NMD) to prevent the accumulation of mutant mRNAs coding for truncated peptides. However, some in-phase PTCs that occur later in the β-sequence, in the 3′ half of exon 2 and in exon 3, escape NMD and are associated with substantial amounts of mutant β-mRNA, leading to synthesis of β-chain variants that are highly unstable and non-functional with a dominant negative effect (see next section). Other mutations of RNA translation involve the initiation (ATG) codon. Nine of these have been described; apart from an insertion of 45 bp, all are single base substitutions and again result in β^0-thalassaemia.

Mutations affecting post-translational stability

Instability of the β-globin gene product is the basis for the dominantly inherited β-thalassaemias (Figure 6.3). As discussed earlier, in-phase PTCs within the 3′ half of exon 2 and in exon 3 cannot efficiently elicit NMD and hence abnormal mRNA containing the PTCs are transported to the cytoplasm and translated. However, these truncated variant β-chains are highly unstable, non-functional and not able to form viable tetramers. They precipitate in the erythroid precursors together with the redundant α-chains, causing premature death of these cells, and accentuating the ineffective erythropoiesis. Severe anaemia and clinical disease results even in the heterozygous state. Highly unstable β-globin chains can also result from single-base substitutions or minor insertions/deletions that affect a critical amino acid of the β-globin peptide that is involved in α/β-dimer formation or haem binding. In other cases, the minor insertions/deletions lead to shifts in the reading frame resulting in long unstable β-globin gene products that form prominent inclusion bodies in red cell precursors.

Deletions restricted to the β-globin gene

β-Thalassaemia is rarely caused by deletions (Figure 6.4). Of these, only the 619-bp deletion at the 3′ end of the β-gene is common, but even that is restricted to the Sind populations of India and Pakistan where it constitutes about 30% of the β-thalassaemia alleles. The other deletions, although extremely rare, are of particular clinical interest because they are associated with unusually high levels of HbA_2 and HbF in heterozygotes. The increase in HbF is adequate to compensate for the complete absence of HbA in homozygotes for these deletions. The mechanism underlying the elevated levels of HbA_2 and HbF appears to be related to removal of the 5′ promoter region of the β-globin gene, which removes competition for the upstream β-LCR and limiting transcription factors, resulting in increased interaction of the LCR with the γ- and δ-genes in *cis*, thus enhancing their expression. This mechanism may also explain the unusually high HbA_2 levels that accompany the point mutations in the β-promoter region.

Unusual causes of β-thalassaemia

These are extremely rare and are mentioned here not just for the sake of completeness, but also to illustrate the numerous molecular mechanisms that downregulate the β-globin gene.

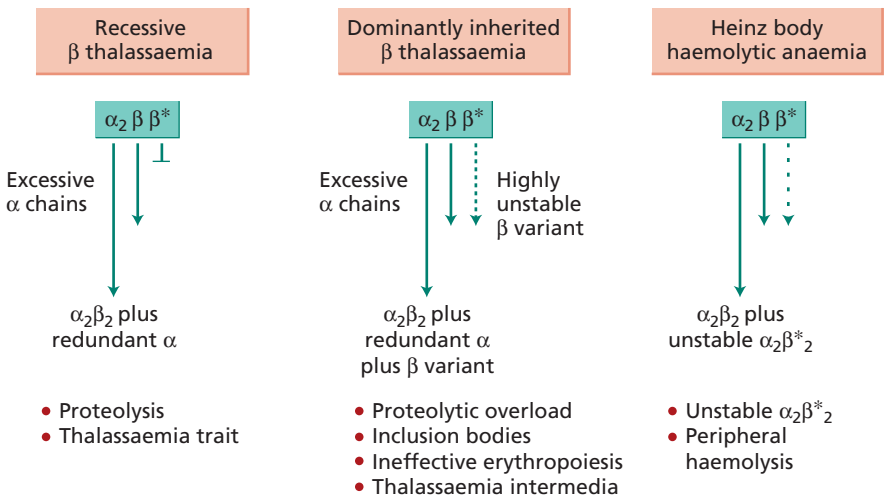


Figure 6.3 Heterozygous mutations in the β-globin gene and the different phenotypes.

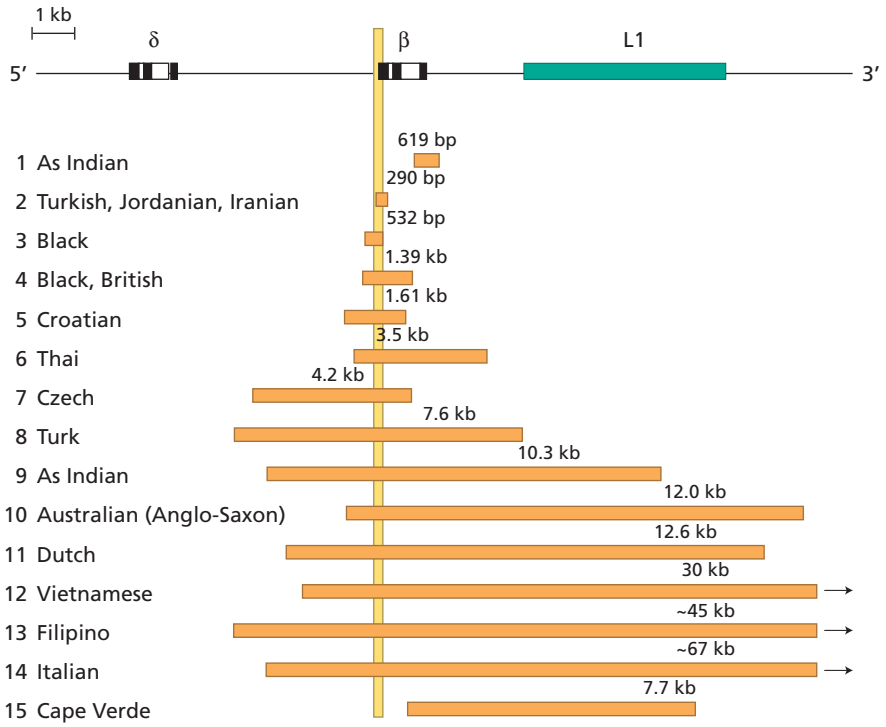


Figure 6.4 Deletions causing β-thalassaemia. The vertical bar indicates the β-globin promoter region that is removed in common by these deletions, except for the 619 bp deletion. The horizontal arrows indicate that the 3′ end of the deletions have not been defined.

Transposable elements may occasionally disrupt human genes and result in their inactivation. The insertion of such an element, a retrotransposon of the LI family, into intron 2 of the β-globin gene has been reported to cause β^+-thalassaemia. Rarely, mutations in other genes distinct from the β-globin complex can downregulate β-globin expression. Such *trans*-acting mutations have been described, affecting the XPD protein, which is part of the general transcription factor TF11H, and the key erythroid-specific transcription factor GATA-1. Somatic deletion of the β-globin gene contributed to thalassaemia intermedia in three unrelated families of French and Italian origins. The affected individuals with thalassaemia intermedia were constitutionally heterozygous for β^0-thalassaemia, but subsequent investigations revealed a somatic deletion of chromosome 11p15, including the β-globin gene complex, in *trans* to the mutation in a subpopulation of erythroid cells. This results in a somatic mosaic: 10–20% of the cells were heterozygous with one normal copy of the β-globin gene, and the rest hemizygous (i.e. without any normal β-globin gene). Unusually severe anaemia can also result from uniparental isodisomy of chromosome 11p, which encompasses the β-globin gene complex. In one case, thalassaemia major in a Chinese patient was caused by homozygosity for a paternal β-thalassaemia allele due to unipaternal isodisomy of chromosome 11p15.5, and in another, a previously healthy Portuguese adolescent carrier for β-thalassaemia became transfusion-dependent due to mosaicism of chromosome 11p14.3-11p15 that contained a β^0 CD15 mutation.

Pathophysiology

The molecular defects in β-thalassaemia result in absent or reduced β-chain production while α-globin synthesis is unaffected. The imbalance in globin chain production leads to an excess of α-chains. The free α-globin chains are highly unstable and precipitate in red cell precursors, forming intracellular inclusions that interfere with red cell maturation (Figure 6.5). There is a variable degree of intramedullary destruction of erythroid precursors (i.e. ineffective erythropoiesis) that characterizes all β-thalassaemias. Those red cells that mature and enter the circulation contain α-chain inclusions that interfere with their passage through the microcirculation, particularly in the spleen. However, the damage to red cell precursors and their progeny in β-thalassaemia is not entirely mechanical. The degradation products of excess α-chains, particularly haem and iron, produce a wide range of deleterious effects on red cell membrane proteins and lipids, manifested by marked abnormalities of electrolyte homeostasis and membrane deformability. The end result is an extremely rigid red cell with a shortened survival.

Thus, the anaemia of β-thalassaemia results from a combination of ineffective erythropoiesis and haemolysis. It stimulates erythropoietin production, which causes expansion of the bone marrow and may lead to serious deformities of the skull and long bones. Because the spleen is being constantly bombarded with abnormal red cells, it hypertrophies. The resulting splenomegaly, together with bone marrow expansion, causes a

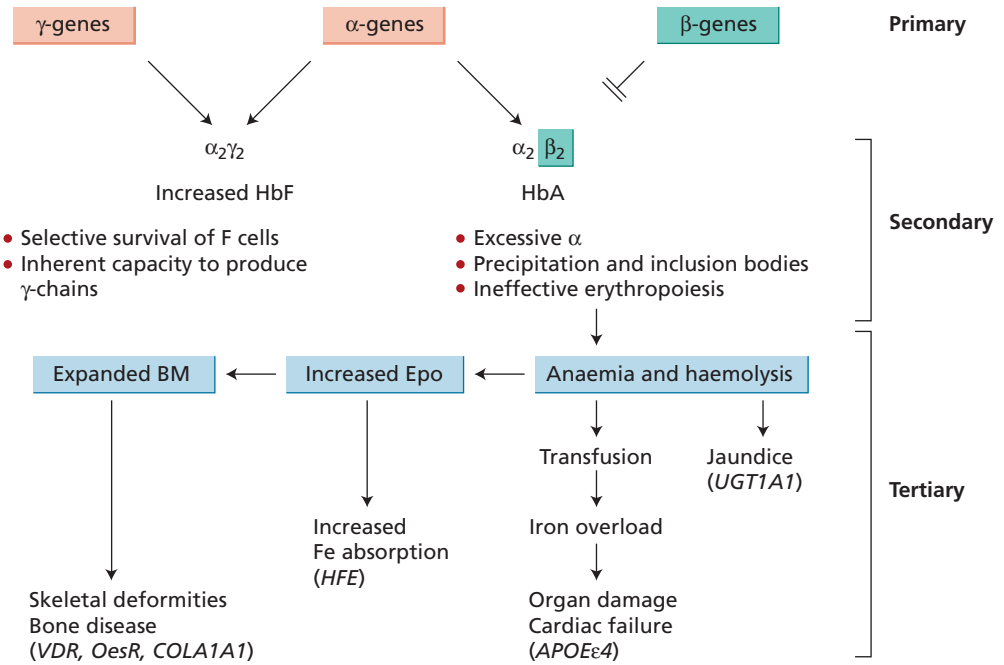


Figure 6.5 The pathophysiology of β-thalassaemia. BM, bone marrow; Epo, erythropoietin. (Source: Thein, 2004 [*Br J Haematol* **124**: 264–74]. Reproduced with permission of Wiley.)

major increase in plasma volume, which also contributes to the anaemia.

As mentioned previously, HbF production almost ceases after birth. However, some adult red cell precursors retain the ability to produce a variable amount of γ-chains. Because the latter can combine with excess α-chains to form HbF, cells which make relatively more γ-chains in the bone marrow are partly protected against the deleterious effect of α-chain precipitation. These F cells come under selection in the marrow and peripheral blood and thus individuals with β-thalassaemia have variable increases in HbF due to selective survival of these F cells. In some cases, there is also a genuine increase in HbF production, as well as selection of F cells due to coinheritance of genetic determinant(s), or quantitative trait loci (QTL), which increase HbF production.

It follows therefore that if the anaemia is corrected with blood transfusion, the erythropoietin drive is shut off, growth and development are normal, bone deformities do not occur and splenomegaly is less marked. On the other hand, each unit of blood contains 200–250 mg of iron, and with regular transfusion there is steady accumulation of iron in the liver, endocrine glands and myocardium. Thus, although well-transfused thalassaemic children grow and develop normally, they die of iron overload unless steps are taken to remove iron.

Genotype–phenotype relationships

The β-thalassaemias show remarkable phenotypic variability, ranging from severe life-threatening anaemia to an extremely mild condition that may be identified only by chance. The molecular basis for this diversity is at least partly understood.

The genetic modifiers of the β-thalassaemia phenotype can be divided into primary, secondary and tertiary (Figure 6.5). Primary modifiers are the different mutations that affect the β-globin gene. These have variable effects on β-globin gene expression that may affect the output of β-globin chains, ranging from zero to a very mild reduction. Secondary modifiers are those that reduce the degree of imbalance of globin chain synthesis. They include the coinheritance of α-thalassaemia and a variety of genetic modifiers of γ-chain production in adult life. Three major QTLs – *Xmn*1-$^{G}\gamma$ site, *HBS1L-MYB* intergenic polymorphisms (*HMIP*) on chromosome 6q, and *BCL11A* gene on chromosome 2 – have recently been mapped, and it seems likely that many remain to be discovered.

Coinheritance of α-thalassaemia reduces chain imbalance and disease severity in individuals who have inherited two copies of β-thalassaemia alleles, while the increased output of α-globin through coinheritance of extra α-globin genes in β-thalassaemia heterozygotes increases chain imbalance, converting a typically asymptomatic state to that of thalassaemia intermedia. The outcome depends on the number of α-globin genes inherited as one or two copies of triplicated (/αααα) or quadruplicated (/αααα) α-globin complexes, and the type of β-thalassaemia mutation (β^0 or β^+). A rarer mechanism of inheriting extra α-globin genes involves segmental duplication of the whole α-globin gene cluster.

Tertiary modifiers are those that affect the complications of disease; the severity of bone disease, iron loading and jaundice may be affected by polymorphisms of genes involved in the metabolic pathways concerned with these complications. Bone mass, like HbF, is a quantitative trait under strong genetic control involving multiple QTLs, those implicated including oestrogen receptor gene, vitamin D receptor (*VDR*) gene, collagen type α1 genes and transforming growth factor β1 (*TGFB1*) gene. Studies have shown that the levels of bilirubin and incidence of gallstones are related to a polymorphic variant (seven TA repeats) in the promoter of the uridine diphosphate glucuronosyltransferase 1A (*UGT1A1*) gene, also referred to as Gilbert syndrome. Iron loading in β-thalassaemia results not just from blood transfusion, but also from increased iron absorption. Variants in the *HFE* gene have a modulating effect on iron absorption, and as other genes in iron homeostasis become uncovered, it is likely that there will be genetic variants in these loci that influence the different degrees of iron loading in β-thalassaemia. Similarly, it seems very likely that the propensity to infection is modified by polymorphisms involving the immune system and its regulation. Finally, it should be remembered that environmental factors, long neglected, may also play an important role in modifying the β-thalassaemic phenotype.

Clinical findings in severe β-thalassaemia

In many high-income countries, neonatal screening programmes will first identify infants with more severe forms of β-thalassaemia, before the development of any symptoms; in some cases, antenatal screening of the parents and possibly prenatal diagnosis will have identified the fetus to be at high risk of β-thalassaemia before birth. Mutations in the β-globin gene almost never cause clinical symptoms *in utero* or neonatally due to the predominance of γ-globin at this stage. However, in many countries, neonatal screening programmes do not exist and diagnosis in the child will depend on their symptomatic presentation. Severe β-thalassaemia usually presents in the first year of life. Typically there is failure to thrive, with poor weight gain and growth with developmental delay. The parents may have noticed that the infant is pale and jaundiced, with a protruding abdomen. There may be a family history of severe anaemia, and typically the family will not be of northern European origin. Examination confirms the pallor and jaundice, with palpable hepatosplenomegaly. There may be evidence of marked erythroid hyperplasia, with typical 'thalassaemic facies' including expansion of the skull vault and maxillary bones. The symptoms and signs are not specific and differential diagnoses include gastrointestinal or hepatic disease, and malignancy.

Laboratory diagnosis of severe β-thalassaemia

In untransfused patients with severe β-thalassaemia, the full blood count shows severe anaemia with the haemoglobin usually less than 50 g/L. Mean corpuscular haemoglobin (MCH) and mean corpuscular volume (MCV) are low, with a very wide red cell distribution width. The nucleated cell count may be very high due to the presence of large numbers of nucleated red cells. A blood film shows marked anisopoikilocytosis, with basophilic stippling and small red cell fragments (Figure 6.6). The reticulocyte count is elevated, but less than expected for the degree of anaemia, due to the ineffective erythropoiesis. Renal function is normal, but liver function tests show elevation of bilirubin, aspartate aminotransferase and lactate dehydrogenase, with a normal alanine aminotransferase. Erythropoietin levels will be high, with soluble transferrin receptor levels up to 30 times greater than normal. White cell and platelet counts should be normal unless there is hypersplenism. A bone marrow aspirate is not essential to make the diagnosis, but if performed shows very marked erythroid hyperplasia, with dyserythropoiesis. Many of the erythroid precursors show inclusions after incubation with methyl violet; similar inclusions are found in the peripheral red cells after splenectomy. Immunoelectron microscopy confirms that the inclusions in β-thalassaemia consist of precipitated α-globin chains.

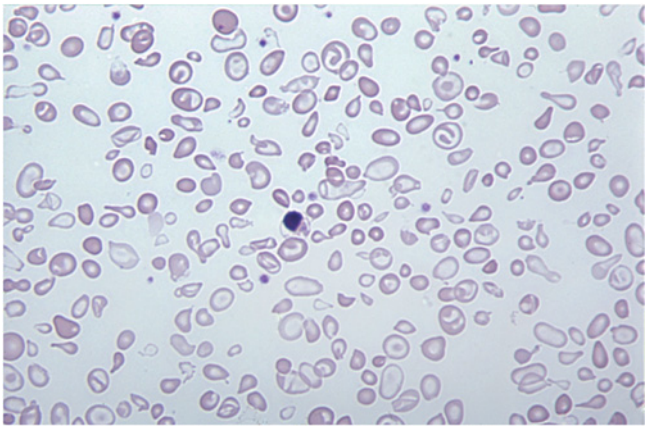


Figure 6.6 The peripheral blood appearances in β-thalassaemia.

Haemoglobin analysis is needed to confirm the diagnosis, typically using either electrophoretic or chromatographic techniques. This will usually show an increased amount of HbA_2, with the vast majority of the remainder consisting of HbF; small amounts of HbA may be present depending on the β-globin mutation, the age of the child and whether the child has been transfused. Absence of HbA confirms a diagnosis of β^0-thalassaemia, while presence of HbA (pretransfusion sample) confirms β^+-thalassaemia. Testing of the parents should confirm the diagnosis, both typically being carriers of β-thalassaemia, with HbA_2 levels greater than 3.5% and MCH below 27 pg. These findings are sufficient to make a diagnosis of severe β-thalassaemia, although where DNA analysis is available it is often used to identify the β-globin mutations and confirm the diagnosis.

Management of severe β-thalassaemia

Any child presenting in the first year of life with the features described above is likely to require regular red cell transfusions to grow and develop normally; this is referred to as thalassaemia major. If such children are not transfused regularly, as happened historically, and as happens currently in many low-income countries, progressive deterioration occurs. Growth and development continue to be severely impaired. The child often has muscle wasting due to increased metabolic demands, and in particular may become folate deficient. The spleen and liver become progressively enlarged; the spleen can become massive with hypersplenism and resulting cytopenias. There is marked erythroid hyperplasia with bony distortion and extramedullary haemopoiesis. Bones become enlarged, and this is most apparent in the face, with maxillary hyperplasia, dental malocclusion and development of a 'tower' skull. Extramedullary haemopoiesis can occur anywhere but is typically paraspinal and may cause compression of spinal nerves with resulting pain and weakness; intracranial haemopoiesis may cause cranial neuropathies and symptoms of raised intracranial pressure. There is an increased tendency to infection, and without transfusion the child typically dies from either infection or high-output cardiac failure. Radiography may show lacy trabecular patterns in long bones and a 'hair-on-end' appearance of the skull (Figure 6.7).

This picture is completely transformed by an appropriate transfusion regimen. The decision to start regular blood transfusions is based on clinical factors, and not dependent on a particular haemoglobin level or molecular diagnosis. Occasional blood transfusions may be necessary because of an acute exacerbation of anaemia, often related to infection, and do not necessarily mean that the child is transfusion dependent. It is a bigger decision to institute a regular transfusion regimen, in that typically transfusions are then continued lifelong. Stopping transfusions in a child or adult who is adapted to a high haemoglobin level inevitably results in a prolonged period of symptomatic anaemia and ill health, which may not be tolerated. Regular transfusions should be started if the child is failing to thrive, or if erythroid expansion is causing bony distortion or hypersplenism; abnormal facial appearances, if allowed to progress, may be irreversible without maxillofacial surgery.

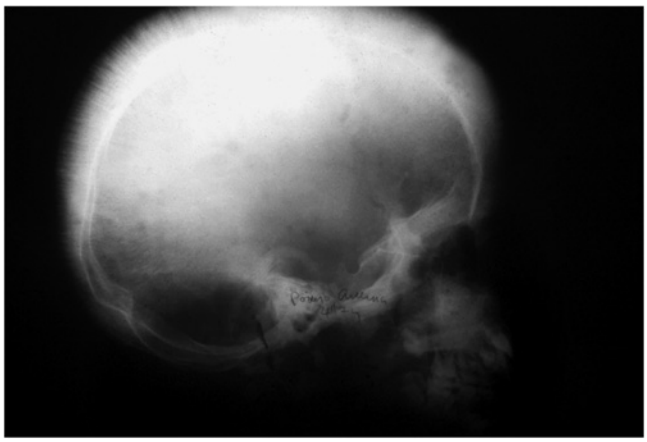


Figure 6.7 Radiograph of a skull of a thalassaemia major patient showing 'hair-on-end' appearance.

Blood transfusions

The aim of regular transfusions is to correct anaemia and suppress the abnormal erythroid hyperplasia. Correcting anaemia improves oxygen delivery to the tissues and facilitates normal growth and development. Suppression of erythropoiesis limits damage to bones, reduces excessive iron absorption and reduces extramedullary haemopoiesis. Measurement of soluble transferrin receptor levels, which are proportionate to the size of the erythron, suggest that suppression of erythropoiesis requires the haemoglobin to be kept above 95 g/L; in practice this means aiming for a pretransfusion haemoglobin of 90–100 g/L. This can usually be achieved by regular red cell transfusions every 2–4 weeks, with a post-transfusion haemoglobin target of 130–150 g/L. If venous access is difficult, as is often the case in young children, it is sometimes beneficial to insert a semi-permanent central venous access device, such as a Portacath or Hickman line. The majority of thalassaemia patients are treated with simple top-up transfusions. An alternative involves regular exchange transfusion in which blood is both removed by venesection and transfused; this can be performed either manually or automatically using an apheresis machine. The main advantages are that transfusion can be less frequent, at intervals of up to 6 weeks, and iron loading is significantly less, as blood is also removed. Disadvantages include increased expense and time, increased donor exposure with risk of infection and alloimmunization, and difficulties with venous access.

Ideally transfusions are with packed red blood cells. Leucodepleted blood reduces the risk of transfusion reactions and cytomegalovirus infection, and should be used where available. When the ethnicity of the blood donor population differs from that of the recipient thalassaemia population, as occurs in northern Europe, the USA and Australia, the transfused red cells should be matched for an extended range of blood groups to minimize the risk of alloimmunization; typically, full matching for all Rh and Kell groups, in addition to ABO, is of benefit. Blood grouping based on DNA analysis offers potential benefits, by identifying variant blood groups that may not be uncovered using serological testing. This is particularly relevant for Rh variants, which are responsible for an increasing proportion of alloantibodies as serological grouping becomes more widespread. Before starting transfusions the patient should be vaccinated against hepatitis B, and ideally also against hepatitis A.

Iron overload

Iron overload inevitably complicates regular blood transfusions and is the source of many serious complications. Each unit of transfused blood contains about 200–250 mg iron, compared with the 1 mg iron normally absorbed each day. Hepcidin is an important negative regulator of iron absorption, but in thalassaemia major and intermedia, hepcidin levels remain inappropriately low, despite increased iron stores. Recently, a new hormone, erythroferrone (ERFE) that inhibits hepcidin synthesis during stress erythropoiesis, has been identified. ERFE is an erythroid regulator secreted by differentiating erythroblasts; greatly increased levels of ERFE were found in mice with thalassaemia intermedia, providing an explanation for the inappropriately low hepcidin and increased iron stores. ERFE is thus a strong candidate for the low hepcidin levels encountered in β-thalassaemia where there is a large component of ineffective erythropoiesis. Growth differentiating factor-15 (GDF-15) which is produced by erythroid precursors and reaches very high levels in thalassaemia major due to erythroid hyperplasia, has previously been suggested as a pathological suppressor of hepcidin in thalassaemia, but, to date, there is no definitive evidence for its role in suppression of hepcidin (Chapters 3 and 4). Iron is initially stored in macrophages within the liver, and is chaperoned around the body bound to transferrin. As transferrin becomes saturated, labile, more toxic forms of iron appear in cells and plasma, referred to as non-transferrin-bound iron (NTBI). It is thought that NTBI is responsible for most of the iron toxicity, including iron loading into cardiac and endocrine tissues. Once in cells, the iron causes oxidative tissue damage, mostly through the generation of free radicals.

Iron overload and chelation is discussed in more detail in Chapter 4. Before iron chelation was available, most regularly transfused patients with thalassaemia died in their late teens, mainly from cardiac iron deposition causing heart failure. Endocrine failure was also inevitable, with diabetes mellitus, hypothyroidism, hypoparathyroidism, hypogonadism and hypopituitarism. With good iron chelation life expectancy is open-ended, and there is an established link between the efficacy of iron chelation and life expectancy.

The body has no mechanism for excreting iron and iron-chelating drugs are necessary to avoid toxic iron accumulation. Iron chelation is usually started after about 1 year of monthly blood transfusions. Ideally, children delay starting chelation until they are 3 years old, as drug toxicity is thought to be highest in the young; it is usually necessary to start chelation earlier in children who start transfusions in the first year of life, although in general low doses are used.

Currently, three drugs are used for iron chelation: desferrioxamine, deferiprone and deferasirox. Desferrioxamine has been in clinical use since the 1970s and is known to be safe and effective; side-effects seem only to occur when the drug is used in high doses or iron stores are low. Important side-effects include ocular and retinal toxicity, growth impairment and cartilaginous dysplasia. *Yersinia enterocolitica* infection is increased in iron overload, particularly if desferrioxamine is also in use. Regular use of desferrioxamine has been found to prolong survival in several observational studies. The main problem is that it has to be given by injection; the most commonly used regimen involves overnight subcutaneous infusions given over 8 hours using a syringe driver or balloon-pump. To achieve negative iron balance in a regularly transfused person requires a dose of 40 mg/kg five times a week. The subcutaneous route is inevitably a cause of poor adherence to treatment regimens because of the pain and inconvenience. Historically, therefore, many patients have developed life-threatening iron overload despite the availability of desferrioxamine, with median age of death being 30–40 years. Deferiprone was developed in the UK about 30 years ago and was licensed in Europe in 1999, and the USA in 2011. Initial trials demonstrated its efficacy, although subsequent studies suggested that there may be individual variation in response requiring higher dosage up to 100 μg/kg/day instead of 75 μg/kg/day. Arthropathy and agranulocytosis are potentially serious side-effects and it is currently recommended that people taking the drug have full blood counts every week. Despite concerns about its safety and efficacy, deferiprone seems to be particularly effective at removing cardiac iron, which is possibly linked to improved cardiac function and reduced cardiac mortality. It is increasingly used in combination with desferrioxamine. Deferasirox is a rationally designed oral iron chelator that has been approved in the USA and Europe since the mid-2000s. Large clinical trials have demonstrated its efficacy in thalassaemia. It effectively removes hepatic iron, with increasing evidence that it also removes cardiac iron. Clinical trials are beginning to address its use in combination with other chelators.

It is important to monitor iron stores regularly in transfusion-dependent patients who are receiving iron chelation. Monitoring the volume of transfused blood allows the iron input to be calculated and can be valuable when choosing the dose of iron chelator. Serum ferritin is proportional to the amount of stored iron in the liver and can be used to effectively monitor iron overload, particularly if serial measurements are used. Ferritin is artificially elevated by any inflammatory process, which can cause misleading results, particularly if there is coincidental hepatitis. Liver iron can be accurately assessed using liver biopsy, but this is invasive with a risk of complications. Increasingly, magnetic resonance imaging (MRI) is used to quantitate liver iron, with the R2 method being approved by both European and North American authorities. Cardiac MRI is an important way of identifying those patients with significant cardiac iron, which allows chelation to be targeted to the heart, using either continuous intravenous desferrioxamine or a combination of desferrioxamine and one of the oral iron chelators. Iron monitoring and chelation are very expensive and not available to most patients in the world with thalassaemia major.

Monitoring and annual review of patients with thalassaemia major

The aim of regular transfusions is to allow the child to grow and develop normally, and for the quality and quantity of life to be as close to normal as possible. This requires the child to be closely monitored, and this is ideally done in a centre with expertise in the condition, often in the form of an annual review. Each year the volume of transfused blood (expressed as mL/kg) should be recorded; if the transfusion requirement is high, exceeding 200–250 mL/kg, this suggests the possibility of hypersplenism, and if the spleen is enlarged splenectomy may significantly reduce the rate of transfusion and iron loading. Splenectomy would not normally be considered before the age of 6 years, and there is emerging concern that it may increase the risk of pulmonary hypertension and other vascular complications in later life.

Growth should be carefully monitored in children, including annual measurement of sitting height to assess spinal growth; desferrioxamine toxicity has been specifically linked to impaired spinal growth. Blood tests should be performed each year to look for endocrinopathy, including fasting glucose, thyroid-stimulating hormone, parathyroid hormone, insulin-like growth factor 1 and sex hormone levels. Hepatic and renal function should be monitored regularly, together with hepatitis serology. Cardiac assessment should include an ECG, echocardiogram and increasingly MRI, which can assess both ejection fraction and iron loading; MRI is not tolerated by children under the age of 7 years without general anaesthesia or sedation, and is not usually justified at this age unless there is strong suspicion of cardiac or other problems. Osteopenia is more common in children and adults with thalassaemia major, and bone densitometry should be monitored regularly from the age of about 10 years. Vitamin D levels should be measured regularly and optimized, with oral supplements of cholecalciferol or ergocalciferol being necessary in most countries. There is evidence that bisphosphonates may be effective in preventing and treating osteopenia in adults with thalassaemia, although their role in children is less established, with concern about their effects on growth. Dietary and psychological support are also often beneficial. In general, the organization and provision of all this care requires a multidisciplinary team, including paediatricians, haematologists, cardiologists, endocrinologists, nurse specialists and psychologists.

Stem cell transplantation in severe thalassaemia

Stem cell transplantation is generally seen as the treatment of choice if there is an HLA-identical sibling and it is clear that the child is transfusion dependent. The success of stem cell transplantation is generally reduced as children get older, iron overload increases and iron-related organ damage increases. Under optimal circumstances the transplant is successful in more than 90% of cases. The main complications are severe infection during the period of transplantation and either acute or chronic graft-versus-host disease. Significantly reduced fertility is also almost inevitable. Recent improvements in the outcome of medical treatment, related to oral iron chelation and better cardiac monitoring, have possibly shifted the balance away from transplantation. One limitation of transplantation is the availability of suitable donors; cord blood transplants are potentially important in this respect. In a small number of cases preimplantation genetic diagnosis has allowed the selection of HLA-matched embryos to produce siblings who can act as stem cell donors. Efforts are also being made to develop mini- and microtransplants, which may be less toxic, preserve fertility and potentially allow the use of alternative donors, such as haploidentical or matched unrelated donors. These regimes involve prolonged use of immunosuppression and are still essentially experimental, although they may lead to much greater availability of transplantation in the relatively near future.

Prognosis in severe thalassaemia

In theory, with adequate transfusion and effective chelation, survival in thalassaemia major should approach that of the normal population, although historically this has never been the case because of the difficulties patients encounter in adhering to subcutaneous iron chelation regimens. In Europe and North America, median survival has improved with each decade since the introduction of desferrioxamine in the 1970s. Most deaths are related to cardiac iron overload and improved survival is primarily related to earlier and more rational use of chelation. Recent studies suggest that death of cardiac origin has become much less common in the last 5–10 years, probably related to improved awareness and treatment of iron-related cardiomyopathy with intravenous desferrioxamine and combination therapy with both desferrioxamine and deferiprone. Earlier detection, using cardiac T2* MRI, and prevention of cardiac iron loading, with oral chelation, seem likely to improve prognosis further.

Heterozygous β-thalassaemia

Carriers for β-thalassaemia are usually symptom-free except in periods of stress such as pregnancy, when they may develop significant anaemia. Transfusion is sometimes necessary during pregnancy, but no other treatment or follow-up is necessary. Palpable splenomegaly is rare. Carriers may have mild anaemia; haemoglobin values are in the range 90–120 g/L. The red cells show hypochromia and microcytosis proportionately greater than the degree of hypochromia seen in iron deficiency. The reticulocyte count is normal. The bone marrow shows moderate erythroid hyperplasia with corresponding increase in soluble transferrin receptor levels. The characteristic finding is an elevated HbA_2 level, usually greater than 3.5%. There is a slight elevation of HbF in the 1–3% range in about 50% of cases. Silent or near-silent carriers of β-thalassaemia occur fairly rarely, in which the HbA_2 level is normal with only slight hypochromia; this is typically associated with mild thalassaemic

mutations in the promoter region of the gene. Failure to identify this state antenatally can result in significant and unexpected thalassaemia in the offspring, if the partner also carries β-thalassaemia. Borderline increases of HbA_2 with variable increases of HbF, but normal red cell indices have also been associated with variants in the *KLF1* gene, but the clinical significance of such variants when coinherited with β-thalassaemia is not clear.

β-Thalassaemia in association with haemoglobin variants

In many populations, because there is a high incidence of both β-thalassaemia and other globin gene mutations, it is quite common for an individual to inherit a β-thalassaemia allele from one parent and a gene for a structural haemoglobin variant from the other. Although numerous interactions of this type have been described, only three are common and significant: HbS/β-thalassaemia, HbC/β-thalassaemia and HbE/β-thalassaemia.

HbS/β-thalassaemia (see also Chapter 7)

HbS/β-thalassaemia causes sickle cell disease, with the severity varying from very mild to severe. This interaction does not result in clinically significant thalassaemia. The principal determinant of the severity of sickle cell disease is the nature of the β-thalassaemia mutation, and to what extent β^A is reduced compared with normal. HbS/β^0-thalassaemia with no HbA present and an average haemoglobin of 70–80 g/L, is broadly similar in severity to HbSS. It differs in that red cells are markedly hypochromic and contain excess α-globin chains. Clinically significant splenomegaly is more common than in HbSS, affecting up to 30% of adults. The severity of HbS/β^+-thalassaemia varies from severe to very mild, depending on how much HbA is produced. Some β^+-thalassaemia alleles produce less than 10% of the normal HbA output and result in a clinical picture identical to HbS/β^0-thalassaemia. At the other end of the spectrum, promoter mutations typically reduce HbA output by only 10% and result in a condition that is only marginally more symptomatic than sickle cell carriers. These mild forms of HbS/β^+-thalassaemia occur most often in populations of African origin, whereas the more severe forms are most often found in Mediterranean countries and India.

HbC/β-thalassaemia

HbC is common in parts of West Africa, and occurs to a minor extent in populations of North African and southern Mediterranean origins. It is largely asymptomatic and characterized by a mild haemolytic anaemia associated with splenomegaly. The peripheral blood film shows numerous target cells and thalassaemic red cell changes with a moderately elevated reticulocyte count. Haemoglobin electrophoresis shows a preponderance of HbC. The diagnosis is confirmed by finding the HbC trait in one parent and the β-thalassaemia trait in the other, and by DNA analysis where appropriate.

HbE/β-thalassaemia

This the commonest severe form of thalassaemia in Southeast Asia and parts of the Indian subcontinent. The β^E allele is mildly thalassaemic due to the activation of a cryptic splice site, and when it is inherited together with β^0-thalassaemia, there is a marked deficiency of β-chain production. HbE is slightly unstable *in vitro*, although it is not clear how clinically significant this is. The clinical and haematological features are very variable. There are thalassaemic red cell changes, and the bone marrow shows marked erythroid hyperplasia. There is nearly always anaemia and splenomegaly, with typical thalassaemic bone changes. Haemoglobin values are in the range 40–100 g/L, with an average of 60–70 g/L; in well-resourced countries about half of the patients with this condition are regularly transfused. There are thalassaemic red cell changes and the bone marrow shows marked erythroid hyperplasia.

Although relatively little is known about the natural history of this disorder, it is clear that in many parts of Southeast Asia and India it causes a very high mortality in early life. Clinically it is indistinguishable from other forms of β-thalassaemia, although the link between genotype and phenotype is less predictable, with some patients having thalassaemia major and others, being non-transfusion dependent, growing and developing with few complications.

The diagnosis is confirmed by finding only HbE and HbF on haemoglobin electrophoresis and by demonstrating HbE trait in one parent and β-thalassaemia trait in the other. In cases of HbE/β^+-thalassaemia, variable quantities of HbA are present and the condition is usually milder. Neonatal diagnosis can be difficult, particularly distinguishing between homozygotes for HbE and HbE/β^0-thalassaemia; family studies and DNA analysis are useful in these circumstances.

Variant forms of β-thalassaemia

Despite the vast heterogeneity of mutations, the increased levels of HbA_2 in β-thalassaemia heterozygotes is remarkably uniform, in the range 3.5–5.5%, and rarely exceeds 6%. Unusually high HbA_2 levels in excess of 6.5% seem to characterize the subgroup caused by lesions (point mutations or small deletions) that affect the regulatory elements in the promoter region of the β-globin gene (see Figure 6.4). The unusually high HbA_2 levels are usually accompanied by higher than usual increases in HbF, resulting in a milder thalassaemia phenotype, despite the absence of HbA_2 in some cases. Otherwise, the haematological picture is identical to the common forms of β-thalassaemia. Some β-thalassaemia heterozygotes have normal HbA_2 levels, despite the typical hypochromic microcytosis. In most cases, this is due to the coinheritance of δ-thalassaemia. Other cases of normal HbA_2 β-thalassaemia are extremely mild forms of β-thalassaemia that is completely silent in heterozygotes and is only identified when it is coinherited with a common form of β-thalassaemia. Heterozygotes do not have any evident

haematological phenotype, the only abnormality being a mild imbalance of globin chain synthesis.

β-Thalassaemia occasionally follows a dominant pattern of inheritance, producing symptoms in heterozygotes (see Figure 6.3). The clinical picture is characterized by a moderate degree of anaemia and splenomegaly with marked thalassaemic changes of the red cells, and ineffective erythropoiesis with intracellular inclusion bodies. Although not usually transfusion dependent, such individuals develop iron overload from hyperabsorption due to ineffective erythropoiesis and may develop liver or endocrine damage. The common denominator of these dominantly inherited β-thalassaemias is the synthesis of hyperunstable β-chain variants, caused by a spectrum of mutations; those resulting from PTCs are typically found in exon 3 of the β-globin gene. Unlike the recessive forms that are prevalent in malarious regions, dominantly inherited β-thalassaemias are rare, occurring in dispersed geographical regions. Most of the dominant β-thalassaemia alleles have been described in single families, many as *de novo* mutations. Dominantly inherited β-thalassaemia should be suspected in any patient with a thalassaemia intermedia phenotype, even if both parents are haematologically normal, and the patient is from an ethnic background where β-thalassaemia is rare. The diagnosis is confirmed by DNA sequence analysis of the β-globin genes.

δβ-Thalassaemia and hereditary persistence of fetal haemoglobin

HPFH and δβ-thalassaemia are much less common than β-thalassaemia and manifest as a range of disorders characterized by decreased or absent HbA production and a variable compensatory increase in HbF synthesis. The distinction between them is subtle and originally made on what appeared to be clear-cut clinical and haematological grounds. However, as more cases became recognized and their underlying mutations delineated, it became evident that there is considerable overlap between the two groups of disorders.

The level of compensatory increase is higher in HPFH compared with δβ-thalassaemia. HPFH heterozygotes have essentially normal red cell indices, normal HbA_2 levels and HbF levels of 10–35%, whereas heterozygotes for δβ-thalassaemia have hypochromic microcytic erythrocytes and normal HbA_2 levels and the HbF increases are lower (5–15%). A distinguishing feature is the heterocellular distribution of HbF in δβ-thalassaemia compared with a pancellular (or homogeneous) distribution in HPFH, although the intercellular distribution of HbF may be a reflection of the magnitude of increase and the sensitivity of the technique used to stain F cells. Clinically, HPFH homozygotes are asymptomatic with slightly reduced MCV and MCH, and compound heterozygotes with β-thalassaemia have very mild disease. Compound heterozygotes of δβ-thalassaemia with β-thalassaemia, and δβ-thalassaemia homozygotes have disease severity that ranges from mild anaemia to transfusion dependence.

Two types of mutations underlie this group of disorders: deletions that remove substantial regions of the β-globin cluster, including the β-globin gene (Figure 6.8), and point mutations in the promoters of either of the γ-globin genes ($^{G}\gamma$ or $^{A}\gamma$). Six deletion forms of HPFH have been described in Africans, Mediterraneans, Indians and Southeast Asians, with deletions ranging from 13 to 86 kb in size. The commonest forms are Black HPFH-1 and HPFH-2, the latter also referred to as Ghanaian HPFH. Included within the HPFH conditions is Hb Kenya, a rare condition found largely in East Africa and characterized by production of a hybrid $^{A}\gamma\beta$-globin chain and increased $^{G}\gamma$-chain (Figures 6.8 and 6.10). The first type of δβ-thalassaemia to be described was Hb Lepore, which consists of hybrid δβ-globin chains produced by misaligned crossing-over between the δ- and β-globin genes (Figures 6.8 and 6.9). Heterozygotes for Hb Lepore have hypochromic microcytic red cells, normal HbA_2 and variably increased HbF. The more common forms of δβ-thalassaemia result from deletions of different lengths in the β-globin gene cluster (Figure 6.8). Deletions that leave both the γ-globin genes intact are called $^{G}\gamma^{A}\gamma$ $(\delta\beta)^0$-thalassaemias, and they all include parts or all of the δ- and β-globin genes. $^{G}\gamma(^{A}\gamma\delta\beta)^0$-thalassaemia deletions include part or all of the $^{A}\gamma$-globin gene and hence produce HbF containing only $^{G}\gamma$-chains. A couple of the $(\delta\beta)^0$-thalassaemias include two deletions separated by an inverted region.

Non-deletion HPFH is caused by mutations within the promoters of either of the γ-globin genes, and lead to variably increased HbF with a preponderance of only one of the γ-chains. The mutations (single-base substitutions or minor deletions) are clustered in three regions of the promoters, around positions −114, −175 and −200. These regions contain binding sites for ubiquitous and erythroid-specific factors. Altered binding patterns of the transcription factors due to the point mutations are thought to be the cause of the elevated HbF levels, which vary from 5 to 35% in heterozygotes. Deletion and non-deletion HPFH and δβ-thalassaemias are clearly inherited in a Mendelian fashion as alleles of the β-globin gene complex. Heterozygotes have significant elevations of HbF ranging from 10 to 40%.

Inherited increases in HbF have also been reported in families with β-thalassaemia and sickle cell disease; these increases had a modulating effect on the severity of disease and the inheritance appeared to segregate independently of the β-globin gene cluster. Healthy members of the families may also have slight increases in HbF, and historically these individuals were said to have coinherited heterocellular HPFH. Heterocellular HPFH was originally recognized in a group of Swiss army recruits; the HbF elevation was modest and unevenly distributed among the erythrocytes, the distinct HbF-carrying red blood cells being termed F cells. Hence Swiss HPFH was also referred to descriptively as heterocellular HPFH. It is now clear that HbF is a highly variable quantitative trait, and that heterocellular HPFH

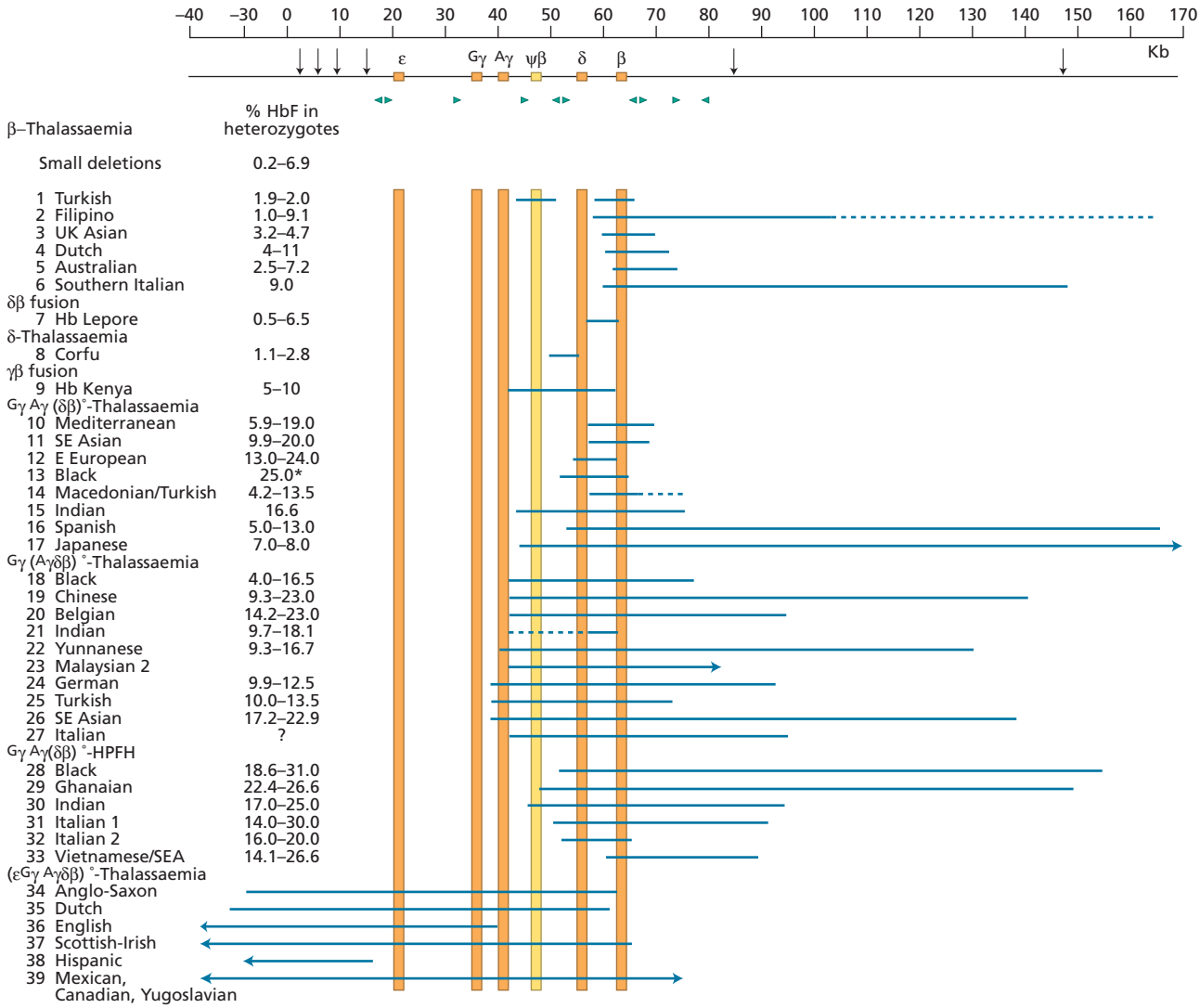


Figure 6.8 The deletions that underlie δβ-thalassaemia and hereditary persistence of fetal haemoglobin. The upper arrows represent DNase I-hypersensitive sites. (Source: Weatherall and Clegg 2001. Reproduced with permission of WHO.)

represents the upper tail of the natural continuous distribution that includes approximately 10% of the population with HbF levels between 0.8 and 5%. Unlike the Mendelian forms caused by major deletions or point mutations in the γ-globin promoters, the inheritance of heterocellular HPFH is complex, with an overwhelming genetic contribution. The sequence variant (C→T) at position −158 of the ^{G}γ-globin gene, also referred to as the *Xmn*1-^{G}γ site (or rs74821440), was the first QTL to be implicated through family studies. Subsequent genetic association studies have confirmed *Xmn*1-^{G}γ as one of the three major QTLs modulating HbF production in adults, the other two being the *HBS1L-MYB* intergenic region on chromosome 6q and *BCL11A* on chromosome 2p. These three QTLs account for a relatively large proportion (20–50%) of the common variation in HbF levels, not only in healthy adults, but also in patients from diverse ethnic groups – Brazilians, African-Americans, African-British, Tanzanians with sickle cell disease. In African-American patients with sickle cell disease, the three loci contribute more than 20% to the HbF variation with a corresponding reduction in frequency of acute pain. Coinheritance of the *Xmn*1-^{G}γ site delays transfusion requirements in β-thalassaemia. A combination of the three HbF QTLs and α-globin genotype have also been shown to be predictors of disease severity and transfusion requirements in patients with β-thalassaemia and HbE/β-thalassaemia. HbF increases secondary to *KLF1* variants, either as a primary phenotype or in association with other red cell disorders, such as congenital dyserythropoietic anaemia (CDA type IV), are increasingly noted.

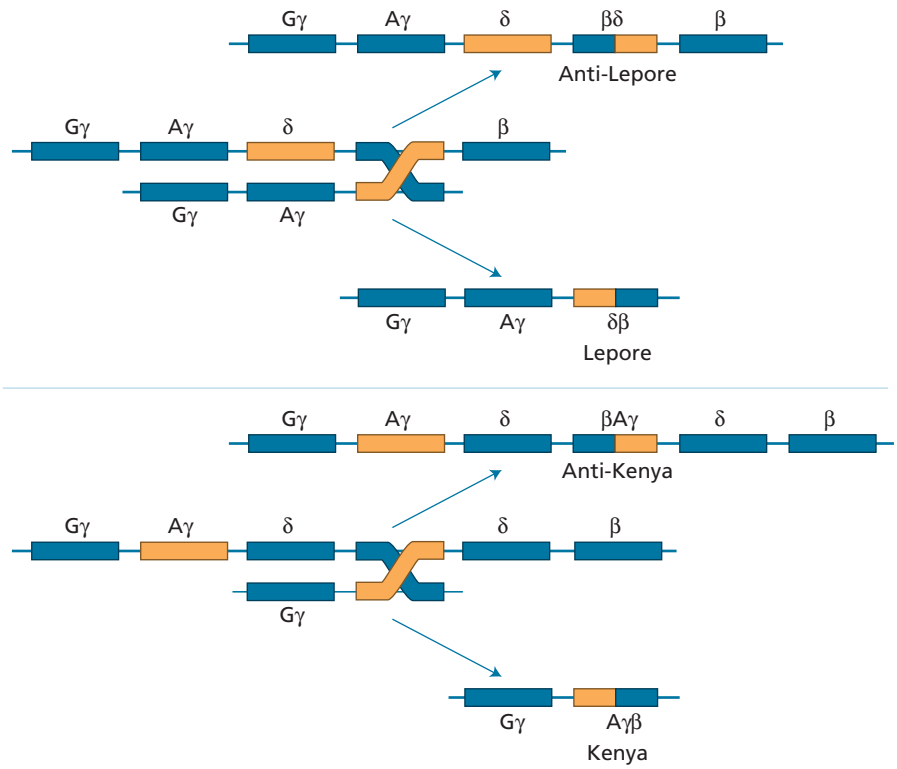


Figure 6.9 The mechanisms for the production of haemoglobin Lepore and related variants.

εγδβ-Thalassaemia

These are rare conditions and result from large deletions of the β-globin gene cluster that involve the β-LCR. The deletions fall into two main categories: group I removes all, or a greater part of, the complex including the β-globin gene and the β-LCR, and group II removes extensive upstream regions including the β-LCR but leaving the β-globin gene itself intact. There is no output from the globin genes of the affected cluster. Clearly, the homozygous state would not be compatible with survival. Heterozygotes have severe haemolytic disease of the newborn, with anaemia and hyperbilirubinaemia. The severity of anaemia and haemolysis is variable, even within a family, and in some cases blood transfusions are necessary during the neonatal period. If they survive the neonatal period, the infants grow and develop normally; in adult life they have the haematological picture of heterozygous β-thalassaemia, with mild anaemia, hypochromic microcytic red cells and a haemoglobin pattern of normal HbA_2 β-thalassaemia. Using next-generation sequencing (NGS) technology, a novel form of εγδβ-thalassaemia due to an inversion-deletion has recently been characterized in three unrelated English families.

The α-Thalassaemias

Distribution

The α-thalassaemias follow a similar distribution to the β-thalassaemias, extending throughout sub-Saharan Africa, the Mediterranean region, the Middle East, the Indian subcontinent and Southeast Asia, in a line stretching from southern China through Thailand, the Malay Peninsula and Indonesia, to the Pacific Island populations.

In some prevalent areas, carrier frequency for the mild form (α^+) reaches 80% or more. The more severe forms (α^0-thalassaemias) reach their highest frequency in Southeast Asia, where carrier frequency can reach 10%.

Genetic basis of disease: molecular pathology

Normal individuals have four α-globin genes arranged as linked pairs, α2 and α1, at the tip of each chromosome 16, the normal α-genotype being represented as αα/αα- (Figure 6.10). The α-thalassaemias can be classified as α^0-thalassaemia, in which no α-chains are produced from the linked pair, and α^+-thalassaemia, in which production of α-chain from the affected chromosome is reduced.

The α^0-thalassaemias are caused by deletion of both α-globin genes (Figure 6.11). The deletions vary in size and tend to be geographically isolated, with two particularly common ones in Southeast Asia ($/- -^{SEA}$) and the Mediterranean region ($/- -^{MED}$). Rarely, α^0-thalassaemia can also arise from deletions of the upstream α-globin regulatory elements (MCS) in which the α-globin genes remain intact, but completely inactivated. In other cases, more extensive deletions are associated with monosomy for a segment of the tip of chromosome 16p that, when greater than about 1 megabase (Mb), result in the syndrome of

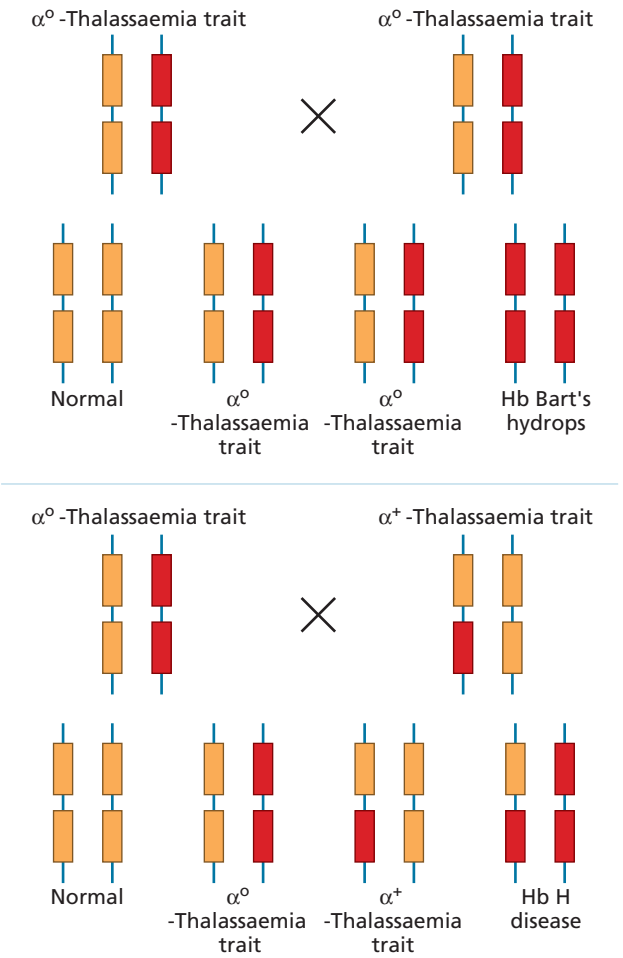


Figure 6.10 The genetics of α-thalassaemia.

α-thalassaemia and mental retardation (ATR-16). The molecular basis of the α^+-thalassaemias is more complicated; the commonest forms result from deletions that remove one of the linked pairs of α-globin genes, leaving the other intact (−α/αα). The linked α-globin genes are embedded in two highly homologous duplicated units of 4 kb within which are three homologous subsegments designated X, Y and Z (Figure 6.12), Misalignment and recombination between the Z segments, which are 3.7 kb apart, produces chromosomes with one α-globin gene ($/-\alpha^{3.7}$ or *rightward* deletion) and others with three α-globin genes ($/\alpha\alpha\alpha^{anti\text{-}3.7}$). Similarly, crossover between the X boxes, which are 4.2 kb apart, produces the *leftward* deletion $/-\alpha^{4.2}$ and the $/\alpha\alpha\alpha^{anti\text{-}4.2}$ allele.

Less commonly, both the α-globin genes are intact and α^+-thalassaemia results from point mutations that partially or completely inactivate one of them ($\alpha^T\alpha/\alpha\alpha$). In contrast to the β-thalassaemias, non-deletional mutations are much less common causes of α-thalassaemia but, like those that cause β-thalassaemia, the non-deletional α-thalassaemia variants act at different stages of gene regulation and expression. The majority of the non-deletional α-thalassaemia (more than two-thirds) are found on the dominant α2 gene, less than one-third on the α1 gene and the others on an /−α-chromosome ($/-\alpha^T$). In general, the non-deletional α^+-thalassaemia variants ($/\alpha^T\alpha$) give rise to a more severe reduction in α-chain output than the single α-gene deletion (/−α) due to the lack of compensatory increase in α-output from the linked α1 gene, as observed in deleted cases. One particularly common form of non-deletional α-thalassaemia found in Southeast Asia is Hb Constant Spring ($/\alpha^{CS}\alpha$), which is due to a single-base substitution (TAA→CAA) in the α_2-globin termination codon. Instead of terminating at the stop codon, mRNA is read through the 3′-UTR until another in-phase termination codon is encountered 31 codons later. This results in an elongated α-globin chain of 172 residues, 31 amino acids from the natural arginine at codon 141. Of the six predicted α_2-chain termination variants, five have been described: Hb Constant Spring (α142 Gln), Hb Icaria (α142 Lys), Hb Koya Dora (α142 Ser), Hb Seal Rock (α142 Glu) and Hb Pakse (α142 Tyr). Hb Constant Spring is by far the most common of these variants, reaching frequencies of up to 4% in Thailand. Finally, non-deletional α-thalassaemia can also arise from single-base substitutions causing structural α-globin variants that are highly unstable, for example Hb Quong Sze α125 Leu→Pro ($/\alpha^{QS}\alpha$).

Unusual causes of α-thalassaemia

This includes a single case report of α^0-thalassaemia arising from a deletion involving the α1 gene that also inactivated the intact linked α2 gene. Subsequent studies showed that the deletion results in juxtaposition of a downstream gene (*LUC7L*) next to the structurally normal upstream α2 gene. Transcription of antisense mRNA from *LUC7L* led to silencing of the linked α2 gene. Another novel form of non-deletional α-thalassaemia results from a single nucleotide substitution in a non-genic region between the α-globin genes and their upstream regulatory elements. The single-base substitution leads to the creation of a new promoter-like element that interferes with normal activation of all the downstream α-like globin genes, resulting in α^0-thalassaemia.

Pathophysiology

The pathophysiology of α-thalassaemia is different to that of β-thalassaemia. A deficiency of α-chains leads to the production of excess γ- or β-chains, which form Hb Bart's (γ_4) and HbH (β_4), respectively. These soluble tetramers do not precipitate extensively in the bone marrow and hence erythropoiesis is more effective than in β-thalassaemia. However, HbH is unstable and precipitates in red cells as they age. The inclusion bodies cause red cell membrane damage and obstruction in the spleen leading to shortened red cell survival. Furthermore, both Hb Bart's and HbH have a very high oxygen affinity and their oxygen dissociation curves resemble myoglobin. Thus, the severe forms of

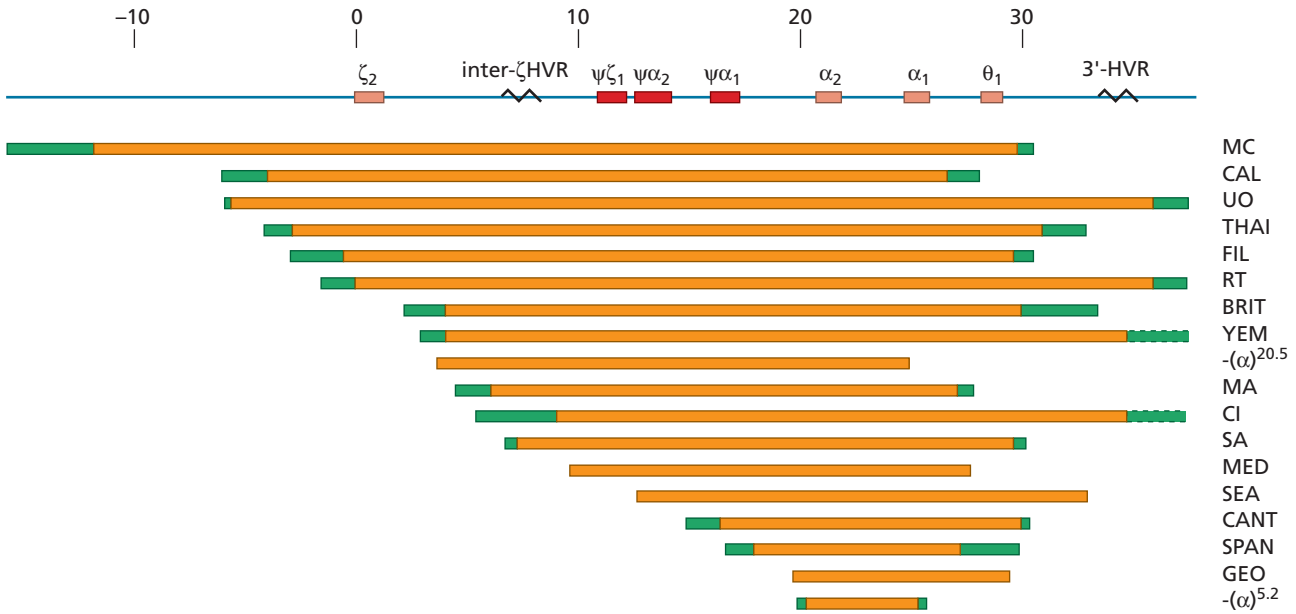


Figure 6.11 The α-globin gene cluster deletions that underlie α^0 thalassaemia. (Source: Weatherall and Clegg 2001. Reproduced with permission of WHO.)

α-thalassaemia are due to defective haemoglobin production, the synthesis of homotetramers that are physiologically useless and a haemolytic component.

Genotype–phenotype relationship

Loss of one functioning α-gene (αα/−α) is almost completely silent, with normal or only slightly hypochromic red cells. Loss of two α-genes (− −/αα-or −α/−α) produces a mild hypochromic microcytic anaemia, the α-thalassaemia trait. Homozygotes for α^0-thalassaemia (− −/− −) have a lethal condition with intrauterine haemolytic anaemia called Hb Bart's hydrops fetalis syndrome (see Figure 6.11). As Hb Bart's hydrops fetalis syndrome follows the homozygous inheritance of α^0-thalassaemia, this condition occurs only in populations in which α^0-thalassaemia is common, notably those of Southeast Asia and the Mediterranean islands. Deficiency of α-chains

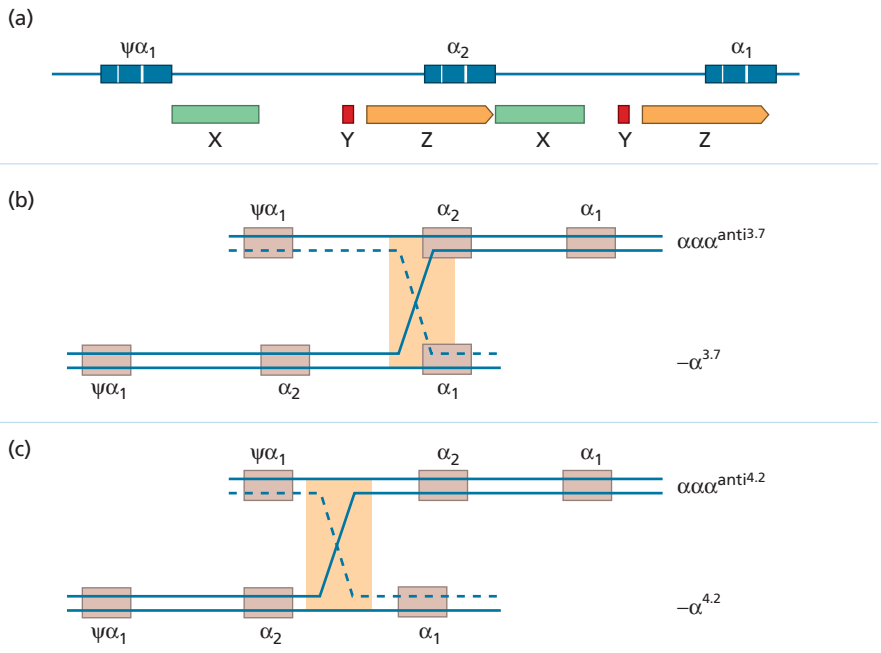


Figure 6.12 The molecular mechanisms that underlie the deletion forms of α-thalassaemia: (a) normal cluster showing X, Y and Z homology boxes; (b) 3.7-kb deletion; (c) 4.2-kb deletion. (Source: Weatherall and Clegg 2001. Reproduced with permission of WHO.)

gives rise to an excess of γ-chains (in fetal life) or β-chains (in adult life), which form γ_4 tetramers (Hb Bart's) or β_4 tetramers (HbH). The presence of Hb Bart's or HbH is thus diagnostic of α-thalassaemia. Due to their very high oxygen affinity, Hb Bart's and HbH are not functional haemoglobins; HbH is unstable and precipitates in older red cells, forming H inclusions.

HbH disease lies between the two ends of the clinical spectrum, the asymptomatic α-thalassaemia trait and Hb Bart's hydrops fetalis. As in β-thalassaemia intermedia, HbH disease spans a wide range of clinical and haematological phenotypes, with equally heterogeneous genotypes, varying with the geographic distribution of the different α-thalassaemia variants. HbH disease most commonly results from the interaction of α^0 and α^+-thalassaemia ($--/-\alpha$) and thus, similar to Hb Bart's hydrops syndrome, is also restricted to populations where α^0-thalassaemia is common. Less often, it can result from the interaction of α^0-thalassaemia with non-deletional forms of α-thalassaemia ($/\alpha^T\alpha$) or from homozygous non-deletional α-thalassaemia ($\alpha^T\alpha/\alpha^T\alpha$). HbH disease in Southeast Asia commonly arises from homozygosity or compound heterozygosity for Hb Constant Spring ($\alpha^{CS}\alpha/\alpha^{CS}\alpha$- or $\alpha^{CS}\alpha/--$).

As the non-deletional forms of α^+-thalassaemia tend to have a more severe phenotype than the deletional forms, in some cases the homozygous state ($\alpha^T\alpha/\alpha^T\alpha$) may be associated with the phenotype of HbH disease. In a small number of cases, HbH can be unusually severe, requiring regular blood transfusion, and rare cases may result in Hb Bart's hydrops syndrome (in particular genotype $--/\alpha^T\alpha$).

Haemoglobin Bart's hydrops syndrome

This is an important cause of fetal loss throughout Southeast Asia and the eastern Mediterranean. There is no production of α-chains and hence neither fetal nor adult haemoglobin. The fetus is usually stillborn between 28 and 40 weeks or, if liveborn, takes a few gasping respirations and then expires within the first hour after birth. Affected neonates show the typical picture of hydrops fetalis, with gross pallor, generalized oedema and massive hepatosplenomegaly. Fetal anaemia may be detected from early in the second trimester, with increased blood flow on the fetal middle cerebral artery, as measured by Doppler ultrasound. There is an increased frequency of congenital abnormalities and a very large friable placenta. All these findings are due to severe intrauterine anaemia. The haemoglobin is in the range 60–80 g/L and there are gross thalassaemic changes of the red cells, with many nucleated erythroid cells in the blood. The haemoglobin consists of approximately 80% Hb Bart's and 20% Hb Portland ($\zeta_2\gamma_2$). Very rarely, Hb Bart's hydrops infants have survived to term because they continue to produce embryonic haemoglobin. Apart from fetal death, this syndrome is characterized by a high incidence of toxaemia of pregnancy and obstetric complications due to the large placenta.

Haemoglobin H disease

This condition is characterized by a variable degree of anaemia and splenomegaly, but it is unusual to find severe thalassaemic bone changes or growth retardation. Patients usually survive into adult life, although the course may be interspersed with severe episodes of haemolysis associated with infection or worsening of the anaemia due to progressive hypersplenism. Parvovirus B19 infection in particular may cause transient reticulocytopenia with severe anaemia and the need for red cell transfusion. In addition, oxidant drugs may increase the rate of precipitation of HbH and exacerbate the anaemia. Haemoglobin values range from 70 to 100 g/L and the blood film shows typical thalassaemic changes. There is a moderate reticulocytosis and, on incubation of the red cells with brilliant cresyl blue, numerous inclusion bodies are generated by precipitation of HbH under the redox action of the dye. After splenectomy, large preformed inclusions can be demonstrated on incubation of blood with methyl violet. Haemoglobin analysis reveals 5–40% HbH, together with HbA and a normal or reduced level of HbA_2.

α-Thalassaemia traits

α^0-Thalassaemia trait is characterized by very mild hypochromic anaemia with red cell indices similar to those of β-thalassaemia trait; the MCH is less than 25 pg and the HbA_2 level is normal. Occasional HbH bodies may be present in the red cell on supravital staining. There are no diagnostic tests with which to identify this condition with certainty except DNA analysis. Deletional α^+-thalassaemia carriers have near-normal haematological findings. The heterozygous states for the non-deletional forms of α^+-thalassaemia are sometimes associated with very mild hypochromic anaemia; the type associated with Hb Constant Spring can be identified by the presence of trace amounts of the variant on haemoglobin electrophoresis at an alkaline pH. α^0-Thalassaemia carriers can be identified with more certainty in the neonatal period, when they have 5–10% Hb Bart's, which disappears over the first few months of life and is not replaced by HbH. Some α^+-thalassaemia carriers have slightly increased levels of Hb Bart's, in the 1–3% range, but its absence does not exclude the diagnosis.

Other forms of α-thalassaemia

There are several other forms of α-thalassaemia that are completely unrelated in their pathogenesis and distribution to the conditions described in the previous sections. They comprise the α-thalassaemia mental retardation syndromes and the association of α-thalassaemia with myelodysplasia.

α-Thalassaemia with mental retardation syndromes

There are two forms of α-thalassaemia associated with mental retardation, one encoded on chromosome 16 (ATR-16), the other on the X chromosome (ATR-X). ATR-16 results from large chromosomal rearrangements and extensive deletion of 1–2 Mb from the subtelomeric end of the short arm of chromosome 16.

Affected children usually have a relatively mild degree of cognitive impairment and no dysmorphic features. On the other hand, ATR-X, which because of its mode of inheritance affects boys, is associated with widespread dysmorphic features and severe learning difficulties. It results from many different mutations of the *ATRX* gene, which is located on the X chromosome. The ATRX protein has many features in common with DNA helicases, transcription factors that are involved in the modelling of chromatin and gene regulation. ATRX also appears to play an important role in the transcription of the α-globin genes and undoubtedly many other genes during early development. The *ATRX* gene has also been reported to be involved in a considerable number of X-linked mental retardation syndromes without an α-thalassaemia phenotype.

α-Thalassaemia associated with myelodysplasia

An α-thalassaemic phenotype is also found in association with forms of myelodysplasia in elderly patients. The blood films of such patients show dimorphic features, with populations of red cells containing HbH inclusion bodies and a variable level of HbH in peripheral blood. Acquired mutations in the *ATRX* gene have been identified in the blood cells of patients with this syndrome. The relationship of the mutations in *ATRX* to the neoplastic phenotype remains to be determined.

Thalassaemia intermedia, non-transfusion-dependent thalassaemia

Definition and molecular pathology

The term thalassaemia intermedia is used to describe patients with the clinical picture of thalassaemia which, although not transfusion dependent, is associated with a much more severe degree of anaemia than is found in carriers for α- or β-thalassaemia. It is increasingly referred to as non-transfusion-dependent thalassaemia (NTDT), particularly in literature discussing iron chelation. Whether a patient is classified as having thalassaemia intermedia or thalassaemia major depends on a doctor deciding that the patient would benefit from regular blood transfusions; this decision is based not only on the clinical factors mentioned in the section on thalassaemia major, but also on non-clinical factors such as the availability of blood transfusions, the experience of the clinician and the wishes of the patient. Many different genotypes may underlie thalassemia intermedia, as mentioned earlier, with HbE/β-thalassaemia perhaps being the commonest (Table 6.2). HbH disease is sometimes considered as a type of thalassaemia intermedia, but its pathophysiology is quite different to that caused by β-thalassaemia.

Table 6.2 Molecular basis of β-thalassaemias intermedia.

Homozygous or compound heterozygous state for β-thalassaemia
Inheritance of mild β-thalassaemia alleles (homozygous or compound heterozygotes)
Compound heterozygosity for a mild and a more severe allele
Coinheritance of α-thalassaemia
Increased HbF response
β-Globin gene promoter mutations (deletional or non-deletional)
Coinheritance of HbF quantitative trait loci
Linked: Xmn1-$^{G}\gamma$ polymorphism
Unlinked: BCL11A gene (chromosome 2p), *HBS1L-MYB* intergenic polymorphisms (chromosome 6q) *KLF1* variants
Heterozygous state for β-thalassaemia
Coinheritance of extra α-globin genes as triplicated (/ααα) or quadruplicated (/αααα) globin complexes or segmental duplication of entire α-globin cluster
Dominantly inherited β-thalassaemia (hyperunstable β-globin chain variants)
Compound heterozygosity for β-thalassaemia and β-chain variants
HbE/β-thalassaemia
Compound heterozygosity for β-thalassaemia and HPFH or δβ-thalassaemia
Homozygosity for δβ-thalassaemia

Clinical and haematological changes

At one end of the spectrum are individuals who, except for mild anaemia, are symptom-free. At the other, there are patients who have haemoglobin values of 50–70 g/L and who develop marked splenomegaly, osteopenia and skeletal deformities due to expansion of bone marrow. Iron absorption is increased because of the expanded erythron producing high levels of GDF-15 or other proteins, with suppression of hepcidin production. A wide range of other problems are particularly associated with thalassaemia intermedia, including pulmonary hypertension, hypercoagulability, pseudoxanthoma elasticum and other connective tissue disorders, hypersplenism, leg ulceration, folate deficiency, extramedullary haemopoietic tumour masses in the chest and skull, gallstones and a marked proneness to infection. Many of these problems worsen with age, due to iron loading and cardiovascular impairment, and increasing numbers of patients become transfusion dependent. Because of the heterogeneity of these disorders, management is highly dependent on the course that evolves in an individual patient; all patients should be followed up very carefully from early childhood. The haemoglobin constitution of the intermediate forms of β-thalassaemia depends on the contributing genotypes, and in many cases is similar to that found in the major forms.

Treatment

Intermittent blood transfusions are often necessary due to falls in haemoglobin caused by fever and infection, particularly with parvovirus B19, which causes reticulocytopenia. It can be difficult to decide who would benefit from more regular blood transfusions and when to start them. In countries with a ready supply of safe blood, there is an increasing tendency to start regular transfusions even in children, maintaining the haemoglobin at greater than 80 g/L in order to avoid the emerging complications of skeletal deformities, pulmonary hypertension and osteopenia. There is also some evidence that this improves the quality of life, particularly with emerging options for oral iron chelation. This is not possible in much of the world and management consists of reserving transfusion for severe symptomatic anaemia.

Even in the absence of regular erythrocyte transfusions, significant levels of iron accumulate through increased gastrointestinal absorption. Clinical trials with iron chelators have shown that it is possible to remove this excess iron safely, and this is increasingly practised (Chapter 4). However clinical complications of iron overload are relatively uncommon in thalassaemia intermedia, and it is less clear when chelation is clinically beneficial.

Pharmacological treatment to increase HbF and total haemoglobin levels is potentially applicable to thalassaemia intermedia, in that relatively small increases in haemoglobin levels with a corresponding reduction in ineffective erythropoiesis could help a patient thrive who would otherwise require regular transfusions. Hydroxycarbamide (hydroxyurea) is the most widely used drug in this context, and while some patients undoubtedly benefit, in general results are disappointing. Butyrate and other short-chain fatty acids also promote HbF synthesis and have been used with limited clinical success in thalassaemia intermedia. A number of newer drugs are being developed which may boost HbF to a greater extent, most notably the new generation of short-chain fatty acid derivatives and immunomodulatory drugs such as pomalidomide and lenalidomide. Mouse studies and early clinical trials of activin receptor IIA ligand traps show promising results, with increased haemoglobin concentrations resulting from reduced oxidative stress and improved erythropoiesis. Gene therapy continues to be studied, but is still a very experimental procedure in thalassaemia.

Screening for thalassaemias

Preconception, antenatal and neonatal screening programmes are important in the clinical care and public health management of haemoglobinopathies. Thalassaemia is prevalent in many middle- and low-income countries, where either treatment is not available or the cost of regular blood transfusions and chelation is a major drain on limited medical resources. In other countries, including parts of Europe, Australia and North America, it is important to identify the antenatal risk of having a baby with thalassaemia to provide informed parental choice. Apart from haemopoietic stem cell transplantation, there is no definitive treatment, and many countries in which the disease is common are putting a major effort into programmes for its prevention.

Premarital and preconception screening

Ideally, individuals know their thalassaemia status before they decide to have children, and potentially use this information to choose a suitable partner. Some programmes therefore concentrate on screening teenagers at school. Other countries insist that a couple are screened for thalassaemia status before they can get married; this latter approach has been successful in Cyprus and parts of the Middle East, such as Bahrain, where social pressures have reduced the number of at-risk marriages.

Antenatal screening, prenatal diagnosis and preimplantation genetic diagnosis

Many screening programmes concentrate on identifying pregnant women who are thalassaemia carriers in the first trimester of pregnancy. This is done by varying combinations of blood tests and the identification of women at high risk of carrying thalassaemia based on their ethnic origin. The latter approach is potentially effective in areas with a low prevalence of thalassaemia in the native population, such as northern Europe, although increasing racial admixture is reducing the feasibility of such selective screening, particularly in cities. If a woman is found to be a carrier, screening is then offered to her partner, and if both are carriers, they are counselled about the risk of the fetus inheriting a severe form of thalassaemia and offered prenatal diagnosis (PND).

PND of the thalassaemias was first carried out by fetal blood sampling, typically for β-thalassaemia, between 18 and 20 weeks of gestation. At about 20 weeks in the normal fetus, β-globin synthesis constitutes about 10% of the total haemoglobin, giving a β/γ ratio of 0.1. A fetus heterozygous for β-thalassaemia has a β/γ ratio of about 0.05, while one with β-thalassaemia major produces either none or traces of β-globin with a β/γ ratio of less than 0.025. Fetal blood sampling has now been replaced by fetal DNA analysis. DNA analysis was applied to PND of the haemoglobinopathies in the late 1970s and early 1980s. Fetal DNA is obtained at 11–12 weeks' gestation by chorionic villus sampling or amniocentesis. Initially, detection of the mutations was indirect, relying on linkage analysis of restriction fragment length polymorphisms (RFLPs) and Southern blot hybridization, which required relatively large amounts of DNA and the whole procedure took 7–10 days. The development of polymerase chain reaction (PCR) for specific amplification of DNA has revolutionized the molecular diagnostic field, and mutations can now be detected directly using PCR-based techniques in 1 day, enabling PND to be carried out rapidly within 3 days.

Couples can then make an informed choice as to whether to terminate the pregnancy if the fetus is affected. Some couples

find PND helpful in preparing for the birth of a potentially ill child, even though they would not contemplate termination. The main complication of invasive PND is the increased risk of miscarriage of about 1%. This has led to research to develop non-invasive methods of PND based on maternal blood sampling. Maternal blood contains small numbers of fetal cells and also cell-free fetal DNA, both of which could potentially be used to diagnose fetal thalassaemia. The low concentration of fetal material relative to maternal has made this technically very difficult, although advances in methods of DNA analysis and sequencing platforms seem likely to make this feasible in the near future.

Some couples want to avoid having a fetus with thalassaemia, but find PND and selective termination unacceptable. Preimplantation genetic diagnosis involves the use of *in vitro* fertilization techniques to generate 5–15 embryos; at the eight-cell stage, one embryonic cell can be removed and tested for thalassaemia alleles; it is then possible to only implant embryos without thalassaemia. Whilst appealing, it is currently a difficult, stressful and expensive procedure, with only 10–20% of couples taking home a baby. Again, advances in reproductive biology and DNA technology are making this more applicable, and it is available in an increasing number of countries.

Neonatal screening

It is possible to detect the majority of babies with severe thalassaemia by neonatal testing, either as cord blood or more commonly from the neonatal blood spot. Many neonatal screening programmes involve testing dried blood spots on pieces of filter paper, usually made from a heel-prick blood sample when the baby is 5–7 days old. This blood spot is also used to screen for other serious neonatal disorders, including phenylketonuria, congenital hypothyroidism and cystic fibrosis. If haemoglobin analysis shows HbF only, with no HbA or other haemoglobin variants, it is likely that the baby has inherited a severe form of β-thalassaemia and may be transfusion dependent; less severe possibilities include thalassaemia intermedia and homozygosity for HPFH. Babies identified in this way can then be followed up closely rather than waiting until they present following a period of prolonged illness. Parents can also be tested and given advice concerning the risk to future pregnancies.

Structural haemoglobin variants related to thalassaemia (Table 6.3)

The unstable haemoglobin disorders

The unstable haemoglobin disorders are a rare group of inherited haemolytic anaemias that result from structural changes in the haemoglobin molecule, which cause its intracellular precipitation with the formation of Heinz bodies. Their true incidence is not known and there have been several well-documented instances in which patients with one of these variants have had no affected relatives, suggesting that they have arisen by a new mutation. Unlike more common globin gene mutations, there is no obvious ethnic or geographical variation in the prevalence of unstable haemoglobins. Haemoglobin Koln (β98, Val→Met) is perhaps the most common.

Table 6.3 Diseases due to structural haemoglobin variants.

Sickle syndromes causing haemolysis and vaso-occlusion
HbSS
Compound heterozygosity for HbS with other β-haemoglobin variants (HbS/C, HbS/D Punjab, HbS/O Arab)
Compound heterozygosity for HbS with β-thalassaemia (HbS/β-thalassaemia)
Haemolytic anaemia
Unstable haemolytic variants
Congenital polycythaemia
High-oxygen-affinity haemoglobin variants
Congenital cyanosis
Low-oxygen-affinity haemoglobin variants
M haemoglobins
Hypochromic microcytic anaemia (thalassaemia phenotype)
Variants with inefficient synthesis due to alternative splicing, e.g. HbE
Lepore haemoglobins
Unstable chain termination variants, e.g. Hb Constant Spring
Drug-induced haemolysis
e.g. Hb Zurich

Molecular pathology and pathogenesis

Most of the unstable haemoglobins result from single amino acid substitutions or small deletions. For example, substitutions in or around the haem pocket can disrupt its anatomy and allow in water, with subsequent oxidative damage to haem, leading to precipitation of haemoglobin. Some substitutions, such as those involving proline residues, cause a marked disturbance of the secondary structure of globin chains. A few variants result from deletions of either single amino acids or several residues. For example, in Hb Gun Hill, five amino acids are missing, including the haem-binding site. As the unstable haemoglobins precipitate in the red cells or their precursors, they produce intracellular inclusions (Heinz bodies) which, together with oxidant damage to their membranes, make the cells more rigid and hence cause their premature destruction in the microcirculation.

Clinical features

Unstable haemoglobins are characterized by a chronic, non-spherocytic haemolytic anaemia and splenomegaly. Like all chronic haemolytic anaemias, there is an increased incidence of pigment gallstones with their associated complications; the

risk is particularly high if there is coinheritance of Gilbert syndrome (polymorphic variant in the promoter of the *UGT1A1* gene). The condition may become worse during periods of infection, causing fever, and, in the more severe forms, such episodes are associated with life-threatening anaemia and the need for blood transfusion. Some oxidant drugs may increase the rate of haemolysis, and parvovirus B19 infection may cause temporary reticulocytopenia. Apart from icterus and splenomegaly, there are no characteristic physical findings.

Laboratory diagnosis

The peripheral blood film shows typical features of haemolysis, but the red cell morphology may be normal. Occasionally, there is mild hypochromia and microcytosis. Heinz bodies are present in the peripheral blood after splenectomy. The most characteristic feature of the unstable haemoglobins is their heat instability. If a dilute haemoglobin solution is heated at 50 °C for 15 min, the unstable haemoglobins precipitate as a dense cloud. A similar effect can be induced by isopropanol at lower temperatures. The isopropanol heat sensitivity test is more sensitive and specific than simple heat instability. Some of these variants can be seen on haemoglobin electrophoresis, but others are electrophoretically silent, because they result from a neutral amino acid substitution; these can be demonstrated only by the heat precipitation test. DNA analysis can provide definitive diagnosis, typically by sequencing of the α- or β-globin genes, and this is increasingly used at an early stage as sequencing gets faster and cheaper.

Treatment

Splenectomy seems to be beneficial in some cases, although experience is inevitably limited. Intermittent blood transfusions may be necessary. If haemolysis is very severe, the patient may benefit from regular blood transfusions and bone marrow transplantation should be considered.

High-oxygen-affinity haemoglobin variants

Some haemoglobin variants cause increased oxygen affinity, which results in varying degrees of polycythaemia. These are rare and occur sporadically; haemoglobin Chesapeake (α92, Arg→Leu) was the first one identified in 1966.

Molecular pathology

Some high-oxygen-affinity haemoglobin variants result from single amino acid substitutions at critical parts of the haemoglobin molecule that are involved in the configurational changes which underlie haem–haem interaction and the production of a sigmoid oxygen dissociation curve at the junctions between the α- and β-subunits. Others involve the amino acids concerned with the binding of 2,3-DPG to haemoglobin. Reduced 2,3-DPG binding moves the oxygen dissociation curve to the left, reducing the P_{50} and causing the haemoglobin to hold on to oxygen more avidly than normal. This causes functional anaemia, with tissue hypoxia, which in turn causes an increased output of erythropoietin and an elevated red cell mass.

Clinical features

Most affected individuals are completely healthy and are identified only when a routine haematological examination shows an unusually high haemoglobin level or haematocrit. There is no splenomegaly and, apart from a raised red cell mass, there are no associated haematological findings. Although it might be expected that a high-oxygen-affinity haemoglobin would cause defective oxygenation of the fetus, none of the reported families has a history of such problems.

Diagnosis

The condition should be suspected in any patient with a pure red cell polycythaemia associated with a left-shifted oxygen dissociation curve. Erythropoietin levels will typically be inappropriately high for the haematocrit. The absolute reticulocyte count may be slightly elevated, although the reticulocyte percentage should be normal. The erythron is expanded, as shown by high plasma concentrations of soluble transferrin receptor. The diagnosis can be confirmed by haemoglobin analysis using chromatography, mass spectrometry or DNA analysis.

Treatment

In asymptomatic individuals, no treatment is necessary. The difficulty arises if there is associated vascular disease, particularly coronary or cerebral artery insufficiency. As these patients require a high haemoglobin level for oxygen transport, venesection should be carried out with great caution. Venesection is undertaken because of increased risk of vascular complications, and typically the aim is to keep the haematocrit below 0.55, although there is little evidence to support this.

Low-oxygen-affinity haemoglobin variants

More than 50 haemoglobin variants with reduced oxygen affinity have been identified, often associated with other abnormal properties such as instability. The first to be described, Hb Kansas (β102 Asn→Thr), was found in a mother and son with unexplained cyanosis. The subjects were asymptomatic and had normal haemoglobin levels without any evidence of haemolysis. Like many of the high-affinity haemoglobins, the amino acid substitution in this variant was at the interface between the α- and β-globin chains. For reasons that are not clear, some substitutions in this region give rise to variants with a relatively low oxygen affinity. This condition should be thought of in any patient with unexplained congenital cyanosis.

Congenital methaemoglobinaemia due to haemoglobin variants

Several α- and β-globin variants associated with methaemoglobinaemia have been discovered. These are usually referred to as M-haemoglobins, for example haemoglobin M-Boston (α58, His→Tyr). These disorders, unlike the genetic methaemoglobinaemias due to enzyme defects, follow a dominant pattern of inheritance. The patients are blue in colour, and may have mild polycythaemia as methaemoglobin does not carry oxygen. Diagnosis is based on detecting high methaemoglobin levels, which are increasingly measured by blood gas analysers and some pulse oximeters. Increased methaemoglobin levels should be confirmed on formal laboratory analysis using spectrophotometry. Subsequently haemoglobin analysis using electrophoresis or mass spectrometry is performed, or sequencing of the α- and β-globin genes.

Acknowledgement

Parts of the present chapter are based on the corresponding previous edition chapter, and for this we acknowledge the contribution of Professor Sir David Weatherall. We also thank Claire Steward for help on the preparation of the chapter.

Selected bibliography

Angelucci E, Matthes-Martin S, Baronciani D *et al.* (2014) Hematopoietic stem cell transplantation in thalassemia major and sickle cell disease: indications and management recommendations from an international expert panel. *Haematologica* **99**: 811–20.

Giardine B, Borg J, Higgs DR *et al.* (2011) Systematic documentation and analysis of human genetic variation in hemoglobinopathies using the microattribution approach. *Nature Genetics* **43**: 295–301.

Harteveld CL, Higgs DR (2010) α-Thalassaemia. *Orphanet Journal of Rare Diseases* **5**: 21.

Higgs DR (2013) The molecular basis of alpha-thalassemia. *Cold Spring Harbor Perspectives in Medicine* **3**: a011718.

Higgs DR, Engel JD, Stamatoyannopoulos G (2012) *Thalassaemia. Lancet* **379**: 373–83.

Kautz L, Jung G, Valore EV *et al.* (2014) Identification of erythroferrone as an erythroid regulator of iron metabolism. *Nature Genetics* **46**(7): 678–84.

Piel FB, Weatherall D (2014) The α-thalassemias. *New England Journal of Medicine* **371**: 1908–1916.

Sankaran VG, Orkin SH (2013) The switch from fetal to adult hemoglobin. *Cold Spring Harbor Perspectives in Medicine* **3**: a011643.

Steinberg MH, Forget BG, Higgs D, Weatherall DJ (eds.) (2009) *Disorders of Hemoglobin.* Cambridge University Press, Cambridge.

Taher AT, Vichinsky E, Musallam KM, Cappellini MD, Viprakasit V (eds.) (2013) *Guidelines for the Management of Non Transfusion Dependent Thalassaemia (NTDT).* Thalassaemia International Federation, Nicosia, Cyprus.

Thein, SL (2013a) Genetic association studies in β-hemoglobinopathies. *Hematology ASH Education Program Book 2013,* 354–61.

Thein SL (2013b) The molecular basis of beta-thalassemia. *Cold Spring Harbor Perspectives in Medicine* **3**, 3/5/a011700.

Weatherall DJ (2010) The inherited diseases of hemoglobin are an emerging global health burden. *Blood* **115**: 4331–6.

Weatherall DJ, Clegg JB (2001) *The Thalassaemia Syndromes.* Blackwell Science, Oxford.

CHAPTER 7

Sickle cell disease

7

Anne Marsh and Elliott P Vichinsky

UCSF Benioff Children's Hospital Oakland, Oakland, California, USA

Introduction

Sickle cell disease (SCD) is an inherited chronic haemolytic anaemia whose clinical manifestations arise from the tendency of the haemoglobin (HbS or sickle haemoglobin) to polymerize and deform red blood cells into the characteristic sickle shape. This property is due to a single nucleotide change in the β-globin gene leading to substitution of valine for glutamic acid at position 6 of the β-globin chain ($\beta^{6glu\rightarrow val}$ or β^{s}). The homozygous state (HbSS or sickle cell anaemia) is the most common form of sickle cell disease, but interaction of HbS with thalassaemia and certain variant haemoglobins also leads to sickling. The term 'sickle cell disease' is used to denote all entities associated with sickling of haemoglobin within red cells (Table 7.1).

Geographic distribution of sickle mutation

Several distinct β-globin gene haplotypes are associated with the sickle mutation, and their distribution provides evidence for origin of the mutation in several locations within Africa (the Senegal, Benin and Bantu haplotypes) and Asia (the Arab–Indian haplotype). The sickle trait bestows survival benefit in areas endemic for *falciparum* malaria, and the distribution of SCD historically paralleled this disease. The sickle haemoglobin-containing red cells inhibit proliferation of *Plasmodium falciparum*, and are more likely to become deformed and removed from the circulation. In recent times, the dissemination of the sickle mutation in different areas of the world took place from the movement of populations via trade routes and the slave trade (Table 7.2). The prevalence of SCD varies tremendously among ethnic and tribal groups within a geographic area. The disease is observed occasionally among the white population: 10% of patients with HbSS identified by the California newborn screening programme are not of African descent.

Pathophysiology

Molecular basis of sickling

Deoxygenation of HbS leads to a conformational change that exposes a hydrophobic patch on the surface of the β^{s}-globin chain at the site of β^{6} valine (Figure 7.1). Binding of this site to a complementary hydrophobic site on a β-subunit of another haemoglobin tetramer triggers the formation of large polymers. The polymers consist of staggered haemoglobin tetramers that aggregate into 21-nm diameter helical fibres, with one inner and six peripheral double strands. The polymerization proceeds after a delay, the length of which is extremely sensitive to the intracellular deoxy-HbS concentration. Even a small increase in deoxy-HbS concentration, such as might occur with cellular dehydration, profoundly shortens the delay time and augments sickling. The haemoglobin tetramers first aggregate into a nucleus, which rapidly expands into a fibre. The newly formed fibre provides nuclei on its surface for aggregation of haemoglobin tetramers to form several more fibres.

The polymerization of HbS in the circulating red cells is influenced by the oxygenation status, the intracellular haemoglobin concentration and the presence of non-sickle haemoglobins. Acidosis and elevated levels of 2,3-diphosphoglycerate (2,3-DPG) promote polymer formation by reducing the oxygen affinity of haemoglobin. The presence of HbA within the red cells,

Postgraduate Haematology, Seventh Edition. Edited by A Victor Hoffbrand, Douglas R Higgs, David M Keeling and Atul B Mehta.

Table 7.1 The sickling syndromes.

Genotype	Mean haemoglobin (g/L)	MCV	Haemoglobin electrophoresis (%)				
			S	A	F	A_2	Other
SS	81	N	80–95	–	2–20	N	–
SS –α/αα, SS –α/–α	86, 92	↓, ↓	80–90, 80–90	–, –	2–20, 2–20	3.3–3.8, 3.3–3.8	–, –
SC	110	↓	40–50	–	1–4		C: 40–50
S/β^0-thalassaemia	88	↓	75–90	–	2–20	4–6	–
S/β^+-thalassaemia	115	↓	50–85	5–30	2–20	4–6	–
SD Punjab	82	N	40	–	2.5–5	2–3	D Punjab: 50
SO Arab	81	N	45	–	4–7		O Arab: 45
S Lepore	110	↓	75	–	3.5–40	2	Lepore: 10
SE	130	↓	60	–	4		E: 30–35
S/HPFH	137	N or ↓	60–70	–	25–35	1.5–2.5	–
AS*	N	N	30–45	50–65	2–5	N	–

*Sickle cell trait is asymptomatic.
MCV, mean corpuscular volume.

as in sickle trait, inhibits polymerization by diluting HbS. The inhibitory effect of HbF on polymerization of HbS is more profound owing to the greater amino acid disparity between the β^s- and γ-globin chains.

Effect on erythrocytes

Red cells acquire the sickle or elongated shape upon deoxygenation as a result of intracellular polymerization of HbS, a phenomenon that is reversible on reoxygenation. Even in the normally shaped red cells, however, the presence of HbS polymer reduces deformability, with consequent increase in blood viscosity. Repeated or prolonged sickling progressively damages the red cell membrane, which is a phenomenon of primary importance in the pathophysiology of SCD. Membrane damage causes movement of potassium ions and water out of the cell by the Gardos pathway and potassium–chloride cotransport, leading to dehydration of red cells. The intracellular haemoglobin concentration rises (producing dense cells), which shortens the delay time to sickle polymer formation. A second key consequence of membrane damage is alteration of the chemistry of the red cell membrane. Perturbation of lipid organization causes negatively charged phosphatidylserine to appear on the red cell surface instead of its normal location in the inner monolayer. In addition, the red cells become abnormally adherent to the vascular endothelium through vascular cell adhesion molecule (VCAM)-1, thrombospondin and fibronectin.

Table 7.2 Areas of high prevalence of sickle mutation.

Geographic region	Heterozygote rate (%)
Africa	
Northern	1–2
Western	10–30
Central	7–37
Southern	0–5
Mediterranean	
Northern Greece	1–27
Southern Italy	
Americas	
United States: African ancestry	8
Caribbean: African ancestry	10
Brazil: non-white	7
Asia	
Saudi Arabia: south-west	5
Saudi Arabia: eastern province	25
India: central India – tribal population	20–30

Vaso-occlusion

Several processes contribute to development of vaso-occlusion in SCD. Slowing of blood flow arises from abnormal regulation of vascular tone as a result of diminished nitric oxide (NO)-induced vasodilatation. This is aggravated by increase in blood viscosity, resulting from less deformable red cells, a phenomenon called abnormal rheology. Vaso-occlusion is initiated by adhesion of young deformable red cells to the vascular endothelium, and is followed by trapping of rigid irreversibly sickled cells (Figure 7.2). Adhesion occurs in the postcapillary venules and is promoted by leucocytosis, platelet activation and inflammatory cytokines.

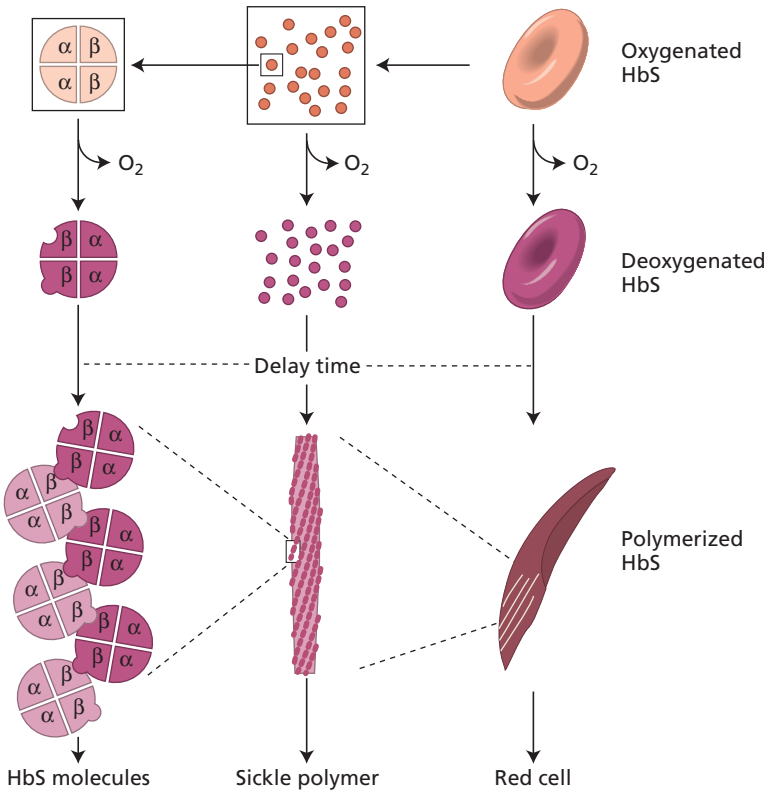


Figure 7.1 Induction of red cell sickling. As red cells traverse the microcirculation, oxygen is released from oxy-HbS (red circles), generating deoxy-HbS (purple circles). Conformational change exposes a hydrophobic patch at the site of the β^6-valine replacement, shown as a projection (left column), which can bind to a complementary hydrophobic site on a subunit of another haemoglobin tetramer, shown as an indentation. Only one of the two β^6-valine sites in each HbS tetramer makes this contact. The middle column shows the assembly of deoxy-HbS into a helical 14-strand fibre, shown as a twisted rope-like structure. The delay time is inversely proportional to the intracellular haemoglobin concentration raised to the 15th power. As deoxy-HbS polymerizes and fibres align, the red cell is distorted into an elongated banana or 'sickle' shape (right column). (Source: Bunn, 1997 [*New Engl J Med* **337**: 762–9]. Reproduced with permission of Massachusetts Medical Society.)

Haemolysis

SCD is characterized by chronic intravascular and extravascular haemolysis. Sickling-induced membrane fragmentation and complement-mediated lysis cause intravascular destruction of red cells. Membrane damage also leads to extravascular haemolysis through entrapment of poorly deformable cells or uptake by macrophages. The red cell survival, measured by ^{51}Cr assay, is 4–25 days, with dense cells surviving for a considerably shorter time than red cells containing some HbF (F cells). Patients have greatly expanded bone marrow space, but the serum erythropoietin level is lower than expected for the extent of anaemia because of the decreased oxygen affinity of HbS. Individuals with concomitant deletion of one or two α-globin genes, or the Senegal or Arab–Indian haplotypes, have higher baseline haemoglobin levels. The significance of chronic intravascular haemolysis in SCD extends beyond anaemia, since the release of free haemoglobin causes depletion of NO in the plasma. This is linked to endothelial dysfunction and the development of several complications, including pulmonary hypertension.

Clinical manifestations

Clinical symptoms vary tremendously between patients with SCD. The disease is more severe in patients with HbSS or HbS/β^0-thalassaemia than in those with HbS/β^+-thalassaemia or HbSC disease. However, for increasingly recognized reasons, the disease severity varies enormously, even within the subgroup of patients with HbSS. Genetic modifiers that influence both laboratory and clinical parameters in part explain the phenotypic heterogeneity of SCD. Some of the genetic modifiers known to impact disease phenotype include coinheritance of α-thalassemia trait, β-globin haplotype, modifiers of endogenous haemoglobin F production, UGT1A1 promoter polymorphisms and genes that have shown protection against stroke. The coinheritance of one or two α-gene deletions results in lower MCV, MCH, total bilirubin and absolute reticulocyte counts. The Senegal and Arab–Indian β-globin haplotypes produce less severe clinical disease than the other African haplotypes. The high HbF level observed in hereditary persistence of fetal haemoglobin (HPFH) is associated with very mild disease. Certain BCL11A single nucleotide polymorphisms known to positively affect endogenous fetal haemoglobin production have been shown to have favourable laboratory profiles and may serve as potential targets for gene editing of haemopoietic stem cells. Specific promoter mutations within the UTG1A1 gene, encoding an enzyme necessary for bilirubin conjugation, are correlated with increased bilirubin levels and a propensity to develop cholelithiasis. Mutations within the GOLGB1 and ENPP1 genes have been shown to have a protective effect against stroke in SCD. These and other genetic modifiers help to explain the phenotypic heterogeneity observed in this monogenic disease. Genome-wide genetic association studies and identification of

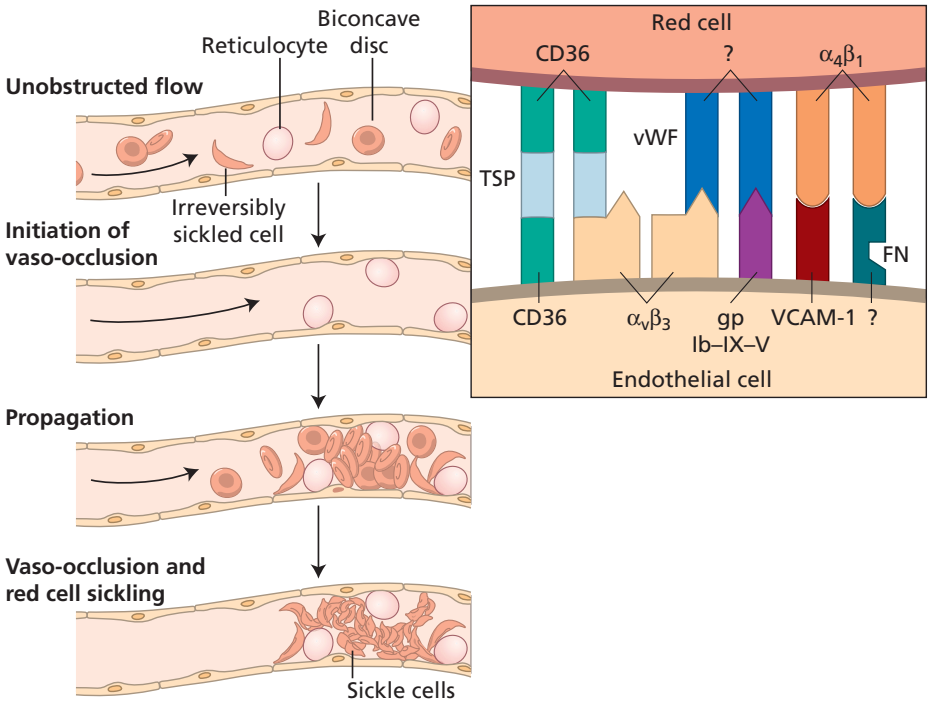


Figure 7.2 Endothelial red cell adhesion and vaso-occlusion in sickle cell disease. Adhesive sickle reticulocytes initiate vaso-occlusion by becoming attached to the endothelium of vessel walls. Thereafter, poorly deformable red cells begin to accumulate behind the site of adhesion, ultimately resulting in an occluded vascular segment containing many sickled red cells. The inset shows the site of red cell attachment to an endothelial cell and several adhesion mechanisms that could participate in the vaso-occlusive process. On the red cell, the relevant adhesion receptors include CD36, which binds thrombospondin (TSP), and integrin $\alpha_4\beta_1$, which binds both fibronectin (FN) and vascular cell adhesion molecule (VCAM)-1. On the endothelial cell, the receptors include CD36; integrin $\alpha_v\beta_3$; the complex of glycoproteins Ib, IX and V (gp Ib–IX–V), which binds von Willebrand factor (VWF); and VCAM-1. Adhesive interactions between the red cell and endothelial cells may be direct ($\alpha_4\beta_1$ to VCAM-1) or mediated by a plasma factor (CD36 to TSP to $\alpha_v\beta_3$). The list of molecules identified as involved in mediating adhesion continues to increase and redundancy in the system is likely. Question marks indicate unidentified receptors. (Source: Hebbel, 2000 [*New Engl J Med* **342**: 1910–12]. Reproduced with permission of Massachusetts Medical Society.)

epigenetic phenomena are likely to identify other potential modifiers of disease severity.

In countries with inadequate healthcare, SCD is associated with high mortality in the first three years of life as a result of sepsis and splenic sequestration. In the developed world, the typical patient with SCD has moderately severe anaemia, leads a relatively normal life interrupted by 'crises' as a result of vaso-occlusion, and has a life expectancy of over 45 years.

Anaemia

The underlying β-globin genotype primarily determines the baseline haemoglobin value in SCD, but exacerbation of anaemia can occur for numerous reasons. Patients with more severe anaemia at baseline have a greater probability of developing stroke and renal dysfunction. On the other hand, a higher haemoglobin level is associated with a higher incidence of painful episodes, avascular necrosis and acute chest syndrome. Infants with SCD have lower than normal haemoglobin levels after the neonatal period, and the decline continues until it reaches a nadir between 12 and 15 months of age. Boys are slightly more anaemic than girls in the first decade, whereas adult men have higher haemoglobin values than women. Gradual exacerbation of anaemia is observed in both sexes beginning in the fifth decade.

A gradual decrease in haemoglobin level from the baseline value may indicate an underlying folate or iron deficiency. In older patients, however, inadequate erythropoietin production due to chronic renal insufficiency is the most important aetiology of worsening anaemia. Many such patients will become transfusion dependent, although recombinant human erythropoietin therapy can improve the anaemia.

Acute exacerbations of anaemia are observed with aplastic crises and splenic sequestration. The transient arrest of

erythropoiesis and the resultant reticulocytopenia in aplastic crisis is most often due to parvovirus B19 infection. The reticulocytopenia begins 5 days after exposure, lasts for 7–10 days and is followed by recovery with reticulocytosis and normoblasts in peripheral blood. Blood transfusion is often required in the short term. Parvovirus B19 infection is followed by development of lifelong protective immunity.

Splenic sequestration is a serious complication in young children whose spleen has not yet undergone fibrosis due to recurrent vaso-occlusion. The peak incidence of the first episode of sequestration is between 6 and 12 months of age and affects 30% of all patients. Approximately 15% of patients die during the acute episode and the condition recurs in half of the survivors. Episodes may be triggered by a viral illness and the rapid acute enlargement of the spleen traps a significant proportion of the blood volume. Clinically, the child presents with acutely worsening anaemia (>20 g/L fall in haemoglobin), reticulocytosis, enlarging spleen and hypovolaemic shock. Prompt restoration of the blood volume and correction of anaemia is required. Splenectomy is recommended following a sequestration crisis due to the risk of recurrences. Chronic transfusion therapy or partial splenectomy is sometimes used in infants with life-threatening anaemic episodes. Parent education to detect splenic enlargement and seek early medical attention significantly reduces the risk of death from sequestration crisis.

Acute painful episode

An acute episode of pain due to vaso-occlusion is the most frequent symptom for which patients with SCD seek medical attention. Painful episodes are more common in young adults and tend to diminish in older patients. One-third of patients with SCD rarely experience pain, whereas a small subgroup of patients suffer from recurrent episodes. When patients maintain a pain diary, painful events are noted on up to half of the days, but are not severe enough to require visit to a physician. Painful episodes vary in intensity and generally last for a few days. The majority of the episodes have no identifiable cause, although some attacks are precipitated by cold, dehydration, infection, stress or menses.

In young children, the initial pain episode typically presents as dactylitis or hand–foot syndrome, with swelling over the dorsal surface of hands and feet. It arises from bone infarction affecting the small bones and the swelling subsides over 1–2 weeks. The radiographs show thinning of cortex and destructive changes of the affected small bones several weeks after onset. In older children and adults, the common sites of pain are the back, chest, extremities and abdomen. Chest pain is of special significance as it can precede development of acute chest syndrome. Frequent incapacitating painful episodes that are inadequately managed have adverse psychosocial consequences and stress the physician–patient relationship.

Growth and development

Children with SCD are born with normal weight, but fall behind other children by the end of the first year. The weight deficit persists through adulthood and imparts a thin habitus to the typical patient, although obesity is seen in some cases. The rate of growth is lower than normal in SCD patients, and the pubertal growth spurt is delayed by 1–2 years, but the final adult height is normal. Delays also occur in skeletal maturation and onset of puberty, and female patients achieve menarche 1–2 years later than their peers.

Infections

Early loss of splenic function from recurrent vaso-occlusion and the inability to make specific IgG antibodies to polysaccharide antigens increases the risk of fulminant sepsis. Pneumococcal infection is a serious problem in SCD, particularly in children under 3 years (Figure 7.3). Meningitis can accompany pneumococcal sepsis, and the overall mortality rate is 20–50%. Patients who have had previous pneumococcal sepsis are at increased risk for recurrent episodes and must remain on lifelong penicillin prophylaxis. *Haemophilus influenzae* type B is the next most common organism and affects older children. There is considerable variation in the relative incidence of bacterial organisms causing sepsis in young children with SCD in various regions of the world. In Africa, *Salmonella* spp., *Klebsiella* spp., *Escherichia coli* and *Staphylococcus* spp. are more commonly isolated from the blood of febrile children than *Streptococcus pneumoniae*. The incidence of pneumococcal and *H. influenzae* sepsis has declined as a result of penicillin prophylaxis and vaccination of infants. The risk of death during septic episodes has decreased considerably owing to empirical use of antibiotics to treat fever in SCD.

Of the other infections, pneumonia is particularly common in SCD and can be difficult to differentiate from non-infective causes of acute chest syndrome. The most frequent organisms responsible for pneumonia are *Mycoplasma pneumoniae*, *Chlamydia pneumoniae*, *S. pneumoniae* and *H. influenzae*. Lung infections can also arise due to respiratory viruses. In adults, bacteraemia and urinary tract infections due to *E. coli* and other Gram-negative organisms are more frequent. Patients with SCD are susceptible to osteomyelitis caused by bone infarction resulting from vaso-occlusion. The infection is typically due to *Salmonella* spp. or *Staphylococcus aureus*.

Neurological complications

Neurological complications are an important cause of morbidity in SCD. Transient ischaemic attacks or stroke due to cerebral infarction or haemorrhage occur in 25% of patients with SCD (Figure 7.4a). The risk of stroke is increased with lower baseline haemoglobin, low HbF level, high leucocyte count or high

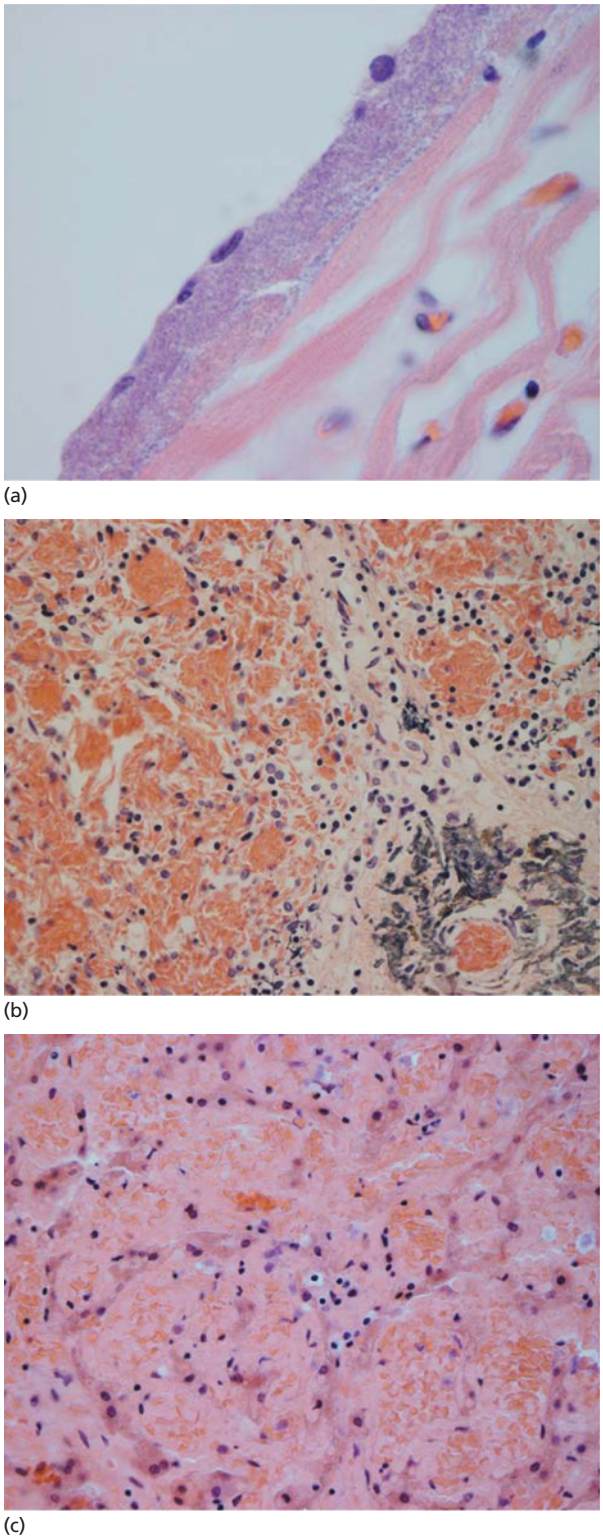
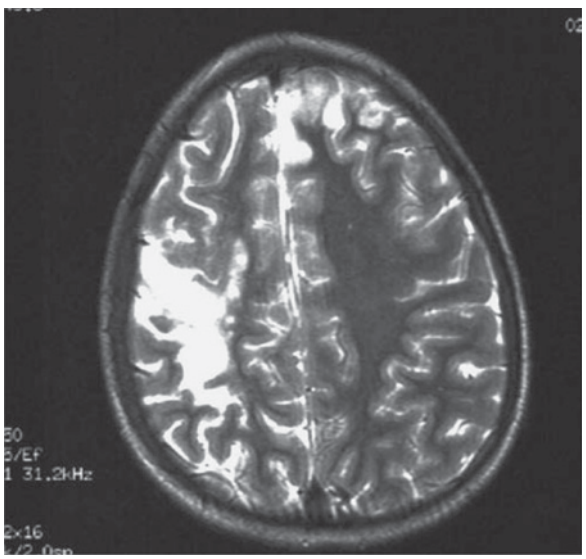


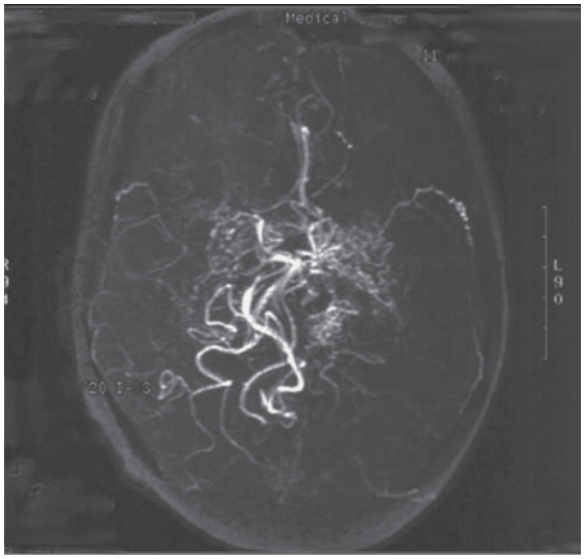
Figure 7.3 Overwhelming pneumococcal sepsis in a 7-year-old child. (a) Numerous bacteria in the blood adjacent to the right ventricular wall. Massive sequestration of the spleen (b) and the liver (c).

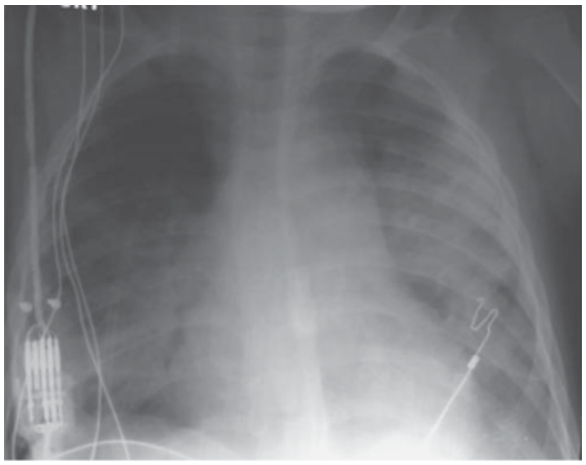
systolic blood pressure. Vascular damage results from elevated cerebral blood flow velocities and interaction of rigid or adherent sickle cells with the vessel wall. Angiography demonstrates stenosis or occlusion of vessels in the circle of Willis and internal carotid arteries, sometimes with aneurysm formation or development of moyamoya disease (Figure 7.4b).

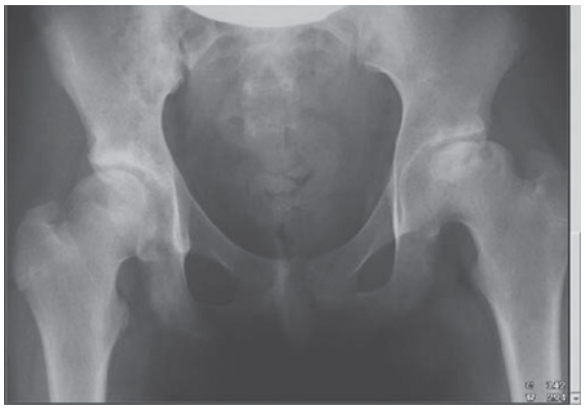
Stroke due to infarction is more frequent in young children and those over 30 years of age, whereas haemorrhage is more common between 20 and 30 years. Stroke is rare in infants, increases to 1 in 100 patients per year between 2 and 9 years, and then diminishes to half that incidence in older patients. The incidence of stroke is 5–10 times greater in HbSS compared with HbSC, HbS/β^+-thalassaemia or HbS/β^0-thalassaemia. Focal seizures or transient ischaemic attacks are common presenting symptoms of stroke, followed by hemiparesis, coma and speech or visual disturbances. The site of bleeding in haemorrhagic stroke is frequently subarachnoid, and these patients present with severe headache, vomiting and coma. Death can occur during the acute event, particularly with haemorrhagic stroke. Patients with neurological symptoms should be evaluated by computed tomography (CT) or magnetic resonance imaging (MRI) to distinguish thrombosis from haemorrhage. Immediate exchange transfusion to lower the HbS level to less than 30% is required. Patients with haemorrhage may require surgical intervention to ligate accessible aneurysms. Surgical vascular bypass procedure with extracranial arteries (encephalodural synangiosis) should be considered in children with moyamoya disease. Even in the absence of overt stroke, silent cerebral infarcts (SCI) are commonly observed on MRI in SCD and are linked to progressive neuropsychiatric and neurological damage, and poor school performance. Acute silent cerebral ischaemic events, which are distinct from overt stroke and SCI, are much more common than previously recognized and are an additional contributor to neurocognitive impairment.

Prevention, early detection and treatment of neurologic complications is important. Transcranial Doppler (TCD) ultrasonography has been shown to be an effective screening tool for identification of stroke risk. TCD screening should begin at 2 years of age and continue annually into the late teenage years. Increased blood flow velocity due to stenosis can be detected by TCD ultrasonography in asymptomatic patients, and flow rates in excess of 200 cm/s correlate with a high risk of stroke. Development of first stroke in children at risk, who are identified by elevated cerebral Doppler blood flow velocity, can be prevented effectively through regular transfusions. A randomized controlled trial comparing the effectiveness of hydroxycarbamide as an alternative to transfusion for primary stroke prevention is ongoing. In patients who have already had a stroke, a prospective randomized controlled trial comparing transfusion to hydroxycarbamide for secondary stroke prevention was undertaken, but was terminated early due to a high prevalence of events in the hydroxycarbamide arm. Patients treated with transfusions had fewer neurologic as well as non-neurologic


(a)


(b)


(c)


(d)

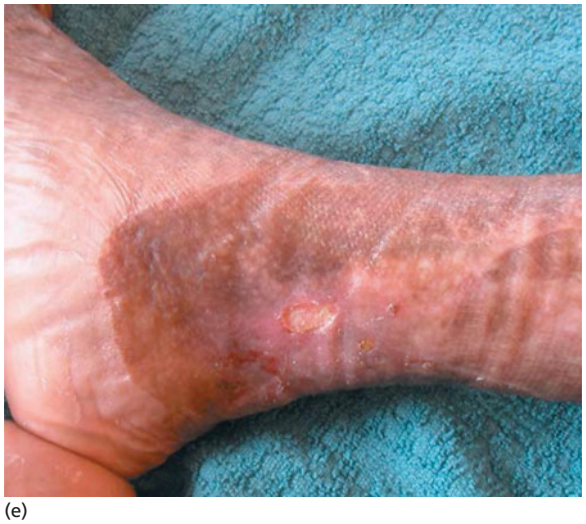

(e)

Figure 7.4 Complications of sickle cell disease: (a) stroke; (b) moyamoya transformation; (c) acute chest syndrome; (d) avascular necrosis of hip joint; (e) chronic leg ulcer.

sickle-cell-related serious adverse events. Though hydroxycarbamide has been shown to lower the incidence of pain and acute chest syndrome, its efficacy in the prevention and progression of cerebral vascular disease is less effective. Blood transfusions are therefore still considered the standard of care for both primary and secondary stroke prevention.

Pulmonary complications

Acute and chronic pulmonary complications are the leading cause of death in older patients. The acute chest syndrome is characterized by hypoxia, tachypnoea, fever, chest pain and pulmonary infiltrate on chest radiography (Figure 7.4c). Acute chest syndrome often follows a painful event, particularly in adults (Table 7.3). The pathogenesis of acute chest syndrome involves vaso-occlusion, infection or embolization of bone marrow fat. Infections due to *Mycoplasma*, *Chlamydia*, *Legionella*, pneumococcus, *H. influenzae* and viruses are more likely in children. Fat-laden pulmonary macrophages in the airways due to fat embolization from the bone marrow are present in half of the cases. Hypoxia due to acute chest syndrome can result in widespread sickling and vaso-occlusion, with risk of multiorgan failure. Patients should receive supplemental oxygen, incentive spirometry and antibiotic therapy directed towards the common organisms. Most patients have a bronchoreactive component and should receive bronchodilator therapy. Patients require close monitoring for persistent hypoxia or worsening lung consolidation for which blood transfusion should be provided. Exchange transfusion and mechanical ventilation is sometimes needed in rapidly progressive cases. NO and steroids may be beneficial in life-threatening cases.

Chronic pulmonary problems seen in SCD are obstructive and restrictive lung disease, pulmonary hypertension and hypoxia. Asthma, a form of obstructive lung disease, has a multifactorial pathophysiology in SCD and is influenced by genetic factors, environmental factors, the presence of a chronic underlying inflammatory state, and haemolysis-induced arginine–nitric-oxide dysregulation. Asthma has been shown to be an independent predictor of morbidity and mortality and may be missed during a physical examination. Screening for asthma with formal pulmonary function testing should be considered. Pulmonary hypertension is another independent predictor of mortality and morbidity in SCD. The diagnosis of pulmonary hypertension is associated with a 10-fold increase in relative risk of dying compared with patients who have normal pulmonary artery systolic pressure. Echocardiography can detect an elevated tricuspid regurgitant velocity (>2.5 m/s), which is observed in 20% of paediatric patients. Right heart catheterization is necessary to confirm the diagnosis (pulmonary artery pressure >25 mmHg). The prevalence of right heart catheter-confirmed pulmonary hypertension is approximately 10%. Pulmonary hypertension is more frequent among patients with high rates of chronic haemolysis, reflected by marked elevation in plasma lactate dehydrogenase (LDH). Hydroxycarbamide therapy should be strongly considered in SCD patients with pulmonary hypertension. Severely affected patients have been treated with regular transfusions, anticoagulation and oxygen inhalation. Current data do not support the use of targeted pulmonary arterial hypertension pharmaceuticals. Lastly, hypoxia, both daytime and nocturnal, has been associated with the development of pain and acute chest syndrome and can be prevented with effective treatment.

Table 7.3 Presenting symptoms of acute chest syndrome.

Symptom	Children (%)	Adults (%)
Fever	86	70
Shortness of breath	31	58
Chest pain	27	55
Extremity pain	22	58
Rib pain	14	30

Adults are more likely than children to have pain preceding the onset of pulmonary symptoms.

Hepatobiliary complications

The liver can be affected by hepatic sequestration, intrahepatic cholestasis, transfusion-acquired infection and transfusional haemosiderosis. Episodes of cholestasis due to intrahepatic sickling can lead to liver failure in rare instances. Pigmented gallstones are seen in two-thirds of patients, particularly those with HbSS, and can occur in young children. Patients with abdominal symptoms attributable to gallstones should undergo cholecystectomy, although the management of asymptomatic gallstones is less clear. Laparoscopic cholecystectomy can be safely performed, but associated common duct bile stones first require endoscopic retrograde cholangiopancreatography. As patients with sickle cell disease are living longer into adulthood, cirrhosis and liver failure are becoming increasingly prevalent disease-related sequelae. Vigilance for hepatic dysfunction, aggressive iron chelation in transfused patients, and prompt recognition and treatment of hepatobiliary disease in children may prevent development of hepatic end-organ dysfunction in adults.

Pregnancy

The steady-state haemoglobin level falls in SCD during pregnancy, similar to the decline in haemoglobin observed in normal pregnant women. Folate deficiency can exacerbate the anaemia and supplements should be provided throughout pregnancy. Painful episodes become more common in the last trimester. The incidence of pre-eclampsia is higher than normal in SCD

patients and there is a slight increase in maternal mortality. Risk to the fetus from abortion, stillbirth, low birth weight and neonatal death is also increased. Prophylactic transfusions during pregnancy or the type of delivery do not alter the outcome for mother or newborn. It is safe to use oral contraceptives for birth control in SCD.

Renal complications

The hypoxic, acidotic and hypertonic renal medulla favours vaso-occlusion, leading to destruction of the *vasa recta* and hyposthenuria in the first year of life. It presents clinically as enuresis or nocturia, and patients are susceptible to dehydration in hot weather. Haematuria as a result of papillary necrosis usually originates from the left kidney. Management is generally by bed rest and hydration, although sometimes blood transfusion and ε-aminocaproic acid are required. The prevalence of essential hypertension in SCD is lower than in the general population, although elevated systolic blood pressure is a risk factor for stroke. Proteinuria due to glomerular injury precedes development of nephrotic syndrome and eventual chronic renal insufficiency in one-quarter of adults. The progression to renal failure can be delayed by angiotensin-converting enzyme inhibitors. Careful control of blood pressure, avoidance of non-steroidal anti-inflammatory drugs (NSAIDs) and aggressive treatment of urinary tract infection and anaemia are important objectives for patients with chronic renal insufficiency. Patients with end-stage renal disease are treated with dialysis and renal transplantation. Some renal complications, such as hyposthenuria and haematuria, are also observed in individuals with sickle trait, as is the rare renal medullary carcinoma.

Priapism

Priapism occurs in two-thirds of males with SCD, with a peak incidence in the second and third decades. It is caused by vaso-occlusion leading to obstruction of venous drainage from the penis. It typically affects the *corpora cavernosa* alone, resulting in a hard penis with a soft glans. Episodes can be brief (stuttering) or prolonged, when they last for longer than 3 hours. Recurrent priapism leads to fibrosis and eventual impotence. Young boys require explanation of symptoms and the need to seek early help for priapism. At the onset of priapism, patients should drink extra fluids and attempt to urinate. An oral dose of pseudoephedrine or terbutaline can be given. Persistent priapism requires intravenous hydration and opioid analgesia. If priapism persists for more than 2–3 hours, aspiration and irrigation of the corpora with dilute phenylephrine or etilefrine solution should be performed. A simple intracavernosal injection with these agents tried early may induce detumescence and avoid the need for aspiration and irrigation. Sildanefil has been proposed as a therapeutic agent for the prevention of recurrent ischaemic priapism, but data on its efficacy is lacking.

Ocular complications

Vaso-occlusion of retinal and other vascular beds in the eye can lead to grave complications. Patients with SCD can develop abnormal (comma-shaped) conjunctival vessels, iris atrophy, retinal pigmentary changes and retinal haemorrhages. However, much more serious is neovascularization causing proliferative retinopathy, appearing as a 'sea fan', with its potential for vitreous haemorrhage and retinal detachment. Such patients are treated with laser photocoagulation or vitrectomy. The incidence of proliferative changes is substantially higher in HbSC and HbS/β^+-thalassaemia patients than in HbSS. All patients with SCD should have annual ophthalmological evaluation, beginning in the second decade.

Sudden change in vision in a patient with SCD is an ocular emergency. Central retinal artery occlusion requires immediate treatment with hyperoxygenation and reduction of intraocular pressure, but the prognosis for vision is poor. Hyphaema, which can arise after minor trauma, leads to glaucoma due to sickling of blood in the anterior chamber. The elevated intraocular pressure causes ischaemic optic atrophy and retinal artery occlusion. Individuals with sickle trait are also vulnerable to this complication. Urgent surgical attention is required to wash out blood from the anterior chamber.

Bone complications

The chronic haemolytic process results in expansion of the medullary space. Bone infarction due to vaso-occlusion produces tenderness, warmth and swelling, which can be difficult to distinguish from osteomyelitis. In such cases, cultures from blood and direct aspiration are negative and radiography later shows patchy sclerosis and cortical thickening. Collapse of vertebral end plates due to infarction produces the codfish appearance. Patients are managed with analgesia and hydration until resolution of symptoms.

Avascular necrosis of the femoral head is a serious complication that is difficult to treat and leads to chronic disability and pain (Figure 7.4d). Patients with coexisting α-thalassaemia have a higher incidence of osteonecrosis at a younger age. The condition also affects the humeral head, but with fewer functional consequences. Aggressive physical therapy or combination of early core decompression followed by physical therapy can postpone the need for additional surgical intervention. Hip arthroplasty may be required for patients with severe symptoms.

Leg ulcers

Chronic leg ulcers are frequent in adult patients with SCD, particularly affecting males with the HbSS genotype (Figure 7.4e). Ulcers arise near the medial or lateral malleolus and may be single or multiple. Occlusion of skin microvasculature from sickle red cells predisposes to ulcers, which are made worse by trauma,

infection or warm climate. Ulcers are always colonized with pathogenic bacteria (*Pseudomonas aeruginosa*, *S. aureus* and *Streptococcus* spp.) and acute infection can occur. The ulcers are painful and resistant to healing, and although bed rest and elevation of the leg are efficacious, they may not be practical, owing to the chronic nature of the problem. Treatment requires debridement, elastic dressings, zinc sulfate and, in some cases, red cell transfusions and skin grafting.

Variant sickle cell syndromes

Sickle cell trait

Sickle cell trait (HbAS) is a benign condition that has no haematological manifestations and is associated with normal growth and life expectancy. The ratio of HbA to HbS is 60:40, owing to the greater affinity of α-globin chains for β^A-globin chains. Sickle cell trait affects 8–10% of African-Americans and up to 25–30% of the population in West Africa. Sickle trait reduces the risk of severe falciparum malaria, but not the prevalence of parasitaemia. There appears to be no effect on infections with other forms of malaria. Impaired urine-concentrating ability and haematuria can occur, and an increased incidence of urinary tract infection is observed in pregnant women with sickle cell trait. Splenic infarction is possible at very high altitudes. A slight risk of sudden death during exercise has been reported. Sickling and vaso-occlusion under extreme circumstances can lead to rhabdomyolysis, acute renal failure and cardiac arrhythmias. Like all people, individuals with sickle cell trait are advised to gradually increase exercise intensity, to avoid dehydration and to stop physical activity with the onset of muscle cramp or fatigue. Treatment of sudden collapse consists of rapid intravenous hydration and oxygen supplementation. Genetic counselling should be provided to individuals with sickle cell trait.

HbSC disease

HbC is found among individuals of African descent and the compound heterozygous state HbSC accounts for 25–50% of patients with SCD. The vaso-occlusive complications seen in patients with HbSC resemble those seen in patients with HbSS, but are less severe. Splenomegaly and the risk of sequestration can persist into adult life. Of particular note is the higher incidence of proliferative retinopathy in HbSC beginning in the second decade. The haemoglobin level (100–120 g/L) is higher than in HbSS, and the red cells are relatively microcytic with a higher mean corpuscular haemoglobin concentration (MCHC). Peripheral blood smear reveals frequent target cells, intraerythrocytic crystals and rare sickle cells. Equal amounts of HbS and HbC are present in the red cells and the solubility test for sickle haemoglobin is positive. The electrophoretic appearance of HbSC, HbSE and HbSO Arab at pH 8.4 is similar, but a distinction can be made based on ethnicity and by performing isoelectric focusing or agar gel electrophoresis at pH 6.5.

Sickle cell/β-thalassaemia

Sickle cell/β-thalassaemia compound heterozygotes account for less than 10% of patients with sickle syndromes. The majority of these patients have the β^+-phenotype, with the proportion of HbA ranging from 3 to 25%. The clinical phenotype is mild and disease severity correlates with the amount of HbA present. The clinical manifestations of the less frequent HbS/β^0-genotype are similar in severity to those of HbSS. The red cells are microcytic and hypochromic, and variable numbers of target cells and sickle cells are observed. Reticulocytosis (10–20%) is present and the level of HbA_2 is elevated.

Sickle cell anaemia with coexistent α-thalassaemia

Coinheritance of αα-thalassaemia (−α/αα or −α/−α) with SCD is common, and such patients have less severe anaemia and demonstrate hypochromia and microcytosis. In general, the clinical severity is similar to that seen in HbSS patients with a normal complement of α-globin genes.

Sickle cell/HPFH

Approximately 1 in 100 patients with HbSS has an elevated HbF level due to deletional or non-deletional mutations that maintain γ-globin gene expression after birth. Such individuals have 20–30% HbF and less than 2.5% HbA_2. The haemoglobin level is normal, with microcytosis, and target cells are observed in the peripheral smear. The clinical course is benign, and vaso-occlusive complications are rare because of the inhibition of sickling by elevated HbF.

Other sickling syndromes

Sickle cell/Hb Lepore disease

Coinheritance of Hb Lepore with sickle cell mutation produces a clinical picture similar to that of HbS/β-thalassaemia, but with a low HbA_2 level.

Sickle cell/HbD disease

Of all the D or G haemoglobins, HbD Punjab (D Los Angeles) alone interacts with HbS to produce moderately severe haemolytic anaemia in compound heterozygotes. Target cells and irreversibly sickled cells (ISCs) are observed in the peripheral smear, and the clinical manifestations resemble mild sickle cell anaemia.

Sickle cell/HbO Arab disease

HbO Arab resembles HbC on alkaline electrophoresis and produces a moderately severe haemolytic anaemia in association with HbS. The disease is more severe than HbSC, and numerous sickled erythrocytes are observed on the peripheral smear.

Sickle cell/HbE disease

HbSE disease causes mild haemolysis and no remarkable abnormality of red blood cell morphology. HbE comprises only 30% of the total haemoglobin because of the thalassaemic nature of the mutation. Patients are generally asymptomatic, although occasionally significant vaso-occlusive complications and anaemia have been observed.

Diagnosis

Peripheral blood findings

The peripheral blood picture depends on the type of sickle cell syndrome. The haemoglobin level is normal in the newborn period, but anaemia develops and sickle or cigar-shaped ISCs can be observed in the peripheral blood by 3–4 months of age as HbF declines. In HbSS disease, the red cells are normocytic and normochromic, with polychromasia, many ISCs and fewer target cells (Figure 7.5a). The average reticulocyte count is 10% (4–20%) and normoblasts may be observed. Red cells are microcytic in the presence of coexisting α-thalassaemia or iron deficiency. In HbS/β-thalassaemia, ISCs, target cells and hypochromic microcytic red cells are prominent. The red cell morphology in HbSC disease is characterized by predominant target cells and rare ISCs. The occasional Howell–Jolly body, indicative of loss of splenic function in SCD, may be observed. The white cell count is elevated ($12–20 \times 10^9$/L) as a result of an increase in mature neutrophils. The platelet count is also elevated to $300–500 \times 10^9$/L as a result of decreased splenic function.

Other laboratory tests

The measurement of clotting factors in SCD indicates mild ongoing activation of the coagulation system, even in the steady state. The erythrocyte sedimentation rate is consistently low. Serum levels of unconjugated bilirubin and LDH are elevated, and haptoglobin is decreased.

Haemoglobin electrophoresis

HbS can be identified by cellulose acetate electrophoresis at pH 8.4 (Table 7.1 and Figure 7.5b). HbD and HbG have the same electrophoretic mobility with this method, but can be distinguished using citrate agar electrophoresis at pH 6.2 or thin-layer isoelectric focusing. Distinction cannot be made between HbSS and HbS/β^0-thalassaemia on electrophoresis. The diagnosis of HbS/β^0-thalassaemia is suggested by microcytosis and elevated HbA_2, and confirmed by finding β-thalassaemia trait in one of the parents. HbA and HbS are observed upon electrophoresis in both sickle cell trait and HbS/β^+-thalassaemia; however, the HbA fraction is greater than 50% in the former, but ranges from 5 to 30% in the latter. The level of HbF is variably elevated, with higher levels observed in patients with the Arab–Indian and Senegal haplotypes.

Other tests to detect sickle haemoglobin

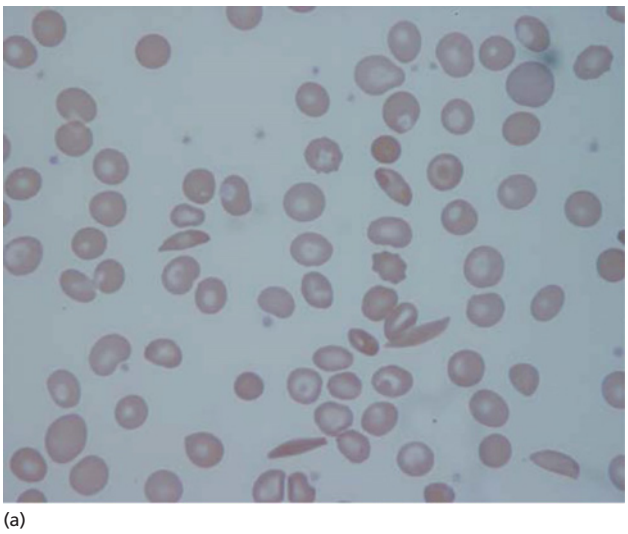
Sickling of red cells can be induced by sealing a drop of blood under a coverslip to exclude oxygen or by adding 2% sodium metabisulfite. The solubility test for HbS utilizes a reducing agent such as sodium dithionite, which is added to the haemolysate. Deoxy-HbS is insoluble and renders the solution turbid (Figure 7.5c). Both these tests are unable to distinguish sickle cell trait from sickle cell anaemia and cannot be used for primary diagnosis. They are useful aids in the identification of an abnormal electrophoretic band as HbS and for identifying sickle cell trait in units of red cells prior to transfusion. High-performance liquid chromatography (HPLC) can be used instead of electrophoresis to identify and quantitate HBS and other haemoglobins.

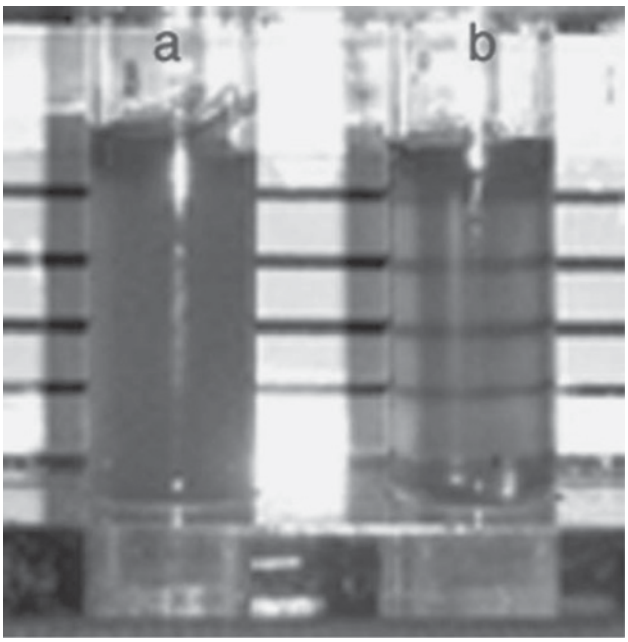
Newborn screening

Universal newborn screening is recommended for identifying SCD in the neonatal period. The efficacy of penicillin prophylaxis in preventing death from early sepsis in SCD provided the rationale for development of screening programmes. Blood samples obtained by heel prick are spotted onto filter paper and tested by electrophoresis or chromatography. Neonates with HbSS disease and HbS/β^0-thalassaemia have an FS pattern (the order of haemoglobins indicates their relative abundance in the sample). In sickle cell trait, the haemoglobin pattern is FAS, whereas newborns with HbS/β^+-thalassaemia have an FSA pattern. Finally, the presence of the FSC pattern suggests HbSC disease. Family studies help to make the definitive diagnosis and, when both parents are unavailable, DNA-based testing is useful.

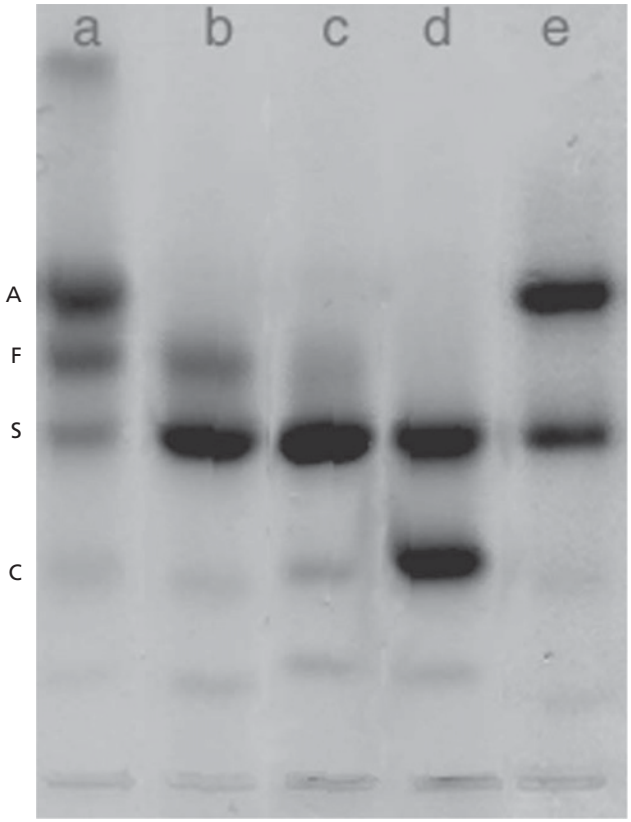
Prenatal diagnosis

Prenatal diagnosis is available through direct detection of the GAG→GTG mutation responsible for SCD in fetal cells. Genetic counselling is difficult owing to the marked variability in clinical manifestations within the same genotype, and the lack of ability at present to predict individual phenotype. Preimplantation diagnosis and selection of healthy embryos may offer a solution to this ethical problem.


(a)



(c)



(b)

Figure 7.5 (a) Peripheral blood smear from an individual with sickle cell anaemia. (b) Haemoglobin electrophoresis showing standard (lane a), HbSS (lanes b and c), HbSC (lane d) and sickle trait (lane e). (c) Sickle solubility assay is positive (tube a) in all three conditions.

Therapy

This section discusses general issues in the management of sickle cell disease. The treatment of specific complications is addressed in the section Clinical manifestations.

Routine healthcare

The majority of children with SCD can be managed by paediatricians or community physicians in coordination with a haematologist. Adults with SCD should also continue to have routine office visits. Patients who suffer from more severe complications or who need therapy to modify the course of SCD require specialized care at experienced centres.

The level of healthcare available to patients with SCD varies tremendously in different countries. Where resources are limited, the primary focus should be on penicillin prophylaxis, vaccination, education and analgesia for painful episodes. Where comprehensive care is available, both medical and psychosocial needs should receive attention (Figure 7.6). Sickle cell centres should have specialists in several fields who are available to address complications that may affect different organs.

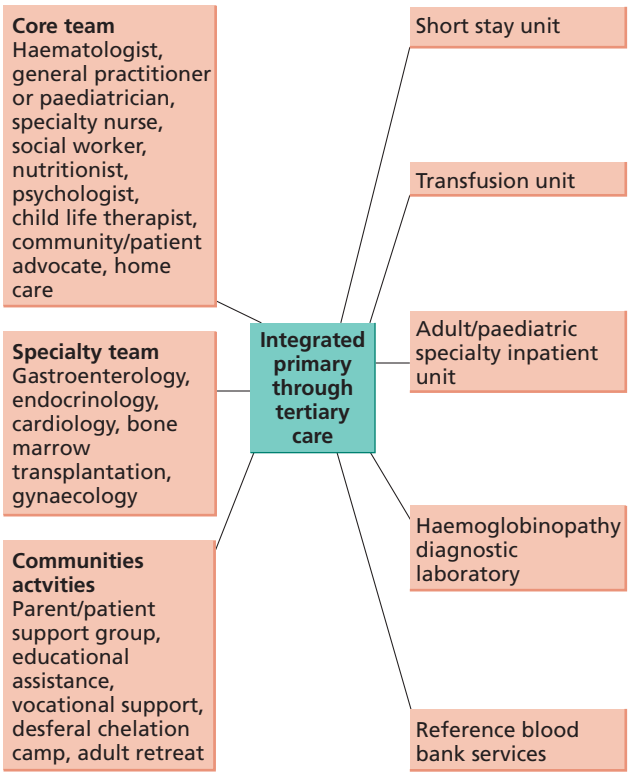


Figure 7.6 Comprehensive care of patients with sickle cell disease.

In cases when diagnosis is made on newborn screening, the infant should be seen within 1–2 months to instruct parents about infections and splenic enlargement. Routine immunization should include pneumococcal, *H. influenzae*, hepatitis B and influenza vaccines. All children receive prophylactic penicillin which may be stopped after the age of 5 years in the absence of any episode of pneumococcal sepsis or splenectomy. Folic acid supplementation (1 mg daily) is recommended. Evaluation of cerebral blood flow by transcranial Doppler should be performed on all children after 2 years to identify those at risk for stroke. Annual retinal examination is begun at 10 years. Sexually active women should have routine pelvic examinations and receive instructions about birth control.

Infections

Fever in children with SCD requires urgent attention in the office or emergency room. A complete blood count, blood and urine cultures and chest radiographs should be obtained, and lumbar puncture should be performed if meningitis is suspected. Very young children with fever or older children who appear septic should be hospitalized for intravenous antibiotics. The choice of antibiotics depends on causative agents prevalent locally and the pattern of resistance. In the USA, cefuroxime or ceftriaxone are preferred, whereas high-dose penicillin is used in several other countries. Many patients older than 2 years who do not look septic or seriously ill can be managed at home after receiving ceftriaxone in the emergency department. Antibiotics should continue for 1 week when bacteraemia is documented. In the presence of pneumonia, a macrolide should be added to cover *Mycoplasma* or *Chlamydia.* Antibiotics for osteomyelitis should provide coverage for *Salmonella* and *S. aureus* and are given for a period of 4–6 weeks.

Transfusion therapy

Blood transfusion in SCD is used to treat severe anaemia or to reduce the amount of circulating sickle haemoglobin. Only sickle-negative blood, which can be identified by negative sickle solubility test, is used for transfusions. The blood should also be leucodepleted, and matched for common minor E, C and Kell antigens. A simple transfusion is used to treat severe anaemia that is often associated with aplastic crisis and splenic sequestration. Older patients with renal failure may also need transfusions for declining haemoglobin level.

Dilution of circulating sickle haemoglobin can be accomplished by simple transfusion if the baseline haemoglobin level is low. Exchange transfusion is required to prevent hyperviscosity from the significant rise in haemoglobin when the patient has high baseline haemoglobin or when a greater reduction in HbS is desired. The final haemoglobin level should not exceed 120 g/L after simple or exchange transfusion. Conditions in which a reduction in the proportion of HbS is required include stroke, progressive acute chest syndrome, persistent priapism or preparation for general anaesthesia. Longer-term reduction in HbS through regular transfusions is advocated to prevent recurrence of stroke and sequestration, and in selected patients with leg ulcers or chronic pain. Routine blood transfusion is not needed for pain episodes, infections, minor surgery or uncomplicated pregnancy.

It is possible to eliminate most complications of SCD with the use of chronic transfusions to suppress endogenous sickle haemoglobin production. However, alloimmunization, iron overload and transmission of viruses are significant risks that limit the use of transfusions for the management of severe complications. In addition, because of the limited availability and decreased safety of blood, criteria for transfusion are more stringent in less developed countries. The high incidence of alloimmunization from minor blood group incompatibility (Rh, Kell, Duffy and Kidd) between donors and recipients can be avoided by use of phenotypically matched units. Genomic analysis of RH alleles can improve the precision of blood matching and decrease the incidence of alloimmunization. Patients on long-term transfusions develop iron overload, which requires chelation with deferoxamine, deferasirox or deferiprone. The prevalence of iron-induced organ damage appears to be lower in SCD compared with thalassaemia, despite similar amounts of iron burden. MRI and, rarely, liver biopsy are necessary to measure

iron burden because the serum ferritin is unreliable. Because iron accumulation can be reduced or prevented by erythrocytapheresis, this technique is now preferred when venous access is available.

Pain management

Prompt management of pain is essential, given its frequent occurrence and potential adverse psychological consequences. Patients with recurrent pain are best managed in a familiar ambulatory setting rather than the emergency department (ED). When ED management is necessary, utilization of pain management protocols can improve the delivery of care. Pain management protocols have been shown to expedite the administration of pain medications and facilitate more timely reassessment of pain. The patient should be evaluated for potential infectious, traumatic or surgical causes of pain. Pain assessment tools are available for young patients and are also helpful in older patients to follow the response to therapy. Adequate hydration should be provided, along with analgesia using narcotics and NSAIDs. Several narcotic agents are available for oral and parenteral use and the choice of medicine depends on local experience, as well as the patient's preference. The use of incentive spirometry reduces the potential for developing hypoxia and acute chest syndrome secondary to hypoventilation. Under-treatment of pain can be avoided by using patient-controlled analgesia, which has the added benefit of reducing apparent drug-seeking behaviour. Narcotic addiction is no more frequent in sickle cell patients than in others requiring analgesia. NSAIDs improve pain control with or without narcotics. Providing psychosocial support and reassurance, and allaying anxiety, are important goals. Chronic pain is rare in SCD and may require long-acting narcotics for management.

Hydroxycarbamide

Hydroxycarbamide is a tremendously important drug in the management of patients with SCD who have severe clinical manifestations. Hydroxycarbamide inhibits ribonucleotide reductase, leading to S-phase arrest of replicating cells, and is used in SCD because of its ability to stimulate production of HbF. Hydroxycarbamide increases HbF as a result of stress erythropoiesis induced by its myelosuppressive effect. Patients show variable response in the degree of rise in HbF, and some experience no change from the baseline value. Other biological effects of hydroxycarbamide play an equally significant role in the beneficial clinical effects observed during therapy. Erythrocytes of patients on hydroxycarbamide have increased water content and deformability and decreased adherence to vascular endothelium. There is elevation of the haemoglobin level, mean corpuscular volume (MCV), HbF and F cells, whereas total white cell and neutrophil count, reticulocyte count and the number of dense sickle cells decrease. Patients on hydroxycarbamide experience 50% reduction in the incidence of acute painful episodes and acute chest syndrome. Transfusion needs, frequency of hospital admissions and the risk of death are also decreased. Hydroxycarbamide does not improve leg ulcers, while the effect on priapism is unclear. Hydroxycarbamide may reduce risk of stroke, and its role in children with elevated cerebral blood flow velocities is under intense study.

The efficacy of hydroxycarbamide therapy is now well established in adults and there is growing evidence to support its use in children. Prospective trials in paediatrics demonstrated a reduction in both painful episodes and acute chest syndrome. In young patients with severe disease, stem cell transplantation should also be considered as an alternative to chronic transfusions or long-term hydroxycarbamide therapy. The overall benefit of hydroxycarbamide to the patient is closely related to compliance, for which monitoring in the clinic and psychological support should be provided. Patients should be made aware that 3–6 months may elapse before clinical benefits are realized. Hydroxycarbamide is offered to patients with frequent pain episodes or acute chest syndrome. It is started at a dose of 20 mg/kg per day and increased by 5 mg/kg every 2–3 months until the absolute neutrophil count is close to 2.5×10^9/L. Compliance with prescribed dose should be ascertained before each dose escalation, up to the maximum dose of 35 mg/kg or 2000 mg/day. Patients require frequent monitoring of blood counts, as well as renal and hepatic function. Myelosuppression is the most commonly encountered side-effect and temporary cessation of therapy and dose reduction is required for neutropenia, thrombocytopenia, reticulocytopenia or fall in haemoglobin. Dose modification is necessary for patients with renal failure. Skin pigmentation affecting the nails, palms and soles is commonly observed. There may be a temporary decline in sperm production among men while on hydroxycarbamide treatment. Despite concerns about the leukaemogenic and teratogenic effects of hydroxycarbamide, no convincing increase has been reported in SCD. Men and women should practise contraception while taking hydroxycarbamide, and women who become pregnant should stop the drug. No decrease in growth rate is observed in children over 5 years using hydroxycarbamide, while the growth effects of the drug in younger children are under evaluation. Overall, the risks associated with hydroxycarbamide therapy appear to be low and are certainly tolerable compared with the perils of untreated SCD in the severely affected patient.

New therapeutic modalities

A better appreciation of the pathophysiology of SCD will make it possible to exploit new therapeutic mechanisms (Table 7.4). Agents under development include membrane-active chemicals that improve hydration of sickle cells by blocking Gardos channels and potassium–chloride cotransport, or inhibit red cell adherence to endothelium. Decreased availability of NO has an important role in vaso-occlusion in SCD, and agents that correct

Table 7.4 Advances in the management of sickle cell disease.

Category	Intervention
Newborn screening	Counselling
	Comprehensive care
Infection	Prophylactic penicillin
	Immunization
Brain injury prevention	Screening with TCD, MRI
	Neurocognitive testing
	Chronic transfusions
Transfusion safety and iron overload prevention	Phenotypically matched red cells
	Erythrocytapheresis
	Iron chelation
Lung injury prevention	Incentive spirometry
	Antibiotics (including macrolides)
	Transfusion
	Echocardiographic screening for PHT
	Prevention with hydroxycarbamide
Surgery/anaesthesia	Preoperative transfusion
Avascular necrosis of the hip	Physical therapy
	Hip joint replacement
Priapism	Adrenergic agonist
	Antiandrogen therapy
Pain	Prevention with hydroxycarbamide
	Patient-controlled analgesic devices
	Non-steroidal anti-inflammatory drugs
Renal	ACE inhibitors for proteinuria
	Improved renal transplantation
Gallbladder disease	Laparoscopic cholecystectomy
Severe disease	Allogeneic bone marrow transplantation
	Chronic transfusions
	Hydroxycarbamide

ACE, angiotensin-converting enzyme; MRI, magnetic resonance imaging; PHT, pulmonary hypertension; TCD, transcranial Doppler.

NO deficiency may have significant therapeutic benefit. Newer agents to induce HbF synthesis that are being studied include orally effective butyrate compounds and decitabine, an analogue of 5-azacytidine. Decitabine can improve HbF levels in patients who do not respond to hydroxycarbamide.

Haemopoietic stem cell transplantation

Allogeneic haemopoietic stem cell transplantation (HSCT) from a matched sibling donor cures 85% of children with SCD less than 16 years of age. Both bone marrow and umbilical cord blood from related donors are suitable sources for the stem cells. Studies are underway to evaluate the use of unrelated donors or reduced-intensity conditioning regimens. However, 5% of patients die of complications related to HSCT and another 10% experience graft rejection with the return of SCD (Figure 7.7). Additional long-term risks after HSCT are infertility and second malignancy. Selection of candidates for HSCT is complex owing to the uncertain long-term course of the disease. HSCT should be considered in children (age <16 years) with SCD (any genotype) who have a human leucocyte antigen (HLA)-identical related donor and evidence of target organ damage involving the brain, lungs, kidneys or eyes. Children who are placed on long-term blood transfusions for any indication should also be evaluated for HSCT. Although it is clear that high-risk patients benefit from this treatment, the role of HSCT in asymptomatic children is not defined. Severe organ dysfunction increases the risks from the procedure, and hence discussion for HSCT should be started early when eligible indications are identified.

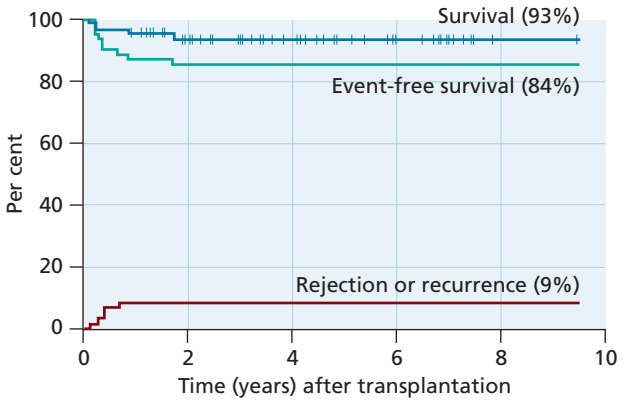


Figure 7.7 Outcome after transplantation for 59 children with advanced symptomatic sickle cell disease. Kaplan–Meier estimates for survival and event-free survival following marrow transplantation are shown. An event is defined as death, graft rejection or recurrence of sickle cell disease. A cumulative incidence curve for graft rejection and return of sickle cell disease is also depicted. (Source: Walters *et al.*, 2001 [*Biol Blood Marrow Transpl* **7**: 665–73]. Reproduced with permission of Elsevier.)

Gene therapy

Correction of SCD by gene therapy requires efficient insertion of a gene into repopulating haemopoietic cells and regulated expression in erythropoietic lineages. An antisickling haemoglobin, constituting 20–30% of the total haemoglobin, would be enough to produce clinical response. Mouse models of sickle cell disease have considerably helped in the effort to develop gene therapy, and correction of the sickling phenotype has been demonstrated in such animals.

Psychosocial issues

Recurrent pain and the unpredictable course of the illness place SCD patients at higher risk of depression and poor family

relationships. Despite this, with integrated medical care and social support most patients with SCD are well adjusted. Addiction to narcotics is an uncommon phenomenon and is the result of social influences rather than analgesic therapy. Attention to psychological wellbeing as well as educational and vocational support are important components of the care provided to SCD patients.

Selected bibliography

Adams RJ (2007) Big strokes in small persons. *Archives of Neurology* **64**: 1567–74.

Angelucci E, Matthes-Martin S, Baronciani D et al. (2014) Hematopoietic stem cell transplantation in thalassemia major and sickle cell disease: indications and management recommendations from an international expert panel. *Haematologica* **99**: 811–20.

Bhatia M, Walters MC (2008) Haematopoietic cell transplantation for thalassemia and sickle cell disease: past, present and future. *Bone Marrow Transplantation* **41**: 109–17.

Brousse V, Buffet P, Rees D (2014) The spleen and sickle cell disease: the sick(led) spleen. *British Journal of Haematology* **166**: 165–76.

DeBaun MR, Gordon M, McKinstry RC *et al.* (2014) Controlled trial of transfusion for silent cerebral infarcts in sickle cell anemia. *New England Journal of Medicine* **371**: 699–710.

Gladwin MT, Vichinsky E (2008) Pulmonary complications of sickle cell disease. *New England Journal of Medicine* **359**: 2254–65.

Hsieh MM, Fitzhugh CD, Weitzel RP, *et al.* (2014) Nonmyeloablative HLA-matched sibling allogeneic hematopoietic stem cell transplantation for severe sickle cell phenotype. *Journal of the American Medical Association* **312**: 48–56.

Sheehan VA, Luo Z, Flanagan JM, *et al.* (2013) Genetic modifiers of sickle cell anemia in the BABY HUG cohort: influence on laboratory and clinical phenotypes. *American Journal of Hematology* **88**: 571–6.

Strouse JJ, Lanzkron S, Beach MC et al. (2008) Hydroxyurea for sickle cell disease: a systematic review for efficacy and toxicity in children. *Pediatrics* **122**: 1332–42.

Vichinsky EP, Neumayr LD, Earles AN et al. (2000) Causes and outcomes of the acute chest syndrome in sickle cell disease. National Acute Chest Syndrome Study Group. *New England Journal of Medicine* **342**: 1855–65.

CHAPTER 8

Hereditary disorders of the red cell membrane and disorders of red cell metabolism

8

Paola Bianchi[1] and Narla Mohandas[2]

[1]Fondazione IRCCS Ca' Granda Ospedale Maggiore Policlinico Milano, Milan, Italy
[2]New York Blood Center, New York, NY, USA

Haemolysis

Definitions

Haemolysis indicates that the destruction of red cells is accelerated. Normally, in adults, the bone marrow output is well below its maximal capacity. Red cell production can be increased more than tenfold in the adult by increasing the cellularity of existing haemopoietic marrow, as well as by expansion of haemopoietic marrow into the long bones. In the newborn, and during infancy, marrow expansion depends on expanding the medullary cavity of bones, leading to thinning of cortical bone. These bony changes are most extreme in the β-thalassaemia syndromes, but some skeletal changes, usually some bossing of the frontal bones, may be seen in more extreme hereditary haemolytic anaemias of other causes. Increased red cell destruction is often completely matched by increased production, resulting in compensated haemolysis. When the rate of haemolysis exceeds the maximum erythropoietic capacity of the bone marrow, or when the latter is limited (e.g. because of inadequate supply of iron or folate or by ineffective erythropoiesis), the result is haemolytic anaemia.

General features of haemolysis

The clinical and laboratory aspects of haemolysis depend on the consequences of increased red cell destruction and production as well as the main process by which destruction takes place. Increased red cell destruction leads to an increase in unconjugated bilirubin from increased haemoglobin turnover. Unconjugated bilirubin does not appear in the urine, although there will be an increase in urinary urobilinogen. The bilirubin level is usually not more than two to three times normal because the normal liver is able to increase excretion to compensate for at least some of the increased production. Jaundice is usually mild in hereditary haemolytic anaemias although there are important exceptions.

In the neonate, particularly premature infants, liver function is not fully developed and more severe jaundice requiring urgent therapeutic intervention may occur. A rare but potentially confusing problem is the coinheritance of Gilbert syndrome, which comprises a group of congenital liver enzyme deficiencies that impair bilirubin conjugation. On its own, Gilbert syndrome does not produce clinical jaundice except when there is inadequate calorie intake, but in conjunction with haemolytic anaemia the hyperbilirubinaemia may be considerable. The increased bilirubin of haemolysis does increase the risk of gallstones and cholecystitis, which in turn may lead to an increase in serum bilirubin.

In the degradation of haemoglobin, the molecule is broken down to two αβ subunits, which are bound to haptoglobin, the complex being rapidly internalized in the hepatocyte after binding to the haptoglobin complex receptor. In the presence of haemolysis, serum haptoglobin levels are greatly reduced or absent. However, haptoglobin is an acute-phase protein and levels will increase in the presence of inflammation. Haemopexin

Postgraduate Haematology, Seventh Edition. Edited by A Victor Hoffbrand, Douglas R Higgs, David M Keeling and Atul B Mehta.

is another haem-binding protein produced by the liver that is decreased in haemolysis. Chronic haemolytic anaemia may increase the iron content of the body through increased iron absorption as a result of anaemia coupled to the retention of the haem iron following binding to haptoglobin and haemopexin. In rare cases of inherited haemolytic anaemia, this iron overload may be sufficient to produce clinically important effects, particularly if there is coinheritance of a haemochromatosis gene. In most haemolytic anaemias, owing to membrane defects, the destruction of red cells takes place extravascularly in the reticuloendothelial system, particularly in the spleen, and the iron is retained. When destruction is intravascular, free haemoglobin will be released into the plasma, producing haemoglobinaemia and methaemalbuminaemia, and will pass through the glomerulus to produce haemoglobinuria and haemosiderinuria. Iron deficiency is thus more likely than overload in intravascular haemolysis.

Increased red cell production leads to expansion of the red cell precursor compartment of the bone marrow, as described above. There are also changes in the structure of the marrow as a consequence of the chronic anaemia, which allows the early release of reticulocytes and, in more marked cases of haemolytic anaemia, nucleated red cells and even myelocytes. In the peripheral blood, the polychromasia and macrocytosis of reticulocytosis are the result of this increased throughput and release. The increased cell production requires an increased supply of folate which, at least theoretically, can produce folate deficiency unless supplements are given. It is usual to give folic acid (400 μg daily or 5 mg once weekly) to people with chronic haemolytic anaemia. The main features of haemolytic anaemia are summarized in Table 8.1.

Table 8.1 Main features of haemolytic anaemia.

Increased red cell destruction
Unconjugated hyperbilirubinaemia
Mild jaundice
Increased risk of gallstones
Increased urinary and faecal urobilinogen
Decreased serum haptoglobin and haemopexin
Extravascular changes
Increased iron stores
Splenomegaly
Intravascular changes
Haemoglobinaemia and haemoglobinuria
Haemosiderinuria
Methaemalbuminaemia
Decreased iron stores
Increased red cell production
Marrow expansion: bone changes
Increased erythropoiesis: ↓ myeloid/erythroid ratio
Reticulocytosis: polychromasia
Increased folate requirements: macrocytosis

Classification

Because of the unique structural and functional specialization of the mature red cell, the impact on it of a wide range of exogenous or endogenous changes is relatively uniform: the cell will be destroyed prematurely. According to the site of the primary change, haemolytic disorders have been traditionally classified as being due either to intracorpuscular or to extracorpuscular causes. According to the nature of the primary change, haemolytic disorders have also been classified as inherited or acquired. These two classifications correlate almost completely with each other, in that extracorpuscular causes are usually acquired, whereas intracorpuscular causes are usually inherited. One notable exception is paroxysmal nocturnal haemoglobinuria, a disease in which an intracorpuscular defect is acquired as a result of a somatic mutation (see Chapter 9).

Although in every cell all molecules and organelles are naturally interdependent, it is convenient to consider the red cell as a conveyance for a large amount of haemoglobin contained in a plasma membrane, the stability of which is maintained by appropriate metabolic machinery and structural organization of the proteins. Unfavourable genetic changes in any of these components may cause haemolysis. Accordingly, inherited haemolytic disorders can be classified into three major groups: (i) genetic disorders of haemoglobin (see Chapter 6); (ii) abnormal membrane (including the cytoskeleton) and (iii) abnormal metabolism (enzymopathies).

Red cell membrane disorders

The red cell membrane

The red cell membrane, like all other cell membranes, consists of a lipid bilayer that is stabilized and given specific properties by the proteins, glycolipids and other specialized molecules and structures with which it is associated. The lipid bilayer consists of approximately equal molar quantities of phospholipids and cholesterol molecules. The charged phosphatidyl groups of the phospholipids are hydrophilic and form the outer and inner surfaces of the bilayer. The interior of the membrane is formed by hydrophobic bonding of the acyl chains and cholesterol, which form the internal parts of the two leaflets (Figure 8.1). The arrangement is energy efficient but the two leaflets are not symmetrical. The outer leaflet consists mainly of phosphatidylcholine and sphingomyelin, the inner leaflet of phosphatidylethanolamine and phosphatidylserine (Figure 8.2). Maintenance of the asymmetry and the proper function of the membrane requires energy. In mature red cells this is provided

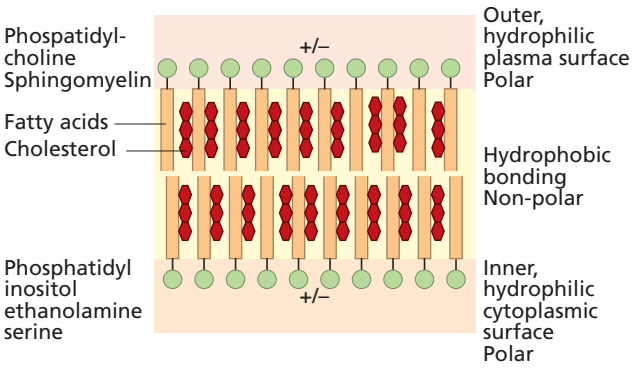


Figure 8.1 Arrangement of membrane lipids. The acyl chains of the diacylphosphatidylglycerides are hydrophobic non-polar domains and they form hydrophobic bonds with the acyl groups of the opposite layer. Cholesterol is present in roughly equimolar amounts and determines the fluidity of the membrane.

by adenosine triphosphate (ATP) from the glycolytic pathway and reducing power mainly in the form of glutathione.

The normal biconcave shape and function of the red cell membrane are determined by the membrane proteins and their interactions with the lipid bilayer and with each other. There are two main sorts of protein–membrane associations. The integral proteins have strong hydrophobic domains that associate with the hydrophobic part of the bilayer. Many of these integral proteins span the membrane and provide channels between the plasma and cytosolic compartments. The cytosolic inner domains of these proteins interact with each other and with the second main group, the proteins of the membrane skeleton. The integral proteins that provide the links between the lipid bilayer

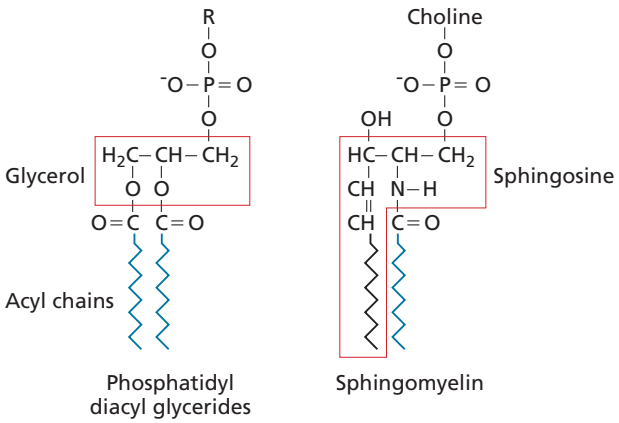


Figure 8.2 Main lipids of the red cell membrane. The outer, plasma, layer contains mostly neutral lipids, sphingomyelin and phosphatidylcholine (lecithin). The inner, cytoplasmic, layer contains mostly acidic groups, phosphatidylserine, phosphatidylethanolamine and phosphatidylinositol. R may be choline, serine, ethanolamine or inositol.

and the membrane skeleton have conveniently been referred to as 'vertical connections', whereas the proteins of the membrane skeleton that comprise the inner network of the cell membrane are characterized as 'horizontal connections'. Genetic abnormalities that produce spherocytosis mainly have mutations affecting the vertical connections. Mutations of the horizontal system usually produce elliptocytosis or more bizarre-shaped changes. The main proteins are listed in Table 8.2, and their arrangement is shown schematically in Figure 8.3.

In addition to the compartments mentioned so far, there are numerous surface proteins that provide the main interface with the plasma, including the blood group systems and other receptors. Many of these molecules are heavily glycosylated, as are the integral proteins, the glycophorins; sialic acid, which comprises the main side-chain of the glycophorins, contributes the most to the negative surface change of the erythrocyte. Some of these surface proteins are linked to the membrane by the glycosylphosphatidylinositol (GPI) anchor, which provides the hydrophobic domain required for association with the inner hydrophobic part of the membrane. Somatic mutations in the gene for phosphatidylinositol glycan A (*PIGA*) leads to a failure to produce the anchor and to paroxysmal nocturnal haemoglobinuria (discussed in Chapter 9).

The integral proteins and vertical interaction

The two major integral proteins that span the lipid bilayer are band 3 (the anion channel protein) and glycophorin C. Band 3 and associated molecules, 4.2 (pallidin) and ankyrin (2.1), form one major vertical interactive pathway with binding to the β-chain of the spectrin tetramer through ankyrin. Glycophorin C and protein 4.1 also provide a vertical interaction, but the association with spectrin is through a link with actin, which is a key part of the horizontal network. The band 3–4.2–ankyrin–spectrin complex is a central part of the organization of the lipid bilayer and loss of part of this complex leads to loss of membrane, thereby reducing the surface-area-to-volume ratio of the red cell and leading to the characteristic spherocytes of hereditary spherocytosis.

The main protein of the membrane skeleton is spectrin, consisting of two subunits, α and β, which associate side by side to produce a heterodimer. The dimers associate head to head to form tetramers about 200 nm long. The tail end of the dimer makes contact with a short actin filament composed of 14 monomers; the interaction between spectrin and actin is stabilized by protein 4.1, adducin and dematin. Binding of spectrin dimers to actin filaments produces the more or less hexagonal network of spectrin tetramers on the inner surface of the membrane associated with the lipid bilayer. Spectrin–actin–4.1 interactions provide much of the flexibility of the red cell membrane. Spectrin qualitative defects that affect these horizontal interactions tend to induce a loss of structural stability of the membrane and elliptocytosis.

Table 8.2 Proteins of the red cell membrane.

Band*	Protein	Gene location	Function	Associated haemolytic anaemias
1	α-Spectrin	*SPTA1*, 1q21	Membrane skeletal network	HE, HS
2	β-Spectrin	*SPTB*, 14q24.1–q24.2	Membrane skeletal network	HPP HE, HS
2.1	Ankyrin	*ANK1*, 8p21.1–11.2	Vertical contact	HS
2.9	Adducin	*ADD1* (α-chain), 4p16.3; *ADD2* (β-chain), 2p13.3	Promotes spectrin binding to actin, binds Ca^{2+}/calmodulin	(HS, HE in mice)
3	Band 3. Solute carrier family 4 (anion exchanger) member 1	*SLC4A1*, 17q12–q21	Anion exchange channel, binds glycolytic enzymes	HS, SAO, HAC
4.1	Protein 4.1	*EPB41*, 1p33–p32	Stabilizes spectrin–actin contact	HE
4.2	Protein 4.2 (pallidin)	*EPB42* (*PLDN*), 15q15–q21	Band 3–ankyrin complex	HS (Japan)
5	β-Actin	*ACTB*, 7p15–p12	Spectrin network junction	?
6	Ga3PD	*GAPDH*, 12p13.31	Links ATP production to membrane	?
PAS-1†	Glycophorin A	*GYPA*, 4q28–q31	MN blood groups	?
PAS-2	Glycophorin C	*GYPC*, 2q14–q21	Gerbich blood groups	HE
PAS-3	Glycophorin B	*GYPB*, 4q28–q31	Ss blood groups	?

*Band numbers refer to the position on SDS-PAGE electrophoresis.
†Periodic acid–Schiff stain: bands seen only on PAS-stained gels.
HAC, hereditary acanthocytosis; HE, hereditary elliptocytosis; HS, hereditary spherocytosis; HPP, hereditary pyropoikilocytosis; SAO, Southeast Asian ovalocytosis; Ga3PD, gyceraldehyde-3-phosphate dehydrogenase.

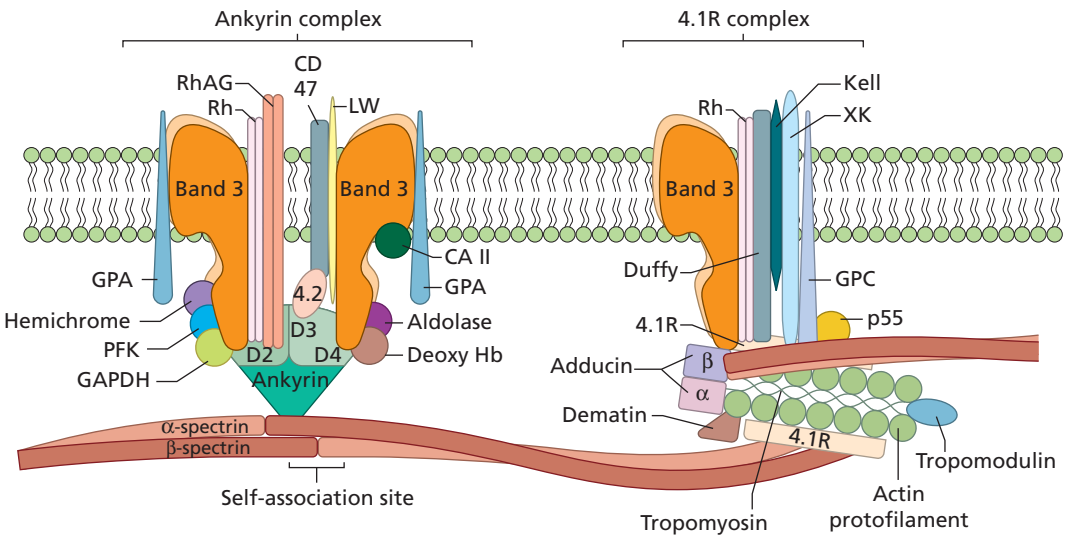


Figure 8.3 A schematic model of the structural organization of the red cell membrane. The membrane is a composite structure in which the lipid bilayer is linked to the spectrin-based membrane skeleton. The linking of the lipid bilayer to the membrane skeleton is mediated by band 3-ankyrin-protein 4.2– β-spectrin interactions and by glycophorin C–protein 4.1R interaction (vertical interactions). The horizontal interactions in the spectrin network involve spectrin dimer–dimer interaction and spectrin–actin–protein 4.1 interaction.

The clinical phenotypes of hereditary membrane disorders

Mutations in the genes that control the proteins of the membrane and their interaction mainly produce changes in the shape of red cells, which is characteristic in any individual. Many of the conditions are inherited as autosomal dominant disorders, homozygosity for major defects mainly being lethal. Severe, bizarre or unexpected red cell morphology is often produced by double heterozygosity or inheritance of more than one defect of the membrane proteins. Mutations affecting the red cell membrane are many and heterogeneous, but the effect on the phenotype can be classified in five main categories: (i) hereditary spherocytosis; (ii) hereditary elliptocytosis and hereditary pyropoikilocytosis (severe form of elliptocytosis); (iii) Southeast Asian ovalocytosis; (iv) hereditary acanthocytosis and (v) hereditary stomatocytosis.

Hereditary spherocytosis

As the name implies, hereditary spherocytosis (HS) is a genetically determined haemolytic anaemia characterized by the spherical shape of the affected red cells. The spherical shape produces a characteristic appearance in the stained blood film of round cells with smaller than normal diameter, which lack the area of central pallor of the normal biconcave discs (Figure 8.4). The disorder is generally inherited as a dominant condition with a wide spectrum of severity. The usual clinical picture is of mild to moderate haemolytic anaemia, but varies from severe neonatal haemolysis with kernicterus (rare) to clinically silent and asymptomatic (usual) haemolysis. Autosomal recessive inheritance occurs in a few mutations, often producing severe haemolysis. In white populations, HS is one of the most common haemolytic anaemias due to membrane defects, with a prevalence of clinically apparent disease of 200–300 per million population. The occurrence of clinically silent cases probably means that the overall prevalence is likely higher.

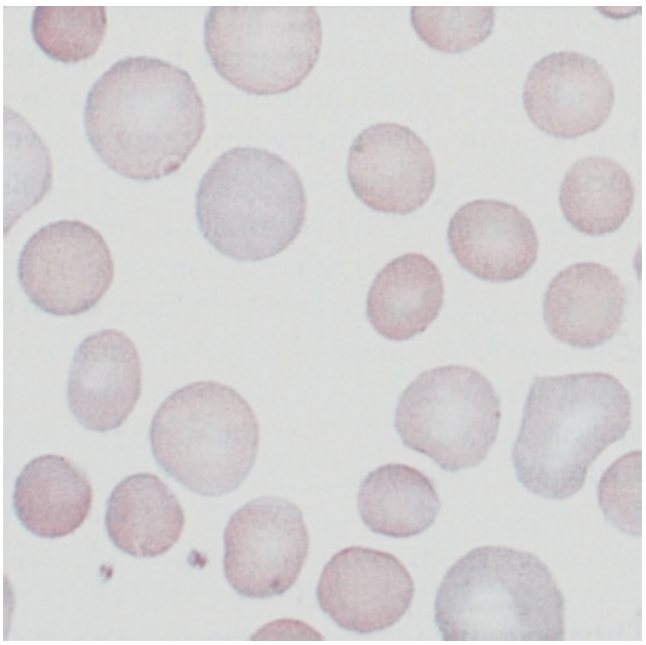


Figure 8.4 Hereditary spherocytosis, peripheral blood. Small spherocytic red cells lack area of central pallor. Large polychromatic red cells (reticulocytes) result in normal MCV, although MCHC may be increased.

Clinical features

The commonest forms of HS present as mild anaemia and jaundice, with a modestly enlarged spleen. However, the genetic heterogeneity of HS (see below) is reflected in the clinical presentation. As the main site of increased red cell destruction in HS is the spleen, it is not surprising that the size of the spleen tends to reflect the severity of the haemolysis, although splenomegaly is rarely marked, enlargement below the umbilicus being very uncommon. When HS presents in adolescence or adult life, it needs to be distinguished from other causes of spherocytosis, particularly warm autoimmune haemolytic anaemia.

HS may present at birth. The functions of the spleen become mature only after birth, so severe anaemia *in utero* is rare. Erythropoiesis is highly active before birth, but enters a phase of reduced activity in the neonatal period. Severe anaemia, developing over 5–30 days post delivery and requiring transfusion, may result from this double physiological development of reduced production and increased destruction, but the anaemia may greatly reduce during the first year of life as compensatory erythropoiesis develops. Decisions about splenectomy do not need to be taken during this time.

Molecular pathology

About 60% of HS cases result from a defect in the ankyrin–spectrin complex, with the genes for ankyrin (*ANK1*) and the α- and β-subunits of the spectrin dimer (*SPTA1*, *SPTB*) being implicated in different genetic types (Table 8.2). A further 25% involve deficiency in band 3, the anion channel. In the remainder of the dominantly inherited HS families there is a deficiency of protein 4.2 or no abnormality has yet been identified. Deficiency of protein 4.2 is particularly common in Japanese families with HS (Table 8.2). The rare cases of Rh null also present with spherocytosis and mild anaemia. These defects involving spectrin–ankyrin–band 3 interactions affect the vertical interactions.

Laboratory diagnosis

The typical findings of extravascular haemolysis are present in HS (Table 8.1). The diagnosis is usually made on the basis of red cell morphology, backed up where possible with a family history. The mean corpuscular haemoglobin concentration (MCHC) is often increased above 350 g/L, but the presence of macrocytic reticulocytes usually results in a low normal mean

corpuscular volume (MCV) rather than true microcytosis. These changes result not only from the reduction of the surface-area-to-volume ratio, but also from the slight dehydration of HS cells. A number of variants of the typical HS features have been described, usually the more severe forms that may have denser and less perfectly round cells in the peripheral blood. In infancy, the morphology may be more difficult to interpret. The effect of immature splenic function and the macrocytosis and anisocytosis of infancy combine with the HS phenotype to produce red cell appearances not typical of the developing HS. Family studies may assist in the diagnosis.

Osmotic fragility test

The osmotic fragility test, which measures the sensitivity of red cells to lysis *in vitro* to swelling caused by incubation in increasingly hypotonic saline solutions, is the most appropriate test to diagnose HS, although its sensitivity decreases in milder cases. In normal red cells with the biconcave disc shape, 50% lysis occurs when the saline solution reaches about 0.5% sodium chloride. The HS cells have less ability to swell and so lyse at higher salt concentrations, producing a right-shifted osmotic fragility curve. One of two patterns may be seen in HS: a generally right-shifted curve, which is the more common finding, and one where there appears to be a 'tail' of lysis-sensitive cells. Incubation of blood for 24 hours at 37 °C accentuates the fragility (Figure 8.5). An alternate to the osmotic fragility test is to measure the extent of membrane surface area loss by flow cytometry following labelling with eosin-5-maleimide (EMA binding test). EMA binds to band 3, which is selectively lost during loss of membrane surface area in red cells from HS patients. The EMA binding screening test has to be used in conjunction with morphology because Southeast Asian ovalocytosis, congenital dyserythropoietic anaemia type II and cryohydrocytosis also give reduced fluorescence because they can create a long-range modulation effect on the dye binding site in band 3 protein.

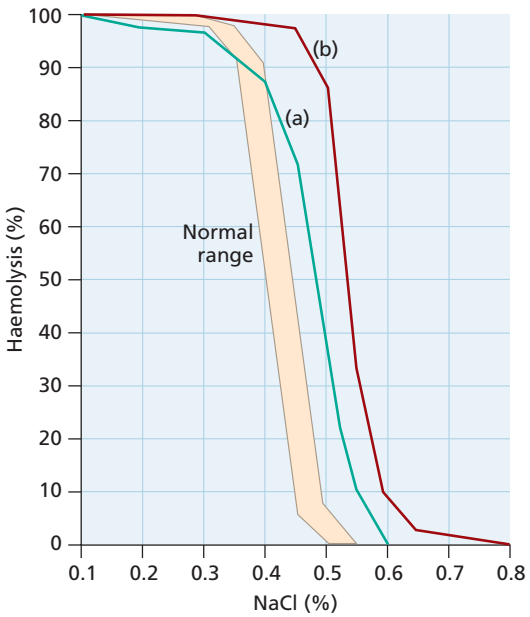


Figure 8.5 Osmotic fragility test in hereditary spherocytosis. Osmotic fragility is increased in the microspherocytes (right shift), but there is also a small population of resistant cells due to increased reticulocytes (a). After splenectomy, the microspherocytes remain, but the proportion of reticulocytes is reduced to normal values and the resistant cells are not seen (b).

Identification of protein abnormalities or gene defects

Methods that identify the defective gene or its product are the most specific for membrane defects, but are beyond the scope of most routine haematology laboratories. The original identification of membrane proteins using sodium dodecyl sulfate-solubilized polyacrylamide gel (SDS-PAGE) electrophoresis has led to the classification according to the banding system indicated in Table 8.2 The identification of specific genetic abnormalities may be important in compound haemolytic syndromes, but requires specialist laboratories.

Clinical course and complications

In most kindreds, the course of the disorder is similar in affected members although, as with most inherited defects, there is some variable penetrance and it is not rare to find a very mildly affected parent with more severely affected offspring. As with all congenital haemolytic anaemias, the anaemia may be aggravated by environmental factors. This may be consequent on an increase in the red cell destruction or a decrease in production. Increased jaundice may occur during viral infections or bacterial sepsis, the anaemia also being aggravated by a decrease in production consequent on the effects of the acute-phase response or the inhibition of erythropoiesis by interferon (IFN)-γ.

Primary infection with parvovirus 19 produces a specific and marked inhibition of erythropoiesis, often characterized as an aplastic crisis. In patients with shortened red cell survival, severe anaemia may be produced by the inhibition, which lasts for some 4–7 days. In normal individuals with a red cell lifespan of 120 days, such an inhibition produces no clinical effect. The anaemia associated with parvovirus infection in HS may require urgent transfusion. The diagnosis is made by finding absent parvovirus antibodies with subsequent appearance of IgM antibodies. The presence of IgG antibodies at the time of the anaemia excludes the diagnosis.

Acute anaemia due to splenic sequestration is a relatively uncommon complication of HS in childhood. The pathogenesis is probably increased splenic size and activity leading to increased trapping of HS cells within the spleen. This complication may also require urgent transfusion. Malnutrition may increase anaemia because of folate deficiency, but also from increased jaundice through the effect of low-calorie input on unconjugated bilirubin levels in the blood. The anaemia of pregnancy may aggravate a haemolytic anaemia and hence bring the

condition to the attention of clinicians and patients. Classical HS is not a risk to mother or child in pregnancy.

Gallstones are an expected complication in HS, as in other chronic haemolytic anaemias. Silent gallstones require no intervention. Recurrent cholecystitis or biliary colic may require cholecystectomy accompanied by splenectomy (see below). Leg ulcers are a rare but well-recognized complication of HS, as with other chronic haemolytic anaemias. Extramedullary haemopoietic masses, usually paravertebral, occur rarely in more severe HS.

Management

Patients with well-compensated haemolysis and no transfusion requirements need no treatment other than reassurance and folic acid supplements (e.g. 400 μg daily or 5 mg weekly). For people with a well-balanced and adequate diet, folic acid supplements are probably unnecessary, but custom dictates the practice should be continued. Radiolucent gallstones, if detected by chance on ultrasound, are common and need no treatment unless complications arise. Gallstones without recurrent inflammation are not a risk factor for carcinoma of the gallbladder. Recurrent cholecystitis or obstruction would be an indication for cholecystectomy, which would also be an indication for splenectomy.

Splenectomy

For the great majority of patients with autosomal dominant forms of HS, splenectomy increases the lifespan of the HS red cells to near normal and alleviates the haemolysis and hyperbilirubinaemia, since the spleen is responsible for the removal of spherocytic red cells with reduced surface area. However, splenectomy carries short- and long-term risks that must be weighed against the benefits in any individual patient. After splenectomy, the blood film continues to show spherocytosis together with common changes associated with splenectomy. The osmotic fragility remains increased.

Risks of splenectomy (see also Chapter 17)

The immediate risks associated with splenectomy include those of any abdominal operation together with an increased risk of thrombosis, associated with a marked rise in platelet count that occurs promptly after splenectomy. In HS, in which the erythropoietic drive returns to normal following splenectomy, the platelet count also returns to normal and the risk diminishes. In conditions where haemolysis persists, the platelet count remains elevated, sometimes markedly, and the increased risk of thrombosis continues.

The major hazard of splenectomy is the long-term susceptibility to severe infection, so-called overwhelming postsplenectomy infection (OPSI) (see also Chapter 17). The spleen plays an important role in filtering and phagocytozing bacteria, and removing parasitized erythrocytes from the blood. The spleen is the major source for mounting the rapid, specific IgM response to organisms that enter through the gut. The main organisms of this class are the encapsulated organisms, *Streptococcus pneumoniae*, *Haemophilus influenzae* type B and *Neisseria meningitidis*. Pneumococcal infection is responsible for about 70% of OPSI and has a 60% mortality. Lack of a spleen greatly increases the virulence of the infection, with progression from the first feeling of fever and non-specific flu-like symptoms to irreversible endotoxic shock occurring in a matter of hours. Patients may present with purpura, evidence of disseminated intravascular coagulation, multiorgan failure, hypotension and peripheral limb ischaemia. Diarrhoea and vomiting are common prodromes. It is this speed of progression that makes the prophylaxis of this fortunately uncommon complication so important. Prophylaxis depends on education and awareness for the patient, specific measures to reduce the risk from particular organisms and the provision of information concerning the splenectomy for healthcare workers (Table 8.3). There is no direct evidence that phenoxymethylpenicillin (e.g. 250 mg twice daily) reduces the risk of OPSI in splenectomized patients, but good evidence that it does so in homozygous sickle cell patients who have functionally inactive spleens. It is on this evidence that such antibiotic prophylaxis (or erythromycin 250 mg b.d. for those sensitive to penicillin) is recommended (see also Chapter 17). The actual incidence of OPSI is difficult to calculate. The overall risk has been stated as 0.04 per 100 patient-years for patients without added immunosuppression, but considerably higher for those immunocompromised by malignancy or chemotherapy. The risk is greatest in the first two years after splenectomy, but continues lifelong. Children under the age of 5 years are particularly susceptible and splenectomy should be avoided in this group if at all possible.

Indications for splenectomy

Patients with marked haemolysis producing symptoms or requiring transfusion should be splenectomized, although preferably not before the age of 5 years (later if possible). Recurrent aplastic crises are also an indication. Attacks of cholecystitis or biliary colic warrant cholecystectomy and splenectomy, but symptomless gallstones are not a necessary indication. In small children with severe anaemia, partial splenectomy has been used to decrease the extent of anaemia, but the performance of this procedure requires special surgical skills and there is need for careful long-term follow-up.

Hereditary elliptocytosis

Deficiency of spectrin tetramers, the horizontal links of the cytoskeleton, produces a wide spectrum of disease from fully compensated haemolysis with mildly elliptocytic red cells to severe and life-threatening anaemia with grossly distorted cells. When the morphological characteristic is a relatively uniform elliptical shape, the condition is referred to as hereditary elliptocytosis (HE). Haemolytic anaemia associated with the more distorted forms, which are also heat labile, was previously called

Table 8.3 Guidelines for prevention and management of infection in the splenectomized patient.

It is essential to educate patients regarding the risk and the importance of prompt recognition and treatment of infections. Leaflet card for patients to alert health professionals to risk of OPSI.
All patients should receive polyvalent pneumococcal, *Haemophilus influenzae* type b and meningococcal vaccination.
Vaccines should ideally be administrated 2 weeks prior to splenectomy or 2 weeks after splenectomy.
Pneumococcal immunization
- Infants <2 yrs*: should be immunized with three doses at the national schedule with polivalent-PCV vaccination. One dose of PPV at 2 yrs should also be offered.
- Children 2–5 yrs: one dose of PPV or PCV13 followed by PPV depending on previous routine childhood vaccinations.
- Children >5 and adults: irrespective of immunization status should continue to receive one dose of PPV by measuring antibody response.
- Responders should be revaccinated with PPV at 5-yearly intervals.

Haemophilis influenzae b (Hib) vaccination
- Infants <2 yrs: should complete their vaccination according to the national schedule.
- Children ≥2 yrs and adults: should be offered one dose of Hib-containing vaccine irrespective of their previous immunization status.

Meningococcal vaccination (Men)
- The quadrivalent MenACWY conjugate vaccine is recommended in preference to the plain polysaccharide meningococcal vaccine for all age groups.
- Infants <2 yrs, if immunized or partially immunized: 3 and 4 mo of age; MenC conjugate vaccine at 12 mo of age as well as quadrivalent MenACWY no earlier than 1 month later. At 2 yrs an additional booster of Hib/MenC should be given at second birthday.
- Children ≥2 yrs and adults: should receive one dose of MenC conjugate vaccine plus a single dose of quadrivalent MenACWY conjugate vaccine one month later.
- Travellers to endemic areas should receive quadrivalent MenACWY coinjugate vaccine before travelling.

All patients should receive yearly influenza vaccination.
Lifelong prophylactic antibiotics (oral phenoxymethylpenicillin or an alternative).
Patients developing infection despite measures should receive systemic antibiotics and be admitted urgently to hospital.
Awareness of risks of malaria and scrupulous prophylaxis if at risk.

Source for further details: Davies et al., *Brit J Haemat* **155** 308–17, on behalf of the Working Party of the Haemato-Oncology Task Force of the British Committee for Standards in Haematology (2011).
*Age at which splenectomy is performed.
PCV, vaccination pneumococcal conjugate vaccines; PPV, pneumococcal polysaccaride vaccine; PCV13, 13-valentPCV

hereditary pyropoikilocytosis (HPP). Recent evidence implies that HPP is a severe form of HE. Within a family, HE and HPP may both be present, the more severely affected individuals having both spectrin deficiency as well as a relative deficiency of spectrin tetramers. This may be caused by coinheritance of a low-expression allele for α-spectrin, compound heterozygosity for two HE alleles or HE homozygosity. A number of families have been described in which mutations involving the protein 4.1 gene result in failure to produce the protein. In heterozygotes with this variant, elliptocytosis occurs without haemolysis; in homozygotes, there is a severe haemolysis with extensive cell fragmentation.

Clinical features

As with HS, haemolytic anaemia in HE has a heterogeneous clinical presentation and molecular basis. The anaemia can range from very mild to severe with splenomegaly and few cases of hydrops fetalis have been reported. The clinical management of HE parallels that of HS.

Mild common hereditary elliptocytosis

Frequently, HE is discovered by chance from a blood film (Figure 8.6) or the presence of marginally raised bilirubin. Some affected people have no evidence of shortened red cell survival, whereas others have a well-compensated haemolytic anaemia. No treatment is required, although the blood film of partners should be examined if there is consanguinity, making homozygosity in offspring possible. For patients with mild haemolysis, anaemia may increase during infections, in pregnancy, with folate deficiency or with other conditions likely to enhance anaemia.

Silent carriers: low-expression genes

Mutations that produce low expression of α-spectrin may lead to no haematological abnormality because of the normal overexpression of α-spectrin in red cells compared with β-spectrin. However, when these defects are inherited in *trans*, on the other allele from an HE gene, HPP may result. Several mutations have been described, particularly commonly in codon 28, which

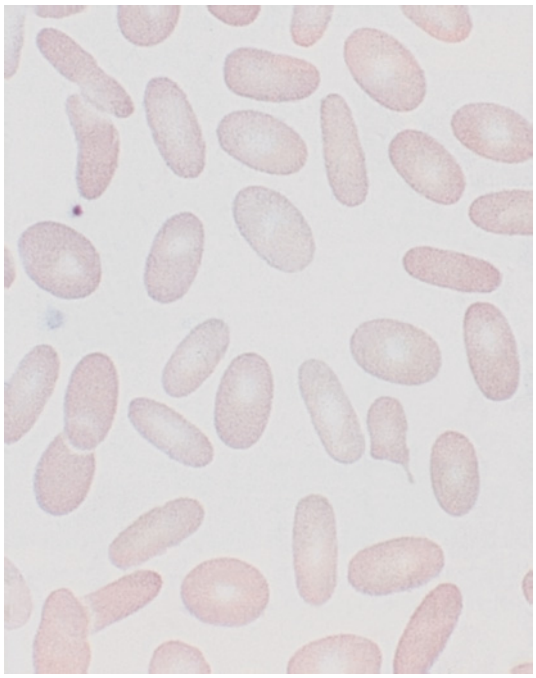


Figure 8.6 Hereditary elliptocytosis, peripheral blood. Characteristic elliptocytes of mild common hereditary elliptocytosis.

produces low-expression genes. A common polymorphism, intron 45C→T, is spectrin α^{LELY} (LELY denoting 'low-expression allele Lyon').

Haemolyic hereditary elliptocytosis

The characteristics of haemolytic HE previously termed HPP are densely contracted and fragmented cells (Figure 8.7), moderate to severe haemolysis and heterogeneity of manifestations within a family. In general, patients with haemolytic HE have spectrin deficiency in addition to the abnormalities of spectrin–spectrin contacts resulting in tetramer deficiency. One parent of an HE propositus may have normal haematology but carry a mutation in *trans*, which leads to spectrin deficiency. The affected cells show thermal lability and fragmentation at lower temperatures than normal, and erythrocytes are mechanically unstable.

Hereditary elliptocytosis and poikilocytosis in the neonate

In the neonate, the manifestations of HE may be a more marked poikilocytosis resembling HPP, with more fragmented red cells. These red cells are susceptible to fragmentation above 46 °C, whereas normal cells only fragment above 50 °C. The morphological changes and haemolysis gradually decrease over the first year until the typical picture of mild HE remains. Treatment of neonatal poikilocytosis is required only if the anaemia is such as to warrant transfusion.

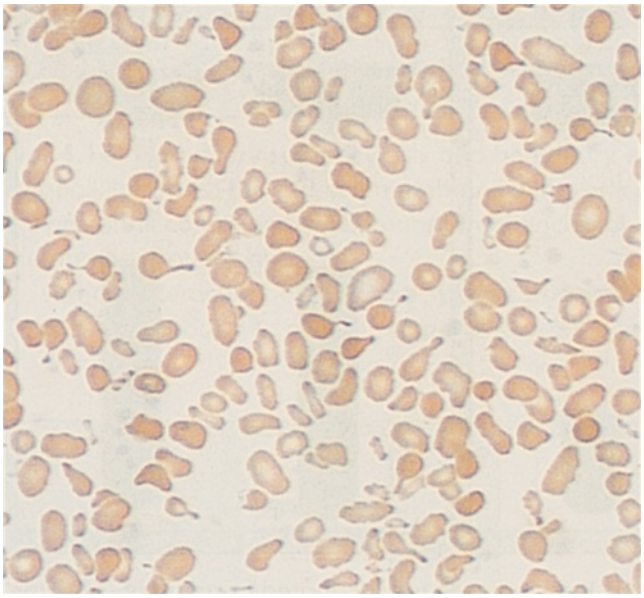


Figure 8.7 Hereditary pyropoikilocytosis, peripheral blood. Marked anisocytosis and poikilocytosis from a child with homozygous hereditary elliptocytosis.

Laboratory investigation

The standard approach to the diagnosis of HE is the identification of haemolysis, coupled to a careful examination of the blood film of the patient and as many first-degree relatives as possible. Examples of blood films are shown in Figures 8.6 and 8.7. Other acquired causes of elliptocytic or fragmented red cells need to be excluded, including iron, folate or vitamin B_{12} deficiency, and the microangiopathic haemolytic anaemias. Congenital dyserythropoietic anaemia and thalassaemia intermedia also need to be excluded.

As with the investigation of HS, appropriate gel electrophoresis assays may reveal protein abnormalities, although more specific identification requires a sophisticated approach beyond the abilities of most haematology laboratories.

Treatment

Patients with chronic haemolysis should be given folate supplements. Splenectomy is indicated for severe haemolytic anaemia in patients with haemolytic HE. Response in HE may not be complete, but the anaemia is usually markedly alleviated. As with HS, there may be an increased risk of increased thrombotic tendency. The precautions against OPSI are the same as for HS.

Hereditary stomatocytosis and related disorders

Stomatocytes are so called from the mouth-like slit or 'stoma' that appears on blood films (Figure 8.8). Stomatocytes are leaky

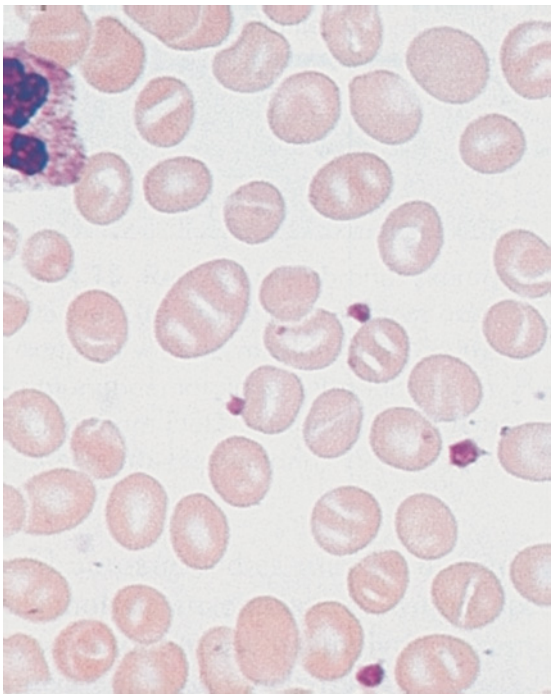


Figure 8.8 Hereditary stomatocytosis, peripheral blood.

to cations. There are other variations with Na^+ or K^+ leaks that are clinically similar to stomatocytosis without the obvious morphological changes. All these conditions are inherited in autosomal dominant fashion, and mostly produce moderate haemolytic anaemia (haemoglobin 100 g/L or above) and macrocytosis. There are two main variants: over-hydrated hereditary stomatocytosis, in which MCHC is low, and dehydrated hereditary stomatocytosis, with an increased MCHC.

The blood film may show stomatocytosis, but more commonly the film is unremarkable, apart from macrocytosis and polychromasia. The group is rare, estimates suggesting that 1 in 10 000 to 1 in 100 000 of the population are affected. However, associated features make the conditions important beyond their rarity (Table 8.4). Pseudohyperkalaemia may occur because K^+ leaks rapidly from the red cells at room temperature. In some individuals, there is no evidence of haemolysis, only macrocytosis and pseudohyperkalaemia. Unless the cause of the apparent hyperkalaemia is diagnosed, unnecessary, and even dangerous, investigation and treatment may be undertaken. In some families, the K^+ leak is greatly increased *in vitro* by cold (cryohydrocytosis). In dehydrated hereditary stomatocytosis, there may be marked perinatal ascites that resolves spontaneously over the first year of life, but which again can lead to extensive unnecessary investigation. The third problem with hereditary stomatocytosis, both over-hydrated and dehydrated varieties, is that splenectomy is followed by very marked thrombotic tendencies such that splenectomy should not be performed.

Table 8.4 Features of hereditary stomatocytosis and related disorders.

Characteristic	Expression	Group affected
Haemolytic anaemia	Mild to moderate	All variants
	Absent	Hereditary pseudohyperkalaemia
Morphology	Macrocytosis	All variants
	Stomatocytosis	Variable
MCHC	Decreased	Over-hydrated HSt (hydrocytosis)
	Increased	Dehydrated HSt (hereditary xerocytosis, desiccocytosis)
Serum $[K^+]$	Raised *in vitro*	Pseudohyperkalaemia
		Dehydrated HSt
		Cryohydrocytosis
Thrombotic tendency	Post splenectomy	All variants
Fluid balance	Perinatal oedema	Dehydrated HSt

Laboratory investigations

Tests for haemolysis and examination of the blood film of the patient and close relatives are the first steps in diagnosis. The finding of a raised serum potassium, together with macrocytosis, especially with some evidence of haemolysis, indicates the pseudohyperkalaemia of dehydrated hereditary stomatocytosis. Definitive studies involve the measurement of intracellular $[Na^+]$ or $[K^+]$ and their flux through the membrane at different temperatures. Four subgroups have been defined according to the intracellular sodium concentration (normal 5–10 mmol/L). Patients with the pseudohyperkalaemia of dehydrated hereditary stomatocytosis may have normal or slightly high sodium concentrations (12–18 mmol/L); in families with cryohydrocytosis (temperature-sensitive leak) the sodium concentration is 20–50 mmol/L, and in over-hydrated hereditary stomatocytosis the sodium concentration is 60 mmol/L or more.

Treatment

There is rarely a need for measures to raise the haemoglobin, and splenectomy should be avoided because of the risk of thrombosis, including hepatic and portal vein thrombosis. If splenectomy is necessary, lifelong anticoagulation should be introduced.

Southeast Asian ovalocytosis

A dominantly inherited ovalocytosis is found in parts of Southeast Asia, where *falciparum* malaria is common, particularly in

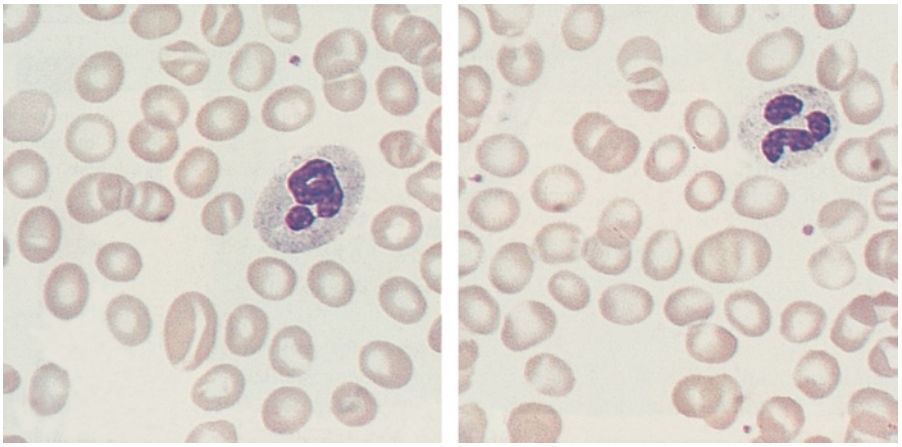


Figure 8.9 Southeast Asian ovalocytosis, peripheral blood films. Mild ovalocytosis and some stomatocytosis. Some cells have apparent transverse ridge.

Papua New Guinea, Borneo, Southern Thailand and the Philippines. The red cell morphology is ovalocytic rather than elliptocytic. Stoma-like slits may be present and transverse banding in the red cells is seen (Figure 8.9). Most individuals have no haemolysis except in the neonatal period but, in a few, mild anaemia may be present. The molecular defect is a deletion of nine amino acids at the transmembrane cytosol junction of band 3; the defect possibly limits the mobility of band 3 within the membrane. Homozygosity is not found and is presumably lethal *in utero*.

Abnormalities of membrane lipids

Acanthocytosis

Acanthocytes, or spur cells, show prominent, somewhat regular projections on the surface, best demonstrated by scanning electron microscopy. They are formed when the outer lipid layer of the membrane acquires additional lipid. Acanthocytosis is an acquired characteristic of severe liver disease, usually end stage, and the result of interaction of altered plasma lipids.

Abetalipoproteinaemia

Abetalipoproteinaemia is a rare inherited defect with absent β-apolipoprotein, which results in low serum cholesterol, but increased sphingomyelin, which enters the cell membrane and produces the acanthocytes. The main clinical features are retinitis pigmentosa, fat malabsorption and hepatic encephalopathy.

McLeod phenotype

In the McLeod phenotype, acanthocytosis occurs (Figure 8.10), together with decreased expression of the Kell antigen. The defective gene is on the X chromosome (Xp21), close to genes for Duchenne muscular dystrophy and retinitis pigmentosa, conditions with which the phenotype has been linked. The gene codes for the Kx protein that carries the Kell blood group protein. There may be mild anaemia.

Defects of red cell metabolism

The main function of the red cell is to carry haemoglobin around the circulation in high concentration and in a functional state so that gas exchange may occur efficiently in the lungs and in the tissue capillaries. In order to fulfil its function, the red cell needs a supply of energy in the form of ATP and a source of reducing power.

Mature red cells contain no DNA or RNA and hence are incapable of protein synthesis, and the only source of energy

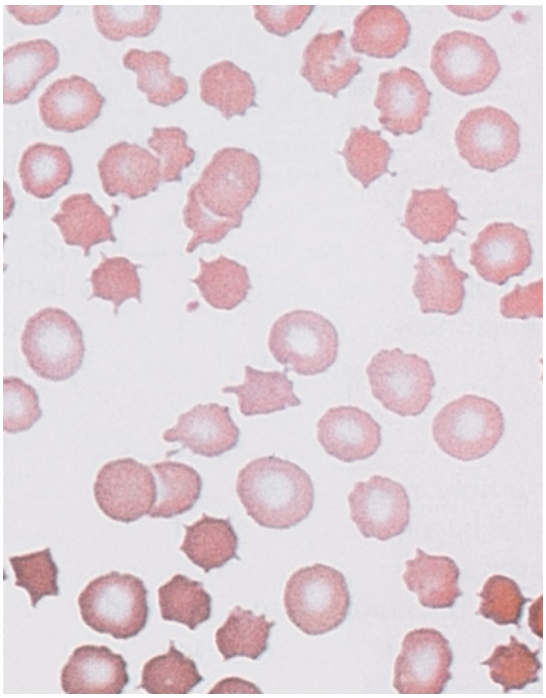


Figure 8.10 McLeod syndrome, peripheral blood. Note the marked acanthocytosis.

as ATP is derived from anaerobic glycolysis, the linked reducing system of the hexose monophosphate shunt (pentose phosphate pathway) and the glutathione cycle. ATP is required to maintain the membrane in its deformable state, with asymmetric lipid layers, and to regulate ion and water exchange. Reducing power is required to reduce methaemoglobin to its functional state of deoxyhaemoglobin and to counteract the strong oxidative stresses, which a circulating cell carrying molecular oxygen is likely to encounter. The main process that reduces methaemoglobin utilizes reduced nicotinamide adenine dinucleotide (NADH), produced from nicotinamide adenine dinucleotide (NAD^+) by the glycolytic pathway. Reduction and detoxification of free oxygen radicals and hydrogen peroxide produced during reactions to infection is provided by reduced nicotinamide adenine dinucleotide phosphate (NADPH) generated by the first steps of the pentose phosphate pathway, catalysed by glucose-6-phosphate dehydrogenase (G6PD) and the linked enzyme 6-phosphogluconate dehydrogenase. NADPH drives the glutathione cycle, glutathione (GSH) being the major reducing agent within the red cell.

The lack of protein synthesis in the mature red cell means that none of the enzymes in the metabolic pathways can be replaced during the red cell lifespan. Over the 120 days of normal red cell survival, enzyme activities decline at variable but predictable rates. This decline probably contributes to the ageing process of the red cell. Many of the abnormalities that affect red cell metabolism provoke haemolytic anaemia.

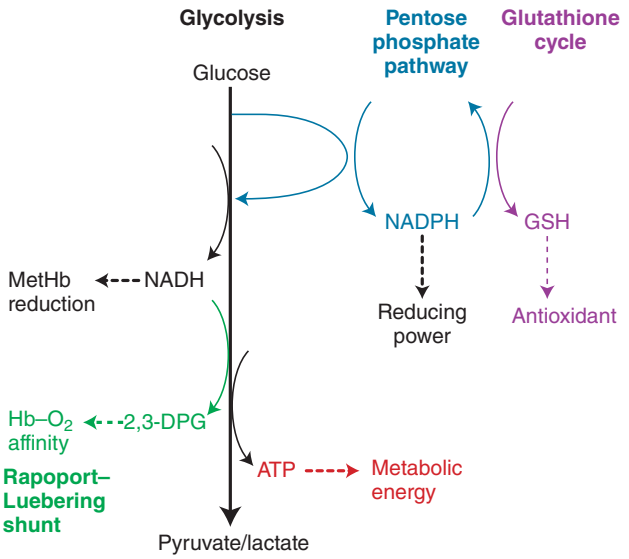


Figure 8.11 Principal pathways of energy production in the mature red cell. The glycolytic pathway provides energy in the form of ATP. Under normal conditions, methaemoglobin (MetHb) is reduced by the coupled reaction with $NADH^+$. The Rapoport–Luebering shunt provides 2,3-diphosphoglycerate (2,3-DPG) for control of haemoglobin oxygen affinity. Reducing power is produced by the pentose phosphate pathway and is linked to redox reactions through the glutathione cycle.

The glycolytic pathway (Embden–Meyerhof pathway)

Glycolysis is the process by which glucose is converted to pyruvate through a number of steps, with a net gain of two moles of ATP generated for each mole of glucose metabolized. Glucose is derived from the plasma by facilitated transfer through the membrane. Pyruvate and lactic acid are in equilibrium determined by the redox potential of the cell (NAD^+/NADH) and can diffuse out of the cell. The internal milieu of the cell, with its high K^+ concentration and presence of other cations, such as magnesium, necessary for efficient glycolysis, is maintained through the activity of various ion channels in the membrane. The glycolytic pathway also provides the redox reaction to convert methaemoglobin to deoxyhaemoglobin by utilizing NADH in a reaction catalysed by methaemoglobin-NADH reductase (cytochrome b_3). The various products and functions of glucose metabolism are shown schematically in Figure 8.11 and in detail in Figure 8.12.

The Rapoport–Luebering shunt

One of the essential roles of erythrocyte metabolism is also to provide sufficient 2,3-diphosphoglycerate (2,3-DPG) to regulate the oxygen affinity of haemoglobin. 2,3-DPG is produced from 1,3-DPG under the influence of the enzyme diphosphoglycerate mutase in linked reactions that form the Rapoport–Luebering shunt: 2,3-DPG is broken down to 3-phosphoglycerate by a phosphatase and thus re-enters the glycolytic pathway. It should be noted that when metabolism takes place via the Rapoport–Leubering shunt, there is bypass of the stage of ATP production (Figure 8.12). The shunt thus not only provides the 2,3-DPG for interaction with the haemoglobin tetramer, but also acts as an energy control mechanism for glycolysis.

Disorders of the glycolytic pathway

Mutations of most of the enzymes in the glycolytic pathway have been described in association with congenital non-spherocytic haemolytic anaemia (CNSHA), as the result of a failure to produce sufficient ATP. However, when the defective enzyme is expressed throughout the body, other non-haematological symptoms may occur.

Pyruvate kinase deficiency

Pyruvate kinase (PK) is one of the dominant controlling steps in glucose metabolism (together with hexokinase and phosphofructokinase). It catalyses the final steps of the glycolytic

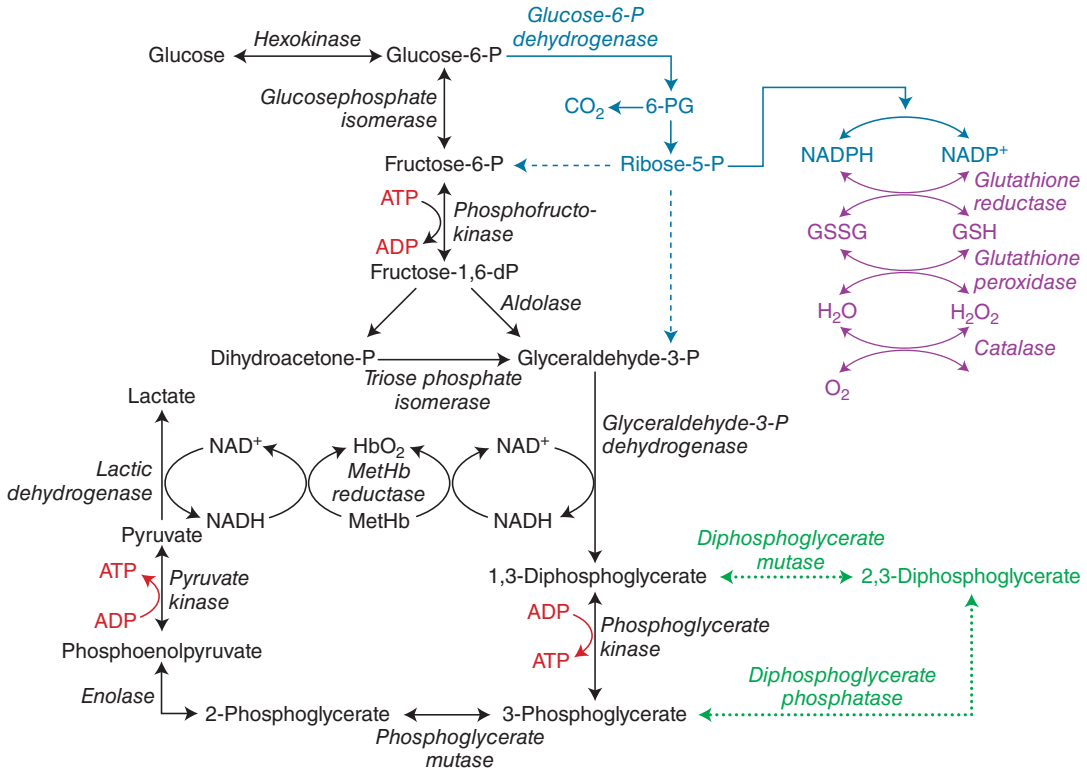


Figure 8.12 The glycolytic pathway and interactions with the other metabolic pathways.

pathway with the conversion of phosphoenolpyruvate to pyruvate, with the concomitant phosphorylation of ADP to ATP, leading to overall net gain of ATP from this pathway. There are four tissue specific isoenzymes codified by two separate genes. The *PKM2* gene, on chromosome 15, produces PKM1 and PKM2 through differences in post-transcriptional splicing. PKM1 is present in skeletal muscles, PKM2 in leucocytes, kidneys, adipose tissue and lungs. The *PKLR* gene, on chromosome 1, gives rise to PKL in the liver and PKR in red cells, controlled by tissue-specific promoters. The active enzymes are homotetramers. PKM2, PKR and PKL demonstrate marked allosteric reactions with several ligands; PKM1, on the other hand, has no allosteric interactions.

PK deficiency is the most common enzymopathy of the glycolytic pathway inherited in an autosomal recessive form. The prevalence of this disease, as assessed by gene frequency studies, has been estimated to be 1:20,000 in the general white population. More than 220 different mutations have been described in *PKLR* gene, the majority involving missense mutations, a few deletions, or mutations in the promoter region. Many individuals are compound heterozygotes. Not surprisingly, there is enormous genetic heterogeneity between affected individuals, reflected in the multiplicity of quantitative and kinetic defects. Because PKR is a homotetramer, the genotype–phenotype correlations, mainly based on analysis of the enzyme's three-dimensional structure and observation of the few homozygous patients, are difficult to predict. The comparison of recombinant mutants of human PKR with wild-type enzymes has revealed the effect of amino acid replacements on the molecular properties of the enzyme. However, the clinical manifestations of red cell enzyme defects are not merely dependent on the molecular properties of the mutant protein, but rather reflect the complex interactions of additional factors, including genetic background, concomitant functional polymorphisms of other enzymes, post-translational or epigenetic modifications, ineffective erythropoiesis and differences in splenic function.

Deficient PK activity leads to accumulation of substrates further up the pathway, including 2,3-DPG. The increased concentration of 2,3-DPG in PK-deficient red cells shifts the oxygen dissociation curve to the right, indicating low oxygen affinity. PK deficient patients tolerate apparent anaemia well because the lower haemoglobin content will deliver the same amount of oxygen to the tissues as normal haemoglobin, at least under normal conditions, though oxygen reserves would be limited.

Reticulocytes have alternative means of producing energy in the form of ATP via the oxidative respiratory pathway of the remaining mitochondria. They can also synthesize enzyme in those variants characterized by enzyme instability in mature red

cells. Reticulocytes thus have a metabolic advantage over mature red cells in PK deficiency.

Clinical features

The genetic heterogeneity is reflected by the wide variation in the phenotype. The presenting features may vary from severe neonatal jaundice and anaemia, rarely even presenting with hydrops fetalis, severe chronic non-spherocytic haemolytic anaemia requiring repeated transfusions, moderate haemolysis with exacerbation during infections or pregnancy, to symptomless compensated haemolysis with only a minor apparent anaemia. The majority of reported cases have presented in childhood. PK deficiency does not usually have an adverse effect on the outcome of pregnancy, although occasionally transfusion may be required to compensate the added dilutional anaemia.

Jaundice, as with other congenital haemolytic anaemias, may be exacerbated by coinheritance of other genes, as in Gilbert's syndrome. The haemolysis is nearly always extravascular, though rare examples with some intravascular haemolysis have been detected. Gallstones are common in PK deficiency and may lead to bouts of cholecystitis and biliary colic. Jaundice may also be increased by administration of drugs that affect bile excretion. As with other cases of congenital haemolytic anaemia with extravascular haemolysis, excess iron accumulation may occasionally develop, even in the absence of transfusion or coinheritance of a haemochromatosis gene.

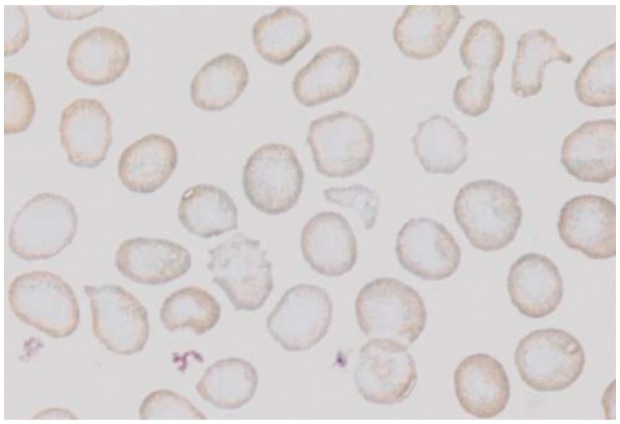
The spleen is usually palpable in cases with significant haemolysis, though in milder cases it may only be evident by ultrasound or other imaging techniques.

Laboratory diagnosis

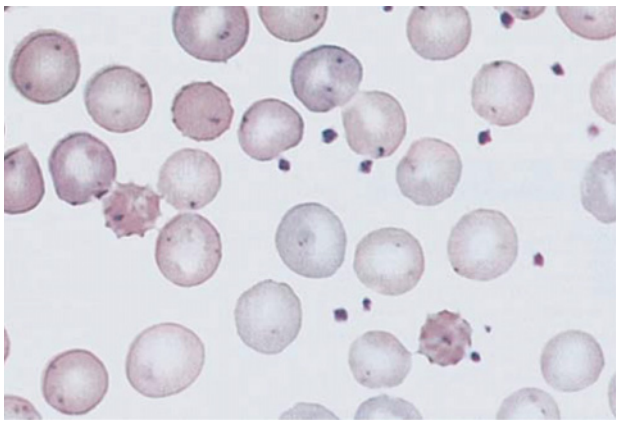
The haematological features of PK deficiency are common to other hereditary non-spherocytic haemolytic diseases; for this reason the diagnosis ultimately depends on the exclusion of other causes of haemolytic anaemia, upon the demonstration of low enzyme activity and the confirmation at DNA level.

The blood count reveals normochromic anaemia with reticulocytosis, sometimes producing a slight macrocytosis. An increased mean corpuscular haemoglobin concentration is occasionally seen in severe cases due to dehydration brought about by ATP deficiency. Red cell morphology is commonly unremarkable, displaying anisocytosis and a variable portion of spur cells or acanthocytes (which are not specific), particularly after splenectomy. A further increase of reticulocytes post-splenectomy is typically observed (Figure 8.13 a–c).

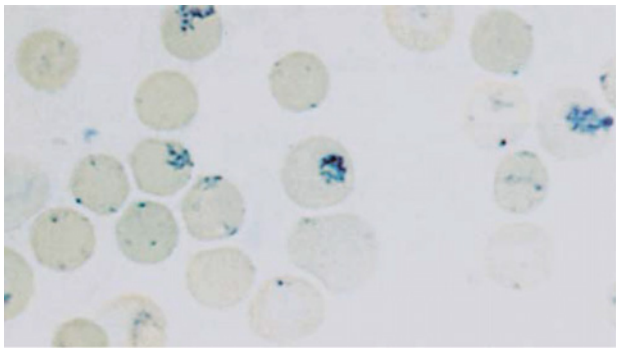
The decreased enzyme activity is associated with an elevated 2,3-DPG level (two to three times normal). Meaningful enzyme levels can only be achieved after total removal of leucocytes, which have up to 300 times the PK activity of red cells; the effect of reticulocytes also has to be taken into account. No correlation has been described among the residual PK activity, degree of haemolysis and clinical severity of PK deficiency.



(a)



(b)



(c)

Figure 8.13 PK deficiency, peripheral blood. (a) Red cell anisocytosis and poikilocytosis before splenectomy. (b) After splenectomy showing acanthocytes or 'prickle' cells. (c) Gross reticulocytosis after splenectomy (supravital new methylene blue stain).

Management

No curative therapy for PK deficiency is available to date, and the treatment is therefore based on supportive measures. Red cell transfusions may be required in severely anaemic cases, particularly in the first years of life; the haemoglobin then tends to stabilize in many cases at about 60–80 g/L, and transfusions are no

longer necessary unless the anaemia is exacerbated by infections, pregnancy or other conditions. Because the delivery of oxygen to tissues is highly efficient due to the high 2,3-DPG content, the decision to transfuse a PK-deficient patient should be based on the clinical conditions rather than on the haemoglobin levels.

Splenectomy does not arrest haemolysis, and should be reserved for severely affected, young patients who need regular blood transfusions, and to patients who do not tolerate anaemia. It usually results in an increase of 10–30 g/L in haemoglobin, reducing or even eliminating transfusion requirement. Indications and risks of splenectomy, particularly the risk of overwhelming sepsis, are the same reported for red cell membrane disorders and have been discussed in Chapter 7. Aplastic or haemolytic crises may still occur after splenectomy. As with all chronic haemolytic anaemias, folic acid 5 mg per week or 400 μg daily is a sensible supplement. Iron chelation may be required since iron overload is not uncommon in PK deficiency, even in non-transfused patients. Bone marrow transplantation has been successfully performed in a few very severely affected children.

Other defects of the enzymes of the glycolytic system

Compared with PK deficiency, other defects of the glycolytic pathway are very rare. The main features of these disorders are summarized in Table 8.5.

Hexokinase deficiency

Hexokinase catalyses the phosphorylation of glucose to glucose-6-phosphate (G6P), the first step in the glycolytic pathway. The erythrocytic isoenzyme, derived from alternative splicing, differs from that in nucleated cells, which have oxidative respiration, by lacking a porin-binding domain that links the enzyme to the mitochondrial membrane. The enzyme provides a major rate-limiting step in glycolysis and has extensive allosteric interactions: it is highly pH sensitive and its activity is regulated by its products, G6P, P_i, 2,3-DPG and disulfide compounds. The enzyme activity decays predictably with age of the normal red cell, and may be used as a comparator for other enzyme activities where absolute levels may be difficult to interpret because of the age distribution of the red cells.

Hexokinase deficiency has been recorded in about 20 cases characterized by molecular and phenotypic heterogeneity. Most patients have moderately reduced activity and complete hexokinase deficiency is probably lethal. The enzyme deficiency results in moderate, mild anaemia, but some cases present with severe anaemia and death in neonatal period. Typically, reduced hexokinase activity is associated with a low concentration of 2,3-DPG within the cells. Patients have less exercise tolerance for a given level of haemoglobin than would be expected because of the left shift in the oxygen dissociation curve.

Glucose phosphate isomerase deficiency

Glucose phosphate isomerase (GPI) catalyses the second step of glycolysis, the interconversion of G6P to fructose-6-phosphate (F6P). The enzyme is also known as phosphohexose isomerase, phosphoglucose isomerase, autocrine motility factor and neuroleukin, indicating that the protein has other actions in other cells.

GPI deficiency is one of the commonest causes of CNSHA after G6PD and PK deficiency. About 30 different mutations have been identified, underlying the molecular heterogeneity and consequently the heterogeneity of the clinical picture. Most reported cases present with mild to moderate haemolytic anaemia; hydrops fetalis may occur. In T lymphocytes, GPI acts as neuroleukin, a lymphokine that induces the formation of antibody-secreting cells. GPI is also expressed in neutrophils, but there is no increase in infections in deficient subjects. In some deficient patients, neurological impairment or mental retardation has been reported, thought to be related to hypoxia or ischaemia *in utero* rather than to direct metabolic effects.

Phosphofructokinase deficiency

Phosphofructokinase (PFK) catalyses a reaction in which F6P is phosphorylated to fructose-1,6-diphosphate, ATP being the donor of the phosphate group. Under normal physiological conditions this may be the major rate-limiting step in glycolysis in the red cell. Erythrocytic PFK is a tetramer composed of M or L subunits encoded by two separate genes; there may be five isoenzymes composed of different numbers of L and M subunits. PFK is a homotetramer of M subunits (M4) in muscle and of L subunits (L4) in liver. A third subunit is found in platelets. Deficiency of the M subunit leads to glycogen storage disease type 7 (Tarui disease). It is characterized by muscle cramps and myoglobinuria on exertion. Shortened red cell viability may be a minor component of this disease. Evidence of haemolysis may be accompanied by mild erythrocytosis as a result of decreased production of 2,3-DPG.

Fructose diphosphate aldolase A deficiency

Fructose-1,6-diphosphate aldolase A (ALDOA) catalyses the conversion of fructose-1,6-diphosphate to Ga3P and dihydroxyacetone phosphate (DHAP). There are three aldolase isoenzymes in human tissues, A, B and C; A isoenzyme is expressed in red cells. ALDOA is produced in the developing embryo and also forms the bulk of the enzyme in muscle, where it may be as much as 5% of the total cellular protein. In the red cell, the reaction catalysed by the enzyme is virtually irreversible.

This enzyme deficiency is extremely rare, with only six patients described so far. The defect presents with anaemia, which may be associated with mental retardation, dysmorphic features or myopathy. In one patient rhabdomyolysis occurred.

Table 8.5 Main features of red cell enzyme deficiencies.

	Transmission	Gene	Chromosome	No. cases	No. mutations	Haematology	Other symptoms
Embden–Meyerof pathway							
Hexokinase	AR	*HK1*	10q22.1	20 cases	5	CNSHA	
Glucosephosphate isomerase	AR	*GPI*	19q13.11	>50 fam	31	CNSHA	Mental retardation?
Phosphofructokinase	AR	*PFK-M* *PFK-L*	12q13.11 21q22.3	50-100 cases	23	Erythrocytosis Minimal haemolysis	Muscle disease, Tarui disase (Glycogenosis type VII)
Aldolase	AR	*ALDOA*	16p11.2	6 cases	4	CNSHA	Mental retardation Dysmorphism
Triosephosphate isomerase	AR	*TPI1*	12p13	50–100 cases	18	CNSHA	Neuromuscular disease Infections
Phosphoglycerate kinase	X-linked	*PGK1*	X13.3	40 cases	20	CNSHA	Neuromuscular disease
Pyruvate kinase	AR	*PKLR*	1q22	>500 fam	>200	CNSHA	
Rapoport-Luebering shunt							
Bisphosphoglycerate mutase	AR	*BPGM*	7q33	6 fam	3	Erythrocytosis, CNSHA	
Hexose-monophosphate shunt							
Glucose-6-phosphate Dehydrogenase	X-linked	*G6PD*	Xq28	>400×10^6 cases	>180	CNSHA–acute Favism	
Glucose-6-phosphate dehydrogenase (class I)	X-linked	*G6PD*	Xq28	> 50 fam	>60	CNSHA- chronic	

(*Continued*)

Table 8.5 (*Continued*)

	Transmission	Gene	Chromosome	No. cases	No. mutations	Haematology	Other symptoms
Glutathione metabolism							
Glutathione synthetase	AR	*GSS*	20q11.22	>50 fam	32	CNSHA	Metabolic acidosis (5-oxoprolinuria) Neurological symptoms Infections
Glutathione reductase	AR	*GSR*	8p21.1	2 fam	3	Induced oxidative HA, Favism, Neonatal jaundice	Cataract
γ-Glutamylcysteine synthetase (glutamate cysteine ligase)	AR	*GCLC* *GCLM*	6p12.1 1p21	12 fam	6	CNSHA, Oxidative HA	Metabolic acidosis (5-oxoprolinuria) Neurological symptoms
Glutathione peroxidase	AR	*GPX1*	3p21.3	1	0	Acute intravascular haemolysis?	
Nucleotide metabolism							
Adenosine deaminase (hyperactivity)	AD	*ADA*	20q13.12	3 fam	0	CNSHA	
Adenylate kinase	AR	*AK1*	9q34.11	12 fam	7	CNSHA	Motor impairment Mental retardation
Pyrimidine-5′-nucleotidase	AR	*NT5C3A*	7p14.3	>60 fam	26	CNSHA	
NADH-cytochrome b5 reductase	AR	*CYB5R3*	22q13.2	50 cases	45	Methemoglobinemia	Neuromuscular disease Mental retardation

AD, autosomal dominant; AR, autosomal recessive; fam, families; CNSHA, congenital non-spherocytic haemolytic anaemia; HA, haemolytic anaemia.

Triose phosphate isomerase deficiency

Triose phosphate isomerase (TPI) catalyses the interconversion of DHAP and Ga3P. In red cells, all DHAP is converted to Ga3P, which is then metabolized down the glycolytic pathway, providing the ultimate two molecules of ATP for each molecule of glucose metabolized. The enzyme is a homodimer present in all tissues that is catalytically active only in its dimeric form. In tissues other than the red blood cell, DHAP is an important precursor for the biosynthesis of lipids. More than 50 cases of TPI have been reported. This produces a severe syndrome, present from birth, consisting of CNSHA, a progressive neurological disorder with spasticity and central nervous system (CNS) degeneration. Cardiac failure and sudden death due to arrhythmias are also features. Death occurs usually at about 5 years of age, likely because of the formation of toxic protein aggregates, induced by misfolded TPI. A number of point mutations have been identified that lead to the syndrome; the most predominant is Glu104Asp, linked by common haplotypes suggesting descent from a common ancestor.

Phosphoglycerate kinase deficiency

Phosphoglycerate kinase (PGK) catalyses the reversible conversion of 1,3-DPG to 3-phosphoglycerate, generating one molecule of ATP for each molecule of 1,3-DPG metabolized. The active enzyme in the red blood cell is encoded by the *PGK1* gene on the X chromosome. The enzyme is monomeric and is expressed in all tissues. PGK deficiency shows a wide clinical phenotype: it may present with chronic haemolysis, neurological impairment (mental retardation, and ataxia), and myopathy (exercise intolerance or muscle weakness). A few affected individuals suffer the full spectrum of symptoms, whereas some cases are described with myopathy without haemolysis.

Defence against oxidative stress: the production of reducing power

As with all cells, but perhaps more urgently, red blood cells need to be protected against the effects of free radicals, hydrogen peroxide and other highly oxidative material to maintain membrane integrity and functional activity. In addition, haemoglobin has to be maintained in its functional state, and the steady production of methaemoglobin reversed to deoxyhaemoglobin, which is able to combine reversibly with oxygen. The major generator of reducing power within the red cell is the pentose phosphate pathway, which generates reducing power in the form of NADPH from $NADP^+$ coupled to the oxidation of G6P to 6-phosphogluconate (6PG) and the subsequent oxidation of 6PG to ribose-5-phosphate, by the action of G6PD and 6-phosphogluconate dehydrogenase (Figure 8.14).

GSH is important for protecting cells from oxidative damage by these free radicals and the effects of infection, maintaining protein sulfydryl groups in the reduced state and maintaining membrane transport. GSH is provided by the glutathione cycle linked to the pentose phosphate pathway through the action of glutathione reductase (Figure 8.15).

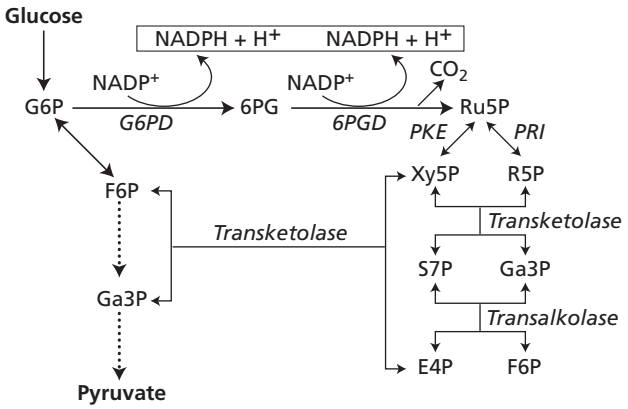


Figure 8.14 The pentose phosphate pathway. Substrates: G6P, glucose-6-phosphate; 6PG, 6-phosphogluconate; Ru5P, ribulose-5-phosphate; R5P, ribose-5-phosphate; Ga3P, glyceraldehyde-3-phosphate; F6P, fructose-6-phosphate; Xy5P, xylose-5-phosphate; S7P, septulose-7-phosphate; E4P, erythrose-4-phosphate. Enzymes: G6PD, glucose-6-phosphate dehydrogenase; 6PGD, 6-phosphogluconate dehydrogenase; PKE, phosphoketoepimerase; PRI, phosphoribose isomerase.

Oxidative stress that exceeds the reducing power of the red cell leads to intravascular haemolysis following denaturation of haemoglobin with precipitation as Heinz bodies and peroxidation of the red cell membrane. Methaemoglobinaemia may occur. The resulting clinical features – acute intravascular haemolysis, Heinz body haemolytic anaemia or methaemoglobinaemia – depend on the nature of the oxidative stress and the size of the imbalance between the stress and the redox potential.

Pentose phosphate pathway (hexose monophosphate shunt)

The pentose phosphate pathway, schematically shown in Figure 8.12 and in detail in Figure 8.14, is essential for the production of ribose and its incorporation into RNA. In red cells its only function is the production of reducing power in the form of NADPH. The first step of the pathway, catalysed by G6PD, utilizes G6P as substrate. G6P is also a substrate for glycolysis. About 10% of glucose is metabolized by the pentose phosphate pathway which is controlled by the availability of NAD^+ and the feedback inhibition of ATP. Under conditions of oxidative stress, the flux through the pentose phosphate pathway can be greatly increased. The products of the pathway re-enter the glycolytic pathway through F6P or Ga3P. The most frequent defect of the pentose phosphate pathway is G6PD deficiency. However,

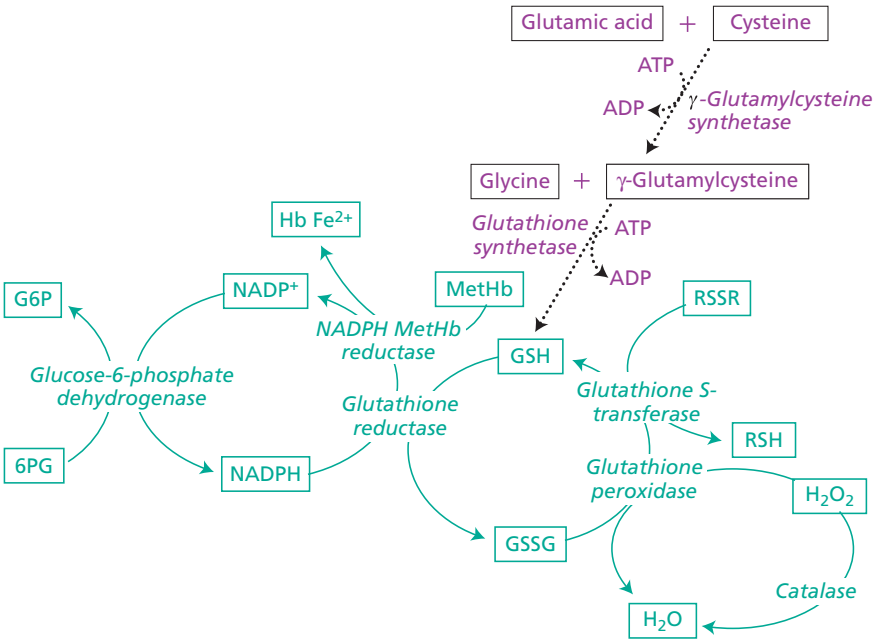


Figure 8.15 The glutathione (GSH) cycle and synthetic pathways. Redox control is exercised by the glutathione cycle linked to the NADPH of the pentose phosphate pathway by glutathione reductase.

other enzymes can be important in acquired disorders, some steps of the pathway being inhibited by drugs, including oestrogen/progesterone contraceptive pills.

Glucose-6-phosphate dehydrogenase deficiency

G6PD is the enzyme controlling flux through the availability of $NADP^+$. Conversion of G6P to 6PG is accompanied by the reduction of $NADP^+$ to NADPH (Figure 8.14); the oxidation of 6PG to ribose-5-phosphate produces a second molecule of NADPH. $NADP^+$ availability is determined by the glutathione cycle, which is linked to the pentose phosphate pathway through the activity of glutathione reductase (Figure 8.15).

The gene encoding for G6PD is located on the X chromosome at Xq28. The protein is assembled to produce an equilibrium of dimers and tetramers. Each monomer has an $NADP^+$-binding domain and a large domain with the active site between the two. The gene is highly conserved throughout evolution and is expressed in all cells, with an essential role in the production of RNA in nucleated cells. Complete inactivity of the enzyme in nucleated cells would be lethal. The clinical consequences of G6PD deficiency are virtually confined to the erythrocyte, with occasional evidence of leucocyte malfunction in some variants. Most mutations affect the stability of the enzyme so that its activity rapidly declines in the mature enucleated red cell as it ages. G6PD deficiency is the commonest genetically determined enzyme deficiency in the world, with an estimated 400 million people being affected. More than 180 different variants associated with deficient enzyme activity have been described and at least 35 of them are polymorphisms. The G6PD variants have been classified into five classes according to their activity relative to the wild-type G6PD type B (Table 8.6). In Africa, the variant G6PD type A^- is the predominant polymorphism, with equivalent activity to G6PD type B.

G6PD deficiency is widely disseminated throughout Africa, the Mediterranean basin, the Middle East, Southeast Asia and indigenous populations of the Indian subcontinent. G6PD-A^- is common in Africa; the Mediterranean variant is common in Southern Italy, Sardinia and other Mediterranean regions, and G6PD-Canton is common in Southern China. These variants are only the most common among many different mutations in these areas and throughout the rest of the affected world. This distribution of the deficiency equates with areas where *Plasmodium falciparum* malaria is common, and this is thought to be the evolutionary drive that has produced such widespread polymorphisms. It has subsequently been confirmed that G6PD deficiency does indeed protect against lethal *falciparum* malaria, particularly in childhood, and this protection, especially in hyperendemic areas, more than outweighs the haematological problems associated with deficiency.

There are four main syndromes associated with G6PD deficiency differing in their clinical presentations: neonatal jaundice, favism, chronic non-spherocytic haemolytic anaemia and drug-induced haemolytic anaemia. In all four, haemolysis is aggravated or promoted by exposure to oxidative stress through infection or ingestion of oxidative foods or drugs. Age always modifies the clinical effects, not always as might be expected.

The neonatal jaundice syndrome has been described in class I, II and III variants, favism mainly, though not exclusively, in class II, CNSHA in class I, and the drug-induced haemolytic anaemias mainly class III. While G6PD deficiency is most

Table 8.6 World Health Organization classification of G6PD deficiency (1989).

Class	Enzyme activity (% normal)	Examples	Clinical effects
I	Severe (usually <2)	Santiago de Cuba (Gly447Arg)	CNSHA, acute exacerbations
II	<10	Mediterranean (Ser188Phe) Canton (Arg459Leu) Orissa (Ala44Gly)	Favism, acute intravascular haemolysis (drug induced), neonatal jaundice
III	Moderate (>10, <60)	A− (Val68Met, Asn126Asp)	Acute intravascular haemolysis (drug induced), neonatal jaundice
IV	100	B (wild type) A (Asn126Asp)	None None
V	150	//	None

CNSHA, congenital non-spherocytic haemolytic anaemia.

common in males, the prevalence of the gene in many parts of the world means that female homozygotes are not uncommon and heterozygous females are often susceptible to oxidative stress because of the effects of X-inactivation and marked lyonization resulting in a significant population of G6PD-deficient red cells.

Neonatal jaundice and G6PD deficiency in infancy

Neonatal jaundice is a severe manifestation of G6PD deficiency and is a major source of potential morbidity from kernicterus. Most common variants have been associated with the syndrome, including the A^- and Mediterranean variants. In parts of the world where G6PD is common, it is the most prevalent cause of neonatal jaundice. The jaundice probably starts *in utero* during the perinatal period, but the clinical problem only becomes apparent on about the second or third day after birth. Phototherapy or exchange transfusion may be required to prevent neurological sequelae. Anaemia is not a feature, and it is thought that jaundice is a manifestation of liver enzyme deficiency coupled perhaps to the physiological underdevelopment of neonatal liver function or the coinheritance of Gilbert's syndrome. Acute haemolytic crises may occur in infants, usually through exposure to oxidative stress, including nitrites or nitrates in water or the ingestion of fava beans by the mother, but in some cases of severe acute haemolysis, even fatal, no cause has been obvious.

Favism

Favism is the term given to the G6PD syndrome where acute intravascular haemolysis may be precipitated by exposure to the broad bean *Vicia fava*, ingested in any form, usually about 24 hours after the exposure (Table 8.7). The offending agent is divicine, or its aglycone isouramil, which can produce free oxygen radicals on autoxidation. Divicine is not present in peas or beans of other types, which may be eaten without effect. The amount of haemolysis is dose-related, which may explain the marked variation in susceptibility not only in different subjects, but also in the same individual at different times. In children, acute haemolysis, sometimes life-threatening, is common, but renal failure is uncommon, although there may be systemic symptoms of fever and loin pain. Renal failure occurs more often in adults, possibly because of comorbidity. Favism is usual in class II variants, for example Mediterranean and Canton, but may occur in others, including the African A^- variant. Acute haemolysis after eating fava beans has also been described in glutathione reductase deficiency. Some other compounds with which affected individuals may come in contact can cause acute haemolysis, including topical henna and some of the pulses used to make local sweetmeats.

Between attacks of favism or exposure to oxidizing substances, the blood count is normal with no evidence of haemolysis. The cooccurrence of infection, which promotes the formation of hydrogen peroxide following the oxygen burst in neutrophils and macrophages with ingestion of oxidizing substances, even mild ones such as chloramphenicol, may promote haemolysis, even though the drug on its own does not.

Chronic non-spherocytic haemolytic anaemia

Sporadic cases of CNSHA are found with underlying G6PD deficiency. Many of the mutations that cause CNSHA occur on exon

Table 8.7 Characteristic features of haemolytic attack in G6PD deficiency.

Phase	Clinical	Laboratory
Acute	Abrupt onset	
	Malaise, prostration	
	Pallor	Anaemia, Heinz bodies, reticulocytosis, G6PD deficient
	(Abdominal pain)	
	Fever	Leucocytosis
	Dark urine	Haemoglobinuria, haptoglobin absent
	Haemoglobinaemia	
	Methaemalbuminaemia	
	Jaundice	Hyperbilirubinaemia
	(Renal failure)	↑ Urea, ↑ creatinine
Recovery	Gradual but rapid cessation of haemolysis	Reticulocytes peak days 5–8
	Urine clears in few days	G6PD increases (rarely to normal range)
	Jaundice clears in 1–2 weeks	

10 and affect the formation of dimers or tetramers. The haemolysis is extravascular, although additional oxidative stress may provoke an acute intravascular episode.

Drug-induced acute haemolysis

The introduction of primaquine and its derivative pamaquine as antimalarials to replace quinine during the Pacific phase of the Second World War and the later Korean War revealed that a proportion of men exposed, particularly in the black population, suffered from severe acute intravascular haemolysis. Studies by Carson and others from the University of Chicago, working at the Statesville Penitentiary Malaria Project, identified the problem as G6PD deficiency and identified that young red cells and reticulocytes had sufficient activity to withstand the oxidative stress so that haemolysis lessened as the reticulocyte level rose. It became apparent that many other drugs could also produce haemolysis. The common African A^- variant is the main example of class III mutations producing this type of haemolysis.

The haemolysis is dose-related and may be self-limiting. Although it is important to recognize which drugs are likely to produce haemolysis (Table 8.8), it is also important to realize that the disease for which the drugs may be needed, for example *falciparum* malaria, may be fatal and thus the concern for haemolysis is a lesser problem.

Laboratory diagnosis

Acute intravascular haemolysis raises the suspicion of G6PD deficiency. Red cell morphology shows contracted haemoglobin in 'ghost' membranes (Figure 8.16). Haemoglobinuria may be gross, producing almost black urine without red cells in the centrifuge deposit. Several screening tests have been devised to identify G6PD deficiency in red blood cells. These tests can reliably distinguish between deficient and non-deficient individuals, but are not reliably quantitative. The gold standard is the quantitative assay of G6PD activity. Since G6PD activity is red blood cell age-dependent the measurement of enzyme activity might yield false-normal results in the presence of reticulocytosis, (i.e. during haemolytic attack). The diagnosis of heterozygous females can be especially difficult; in extreme cases, it can only be done by family studies. However, the probability of

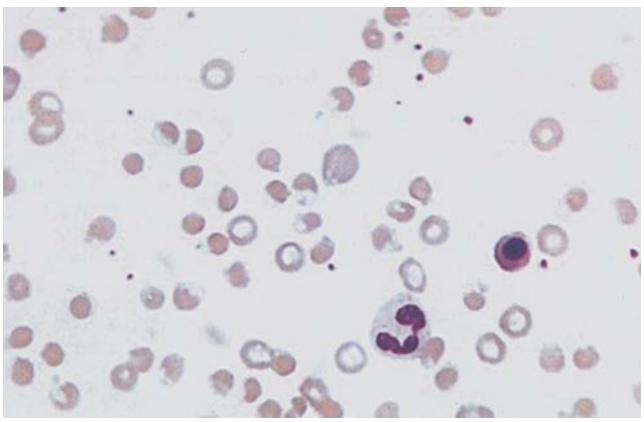


Figure 8.16 G6PD deficiency. Peripheral blood film following acute oxidant drug-induced haemolysis shows an erythroblast and damaged red cells, including 'blister' and 'bite' cells.

Table 8.8 Drugs to be avoided in G6PD deficiency.

Category of drug	Predictable haemolysis	Possible haemolysis
Antimalarials	Dapsone Primaquine Methylene blue	Chloroquine Quinine
Analgesics/Antipyretic	Phenazopyridine	Aspirin (high doses) Paracetamol (Acetaminophen)
Antibacterials	Cotrimoxazole Sulfadiazine Quinolones (including nalidixic acid, ciprofloxacin, ofloxacin) Nitrofurantoin	Sulfasalazine
Other	Rasburicase Toluidine blue	Chloramphenicol Isoniazid Ascorbic acid Glibenclamide Vitamin K Isosorbide Dinitrate

Similar tables have also been published previously (e.g., Betke *et al.*, 1967; Cappellini & Fiorelli, 2008; Luzzatto, 2012). The *Predictable haemolysis* column includes all 7 drugs listed in a recent evidence-based review (Youngster *et al.*, 2010). We consider that even a single case of AHA must be taken seriously if the clinical picture is well documented and if AHA is unlikely to have had another cause (e.g. infection): therefore, in addition to those seven drugs, we have retained in this column the following. (a) *Cotrimoxazole*. This drug is used widely in patients with human immunodeficiency virus (HIV) and in other conditions. AHA developed in at least two HIV-acquired immunodeficiency syndrome patients (Tungsiripat *et al.*, 2008), one of whom was G6PD-deficient. Additional cases have been reported (reviewed in Ho & Juurlink, 2011). It is presumed that, of the two chemicals present in cotrimoxazole, sulfamethoxazole rather than trimethoprim is likely to be the culprit. (b) *Sulfadiazine:* we found at least one case report (Eldad *et al.*, 1991) entirely convincing. (c) *Quinolones.* Although some cases may be regarded as anecdotal, in our view there are at least three convincing ones: one with nalidixic acid (Alessio & Morselli, 1972), one with ciprofloxacin (Sansone *et al.*, 2010), and one with ofloxacin (Carmoi *et al.*, 2009). Interestingly, the last two patients were (presumably heterozygous) women; the first was unique because she had no illness, but was exposed to nalidixic acid by virtue of working in a chemical factory that produced it. We also feel that if several reports are about chemically related compounds they tend to strengthen each other. The *Possible haemolysis* column is based on older literature (see for instance Burka *et al.*, 1966; WHO Working Group, 1989). As examples of the causal role of some of these drugs in causing AHA in G6PD-deficient persons see Sicard *et al.*, (1978) for chloroquine, Meloni *et al.*, (1989) for aspirin, Mehta *et al.*, (1990) for ascorbic acid, Meloni and Meloni (1996) for glibenclamide. *Source*: Luzzatto & Seneca, 2014. Reproduced with permission of Wiley.

clinically significant haemolysis in a heterozygote roughly correlates with the proportion of G6PD-deficient red cells in her blood. Therefore, if a normal level of G6PD activity is found in a heterozygote, she is unlikely to be at risk of G6PD-related haemolysis. DNA analysis is the most effective way of identifying heterozygotes, in particular in regions where G6PD has a high prevalence and the main variants are known.

Management

Management is mostly dictated by the symptoms and signs in patients, although education in the avoidance of oxidizing substances is important (Table 8.8). In many populations the condition is well known and the need for avoidance recognized. Neonatal jaundice may need urgent therapy to prevent neurological damage. Extreme hyperbilirubinaemia can be prevented by administration of Sn-mesoporphyrin if the diagnosis is known at birth. Acute intravascular haemolysis may require transfusion, but the anaemia is often self-limiting. High fluid intake should be encouraged to prevent renal damage. CNSHA may be severe enough to warrant active treatment and splenectomy may be helpful.

Glutathione

GSH is the major intracellular thiol in aerobic cells, and is equally important in the red cell. It has a number of critical functions: protecting cells against oxidative damage, participation in detoxification of foreign compounds, maintenance of protein sulfydryl groups in a reduced state, and possibly transport

of amino acids. In red cells, its main function is as an antioxidant. GSH is synthesized from glutamate, cysteine and glycine by the link reactions of two enzymes, γ-glutamylcysteine synthetase and glutathione synthetase (Figure 8.15). GSH exerts its function in preserving thiol groups and reducing hydrogen peroxide and free oxygen radicals through reactions catalysed by glutathione-*S*-transferase and glutathione peroxidase, respectively. Oxidized glutathione (GSSG) is reduced to GSH by the action of glutathione reductase, the hydrogen donor being NADPH. Failure to maintain the GSH level leads to chronic haemolytic anaemia and increased susceptibility to oxidative stress.

Complete loss of GSH synthesis is probably lethal. Severe deficiency leads to 5-oxoprolinuria, metabolic acidosis and mental retardation. A milder deficiency limited to red cells is associated with haemolytic anaemia aggravated by oxidative stress. Low levels of GSH caused by γ-glutamylcysteine synthetase or glutathione synthetase deficiency, with CNSHA, have been described (Table 8.5). Haemolytic anaemia due to deficiency of glutathione reductase has also been reported.

Nucleotide metabolism

Adenosine nucleotides, ATP (85–90%), ADP (10–12%) and AMP (1–3%), comprise the main nucleotide pool in mature red cells. The cell has no mechanism for making nucleotides once the RNA of the reticulocytes has been degraded, and it has an effective salvage mechanism for maintenance of the adenine pool, with the enzymes adenosine deaminase (ADA) and adenylate kinase involved in regulation. Deficiency of the enzymes in the salvage pathway does not usually lead to haemolysis. ADA deficiency is associated with severe combined immunodeficiency and excess activity is found in Diamond–Blackfan anaemia. During maturation of reticulocytes the RNA is broken down to pyrimidine and purine nucleotides, which are dephosphorylated to nucleosides that can diffuse out of the cell. The purine nucleotides enter the salvage pathway. Deficiency of pyrimidine 5′-nucleotidase leads to accumulation of pyrimidine nucleotides that interferes with the adenine nucleotide pool, producing haemolysis (Table 8.5).

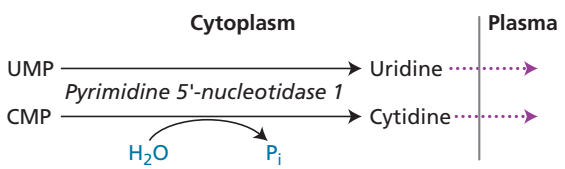


Figure 8.17 Pyrimidine nucleotide catabolism in the reticulocyte. Pyrimidine 5′-nucleotidase converts the nucleotides to monophosphates, which diffuse out of the cell.

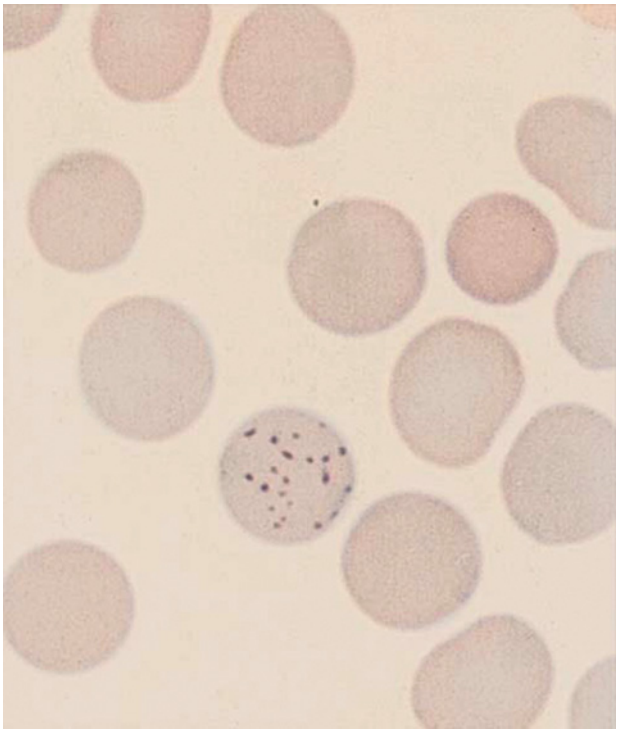


Figure 8.18 Peripheral blood film in pyrimidine 5′-nucleotidase deficiency showing basophilic stippling.

Pyrimidine 5′-nucleotidase

Pyrimidine 5′-nucleotidase (P5N), catalyses the dephosphorylation of the pyrimidine 5′-monophosphates uridine monophosphate (UMP) and cytidine monophosphate (CMP) to corresponding nucleosides (Figure 8.17). In red cells, there are two isoforms, type 1 (P5N1), which has a high affinity for UMP and CMP, and type 2 (P5N2), which hydrolyses deoxypyrimidine nucleotide monophosphates. P5N1 deficiency leads to haemolytic anaemia. The gene (NT5C3, also named P5′N1), codifies for the two isoforms by exon 2 alternative splicing. The enzyme is strongly inhibited by lead.

P5N1 deficiency is not uncommon, it is probably the third most common enzyme deficiency causing CNSHA. It is characterized by marked basophilic stippling in the red cells (Figure 8.18), similar to that seen in lead poisoning, and accumulation of high concentrations of pyrimidine nucleotides. The haemolysis is usually mild to moderate, although more severe cases have been reported. Splenectomy usually results in stabilization of the haemoglobin to higher levels (median increase 32 g/L).

Selected bibliography

Cappellini MD, Fiorelli G (2008) Glucose-6-phosphate dehydrogenase deficiency. *Lancet* **371**: 64–74.

Gallagher PG, Forget BG (1996) Hematologically important mutations: spectrin variants in hereditary elliptocytosis and hereditary pyropoikilocytosis. *Blood Cells, Molecules and Diseases* **22**: 254–8.

Gallagher PG (2013) Disorders of red cell volume regulation. *Current Opinion in Hematology* **20**: 201–7.

Luzzatto L, Seneca E (2014) G6PD deficiency: a classic example of pharmacogenetics with on-going clinical implications. *British Journal of Haematology* **164**(4): 469–80.

Mohandas N, Gallagher PG (2008) Red cells: past, present and future. *Blood* **112**: 3939–48.

Perrotta S, Gallagher PG, Mohandas N (2008) Hereditary spherocytosis. *Lancet* **372**: 1411–26.

Schilling RF (2009) Risks and benefits of splenectomy versus no splenectomy for hereditary spherocytosis: a personal view. *British Journal of Haematology* **145**: 728–32.

van Wijk R, van Solinge WW (2005) The energy-less red blood cell is lost: erythrocyte enzyme abnormalities of glycolysis. *Blood* **106**: 4034–42.

Zanella A (2000) Inherited disorders of red cell metabolism. *Baillières Best Practice and Research. Clinical Haematology* **13**: 1–148.

Zanella A, Bianchi P, Fermo E (2007) Pyruvate kinase deficiency. *Haematologica* **92**: 721–3.

CHAPTER 9

Acquired haemolytic anaemias

9

Modupe O Elebute[1] and Rachel Kesse-Adu[2]

[1]King's College Hospital, London, UK
[2]Guy's and St Thomas' Hospital, London, UK

Introduction

The acquired haemolytic anaemias are usually divided into two main categories, depending on the mechanism by which the premature destruction of red blood cells is produced. In the immune haemolytic anaemias, antibodies are the main agents of destruction, while the non-immune acquired haemolytic anaemias result from diverse causes and mechanisms of haemolysis. In haemolytic anaemias, significant shortening of red cell survival is the major cause of the anaemia and produces the classical features of haemolysis.

Immune haemolytic anaemias

Antibody-mediated haemolysis is an important cause of acquired haemolytic anaemia. Antibodies may be autoantibodies produced by a patient's own immune system and directed against epitopes of his/her own red cell antigens or they may be alloantibodies. Alloantibodies may be produced by the patient and directed against antigens not present on that person's own red cells, but either introduced as foreign red cell antigens by blood transfusion or secondarily acquired by the patient's red cells, as in drug-induced haemolysis. Alloantibodies directed against the patient's red cell antigens might also be introduced from outside the patient, most notably from the mother in haemolytic disease of the newborn. A simple classification of immune haemolytic anaemias is given in Table 9.1. Typically, the immune haemolytic anaemias are distinguished from the non-immune by detecting antibody on the surface of red cells by the direct antiglobulin test (DAT), also known as the Coombs test.

Autoimmune haemolytic anaemia

Autoimmune haemolytic anaemia (AIHA) is characterized by a positive Coombs test or DAT, which detects antibody, with or without complement, on the red blood cell surface. In the majority of cases, the pathogenesis involves a disturbance of the immune system in which T-lymphocyte control of autoreactive B lymphocyte clones is reduced.

AIHA is classified into warm and cold types depending on the characteristic temperature at which the red cell autoantibody is mostly reactive. Warm (IgG) autoantibodies bind to red cells more strongly at 37 °C and have decreased affinity at lower temperatures, while cold (IgM) autoantibodies bind to red cells more strongly at 4 °C, with little affinity at physiologic temperatures. Occasionally, patients have a combination of warm and cold autoantibodies. Haemolysis is mainly extravascular in AIHA; the red cells are more commonly coated with IgG only and IgG-coated red blood cells are preferentially destroyed in the spleen. In cases where the red blood cells are coated with a combination of IgG and complement or with complement alone, the main site of destruction is the liver, but intravascular haemolysis may also occur. The site and severity of red cell destruction depend on the structural and functional characteristics of the antibody and the efficiency of the mechanism of destruction, while the degree of anaemia depends on the rate and acuteness of the destruction and the capacity of the bone marrow to compensate for the red cells lost.

AIHA may occur without any underlying cause, primary or idiopathic AIHA, or may be associated with other disorders including systemic autoimmune disorders (such as systemic lupus erythematosus, SLE, and rheumatoid arthritis);

Postgraduate Haematology, Seventh Edition. Edited by A Victor Hoffbrand, Douglas R Higgs, David M Keeling and Atul B Mehta.

Table 9.1 Classification of immune haemolytic anaemias.

Antigen type	Antibody	Diseases	Associations
Autoimmune	Warm antibody	Primary	Idiopathic
		Secondary	Autoimmune diseases (ITP, SLE, Rheumatoid arthritis)
			Lymphoproliferative disorders
			Infections (EBV)
			Ovarian cysts
			Ovarian carcinoma and some other cancers
			Drugs
	Cold antibody	Cold haemagglutinin disease (CHAD)	
		Cold antibody syndromes	Infections (*M. pneumoniae*), lymphoproliferative disorders
	Donath–Landsteiner antibody	Paroxysmal cold haemoglobinuria (PCH)	Post viral, syphilis
Alloimmune	Induced by red cell antigens	Haemolytic transfusion reactions	
		Haemolytic disease of the newborn (HDN)	
		Post-stem-cell allografts	
	Drug dependent	Antibody/macrophage mediated	
		Antibody/complement mediated	
		Membrane modification	
		Autoimmune	

ITP, immune thrombocytopenia; SLE, systemic lupus erythematosus; EBV, Epstein–Barr virus.

malignancy (lymphoma, thymoma and chronic lymphocytic leukaemia, CLL) or drug exposure.

Antibody characteristics and specificity of red cell autoantibodies

Autoantibodies typically show reactivity against a number of antigens and are not as specific as alloantibodies. Antibody characteristics that influence the site and intensity of red cell destruction in AIHA can be evaluated using the DAT and the thermal range of antibody binding. Monospecific antihuman globulin reagents for the DAT are routinely available for the detection of IgG, IgM and IgA and for complement components C3c and C3d. Multispecific reagents are also available, as are reagents specific for IgG subclasses, but the latter are difficult to standardize.

Warm-acting antibodies

Warm-acting antibodies are mostly active *in vitro* at 37 °C; they are polyclonal and IgG antibodies predominate. In most cases of warm-type AIHA, the antibody detected in the patient's serum is pan-reacting with all cells in a routine group O panel. Where the specificity of the antibody can be determined, it is most commonly in the Rh blood group complex, with 10–15% showing specificity for either anti-e, anti-D or anti-c. A greater proportion show specificity by reacting with all cells except –/– Rh^{null} cells. Other rare specificities against high-frequency antigens include anti-En^a, anti-Wr^b or anti-U.

The most frequent patterns detected by the DAT on the red cell surface are as follows: IgG alone, IgG and complement, and complement alone. Antibodies may be detected in the serum at 37 °C by the indirect antiglobulin test (IAT) in about 50–60% of patients; rising to more than 90% of cases when the red cell membrane of the reagent cells is modified with papain or another proteolytic enzyme. Antibody may also be eluted from the red cell membrane in a majority of cases and the specificity determined.

A subtype of warm AIHA has been defined in which both warm- and cold-type antibodies are found, both tend to be lytic, and this 'mixed type' AIHA tends to produce a more severe haemolysis with an intravascular component. It is most commonly associated with SLE or lymphoproliferative disease (LPD).

Cold-acting antibodies

Cold-acting antibodies are predominantly IgM and are most actively bound to antigen in the cold (4 °C). Cold antibodies act both as agglutinins and lysins *in vitro*; the two functions may have different thermal ranges. The clinical significance of the antibody is not related to the titre of the cold antibody, but depends purely on the thermal range of its activity and how the

antibody binds to red cells at or near 37 °C. IgM antibodies active against I (auto) antigen at 4 °C can be a normal finding and although they have a wide range of thermal activity, they produce no clinical disturbance above 32 °C. *In vitro*, the IgM antibodies elute off the red cell membrane, leaving bound complement to be detected by anti-C3d in the DAT. In idiopathic cold haemagglutinin disease (CHAD) and most cases associated with LPD, the cold antibodies are monoclonal (IgM κ) whilst cold antibodies associated with viral infections are polyclonal. The IgM antibodies have specificity mainly for the I antigen although anti-i specificity may be found in Epstein–Barr virus (EBV)-associated infection and some cases of LPD. In cold agglutinin syndrome, anti-i specificity is usually seen. However, occasionally, there is specificity for anti-I, rarely for anti-Pr, anti-P, anti-M or anti-N and even cold reacting anti-A or anti-B. The antibodies in a subtype of cold AIHA, paroxysmal cold haemoglobinuria (PCH), are IgG and bind to antigen below 20 °C. When the temperature is raised to 37 °C in the presence of complement, lysis occurs. This biphasic reaction is the basis of the Donath–Landsteiner reaction. In PCH, the specificity of the autoantibody is anti-P.

Complement activation

Complement activation on the red cell membrane may be caused by autoantibodies against red cell antigens. Antibody binding to two adjacent sites on the red cell membrane is required to activate the C1 complement component by the classical pathway (Chapter 16). IgM molecules are pentameric and a single molecule can bind adjacent sites whilst IgG molecules can only activate complement if they form a 'doublet'. IgG1, IgG2 and IgG3 can activate complement, whereas IgG4 and IgA do not. In AIHA, complement activation usually stops at the C3 stage where C3b is bound to the membrane and further proteolysed to form the inactive component C3d, which is detected by the appropriate DAT. Complement activation beyond the C3 stage may lead to the formation of the membrane attack complex resulting in intravascular haemolysis. AIHA due to IgG2 alone is very rare and that due to IgG4 or IgA alone is uncommon. It is interesting that in the rare case of IgA AIHA, complement does become activated on the cell surface, although the mechanism is obscure as IgA does not itself fix complement. In general, when there is more than one class or subclass of antibody on the cell surface, the haemolysis is more intense and may be intravascular.

Mechanisms for immune red cell destruction

There are two main effector mechanisms *in vivo*: (a) cell-mediated, predominantly extravascular, immune destruction and (b) complement-mediated intravascular haemolysis.

Cell-mediated immune destruction

Cell-mediated immune red cell destruction of antibody-coated red cells is carried out by macrophages and monocytes which have cell-surface receptors for the Fc portion of IgG and for antigenic determinants present on activated C3. Cellular immune destruction is mediated through these receptors. Neutrophils and lymphocytes also have these receptors, but macrophages of the reticuloendothelial system within the spleen, liver and bone marrow are the main site of destruction *in vivo* (Chapter 16).

Fc receptor mechanism

Macrophages have Fc receptors for IgG1 and IgG3 molecules, but not for IgG2, IgG4, IgM or IgA (see also Chapter 14). Only IgG-coated red cells are destroyed in this way, the main method of destruction in warm AIHA where 70–75% of autoantibodies are IgG. Phagocytosis and antibody-dependent cell-mediated cytotoxicity are the major Fc receptor-dependent modes of antibody-coated cell destruction.

Role of the spleen

The splenic vasculature is adapted to an efficient filter for particles such as effete red cells, bacteria and immune complexes (Chapter 17). As blood passes through the central arteries towards the red pulp, the branches of these arteries have a plasma-skimming effect that raises the haematocrit of the blood as it passes towards the splenic cords. There, red cells come into close contact with splenic macrophages. The low plasma content and the relative lack of free plasma IgG molecules allow red-cell-bound IgG to interact preferentially with macrophage Fc receptors, leading to phagocytosis of coated red cells. When phagocytosis is partial, so that only portions of the cell membrane are removed, the remaining circulating red cell becomes spherocytic, although the somewhat rigid spherocytes may themselves be trapped in the splenic sinusoids and destroyed. The spleen is therefore the major site of red cell destruction when IgG alone is the main Fc-binding protein on the red cell surface (Chapter 17).

Role of the liver

Kupffer cells are macrophages that are present in the liver sinusoids and which express Fc receptors on their surface. Blood flow through the sinusoids is rapid compared with the spleen and there is no plasma-skimming effect so IgG-coated red cells are not preferentially destroyed in this situation. Instead, there is competition for the Fc receptors from circulating IgG and red cell destruction is more dependent on cells being coated with C3.

C3 receptor mechanism

Two types of C3 receptor have been identified on macrophages: CR1 and CR3. CR1 is specific for an antigenic site in the C3c region of activated C3b that is not exposed on native C3 and the breakdown product of C3b (iC3b) is also a major ligand for CR1 and the only ligand for CR3. Immune adherence of C3b-coated red cells to macrophages occurs mainly through CR1,

whereas CR3 binding triggers phagocytosis. The largest concentration of C3b-binding macrophages is found in liver sinusoids and the liver becomes the major site for trapping and phagocytosing C3b-coated red cells. There is no competition for complement receptor sites from non-activated C3 in the plasma. When markedly enlarged, the spleen also becomes an important site of cell destruction. In cold AIHA, IgM agglutinins bind most avidly to red cells in the peripheral circulation, where the temperature may be as low as 10–20 °C. Complement activation occurs and leads to C3b and iC3d expression on the red cell membrane. Macrophage destruction of red cells is the main mechanism of haemolysis in IgM cold-antibody-coated cells and also occurs in warm AIHA as a result of complement-fixing IgG and IgM antibodies.

Complement-mediated intravascular haemolysis

Complement-mediated intravascular haemolysis is a minor mechanism for red cell destruction in most patients with AIHA, however in a small proportion of patients when more than one class or subclass of immunoglobulin is present on the red cell surface, such a mechanism may predominate and produce severe intravascular haemolysis.

In warm AIHA, complement-induced intravascular haemolysis class of immunoglobulin is present on the red cell surface. Intravascular haemolysis has been reported with IgA-coated red cells, although the mechanism is obscure, as IgA does not itself fix complement.

In cold AIHA syndromes, intravascular haemolysis may be precipitated by exposure to cold. In such cases, lytic as well as agglutinating antibodies with a high thermal range may be demonstrated *in vitro*; a pattern that may be found in cold AIHA associated with *Mycoplasma pneumoniae* infection. Acute intravascular haemolysis is the usual presentation in PCH, where the antibody is IgG in type.

Other factors influencing red cell destruction and production

Bone marrow function and capability to provide a compensatory increase in erythropoiesis following significant red cell destruction may be impaired by autoantibodies that bind to and destroy reticulocytes and erythroblasts as well as mature red cells. Red cell production may also be reduced by acute folate deficiency secondary to increased demand. In lymphoproliferative disorders, red cell production can be impaired by bone marrow infiltration.

Reticuloendothelial function plays an important role because the severity of cellular immune red cell destruction depends largely on macrophage function. Reticuloendothelial function may be reduced in SLE by the clearance of immune complexes, a process known as reticuloendothelial blockade. In methyldopa-induced AIHA, the drug has been shown to reduce reticuloendothelial clearance of IgG-coated red cells, a possible explanation for the fact that many patients with a strongly positive DAT due to methyldopa have little or no haemolysis.

Hypocomplementaemia or a relative complement deficiency may result from continuous complement activation such as occurs in patients with chronic CHAD. Hypocomplementaemia is common in SLE and may also be caused by chronic activation of the complement pathway. There is a strong association between SLE and the occurrence of null alleles for the C2 and C4 genes, which causes a genetically determined complement deficiency.

Warm-type autoimmune haemolytic anaemias

Clinical Features

Warm-type AIHA affects all age groups and accounts for up to 70% of AIHA cases. Haemolysis is caused by high-affinity IgG antibodies produced by reactive polyclonal B cells. An underlying or associated disorder can be identified in 50–70% of cases and there is a preponderance of female patients in both idiopathic and secondary AIHA.

Presentation is variable and depends on the speed with which anaemia develops, the capacity of the bone marrow to compensate and the effects of any associated disease. Typically, the onset is insidious, with gradual awareness of symptoms of anaemia or observation of pallor or mild icterus by friends or relatives. Occasionally, the onset is acute, with rapidly developing anaemia and, in older patients, the risk of heart failure. Rarely, severe fulminating haemolysis may occur, resulting in life-threatening anaemia. A moderate increase in unconjugated serum bilirubin and excess urinary urobilinogen occurs as a result of extravascular haemolysis. Where intravascular haemolysis occurs, patients pass very dark urine caused by haemoglobinuria and haemosiderinuria. More marked icterus (bilirubin >90 μmol/L) suggests coexisting liver disease or biliary tract obstruction due to pigment gallstones or biliary sludge. Mild splenomegaly is common, rarely more than 2–3 cm below the costal margin; marked splenomegaly suggests the possibility of an underlying LPD.

Anaemia with marked reticulocytosis is present and the peripheral blood film is characterized by polychromasia, spherocytes (Figure 9.1), circulating nucleated red cells and, in some cases, red cell agglutination. Rarely, there may be reticulocytopenia associated with a positive DAT. There is an increase in lactate dehydrogenase (LDH) due to red cell lysis, but other liver function tests are normal unless there is associated liver or biliary tract disease. The DAT is positive, commonly with IgG alone (mainly IgG1 and IgG3 subclasses) or with IgG and complement (the latter pattern commonly seen in SLE). Autoantibody in the serum may show specificity within the Rh system (e.g. auto-anti-e), but in most cases is pan-reactive with all red blood cells. The autoantibody reacts at 37 °C. In very rare cases, the amount of antibody remaining on the red cell surface is insufficient to be detected by the conventional DAT.

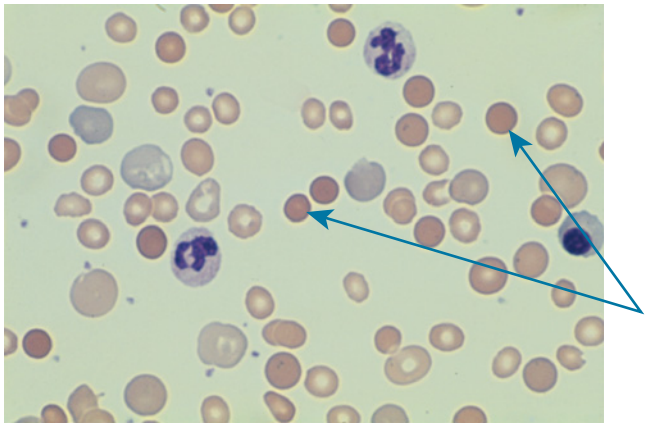


Figure 9.1 Warm autoimmune haemolytic anaemia. Blood film showing spherocytosis (arrows), polychromasia and a nucleated red blood cell (×40).

Idiopathic warm AIHA

Idiopathic warm AIHA with no underlying cause or associated disorder accounts for approximately 30% of patients with a DAT-positive haemolytic anaemia. It may occur at any age; peak incidence occurs in infancy and early childhood with a second rise during the third decade and the majority of cases occurring after the fifth decade. In infancy, the onset is often acute, typically precipitated by infection and transient. However, anaemia may be profound and difficult to control. There is a male preponderance in the childhood setting in contrast to females in adult cases. IgG antibodies can be transferred from an affected mother across the placenta, producing haemolysis in the newborn. A careful drug history should be taken in all cases to exclude haemolysis induced by drugs or by chemical exposure. In girls, AIHA may precede clinical or immunological evidence of SLE which should not be excluded on account of initial negative serology. Clinical presentation may vary from gradual onset of symptomatic anaemia to an acute, fulminating and life-threatening haemolytic process. Systemic symptoms include pallor, jaundice and mild splenomegaly (1.5–5 times normal size). Marked splenomegaly suggests a secondary cause.

Evans syndrome

Evans syndrome is defined as the combination of AIHA and immune thrombocytopenia (ITP) in the absence of any underlying disorder. The onset of thrombocytopenia may coincide with haemolysis or may arise separately. The platelet and red cell antibodies are distinct and do not cross-react. Rarely, episodes of immune neutropenia or pancytopenia have also been described in association with a positive DAT. The diagnosis is important because there appears to be a higher incidence in these cases of underlying conditions including immunodeficiency or autoimmune lymphoproliferative disease (ALPS) in children and of SLE or T-cell lymphoma in adults. Typically, the condition runs a chronic course characterized by relapses and remissions. Management is as for warm AIHA or ITP (Chapter 42), but patients with Evans syndrome tend to be more resistant to initial therapy with steroids. Second-line therapeutic options include immunosuppressive drugs such as vincristine, danazol, ciclosporin and mycophenolate mofetil as single agents or in combination regimens. Response to splenectomy is inferior in comparison with cases of uncomplicated ITP. The monoclonal anti-CD20 antibody, rituximab, has produced significant remission rates of both cytopenias in children and adults with steroid-refractory cases. Stem cell transplantation offers the only hope of cure in severe, refractory patients, but unfortunately carries a significant risk of transplant-related morbidity and mortality.

Warm AIHA associated with other autoimmune diseases

This is not uncommon and is well described in patients with SLE, especially in young women, and in other autoimmune or presumed autoimmune diseases, notably rheumatoid arthritis, Sjögren syndrome and ulcerative colitis. Warm AIHA is also part of the spectrum of autoimmune diseases associated with agammaglobulinaemia. Autoantibodies are usually IgG, and both IgG and C3d are found on the red cell surface. Occasionally, the DAT may be positive because of immune complexes adsorbed onto the red cell surface. The spleen is important for clearing such coated cells and splenectomy should be avoided if possible. Otherwise, treatment is as for idiopathic AIHA (see below).

Lymphoproliferative diseases

Lymphoproliferative diseases, including B-cell chronic lymphocytic leukaemia (CLL), low-grade B-cell non-Hodgkin lymphoma and Hodgkin lymphoma are well described in association with cases of warm AIHA. The AIHA may precede the diagnosis of lymphoma by months or years, may occur simultaneously with onset of the LPD or occur afterwards. Antibody response is thought to be due to immune dysregulation rather than direct production by the malignant clone. Antibodies are polyclonal and have no distinct pattern of type or specificity. Refractory autoimmune cytopenias are also well described following the use of therapeutic agents for LPDs including fludarabine and alemtuzumab, probably the result of their powerful effect on the immune system.

Drug-related warm AIHA

Drug-related warm AIHA caused by antibodies directed against self-antigens has been reported in the literature over many years including agents such as mefenamic acid, levodopa and procainamide. The mechanism by which AIHA is produced by exposure to drugs is not known; however, alteration of the red cell membrane or modulation of the immune response by the drug have been suggested. Treatment of CLL with fludarabine and other purine analogues may provoke a very severe and

life-threatening AIHA and, less commonly, other autoimmune cytopenias. With fludarabine, the mechanism is thought to be related to a decrease in auto-regulatory T cells.

Carcinoma-related warm AIHA

Carcinoma-related warm AIHA has been recorded with a number of malignancies, but it remains unclear whether there is a true association between the tumour and onset of haemolysis. Reports exist in patients with ovarian cysts where the cystic fluid has been found to contain the agglutinin and in patients with ovarian carcinoma.

Viral infections

Viral infections may precede warm AIHA in children, but rarely in adults; the haemolysis is usually brisk, but self-limiting. Hypotheses include alteration of the red cell membrane by the virus, thereby generating 'auto' antibodies against altered antigens or formation of antiviral antibodies that cross-react with membrane antigens. A third possibility involves the formation of immune complexes between the virus and specific antibodies, which are secondarily adsorbed onto the red cell surface, leading to immune destruction.

Management

Treatment is not evidence based, with only one reported randomized study, a few prospective Phase II trials and small case series reported in the literature.

Supportive therapy

Blood transfusion must be given if the clinical situation demands it, if the haemoglobin continues to fall or heart failure develops, despite the impossibility of achieving a satisfactory cross-match in the presence of a positive DAT. The least incompatible grouped blood should be used and transfused slowly. Some authors recommend the use of blood lacking antigens to which the autoantibodies react, but others point out that specificity is rarely absolute and that there is a risk of provoking an alloantibody response. Intravenous immunoglobulin (IVIg) has been used in AIHA, particularly when IgG is the main component on the red cell surface. The dose used is the same as for ITP, 0.4 mg/kg daily for 5 days or 1 g/kg daily for 2 days. AIHA responds less frequently to IVIg than does ITP. Side-effects include headache, fever–chill reactions and a small but present risk of viral transmission. Folic acid supplements should be given to all patients with chronic haemolysis.

First-line treatment

Corticosteroids are effective in warm AIHA; prednisolone is given at a dose of 1–2 mg/kg body weight once daily for 10–14 days. In patients who respond, the dose should be reduced steadily using laboratory tests, including Hb and LDH as a guide. It is important not to stop steroids too quickly as this may result in relapse. About 70–80% of patients improve their Hb initially, but only 30% maintain this response long term. Response rates may be improved when steroids are used in combination with rituximab, a chimeric murine–human monoclonal antibody that binds to the CD20 antigen on B cells and immature plasma cells. A randomized controlled study reported improved responses in 70% of patients in the combination arm, compared to 40% in the steroid-only group at 1 year.

Second-line treatment

Splenectomy is usually considered if there is no response after 3 months' trial of corticosteroids. Patients with predominantly IgG on the red cell surface respond best, and those with complement respond poorly. Of a selected group of patients, about 30% achieve a complete remission and do not require steroids, 30% have a significantly reduced steroid requirement and the remainder show no or only transient response. There is no certain way to determine clinical response to splenectomy. There is an increased risk of overwhelming sepsis after splenectomy, therefore all patients should receive pneumococcal, meningococcal and *Haemophilus influenzae B* vaccination at least 2 weeks before the procedure, and prophylactic penicillin indefinitely. The complications of splenectomy are dealt with in more detail in Chapter 17.

Monoclonal antibody therapy has been introduced with variable success in warm AIHA. Rituximab has been shown to be effective in idiopathic and secondary warm AIHA as well as in cold agglutinin disease. It is usually given at the standard dose and treatment schedule as for CLL of 375 mg/m^2 i.v. weekly for 4 weeks, however recent studies suggest that a lower dose (100 mg/m^2) may be as effective. Response has been shown in patients who are steroid-refractory and who have received multiple treatment modalities, and some patients who relapse after rituximab have been found to respond to retreatment. Overall response rates of 45–90% have been reported, with durable responses of up to 3 years in patients of all ages. The emergence of rituximab therapy as an effective alternative treatment for AIHA has led to a trend towards delaying splenectomy as an option for more resistant/refractory cases. Immediate side effects are minimal and mainly infusion related. Longer-term concerns are related to the fact that rituximab causes a rapid depletion of CD20-expressing B cell precursors and mature B cells which remain at undetectable or at very low levels for up to 1 year post-therapy. Significant B-cell depletion causes hypogammaglobulinaemia and prolonged immunosuppression, exposing patients to increased risk of severe infections such as hepatitis B, cytomegalovirus reactivation and multifocal leucoencephalopathy, a rare and usually fatal condition caused by uncontrolled activation of the JC virus. Results with alemtuzumab (Campath 1H), the anti-CD52 antibody have been disappointing; however, newer agents such as ofatumumab, an anti-D20 monoclonal antibody with enhanced Fc effector function may be more promising.

Table 9.2 Serological characteristics of cold-acting antibodies in the cold agglutinin syndromes.

Disorder		Antibody characteristics	
		Specificity	**Clonality and subtype**
Primary	Idiopathic (CHAD)	Anti-I Rarely Anti-P	Monoclonal IgM κ Monoclonal
Secondary	Lymphoproliferative disease	Anti-I Rarely Anti-I	Monoclonal IgM κ/λ (IgG) Monoclonal
	Mycoplasma pneumoniae	Anti-I	Polyclonal
	Infectious mononucleosis	Anti-I Anti-i	Polyclonal
Paroxysmal cold haemoglobinuria		Anti-P Donath–Landsteiner (biphasic)	Polyclonal IgG

CHAD, cold haemagglutinin disease.

Third-line treatment

Treatment with cytotoxic immunosuppressive drugs is reserved for steroid-refractory patients, for relapse following steroid withdrawal and non-response to splenectomy. The mechanism of action of these agents is relatively non-selective, resulting in significant systemic toxicity, including bone marrow suppression, renal toxicity and the potential long-term risk of malignancy. It should be emphasized that there are no controlled trials to prove the worth of cytotoxic drugs in AIHA, but there are small case reports of success with each agent. Response to these drugs is not usually seen for 4–6 weeks and they should be continued for at least 3 months before being deemed ineffective.

On its own, azathioprine is largely ineffective; however, it can have significant steroid-sparing action at a dose of 1.5–2.0 mg/kg daily. Danazol has been used, both in combination with corticosteroids as a first-line treatment and in the refractory setting, with mixed results in AIHA. Other effective agents include cyclophosphamide 1.5–2.0 mg/kg daily and ciclosporin, starting at 5 mg/kg daily in two divided doses to achieve trough plasma levels of 100–200 mg/L. There are small case reports of the use of mycophenolate mofetil, an immunosuppressive agent, and of vincristine, a cytotoxic agent with powerful antitumor effect, with some success in severe refractory AIHA.

Prognosis

Prognosis depends on a number of variables, including age, associated diseases and severity of haemolysis. AIHA can be a serious and potentially life-threatening disease with reported estimates of mortality of idiopathic AIHA in adults of 10% at 5 years to 40% at 7 years. Increased mortality typically occurs in patients over 50 years with deaths in the first 2 years after diagnosis usually related to underlying clinical disorders or treatment side-effects. Most children with AIHA recover completely, and mortality is much lower, probably about 5%. Treatment-related side-effects, particularly prolonged high-dose steroid therapy, can result in considerable morbidity.

Cold-type autoimmune haemolytic anaemias

The clinical features of the cold haemagglutinin syndromes vary with the pathogenesis of the disorder. Serological tests are useful in identifying the cause and in determining treatment. The serological characteristics of the antibodies found in these syndromes are shown in Table 9.2.

Clinical features

Idiopathic cold haemagglutinin disease

Idiopathic cold haemagglutinin disease (CHAD) is a relatively uncommon disorder accounting for only 15% of AIHA; it occurs mainly in the elderly and typically runs a chronic course. CHAD is mostly benign, but the clinical features can be very disabling and distressing. Purplish skin discoloration, maximal over the extremities (acrocyanosis), may be present in cold weather. Acrocyanosis is due to stasis in the peripheral circulation secondary to red cell agglutination. On warming the skin, the colour returns to normal or there is transient erythema, thus distinguishing acrocyanosis from Raynaud's syndrome. Haemolysis results in anaemia and the patient may be mildly icteric. Occasionally, haemolysis dominates the clinical picture, depending on the ability of the antibody to activate complement on the red cell surface. The cold agglutinins are monoclonal IgM κ, with heavy chain variable regions encoded by the IGHV4–34 gene segment, but serum electrophoresis may not reveal a monoclonal band because the concentration of the protein is too low. Traditionally defined by the absence of an underlying disorder, recent studies using sensitive tests including flow-cytometry and immunohistochemical assessments have demonstrated a monoclonal CD20-positive κ-positive B-lymphocyte population

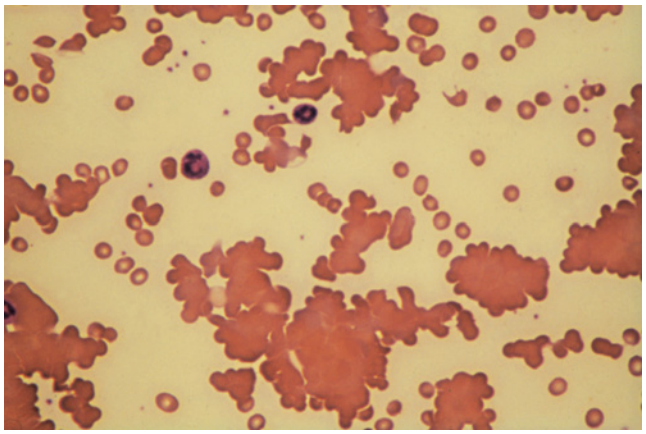


Figure 9.2 Cold haemagglutinin disease. Blood film showing gross haemagglutination (×20).

in the bone marrow of 90% of patients with CHAD. In addition, lymphoplasmacytic lymphomas are frequently associated with CHAD; therefore some authorities consider CHAD a pre-malignant B-cell disorder that becomes clinically overt due to the specificity of the antibody for red cell surface antigens.

In the laboratory, spontaneous agglutination of red cells is frequently observed, both macroscopically and on the peripheral blood film if made at room temperature (Figure 9.2). Automated blood cell counters detect agglutinates and record erroneously high mean corpuscular volume and low Hb values, unless the sample is tested at 37 °C. DAT is positive with only C3d on the red cell surface; IgM cold agglutinins are not detected because they elute from the cell surface *in vitro*. They are usually anti-I, although anti-Pr, anti-P and other rarer specificities have been described.

B-cell lymphoproliferative disorders may precede or accompany CHAD. In these cases, the antibody is monoclonal and a product of the malignant clone. Serological specificity is either anti-I or anti-i. Haemolysis is often more troublesome than symptoms of agglutination. The prognosis is usually that of the underlying lymphoproliferative disease.

Infections

Infections, almost always *M. pneumoniae*, or infectious mononucleosis, may be followed by haemolysis due to cold agglutinins. Rare cases following *Listeria* or *Toxoplasma* infections have also been reported. The antibodies are mostly polyclonal IgM in type, but occasional IgG cold antibodies are found. The antibodies develop in response to the infectious organism and cross-react with red cell antigens; haemolysis occurs 2–3 weeks after the infection and is usually mild and self-limiting. Acute and potentially fatal episodes of intravascular haemolysis have been reported in association with *M. pneumoniae* infection. Urgent blood transfusion through a blood warmer may be required.

Management

Management of cold haemagglutinin syndromes is difficult.

Supportive therapy

All patients should avoid exposure to cold; the use of electrically heated gloves and socks should be encouraged in the winter months. Patients should be nursed in a warm environment, preferably at 37 °C and kept warm using a number of techniques, including special blankets during surgical procedures. Folic acid supplements should be given to patients with chronic haemolysis. Blood transfusions should be given with due regard to the difficulty in cross-matching in the presence of cold haemagglutinins. Blood should be given through an in-line blood warmer. Plasma exchange (therapeutic apheresis) can be used to control severe symptoms by temporarily lowering the titre of cold agglutinins, but it has no effect on the underlying disease.

Definitive treatment

Unlike warm-type AIHA, corticosteroids are rarely of use in CHAD and should only be used in exceptional circumstances when antibodies are present in low titres and have a high thermal range. Alkylating agents such as chlorambucil can be effective in reducing antibody production in the context of an underlying B-cell neoplasm such as CLL. Intermittent regimens such as 10 mg/day for 14 days every 4 weeks or continuous treatment of 2–4 mg/day are both effective. Long-term treatment carries the risk of bone marrow suppression and the development of secondary haematological malignancies. Purine analogues such as cladribine and interferon-α (IFN-α) have been tried without significant success in CHAD. Rituximab, the anti-CD20 monoclonal antibody used with success in warm AIHA and in ITP has also been shown to be effective in idiopathic CHAD and in cold agglutinin syndromes associated with LPD. Response rates of up to 50% have been reported in Phase II studies, but with short duration of only 11 months. Longer response is seen with a combination of rituximab with fludarabine, with remission in 75% of patients and median duration of over 66 months. However, significant side-effects including grade 3–4 neutropenia makes this regimen unsuitable for elderly patients. Splenectomy is ineffective as the red cells are coated with C3b and destruction occurs mainly in the liver.

Paroxysmal cold haemoglobinuria (PCH)

This rare syndrome occurs in children following acute viral infections or childhood immunisations. A history of cold exposure is not always present and presentation is acute with significant intravascular haemolysis resulting in pallor, dark urine (haemoglobinuria), abdominal pain and sometimes collapse. The original cases were described by Donath, Landsteiner and Ehrlich in congenital and tertiary syphilis. The cold antibody is a biphasic polyclonal IgG antibody (Donath–Landsteiner antibody) that reacts with the P antigen on red cells below 20 °C

in the peripheral circulation, with subsequent lysis by complement activation on rewarming to 37 °C in the central vessels. Treatment involves strict avoidance of cold and keeping the patient warm, preferably at an ambient temperature of 37 °C until haemolysis, which is usually self-limiting, subsides. This is usually sufficient, but on rare occasions, brisk haemolysis causing symptomatic anaemia occurs and blood transfusion is required. The P antigen is a high incidence antigen and therefore most blood donors are P positive, and obtaining compatible P negative or pp donor red blood cells for transfusion is difficult. Despite this *in-vitro* serologic incompatibility, there are several case reports of successful transfusion of P positive red cells to patients with PCH who achieve a rise in Hb level without precipitation of further haemolysis. In emergency cases associated with symptomatic anaemia despite supportive measures, transfusion of ABO- and Rh-compatible P-positive blood should be used. It is important that the blood is transfused using a blood warmer or where this is not available, that the patient is kept warm and transfused slowly.

Alloimmune haemolytic anaemia

Drug-induced immune haemolytic anaemia

Antibody-induced haemolytic anaemia caused by drugs is rare but, when it occurs, it can result in acute, brisk and potentially life-threatening haemolysis. Drugs typically have a molecular weight that is too low to be immunogenic (hapten), unless they are conjugated with a larger carrier molecule such as a protein, which then allows them to elicit an immune response. The diagnosis of drug-induced immune haemolytic anaemia should be made in three stages: (i) diagnosis of a DAT-positive haemolytic anaemia, (ii) careful drug history and (iii) serological demonstration of drug-specific antibody, which interacts with red cells.

Pathogenesis

Four main mechanisms have been proposed for how drugs induce antibody-dependent haemolytic anaemia; however, the same drug at different doses or repeated usage may activate different mechanisms and there are recent suggestions that membrane modification may underlie most of the mechanisms.

Drug adsorption and extravascular haemolysis

Penicillin is the prototype drug, although other penicillin derivatives and cephalosporins have also been implicated. These drugs form hapten–carrier complexes with plasma proteins enhancing drug-specific antibody production such that an estimated 90% of individuals receiving penicillin will produce clinically insignificant IgM antipenicillin antibodies. When high-dose intravenous penicillin is administered, the drug is adsorbed onto the red cell surface and becomes non-specifically attached to cell-surface proteins. A minority of patients on high-dose intravenous penicillin therapy (>1 million units daily) develop high-titre IgG antipenicillin antibodies that attach to the drug bound to the red cell surface causing predominantly extravascular haemolysis. Typically, mild to moderate haemolysis occurs, but if unrecognized and large doses of drug are continued in the presence of increasing antibody levels, complement fixation and acute intravascular haemolysis may occur. The DAT, due to IgG on the red cell surface becomes positive some weeks following treatment. When the drug is stopped, haemolysis stops and the DAT rapidly becomes negative. Antibody in the patient's serum or eluate from the red cells will react with normal red cells only in the presence of the drug. The clinical and serological features are shown in Table 9.3.

Immune complex and complement-activated acute intravascular haemolysis

Several drugs have been reported to cause immune haemolytic anaemia by this mechanism. Most frequently reported are rifampicin, phenacetin, quinine, quinidine, hydrochlorothiazide, chlorpropamide and, more recently, intravenous cephalosporins and diclofenac (see below). Hapten–carrier complexes are formed between these drugs and plasma proteins, leading to the production of drug-specific antibodies. Once drug antibodies are present, reintroduction of the drug causes formation of immune complexes that are adsorbed onto the red cell membrane and activated complement. Classically, haemolysis occurs on the second or subsequent exposure to the drug and may develop within minutes or hours of drug ingestion. Severe intravascular haemolysis may occur with fever, rigors and, in extreme cases, acute renal failure. Several groups have reported fatal immune haemolysis with the third-generation cephalosporin, ceftriaxone, whilst others from the same group such as cefotaxime and ceftazidime have also been implicated. There are far fewer reports with the second-generation cephalosporins. Diclofenac can cause an immune haemolytic anaemia with intravascular haemolysis, thought to be mediated by both immune complex and drug adsorption mechanisms.

Membrane modification

Cephalosporin, in addition to the drug adsorption mechanism, can cause a positive DAT by modifying red cell membrane components. By this mechanism, a variety of plasma proteins, including immunoglobulin and complement may attach via a non-immune mechanism to the red cell membrane, resulting in a positive DAT, but rarely causing haemolysis. Cisplatin, carboplatin and ribavirin in combination with interferon-α (IFN-α) or with pegylated IFN-α can all cause haemolytic anaemia by membrane modification. Ribavirin enters cells, is phosphorylated to ribavirin triphosphate and hydrolysed back to ribavirin in nucleated cells, which leaves the cells as the serum level declines. Red cells lack hydrolyzing enzymes so the triphosphate accumulates, alters the membrane and extravascular haemolysis occurs. With Ribavirin monotherapy, haemolysis may be compensated by increased marrow activity, but in combination with IFN-α,

Table 9.3 Drug-induced immune haemolytic anaemias: clinical and serological features.

	Drug adsorption mechanism	Immune complex mechanism	Autoimmune mechanism	Membrane modification mechanism
Examples	Penicillin Cephalosporins	Third-generation cephalosporins Quinidine Diclofenac	Methyldopa Procainamide Mefenamic acid Fludarabine* Cladribine*	Cephalosporins Cisplatin Carboplatin
Dose/duration	Large therapeutic doses/prolonged	Very low dose on second or subsequent exposure/short	Therapeutic about 6 weeks	Therapeutic
Haemolysis	Extravascular Subacute	Intravascular Acute	Extravascular Mild/subacute	Rare
DAT	IgG ± C′3	C′3 only	IgG only	IgG
Serum reaction	To drug-treated cells	Only in presence of drug or metabolite	To normal cells	To drug-treated cells
Eluate reaction	To drug-treated cells	Non-reactive	To normal cells	To drug-treated cells

*May change T-cell regulation.

bone marrow suppression by the latter agent inhibits the compensatory response. In this setting, management involves dose reduction (<800 mg/day) or supportive therapy with recombinant erythropoietin (EPO), which may allow continuation of therapy in 10–30% of patients.

Autoimmune mechanism

Methyldopa was the paradigmatic drug in drug-induced AIHA. There was a delay of some 6 weeks before the DAT became strongly positive due to IgG on the red cell surface. Haemolysis was absent or trivial, although this is not true with some other drugs that produce haemolysis by this route, notably mefenamic acid (Ponstan). The antibodies usually show no Rh specificity when tested against Rh_{null} cells. It is important to note that some drugs produce haemolysis by both the immune complex and autoimmune mechanisms, depending on the circumstances.

Diagnosis and treatment of drug-induced immune haemolytic anaemia

Patients present with clinical signs and symptoms similar to any other immune haemolytic anaemia; with acute haemolysis soon after receiving the drug or a mild haemolysis months after receiving the putative drug, e.g. methyldopa-induced haemolysis. Patients will usually have received the drugs previously without any clinical problem. The DAT is typically positive, but may be negative if performed immediately after a brisk episode of haemolysis or after transfusion. When haemolysis is dependent on drug adsorption and membrane modification mechanisms, the DAT is usually positive with IgG1, or IgG and C3 and the red cell eluate and serum do not react against normal or enzyme-modified red cells. The warm-reacting drug-specific antibodies in the eluate will only react to red cells that have been preincubated with the appropriate drug. The DAT is also positive where drug-induced haemolysis occurred via the immune complex mechanism. The red cell eluate here is not reactive in the presence of the drug and the drug-specific antibody is best detected by preincubating the patient's serum with the drug in solution to allow immune complexes to form. The preincubated serum is then tested against normal and enzyme-modified groups of cells in the presence of fresh complement. In some cases the antibodies may be specific for metabolites rather than for the parent drug. Drug metabolite antibodies may be detected by preincubating drug metabolite obtained from the serum or urine of a volunteer (who has taken the drug) with the patient's serum. A simplified summary of the serological investigation of a patient with suspected drug-induced immune haemolysis is shown in Table 9.3. The mainstay and often only required treatment is to stop the suspect drug and to consider transfusion where clinically indicated.

Alloimmune haemolytic anaemia with anti-D

Severe, acute intravascular haemolysis has been reported in some patients, 1–24 hours following treatment with a number of intravenous anti-D immunoglobulin preparations, which have been licensed for treatment of ITP in unsplenectomized RhD-positive patients. In this setting, higher doses are used than for prevention of haemolytic disease of the newborn. The mechanism of action is thought to be inhibition of platelet destruction by the spleen due to phagocytosis of coated red cells. Intravascular haemolysis may lead to oliguria, disseminated

Table 9.4 Non-immune acquired haemolytic anaemias.

Cause	Mechanisms	Examples
Infections	Intracellular organisms	*Falciparum* malaria Babesiosis *Bartonella*
	Endotoxin-induced DIC	Meningococcal sepsis Pneumococcal sepsis Gram-negative sepsis
	Haemophagocytic syndromes	Atypical mycobacterial infections HIV Viruses
	Enzyme toxins	*Clostridium perfringens* Snake, spider bites
Chemical and physical agents	Oxidative damage	Drugs Industrial/domestic substances
	Heat	Burns
	Osmotic lysis (fresh water), dehydration of red cells (salt water)	Drowning
	Enzyme inhibition	Lead poisoning Copper (Wilson's disease)
Fragmentation (mechanical)	Lysis on prosthetic surfaces	Cardiac haemolysis Perivalvular leak
	Vasculitis, endothelial cell swelling, fibrin shear	Microangiopathic haemolytic anaemia March haemoglobinuria
Acquired membrane disorders	Lipid or cholesterol changes Somatic mutation	Liver disease Paroxysmal nocturnal haemoglobinuria (PNH)

DIC, disseminated intravascular coagulation.

intravascular coagulation (DIC) and a number of deaths have been reported. The reaction is similar to that seen following mismatched transfusion reactions (see Chapter 13), but so far no culpable antibodies other than anti-D have been detected in different batches of the anti-D preparations tested. The reaction does not happen at the doses used to prevent Rh sensitization in pregnancy.

Non-immune acquired haemolytic anaemias

Haemolysis and haemolytic anaemia may be the consequence of a wide variety of acquired conditions that do not lend themselves to a precise and logical classification. Classification tends to be based on causes rather than mechanisms, although there are some common pathogenetic mechanisms that lead to red cell destruction. The main groups of agents causing haemolysis are infections, vascular disorders (mechanical disorders), chemical and physical agents, and disorders affecting the red cell membrane. A classification is shown in Table 9.4.

Infections causing haemolytic anaemia

A variety of infections may produce haemolysis through several different pathways. Haemolysis may be a consequence of direct invasion of the red cell by a microorganism or may arise from alterations in the microcirculation, leading to mechanical haemolysis. The intracellular organisms tend to produce the more severe haemolysis.

Malaria

Some degree of haemolysis is seen in all types of malarial infection, but the most severe abnormalities are found in *Plasmodium falciparum* infection (Figure 9.3; Chapter 49). *Plasmodium falciparum* infection is one of the most common causes of anaemia in the world. Many factors may contribute to the anaemia, including marrow suppression, dyserythropoiesis,

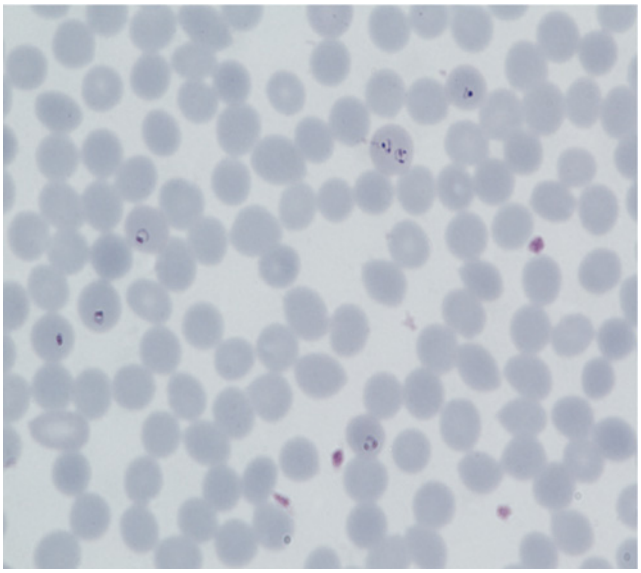


Figure 9.3 *Falciparum* malaria infection (Giemsa × 100) (Source: Dr Robin Ireland. Reproduced with permission).

folate deficiency, hypersplenism and red cell sequestration, as well as haemolysis. The condition has two main components: extravascular destruction of parasitized cells in the reticuloendothelial system, particularly the spleen, and intravascular lysis when the sporozoites break out of the red cells into the circulation. In most patients, the systemic symptoms of malaria dominate the clinical picture, but occasionally, acute intravascular haemolysis is the presenting problem. Haemolysis in malaria is often associated with a positive DAT.

Blackwater fever

This is an uncommon, but well described and feared, complication of *falciparum* malaria with an acute and potentially fatal episode of intravascular haemolysis leading to the passing of black or dark-red urine (see also Chapter 49). Occurring after a few days of malaria fever, the patient presents with further fever and rigors, loin pain and appearance of black urine. Acute hypovolaemia and hypotension may result in oliguric renal failure if not promptly managed and in some cases, pulmonary and cerebral symptoms may develop. Blackwater fever was first described in the Caucasian population, most of whom had been treated with quinine and it was thought that this was a drug-induced complication. The condition has since been reported in all ethnic populations in endemic areas and does not seem to be confined to non-immune individuals. In indigenous populations, glucose-6-phosphate dehydrogenase (G6PD) deficiency may play a part in the pathogenesis in addition to quinine exposure. The spread of chloroquine-resistant malaria in the Far East led to increased use of quinine and subsequently, to an increase in the incidence of blackwater fever.

The degree of parasitaemia is very variable. In about half the cases, the parasite count may be high, whereas in others the count may be low, perhaps following the intense intravascular haemolysis. The red cell count may fall to 1×10^{12}/L within 24 hours of the start of the haemoglobinuria. There is usually a rise in fibrin degradation products in the serum, but this rise is not often marked and is compatible with a degree of renal failure. Intravascular coagulation does not seem to play a major role in the pathogenesis. Immediate treatment includes correction of fluid and electrolyte loss, transfusion support and eradication of the parasite. Renal dialysis may be required and it is important to note that renal supportive therapy may need to continue for a period of weeks before renal function returns. Subsequent attacks of *falciparum* malaria may produce further episodes of blackwater fever making adherence to prophylaxis mandatory in susceptible individuals.

Babesiosis

Infection with the intracellular protozoan *Babesia* is uncommon and symptomatic disease is mostly confined to splenectomized patients. *Babesia* is a tick-borne organism; the tick in Europe, *Ixodes ricinus*, is associated with cattle, whilst in North America, *Ixodes dammini* is carried by rodents and deer. Transmission can also occur by blood transfusion. In splenectomized patients, the disease has an acute onset and in the European variety, can be fatal. A period of malaise for 1–3 days may be followed by vomiting, diarrhoea, high fevers, rigors, jaundice, acute intravascular haemolysis, haemoglobinuria, renal failure and death. In North America, unsplenectomized patients tend to experience a milder self-limiting haemolytic anaemia. The diagnosis is made from identification of the parasites, very similar to *P. falciparum*, within the red cells in the peripheral blood film. A history of tick bite or of exposure to potential vectors should be sought. Treatment is difficult. Clindamycin and quinine are standard therapy, whilst red cell exchange transfusion and renal support may be required in severely affected patients.

Bartonella (Oroya fever)

Bartonella bacilliformis is an arthropod-transmitted infection found only in the western Andes of Peru and neighbouring countries. The diagnosis is made from the peripheral blood film. The organism is an intracellular gram-negative rod during the acute attack, which becomes coccoid in recovery. In non-immune individuals presentation is with anaemia, which occurs partly through intravascular haemolysis and partly through erythrophagocytosis as well as splenomegaly. The organism is rapidly killed by chloramphenicol, tetracyclines, penicillin and aminoglycosides.

Clostridium perfringens

Clostridium perfringens septicaemia causes an intense intravascular haemolysis with prominent microspherocytosis and ghost cells in the peripheral blood film. The spherocytosis

is the result of membrane destruction by lipases and proteases produced by the organism. Although the organism is sensitive to a variety of antibiotics, the appearance of intravascular haemolysis is usually a harbinger of death because of the toxaemia.

Toxoplasmosis

Infection with *Toxoplasma gondii* acquired *in utero* may produce haemolysis and a syndrome similar to haemolytic disease of the newborn. In adults, toxoplasmosis is not associated with haemolysis, except perhaps in the immunocompromised host.

Bacterial infections

Intravascular coagulation produced by bacterial infection may be accompanied by some degree of intravascular haemolysis with fragmentation of red cells. Septicaemia from meningococcal or pneumococcal infection may result in haemolysis, which is not as clinically significant as the other effects of the septicaemia.

Haemorrhagic fevers

Haemorrhagic fevers may be accompanied by haemolysis. These include Dengue fever, which is present worldwide; Yellow fever and West African haemorrhagic fevers, which are mainly found in Africa.

Haemophagocytic syndrome

Haemophagocytic syndrome (HPS) is characterized by proliferation of macrophages in the bone marrow, spleen, liver and lymph nodes which cause inappropriate phagocytosis of erythroid precursors, granulocytes, platelets and in some variations, the skin. It has been reported in severe systemic infections including cytomegalovirus, tuberculosis and some fungal infections (Chapter 14). Onset may be acute or subacute. Clinically, patients present with malaise, persistent fever, anaemia, jaundice, weight loss and there is laboratory evidence of pancytopenia, liver dysfunction and coagulopathy. The jaundice results from the destruction of red cells and their precursors in the marrow, spleen or liver and is associated with a marked rise in LDH. There is a marked acute-phase response with greatly elevated serum ferritin levels (>20 000 μg/L) and increases in interferon (IFN)-γ and tumour necrosis factor (TNF)-α, with variable changes in other cytokines. HPS is associated with abnormal T-cell activation, which triggers the macrophage response and may be the consequence of an underlying T-cell lymphoma or a yet unmasked infection. Clinically, the distinction between malignant- or infection-related HPS may be difficult because in the lymphomas proliferation may be trivial, the syndrome being derived from the release of cytokines, and because in malignant HPS superadded infection is common. It may be impossible to identify the underlying infection in infection-associated HPS. As a general rule, infection-associated HPS tends to be more common in children, whilst it is more likely to be associated with lymphomas in adults.

Fragmentation haemolysis: mechanical haemolytic anaemias

Fragmentation haemolysis may occur in the presence of prosthetic materials, due to altered flow following cardiovascular surgery, due to trapping or adherence of red cells in arteriovenous malformations and as a result of the destruction of red cells in pathologically altered small blood vessels (microangiopathic haemolytic anaemia). Characteristic features of fragmentation haemolysis are the appearance of the blood film (Figure 9.4) and the presence of intravascular haemolysis. Depending on the underlying vascular pathology, there may be a reduction in the platelet count and evidence of DIC. The rate of red cell destruction also varies according to the pathogenesis, so the signs of intravascular haemolysis vary from absence of haptoglobin, elevated LDH and minimal haemosiderinuria to acute intravascular destruction with haemoglobinaemia and haemoglobinuria. The major causes of fragmentation haemolysis are shown in Table 9.5.

Haemolysis associated with cardiac surgery

Cardiac haemolytic anaemia was a term coined to describe haemolysis following cardiac surgery that involved the insertion of prosthetic valves, patches or grafts. Mechanical trauma to red blood cells is the primary cause of haemolysis in this setting and is mainly due to increased turbulent flow resulting in excessive shearing forces on the surface of the red cells. Secondary physiologic mechanisms include pressure fluctuations, intrinsic abnormalities of the red cell membrane (largely due to fragile, iron-poor red cells in iron-deficient patients), interaction with foreign surfaces and unfavourable flow characteristics of valves. Clinical

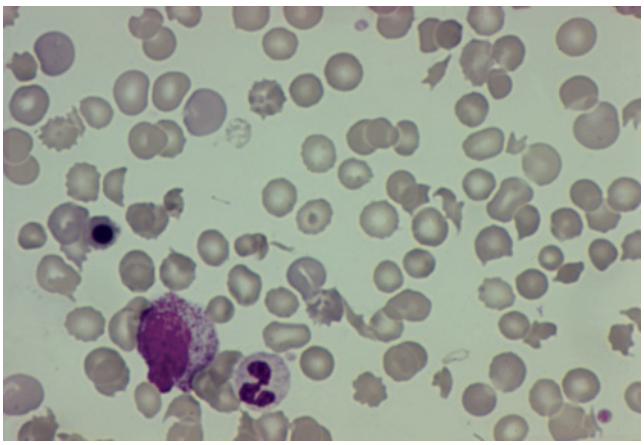


Figure 9.4 Microangiopathic haemolytic anaemia. Blood film from a patient with carcinoma and bone marrow metastases. Note fragmentation of red cells, low platelets and leucoerythroblastic changes (circulating nucleated red cell and metamyelocyte) suggestive of underlying disseminated intravascular coagulation (×40).

Table 9.5 Classification of anaemias caused by fragmentation haemolysis.

Vascular origin	Pathogenesis	Platelets
Cardiac haemolysis	Prosthetic heart valves	Normal
	Patches, grafts	
	Periprosthetic or perivalvular leaks	
Arteriovenous malformations	Kasabach–Merritt syndrome	Very low
	Malignant haemangio-endotheliomas	
Microangiopathic (see also Table 9.6)	TTP/HUS	Low
	Malignant disease	Normal/low
	Vasculitis	Normal/low
	Pre-eclampsia, HELLP	Low
	Renal vascular disorders	Normal/low
	Disseminated intravascular coagulation	Low

HELLP, haemolysis with elevated liver function tests and low platelets; HUS, haemolytic–uraemic syndrome; TTP, thrombotic thrombocytopenic purpura.

situations in which cardiac haemolysis may be of considerable importance include:

- *Periprosthetic or perivalvular* following insertion or repair of a heart valve, a leak occurs around the valve or through a suture track, which may cause severe intravascular haemolysis without haemodynamic distress. Laboratory diagnosis may be difficult because red cell fragmentation may not be prominent on the blood film, although spherocytes may be present. The diagnosis is made by exclusion of an autoimmune haemolysis in a patient who experiences intravascular haemolysis following recent cardiac surgery. The haemolysis can only be cured by further surgery.
- *Ambulatory haemolysis* occurs when a patient who has undergone valve replacement shows evidence of mild haemolysis while in hospital but experiences significant anaemia after discharge. The higher cardiac output associated with greater exercise following discharge from hospital produces greater turbulence and results in red cell fragmentation. A similar mechanism may occur if the patient becomes iron deficient as a result of chronic intravascular haemolysis. Iron replacement and advice about slowly increasing exercise levels may prevent the need for further surgery.
- *Cardiopulmonary postperfusion syndrome*: Acute intravascular haemolysis may occur in patients who have undergone cardiopulmonary bypass surgery. The haemolysis may be accompanied by neutropenia and pulmonary distress. The syndrome does not strictly belong in this section as the haemolysis seems to be caused by complement activation and binding of the membrane attack complex to the red cell surface. The blood film shows ghost red cells rather than fragmentation. The condition is self-limiting and the patient requires only supportive care.

Arteriovenous malformation

Fragmentation of red cells may be seen in Kasabach–Merritt syndrome, in which platelets are trapped in the vascular network of giant arteriovenous malformations, sometimes with evidence of a consumption coagulopathy. The bleeding disorder that ensues is of greater significance than haemolysis in these patients. A similar pattern is seen in malignant haemangioendothelioma, in which the tumour invades and grows alongside veins resulting in a consumptive coagulopathy with laboratory evidence of DIC.

Microangiopathic haemolytic anaemias (MAHA)

A condition in which intravascular haemolysis with fragmentation of red cells is caused by their destruction in an abnormal microcirculation. Proof of microangiopathy may be lacking in those not subjected to a post mortem, and MAHA should be considered a clinical syndrome. The three main pathological lesions that give rise to MAHA are deposition of fibrin strands, often associated with DIC, platelet adherence and aggregation, and vasculitis. The vessel abnormalities may be generalized or confined to particular sites or organs. In most cases, haemolysis is of less consequence than the underlying cause of the microangiopathy, but fragmentation of red cells helps to confirm the diagnosis. Some of the disorders producing MAHA are given in Table 9.6 and the more well-defined clinical syndromes are described in detail here.

MAHA and malignant disease

Fragmentation of red cells with chronic intravascular haemolysis may occur in malignant disease. Clinically significant anaemia may occur, especially when the tumour invades large blood vessels, as occurs in haemangiopericytoma, but more commonly haemolysis is trivial or well compensated. Other malignancies closely associated with MAHA include mucin-secreting tumours and acute leukaemia, particularly, but not exclusively, promyelocytic leukaemia, where acute intravascular coagulation occurs and dominates the clinical picture. A blood film showing evidence of MAHA together with leucoerythroblastic changes is virtually diagnostic of malignant disease with secondary deposits in the bone marrow (Figure 9.4).

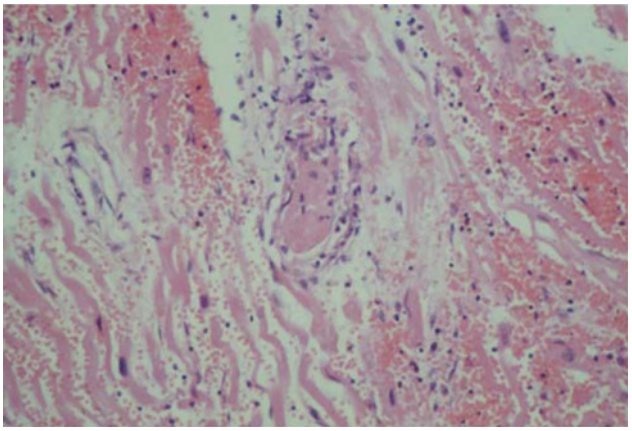
MAHA and infection

Infections, particularly septicaemia, may provoke intravascular coagulation and MAHA. Generally, the coagulation changes and septic shock overshadow the mild fragmentation, but, occasionally, infections produce a chronic state of partially compensated intravascular haemolysis and marked red cell fragmentation.

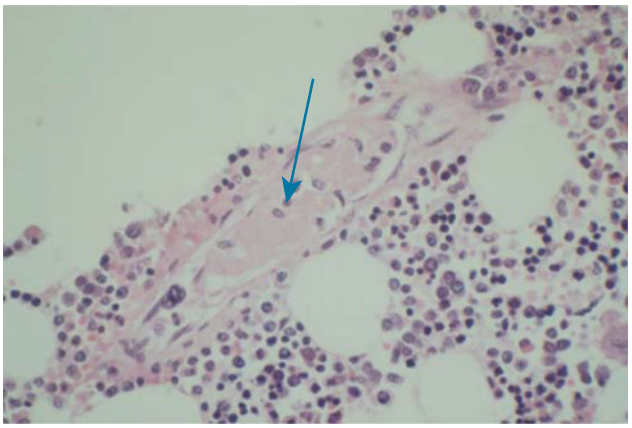
Table 9.6 Causes of microangiopathic haemolytic anaemia.

Disease	Microangiopathy
Haemolytic–uraemic syndrome	Endothelial cell swelling, microthrombi in renal vessels
Thrombotic thrombocytopenic purpura	Platelet plugs, microaneurysms, small-vessel thrombi
Renal cortical necrosis	Necrotizing arteritis
Acute glomerular nephritis	
Malignant hypertension	
Pre-eclampsia	Fibrinoid necrosis
HELLP	
Polyarteritis nodosa	Vasculitis
Wegener granulomatosis	
Systemic lupus erythematosus	
Homograft rejection	Microthrombi in transplanted organ
Mitomycin C	Uncertain
Ciclosporin	Renal vessel anomalies
Carcinomatosis	Abnormal tumour vessels, intravascular coagulation (disseminated or localized)
Primary pulmonary hypertension	Abnormal vasculature
Cavernous haemangioma (Kasabach–Merritt)	Local vascular changes, thrombosis

HELLP, haemolysis with elevated liver function tests and low platelets.



(a)



(b)

Figure 9.5 Thrombotic thrombocytopenic purpura. Microthrombi in capillaries: (a) section from the myocardium; (b) thrombus in a bone marrow capillary indicated by blue arrow (haemotoxylin and eosin ×100) (Source: Dr Margaret Burke. Reproduced with permission).

Haemolytic–uraemic syndrome (HUS) may be precipitated by infection, particularly *Escherichia coli* 0157. HUS is discussed in Chapter 44.

Thrombotic thrombocytopenic purpura

Thrombotic thrombocytopenic purpura (TTP) is an acute syndrome characterized by fever, neurological signs, haemolytic anaemia with fragmented red cells and profound thrombocytopenia. There is severe deficiency of von Willebrand factor cleaving protease (VWFCP; also known as ADAMTS13) to less than 10% of normal activity due to formation of auto-antibodies directed against it. The diagnosis is made on the basis of the clinical presentation and evidence for haemolytic anaemia with fragmented red cells and thrombocytopenia. It can be confirmed with an assay which confirms low ADAMTS13 level. Bilirubin and serum LDH are elevated, indicating intravascular haemolysis. LDH and persistence of red cell fragments in the blood film are useful markers for the microangiopathic process. It may take up to 1 week after the haemolytic process stops for all fragments to be removed from the circulation. The destruction of red cells occurs at the site of intravascular occlusions; at post mortem, platelet and fibrin plugs are found in capillaries (Figure 9.5). TTP requires rapid diagnosis and urgent commencement of plasmapheresis, otherwise it is a potentially fatal condition. Increasingly rituximab is used in patients with suboptimal response to standard therapy. This condition, including its pathophysiology, is discussed in detail in Chapter 43.

March haemoglobinuria

Haemoglobinuria following running has been documented for about 100 years. Its origin is mechanical, with destruction of red cells occurring in the feet. It can be cured by wearing soft shoes or running on soft ground. The disorder may arise in joggers and is benign except that it may lead to extensive invasive investigations unless recognized. The blood film does not show any red cell fragmentation or consistent abnormality. Occasionally, haemoglobinuria after running is accompanied by nausea,

Table 9.7 Substances causing oxidative haemolysis and/or methaemoglobinaemia in normal people.

Substance	Use	Remarks
Dapsone	Leprosy, dermatitis herpetiformis	Chronic haemolysis; slow acetylators more susceptible
Maloprim	Antimalarial	Methaemoglobinaemia in NADH methaemoglobin reductase-depleted subjects
Sulfasalazine	Ulcerative colitis	Chronic intravascular haemolysis
Phenazopyridine	Analgesic in urinary tract infections	Methaemoglobinaemia
Menadiol	Water-soluble vitamin K analogue	Haemolysis/kernicterus in infants
Nitrites	Fertilizer; present in well water, some vegetable juices	Methaemoglobinaemia in infants
Nitrates	Amyl nitrate, butyryl nitrite; abused recreationally	Methaemoglobinaemia/haemolysis
Chlorate	Weed-killer	Acute intravascular haemolysis; renal failure; >30 g fatal
Arsine	Gas produced in smelting and other industrial processes	Acute intravascular haemolysis; renal failure

abdominal cramps and aching legs, and enthusiastic athletes with this condition may exhibit mild splenomegaly and jaundice.

Chemical and physical agents

Oxidative haemolysis

Oxidative substances may cause haemolysis in people with normal red cell metabolism and normal HbA if the oxidative stimulus is large enough. The major causes of oxidative haemolysis in normal subjects are shown in Table 9.7. The clinical features of this condition are dependent on the main sites of oxidative attack: whether on the membrane of the red cell (intravascular haemolysis), the globin chains (Heinz body formation) or the haem group (methaemoglobin accumulation).

Dapsone and sulfasalazine

Dapsone and sulfasalazine (Salazopyrin) will cause oxidative intravascular haemolysis in normal subjects if taken in sufficiently high dosage. Normal doses will cause significant oxidative haemolysis when given to G6PD-deficient subjects with leprosy. Blood film shows polychromasia, macrocytosis and the classical 'bite' abnormality of the chemically damaged cell (Figure 9.6). When used in the treatment of dermatitis herpetiformis, a chronic blistering skin condition closely associated with coeliac disease and functional hyposplenism, Heinz bodies appear on the blood film. In contrast, Heinz bodies are absent or scanty in the blood film of patients with an intact spleen. Haemolysis is usually well compensated and there is no need to stop the treatment unless the anaemia is severe. Methaemoglobinaemia is uncommon unless the patient is partially deficient in NADH methaemoglobin reductase. The affected gene may not be that rare which may account for cyanosis in a number of individuals treated with dapsone-containing antimalarial preparations.

Methaemoglobinaemia

Methaemoglobinaemia may be caused by nitrites in water and by vegetable juices in infants with physiological impairment of the reducing systems. Well-water that comes from land with an excess of nitrites and which is used to reconstitute artificial feeds has produced cyanosis in infants. Cases have also been described following the enthusiastic feeding of juice from carrots grown on organically fertilized land and of spinach juice (spinach has a high concentration of nitrogen-fixing bacteria on its leaves). Nitrate drugs, for example amyl nitrate, also produce methaemoglobinaemia and have proved fatal when taken in sufficiently high dosage for 'recreational' purposes. Water-soluble analogues of vitamin K (menadiol sodium diphosphate) cause haemolysis with or without methaemoglobinaemia in infants and *in-utero* if given to the mother during the third trimester. Fat-soluble vitamin K preparations must be used

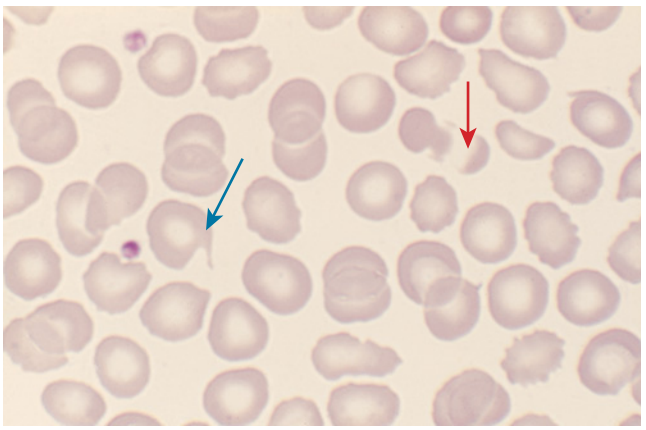


Figure 9.6 Oxidative haemolysis caused by drug (phenacetin). Note red cells with bite cells (blue arrow) and contracted haemoglobin (red arrow).

if required in these situations. Methaemoglobinaemia due to oxidative drugs rarely causes problems with oxygen delivery, but severe cases (>40% methaemoglobin) may be treated with intravenous methylene blue 1–2 mg/kg. Ascorbic acid by mouth may also be used. These measures are ineffective in G6PD-deficient patients and when very strong oxidant substances are implicated. In these circumstances, methylene blue should be avoided because it acts as an oxidant and makes the condition worse.

Methaemoglobinaemia following exposure to strong oxidizing substances may cause acute intravascular haemolysis and renal failure. These substances are found mainly in industrial or horticultural settings and include sodium chlorate, a popular weed-killer, and arsine, a gas produced in various industrial situations such as dross removal in smelting procedures and drain clearance where arsenic is a contaminant. Acute intravascular haemolysis and haemoglobinuria develop 1–24 hours after exposure, depending on the dose. The serum becomes brown, often very dark, so that blood cells cannot be seen in anticoagulated preparations, due to the presence of methaemalbumin, methaemoglobin and free haemoglobin. Vomiting, abdominal pain and oliguric renal failure usually develop over about 24 hours. The blood film shows microspherocytosis, ghost cells and bizarre forms. Urgent treatment is required with plasma exchange and renal dialysis being the mainstay of treatment, however methylene blue is ineffective. Poisoning with arsine is usually reversible with these measures. The management of chlorate poisoning is more difficult with doses of 30 g or more being fatal. Chlorate is mostly ingested deliberately in suicide attempts.

Thermal injury

Normal red cells when heated *in vitro* to 46 °C for 1 hour show no changes, however they show temperature- and duration-dependent changes above 47–50 °C. Some hereditary membrane defects produce red cells that have increased thermal fragility (see Chapter 8).

Severe burns

Severe burns may be accompanied by intravascular haemolysis with haemoglobinuria. The intravascular haemolysis is related to the extent and severity of the burns. The gross haemoglobinuria occurs over the first 24 hours after the burns and ceases thereafter. The blood film shows spherocytosis and schistocytes, the morphological abnormalities reflecting the thermal damage and the amount of lysis. Prolonged anaemia after burning is related to inflammation, occult blood loss and infection rather than haemolysis.

Lead poisoning

Haemolysis is an important contributor to the anaemia associated with excessive lead exposure. Lead is a potent inhibitor of pyrimidine 5′-nucleotidase 1 (P5N1). This is the prime mechanism of lead-induced haemolysis, producing the same gross basophilic stippling of the red cells as seen in the inherited deficiency of the enzyme (see Chapter 9), P5N1 activity is a good surrogate marker for lead exposure.

Acquired disorders of the red cell membrane

The most common acquired disorder is paroxysmal nocturnal haemoglobinuria (PNH), caused by somatic mutation of the phosphatidylinositol glycan A (*PIGA*) gene on the X chromosome, which leads to failure to produce the glycosylphosphatidylinositol (GPI) anchor needed to transport and attach many proteins to the red cell membrane. Intravascular haemolysis occurs through the unchecked action of activated complement. PNH is described in detail in Chapter 11.

The mature red cell does not have the capacity to repair its membrane. The lipids of the membrane are in equilibrium with the lipids of the plasma and changes in the ratio of free cholesterol to phospholipids in plasma may affect red cell shape and, in some instances, lead to haemolysis. This is most commonly seen in liver disease, but other inherited lipid disorders may affect the red cell secondarily.

Liver disease

Some degree of shortening of red cell survival occurs in most cases of acute hepatitis, cirrhosis and Gilbert disease, but anaemia is not present and there is only a slight rise in reticulocytes. Biliary obstruction is associated with the appearance of target cells and fulminant hepatitis with acanthocytosis, both due to changes in plasma lipid composition.

Zieve's syndrome

Zieve's syndrome is an uncommon disorder seen mainly in alcoholics. It comprises intravascular haemolysis and acute abdominal pain in patients with jaundice and cirrhosis. The cause is unknown, but is probably related to lipid changes in the blood. Spherocytes are seen in the peripheral blood.

Wilson's disease

Wilson's disease may present as acute intravascular haemolysis. This is probably not a membrane disorder, but is consequent on the high levels of copper ions in the blood. The haemolysis may antedate development of hepatic or neurological features, but Kayser–Fleischer rings are usually present. The blood film may show spherocytosis. The diagnosis is made once the condition is suspected. Apart from caeruloplasmin deficiency, patients have a specific aminoaciduria.

Hereditary acanthocytosis (abetalipoproteinaemia)

This rare inherited deficiency of low-density lipoproteins is characterized by retinitis pigmentosa, steatorrhoea, ataxia and mental retardation. The haemolysis that occurs is of minor

importance to such patients, but the blood film may indicate the diagnosis, with the red cells showing marked acanthocytosis.

Vitamin E deficiency

Deficiency of vitamin E may occur in premature and other infants who are fed a commercial formula that is rich in polyunsaturated fatty acids. Vitamin E is an antioxidant and oxidative damage to the red cell membrane is thought to be the cause of the haemolysis. Clinically, generalized oedema associated with haemolytic anaemia is reported in association with low levels of vitamin E in premature infants aged between 4–8 weeks. There is laboratory evidence of haemolysis with contracted red cells and a thrombocytosis, all of which are corrected following adequate vitamin E replacement therapy.

Selected bibliography

General

Dacie JV (1992) *The Autoimmune Haemolytic Anaemias: The Haemolytic Anaemias*, 3rd edn, Vol. 3. Churchill Livingstone, Edinburgh.

Petz LD, Garraty G (2004) *Acquired Immune Haemolytic Anemias*, 2nd edn. Churchill Livingstone, Philadelphia, PA.

Autoimmune haemolytic anaemias

Warm type

Birgens H, Frederiksen H, Hasselbach H *et al.* (2013) A phase III randomised trial comparing glucocorticoid monotherapy versus glucocorticoid and Rituximab in patients with autoimmune haemolytic anaemia. *British Journal of Haematology* **132:** 125–37.

Jeffries LC (1994) Transfusion therapy in autoimmune haemolytic anaemia. *Hematology/Oncology Clinics of North America* **8**: 1087–104.

Garvey B (2008) Rituximab in the treatment of autoimmune haematological disorders. *British Journal of Haematology* **141**: 149–69.

Cold type

Gertz M (2007) Management of cold haemolytic syndrome. *British Journal of Haematology* **138**: 125–37.

Berentsen S, Klaus B, Geir ET (2007) Primary chronic cold agglutinin disease: an update on pathogenesis, clinical features and therapy. *Haematology* **12**: 361–70.

Berentsen S (2013) Therapy for chronic cold agglutinin disease: perspective for further improvements. *Blood Transfusion* **11**: 167–8.

Wolach B, Heddle N, Barr RD *et al.* (1981) Transient Donath-Landsteiner haemolytic anaemia. *British Journal of Haematology* **48**: 425–34.

Drug-induced

Wright MS (1999) Drug induced haemolytic anaemias: increasing complications to therapeutic interventions. *Clinical Laboratory Science* **12**: 115–18.

Garratty G (2010) Immune hemolytic anemia associated with drug therapy. *Blood Reviews* **22**: 143–50.

Salama A (2009) Drug-induced immune haemolytic anaemia. *Expert Opinion on Drug Safety* **8**: 73–9.

Non-immune haemolytic anaemias

Davidson RJL (1969) March or exertional haemoglobinuria. *Seminars in Hematology* **6**: 150.

Bruneel F, Gachot B, Wolff M *et al.* (2001) Resurgence of blackwater fever in long-term European expatriates in Africa: report of 21 cases and review. *Clinical Infectious Diseases* **32**: 1133–40.

McArthur HL, Dalal BI, Kollmannsberger C (2006) Intravascular hemolysis as a complication of *Clostridium perfringens* sepsis. *Journal of Clinical Oncology* **24**: 2387–8.

Vannier E, Gewurz BE, Krause PJ (2008) Human babesiosis. *Infectious Disease Clinics of North America* **22**: 469–88.

Shapira Y, Vaturi M, Sagie A (2009) Hemolysis associated with prosthetic heart valves: a review. *Cardiology in Review* **17**: 121–4.

Sadler JE (2009) Von Willebrand factor, ADAMTS13, and thrombotic thrombocytopenic purpura. *Blood* **112**: 11–18.

Chemical and physical agents

Kim Y, Yoo CI, Lee CR *et al.* (2002) Evaluation of activity of erythrocyte pyrimidine 5′-nucleotidase (P5N) in lead exposed workers: with focus on the effect on hemoglobin. *Industrial Health* **40**: 23–7.

CHAPTER 10

Inherited aplastic anaemia/bone marrow failure syndromes

10

Inderjeet S Dokal

Barts and The London School of Medicine and Dentistry, Queen Mary University of London, Barts Health NHS Trust, London, UK

Introduction

A number of inherited (constitutional/genetic) disorders are characterized by aplastic anaemia (AA)/bone marrow failure (BMF), usually in association with one or more somatic abnormalities (Table 10.1). The features of some of these are summarized in Table 10.2. The precise incidence/prevalence of these remains unclear but, collectively, they represent approximately 20–30% of patients presenting with AA and constitute a significant clinical burden, as many are associated with premature mortality. The BMF may present at birth or at a variable time thereafter, including in adulthood in some cases. The BMF may involve all lineages or a single lineage; in some cases it may initially be associated with a single cytopenia and then progress to pancytopenia. Scientifically, they constitute an important group of diseases as advances in understanding the genetics of some of these are not only unravelling their pathophysiology, but are also providing important insights into normal haemopoiesis.

The two syndromes that are frequently associated with generalized BMF/AA are Fanconi anaemia (FA) and dyskeratosis congenita (DC). These two syndromes are now also two of the best characterized, and are discussed in some detail in this chapter (followed by sections on Shwachman–Diamond syndrome, Diamond–Blackfan anaemia, congenital dyserythropoietic anaemia, congenital neutropenia, thrombocytopenia with absent radii and congenital amegakaryocytic thrombocytopenia). This chapter discusses the clinical and genetic heterogeneity of these BMF syndromes, their management and their impact on our understanding of the pathophysiology of more common forms of 'idiopathic aplastic anaemia'. Indeed, patients with both FA and DC can sometimes present with AA alone as their initial manifestation and can thus pose a diagnostic/management challenge.

Fanconi anaemia

Clinical features

Since the first description by Guido Fanconi in 1927, FA has come to be recognized as an autosomal recessive (AR) disorder (X-linked in a rare subset) in which there is progressive BMF and an increased predisposition to malignancy, especially acute myeloid leukaemia (AML). Most, but not all, affected individuals also have one or more somatic abnormalities, including skin (café-au-lait spots), skeletal (absent thumbs, radial hypoplasia, scoliosis), genitourinary (underdeveloped gonads, horseshoe kidneys), gastrointestinal, cardiac and neurological anomalies (Table 10.3). Some of these somatic abnormalities are shown in Figure 10.1. The course of the disease and the pattern of somatic abnormalities show considerable variation, with approximately one-third of patients having no physical abnormalities. This makes diagnosis based on clinical criteria alone difficult and unreliable.

The cumulative incidence of BMF by the age of 40 years in patients with FA is ~90%. At birth, the blood count is usually normal. Pancytopenia develops insidiously and presents in most cases between the ages of 5 and 10 years (median age 7 years). However, in some cases the pancytopenia develops in adolescence or even in adult life. The haemoglobin and platelet count

Postgraduate Haematology, Seventh Edition. Edited by A Victor Hoffbrand, Douglas R Higgs, David M Keeling and Atul B Mehta.

Table 10.1 The inherited BMF syndromes.

Pancytopenia (usually associated with a global haemopoietic defect)
Fanconi anaemia
Dyskeratosis congenita
Shwachman–Diamond syndrome
Reticular dysgenesis
Pearson syndrome
Familial aplastic anaemia (autosomal and X-linked forms)
Myelodysplasia
Non-haematological syndromes (Down, Dubowitz syndromes)
Single cytopenia (usually)
Anaemia
Diamond–Blackfan anaemia
Congenital dyserythropoietic anaemia
Neutropenia
Severe congenital neutropenia including Kostmann syndrome
Thrombocytopenia
Congenital amegakaryocytic thrombocytopenia
Amegakaryocytic thrombocytopenia with absent radii

Table 10.3 Somatic abnormalities in FA.

Abnormality	Percentage of patients
Skeletal (radial ray, vertebral, scoliosis, rib)	71
Skin pigmentation (café-au-lait, hyper- and hypopigmentation)	64
Short stature	63
Eyes (microphthalmia)	38
Renal and urinary tract	34
Male genital	20
Mental retardation	16
Gastrointestinal (e.g. anorectal, duodenal atresia)	14
Heart	13
Hearing	11
Central nervous system (e.g. hydrocephalus, septum pellucidum)	8
No abnormalities	30

Source: Auerbach *et al*. Fanconi anaemia. In *The Metabolic and Molecular Basis of Inherited Disease*, 2001. Adapted with permission of McGraw-Hill.

are usually first to fall; the granulocytes are usually well preserved in the early stages. As the pancytopenia develops, the bone marrow (BM) becomes progressively hypocellular. There is often a marked increase in macrophage activity with evidence of haemophagocytosis. BMF, leading to fatal haemorrhage or infection, is the main cause of death. In an analysis of the International Fanconi Anaemia Registry (IFAR), the median survival time was 24 years; this is changing with improvements in clinical care.

FA is associated with an increased risk of leukaemia and other malignancies. The leukaemias are usually of the acute myeloid type, particularly FAB types M4 and M5 (see Chapter 20). In some cases, leukaemia may be the initial event leading to the diagnosis of FA. The cumulative incidence of

Table 10.2 Characteristics of the BMF syndromes.

	FA	DC	SDS	DBA	CDA	TAR	SCN	IAA	Other[#]
Inheritance pattern	AR, XLR	XLR, AR, AD	AR	AD, XLR	AR, AD	AR	AD, AR	?	AD and AR
Somatic abnormalities	Yes	Yes	Yes	Yes	Rare	Yes	Rare	?None	Yes
Bone marrow failure	AA (>90%)	AA (~80%)	AA (20%)	RCA	Eryth	Megs	Neutropenia	Yes (100%)	Yes
Short telomeres	Yes	Yes	Yes	?	?	?	?	Yes	No
Malignancy	Yes	Yes	Yes	Yes	?No	?No	Yes	Yes	?
Chromosome instability	Yes	Yes	Yes	?	?	?	?	Yes	?
Genes identified	16	10	1	14	4	1	5+	*	3

*Heterozygous mutations in TERC and TERT are risk factors for some cases of AA.
AA, aplastic anaemia; AD, autosomal dominant; AR, autosomal recessive; CDA, congenital dyserythropoietic anaemia; DBA, Diamond–Blackfan anaemia; DC, dyskeratosis congenita; FA, Fanconi anaemia; IAA, idiopathic aplastic anaemia; Eryth, ineffective erythropoiesis; Megs, low megakaryocytes; RCA, red cell aplasia; SCN, severe congenital neutropenia; SDS, Shwachman–Diamond syndrome; TAR, thrombocytopenia with absent radii; XLR, X-linked recessive
[#]Patients in whom genetic defects have been identified (e g. in *SRP72*, *ERCC6L2*) but they do not fit into previously recognized categories.

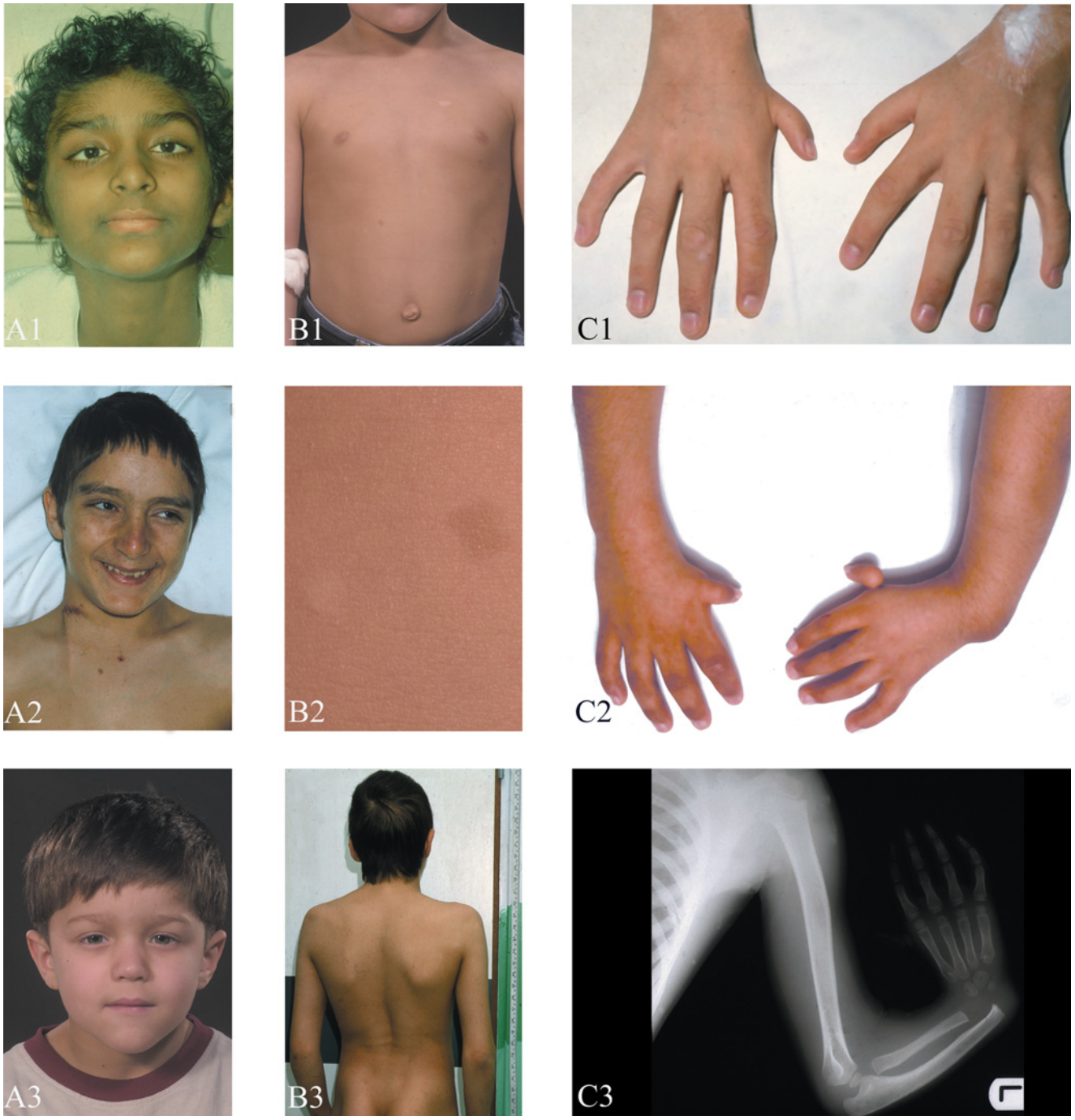


Figure 10.1 FA. (a) Photographs of patients with FA (A1–A3) with small mouth and chin ('Fanconi facies'). (b) Abnormalities of pigmentation (hyper- and hypopigmentation) on the abdomen (B1) with a close-up (B2) of a café-au-lait spot and a hypopigmented patch. The bottom photograph (B3) shows the back of an FA patient demonstrating lumbar scoliosis. (c) Hands/forearms of FA children showing hypoplastic thumbs (C1), rudimentary ('dangling') thumbs (C2) and a radiograph (C3) showing rudimentary thumb (skeletal) development.

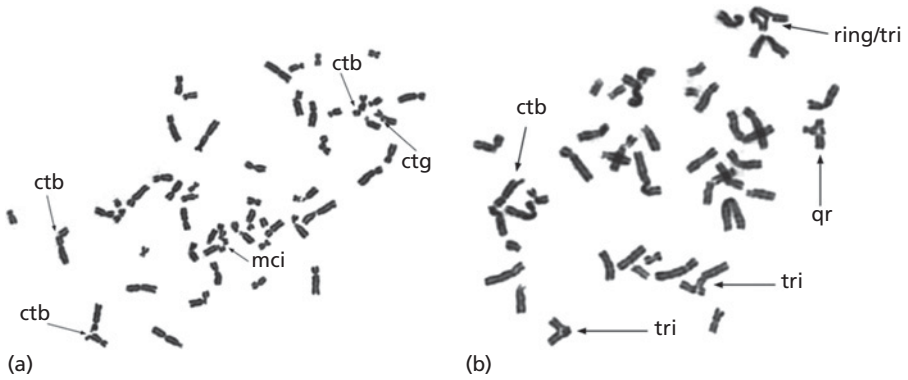


Figure 10.2 (a, b) Chromosomal abnormalities seen in FA lymphocytes following incubation with diepoxybutane. ctb, chromatid break; ctg, chromatid gap; mci, multiple chromatid interchanges (complex rearrangement); tri, triradial; qr, quadriradial. (Source: Nicola Foot, Hammersmith Hospital. Reproduced with permission.)

haematological malignancy by the age of 40 years is 33%. Besides these haematological malignancies, there is a significant risk of hepatic tumours and squamous cell carcinoma of the vulva, oesophagus, head and neck. The cumulative incidence of solid tumours is calculated to be 28% by the age of 40 years. The impression is that malignancies occur mainly in patients with late-onset BMF and longer survival, with a median age of 13 years for leukaemia and 25 years for solid tumours. Recent analysis from the German FA registry largely concurs with these earlier observations. Furthermore, long-term follow-up in FA patients who have been treated by haemopoietic stem cell transplantation (HSCT) is showing a higher incidence of non-haematological malignancies in patients with FA than patients with other types of BMF who underwent HSCT.

Cell and molecular biology

Over the last three decades many advances have been made in our understanding of the pathophysiology of FA. FA cells characteristically display a high frequency of spontaneous chromosomal breakage and hypersensitivity to DNA cross-linking agents such as diepoxybutane (DEB) and mitomycin C (MMC). This genomic instability (Figure 10.2) led to the development of a diagnostic test (i.e. increased chromosomal breakage in FA cells compared with normal controls after exposure to DEB/MMC) over two decades ago and this remains a useful FA screening test today. This 'FA cell phenotype' has also facilitated many advances in our understanding of FA, including elucidation of the complex genetics of this disease, with 16 subtypes/complementation groups currently characterized. The genes responsible for these subtypes are designated *FANCA*, *FANCB*, *FANCC*, *FANCD1*, *FANCD2*, *FANCE*, *FANCF*, *FANCG*, *FANCI*, *FANCJ*, *FANCL*, *FANCM, FANCN, FANCO*, *FANCP* and *FANCQ* (Table 10.4).

Studies from several groups have demonstrated that the proteins encoded by the FA genes participate in a complicated network important in DNA repair. Specifically, eight of the FA proteins (FANCA, FANCB, FANCC, FANCE, FANCF, FANCG, FANCL and FANCM) interact with each other and form the FA core complex (Figure 10.3). The FA core complex is required for the activation of the FANCI-FANCD2 protein complex to a monoubiquitinated form (FANCI-FANCD2-Ub). FANCI-FANCD2-Ub then interacts with DNA repair proteins (including BRCA2, BRCA1 and RAD51) leading to repair of the DNA damage. FA-D1 patients have biallelic mutations in BRCA2. These observations have linked the FA proteins with BRCA1 and BRCA2 (FANCD1) in a DNA damage response pathway 'The FA/BRCA pathway'. The BRCA2 protein is important in the repair of DNA damage by homologous recombination. Cells lacking BRCA2 inaccurately repair damaged DNA and are hypersensitive to DNA cross-linking agents. It has been established that FANCJ is BRIP1 (partner of BRCA1), FANCN is PALB2 (partner of BRCA2) and that SLX4 is also an FA protein. These findings further strengthen the connection between FA and DNA repair; specifically it appears that the FA pathway orchestrates incisions at sites of cross-linked DNA. Recent studies suggest the FA proteins may be important in counteracting aldehyde-induced genotoxicity in haemopoietic stem cells.

The FA–BRCA pathway is activated in response to DNA damage (e.g. replication fork arrest) and involves ATR (ataxia telangiectasia and RAD3-related protein) (Figure 10.3). The pathway is inactivated by the de-ubiquitinating enzyme USP1. ATR appears to directly regulate the FA pathway as it is required for the monoubiquitination of FANCD2 and FANCI. ATR is mutated in a subset of patients with Seckel syndrome, a disease exhibiting some clinical similarity to FA. ATR and ATM (ataxia telangiectasia mutated) are known to phosphorylate FANCD2 and FANCI. Interestingly, ATR-Seckel cells also exhibit defects in FANCD2 monoubiquitination. Furthermore, Nijmegen breakage syndrome (NBS) cells (mutated in *NBN*) show defects in FANCD2 monoubiquitination. This suggests that as well as clinical overlap between patients with FA, NBS and Seckel syndrome,

Table 10.4 FA complementation groups/genetic subtypes.

Complementation group/gene	Approximate percentage of FA patients	Chromosome location	Protein (amino acids)	Exons
A (*FANCA*)	65	16q24.3	1455	43
B (*FANCB*)*	<1	Xp22.2	859	10
C (*FANCC*)	12	9q22.3	558	14
D1 (*FANCD1*)†	<1	13q12.3	3418	27
D2 (*FANCD2*)	<1	3p25.3	1451	44
E (*FANCE*)	4	6p21.3	536	10
F (*FANCF*)	4	11p15	374	1
G (*FANCG*)	12	9p13	622	14
I (*FANCI*)	<1	15q26.1	1328	35
J (*FANCJ/BRIP1*)‡	<5	17q23.2	1249	20
L (*FANCL*)	<1	2p16.1	375	14
M (*FANCM*)	<1	14q21.3	2048	23
N (*FANCN/PALB2*§	<1	16p12.1	1186	13
O (*FANCO/RAD51C*)**	<1	17q25.1	376	9
P (*FANCP/SLX4*)#	2	16p13.3	1834	15
Q (*FANCQ/ERCC4*)##	<1	16p13.12	916	11

**FANCB* is on the X chromosome.
†*FANCD1* is *BRCA2*.
‡FANCJ is BRIP1 (BRCA1 interacting protein).
§FANCN is PALB2 (partner of BRCA2).
***FANCO* is *RAD51C*.
#*FANCP* is *SLX4*.
##*FANCQ is ERCC4*.

there is also overlap in the biological defects observed in cells from these patients. This highlights the complexity of physical interactions between the different molecules involved in this matrix of pathways. Equally it is clear that FA cells also display other abnormalities in addition to DNA repair (Figure 10.3). This includes hypersensitivity to oxygen, accelerated telomere shortening, abnormal cell cycle kinetics, upregulation of p53 and over-activation of the mitogen-activated protein kinase (MAPK) pathways leading to over-production of tumour necrosis factor (TNF)-α. These observations suggest that our understanding of the molecular events responsible for all the FA pathology is currently incomplete.

It is noteworthy that the phenotypes associated with biallelic *BRCA2* (FA-D1) and *PALB2* (FA-N) mutations are markedly similar to each other, but different from the other FA genes. Specifically, FA-D1 and FA-N are associated with high risks of solid childhood malignancies (e.g. Wilms tumour and medulloblastoma), which are not usually seen in the other FA subtypes. Furthermore, heterozygous mutations in *BRCA2* (FA-D1), *PALB2* (FA-N) and *BRIP1* (FA-J) confer an elevated risk of breast cancer yet this is not the case for the other FA genes. These differences highlight that the relationship between the FA proteins and their interactions with other molecules is complex at both the clinical and molecular level.

In vitro gene transfer studies have demonstrated that introduction of the appropriate wild-type FA gene into FA human lymphoid and haemopoietic cells markedly enhances their growth and normalizes their response to MMC; in lymphoid lines, cell kinetics (G_2 phase) and chromosomal breakage are normalized. Thus the transfer of the wild-type FA genes corrects the extreme sensitivity to DNA cross-linking agents, the hallmark of the FA cell phenotype. These studies provide the rationale for haemopoietic gene therapy (discussed below).

Murine models of FA have shown that haemopoietic progenitors are hypersensitive to TNF-α and interferon (IFN)-γ. This differential hypersensitivity to IFN-γ is thought to be mediated by *fas*-induced apoptosis and may turn out to be an important mechanism in the development of progressive BMF in FA. It is noteworthy that patients with idiopathic AA usually have raised IFN-γ levels, thus providing a possible link in the pathophysiology of BMF in both idiopathic and FA-associated AA. The presence of short telomeres in cells from patients with both FA and idiopathic AA can also be expected to be important in the

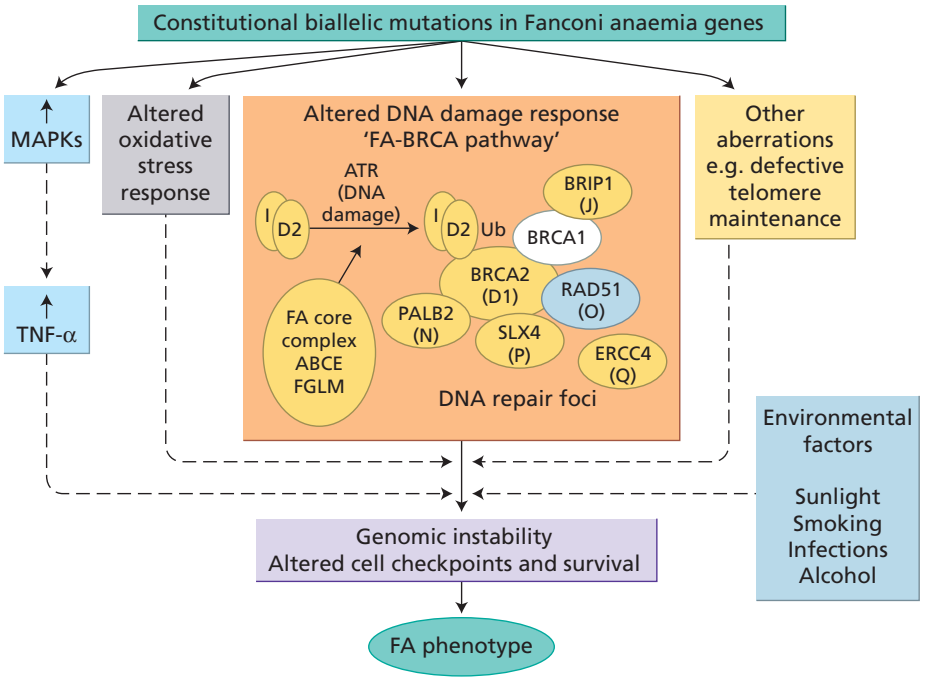


Figure 10.3 Schematic representation of the FA–BRCA pathway and related networks. The diagram shows that the constitutional mutations in FA cells lead to aberration of the FA–BRCA pathway, abnormal handling of oxidative stress, aberrant activation of mitogen-activated protein kinases (MAPKs) and defective telomere maintenance, as well as other biological aberrations. The net impact of these is increased genomic instability and altered cell survival/checkpoints. The diagram also highlights the potential role of environmental factors such as smoking in adding to the effect of the FA mutations. Within the FA–BRCA pathway, the proteins shown in yellow are those mutated in different FA patients. The FA core complex consists of eight FA-proteins (A, B, C, E, F, G, L and M) and this, together with ATR (ataxia telangiectasia and RAD3-related protein), is essential for activation (ubiquitination) of the I–D2 complex after DNA damage. Activated I–D2-Ub translocates to DNA repair foci where it associates with other DNA damage response proteins including BRCA2, RAD51 and SLX4 and participates in DNA repair. TNF, tumour necrosis factor; Ub, ubiquitination.

pathophysiology of BMF in both diseases. This feature is discussed further in relation to DC.

Treatment

The major cause of mortality in FA is the development of BMF. Anabolic steroids such as oxymetholone and danazol can produce useful trilineage haematological responses in 50–70% of patients but many will become refractory after a variable time. These can be associated with side-effects such as liver dysfunction; patients therefore need close monitoring. HSCT is the only curative treatment for BMF. From the *in vitro* and *in vivo* studies it has become clear that cells from FA patients are hypersensitive to agents such as cyclophosphamide and irradiation compared with non-FA patients. Therefore, HSCT conditioning regimens have been modified by reducing the dose of cyclophosphamide and radiation. Initially, using low-dose cyclophosphamide (20 mg/kg) and 4.5–6 Gy of thoraco-abdominal irradiation, the actuarial survival for patients transplanted using HLA-identical sibling donors was ~70% at 2 years. The results using unrelated donors were less good, with 2-year survival between 20 and 40%. Long-term follow-up of patients who have survived after HSCT shows a much higher incidence of malignancies (compared to idiopathic BMF), particularly of the head and neck, usually 8–10 years after the transplant. This partly relates to the inherent predisposition of FA patients to malignancy (which can perhaps now be explained given the link to defects in DNA repair) and partly to factors such as the use of radiotherapy in the conditioning. Transplant groups are therefore exploring low-dose cyclophosphamide (20–40 mg/kg) HSCT protocols that avoid the use of radiotherapy. Results using fludarabine (120–150 mg/m^2) in association with low-dose cyclophosphamide are very encouraging for both sibling and unrelated HSCT. Longer follow-up is necessary to determine if such protocols will be associated with a lower risk of malignancy.

In addition to HSCT, alternative treatment strategies are being explored. The identification of the FA genes, combined with the *in vitro* gene transfer data, which show that FA haemopoietic stem cells rescued by gene therapy should have a selective growth

advantage within the hypoplastic BM environment, has resulted in studies aimed at developing gene therapy for FA patients.

The identification of FA *mosaic* patients strengthens the case for future trials of gene therapy. In such cases, the DEB/MMC test may be negative or only demonstrate chromosomal instability in a subgroup of cells. Somatic mosaicism is due to reversion of a pathogenic allele to 'wild' type in a single haemopoietic (somatic) cell. The mechanism of how this occurs can vary, but in each case it generates one 'normal' FA allele and the resulting cell effectively becomes a 'heterozygous cell', which would be expected to have a growth/survival advantage in the background of FA cells. These mosaic patients can have an improvement in their haematological profile, suggesting that a single pluripotent stem cell may be sufficient to restore adequate haemopoiesis. FA patients with somatic mosaicism can thus be regarded as having undergone natural haemopoietic gene therapy.

Over the last 20 years there has been significant progress in treating the BMF associated with FA. In future, the major challenge in Fanconi anaemia may relate to the treatment of the associated malignancies and the management of complications (e.g. pulmonary disease) in adulthood.

Dyskeratosis congenita

Clinical features

Classical DC is characterized by the mucocutaneous triad of abnormal skin pigmentation, nail dystrophy and mucosal leucoplakia (Figure 10.4). Since its first description by Jacobi in 1906 and Zinsser in 1910, a variety of non-cutaneous (dental, gastrointestinal, genitourinary, neurological, ophthalmic, pulmonary and skeletal) abnormalities have also been reported (Table 10.5). BMF is the principal cause of early mortality, with an additional predisposition to malignancy (haematological and non-haematological) and fatal pulmonary complications. X-linked recessive, AD and AR forms of the disease are recognized. DC is a very heterogeneous disorder, both clinically and genetically.

Clinical manifestations in DC often appear during childhood. The skin pigmentation and nail changes typically appear first, usually by the age of 10 years. BMF usually develops below the age of 20 years; 80–90% of patients will have developed BM abnormalities by the age of 30 years. In some cases, the BM abnormalities may appear before the mucocutaneous manifestations and the patients may be categorized as having 'idiopathic aplastic anaemia'. The main causes of death are BMF/immunodeficiency (~60–70%), pulmonary complications (~10–15%) and malignancy (~10%).

The minimal clinical criteria for diagnosis of DC includes the presence of at least two out of the four major features (abnormal skin pigmentation, nail dystrophy, leucoplakia and BMF) and two or more of the other somatic features known to occur in DC.

Cell biology and link to other diseases

Since 1998, ten DC genes have been identified and these account for ~60% of DC (Table 10.6).

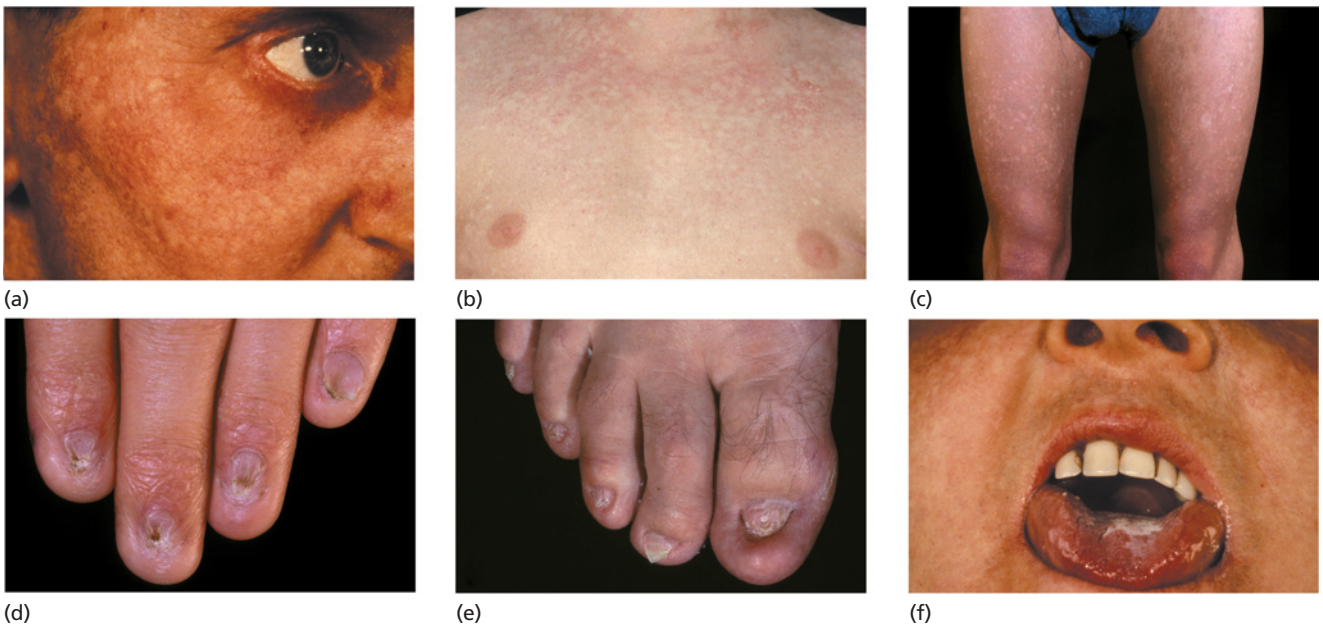


Figure 10.4 Photographs of patients with DC showing abnormal skin pigmentation (a, b, c), nail dystrophy (d, e) and leucoplakia of the tongue (f).

Table 10.5 Somatic abnormalities in dyskeratosis congenita.

Abnormality	Percentage of patients
Abnormal skin pigmentation	89
Nail dystrophy	88
BMF	85.5
Leucoplakia	78
Epiphora	30.5
Learning difficulties/developmental delay/mental retardation	25.4
Pulmonary disease	20.3
Short stature	19.5
Extensive dental caries/loss	16.9
Oesophageal stricture	16.9
Premature hair loss/greying/sparse eyelashes	16.1
Hyperhidrosis	15.3
Malignancy	9.8
Intrauterine growth retardation	7.6
Liver disease/peptic ulceration/enteropathy	7.3
Ataxia/cerebellar hypoplasia	6.8
Hypogonadism/undescended testes	5.9
Microcephaly	5.9
Urethral stricture/phimosis	5.1
Osteoporosis/aseptic necrosis/scoliosis	5.1
Deafness	0.8

X-linked DC and the Hoyeraal-Hreidarsson (HH) syndrome

The gene (*DKC1*) responsible for X-linked DC was mapped to Xq28 in 1986 and identified through positional cloning in 1998. The *DKC1* gene is highly conserved and encodes the protein dyskerin. With the identification of mutations in *DKC1* the first diagnostic tests became available. It also provided the first firm evidence that DC was not a homogeneous disorder and that other syndromes with overlapping presentation can share the same genetic mutations. The first such example was the HH syndrome, a severe multisystem disorder which is characterized by intrauterine growth retardation, microcephaly, cerebellar hypoplasia, AA and immunodeficiency. Due to the overlap in features it was suggested and subsequently proven that HH is a severe variant of DC due to the presence of *DKC1* mutations in males with the classical presentation of HH. However *DKC1* mutations are not the only cause of HH; mutations in other genes can also lead to a phenotype of HH (see below).

Autosomal dominant (AD)-DC and its link to telomeres and other diseases

AD-DC is heterogeneous; to date heterozygous mutations in three genes (*TERC*, *TERT* and *TINF2*) have been characterized. The identification of heterozygous mutations in *TERC* (telomerase RNA component) in 2001 was a major advance in the DC field, as it provided a direct link between DC and telomerase. Telomerase is a ribonucleoprotein composed of two core components: a catalytic component which adds the repeats, telomerase reverse transcriptase (TERT) and *TERC*, which acts as the template. It functions as a specialized polymerase that adds the telomeric repeat (TTAGGG) to the end of the 3′ lagging strand of DNA after replication. Due to the semi-conservative nature of DNA replication, telomerase is essential to maintain telomere length in rapidly dividing cells, as found in the haemopoietic system. Without telomerase the telomeres shorten with each successive round of replication, when they reach a critical length the cells enter senescence. Telomerase is mainly restricted to cells such as germ cells, stem cells and their immediate progeny, activated T cells and monocytes, but in cells where

Table 10.6 Genetic subtypes of dyskeratosis congenita.

DC subtype	Approximate percentage of patients	Chromosome location	RNA/protein product	Exons
X-linked recessive	25	Xq28	Dyskerin	15
Autosomal dominant	5	3q26.2	TERC	1
	3	5p15.33	TERT	16
	12	14q11	TIN2	6
Autosomal recessive	<1	15q14–q15	NOP10	2
	<1	5p15.33	TERT	16
	<1	5q35.3	NHP2	4
	<1	17p13.1	TCAB1	13
	<2	16q13	USB1	7
	<1	17p13.1	CTC1	23
	2	20q13.3	RTEL1	35
Uncharacterized	>30	?	?	?

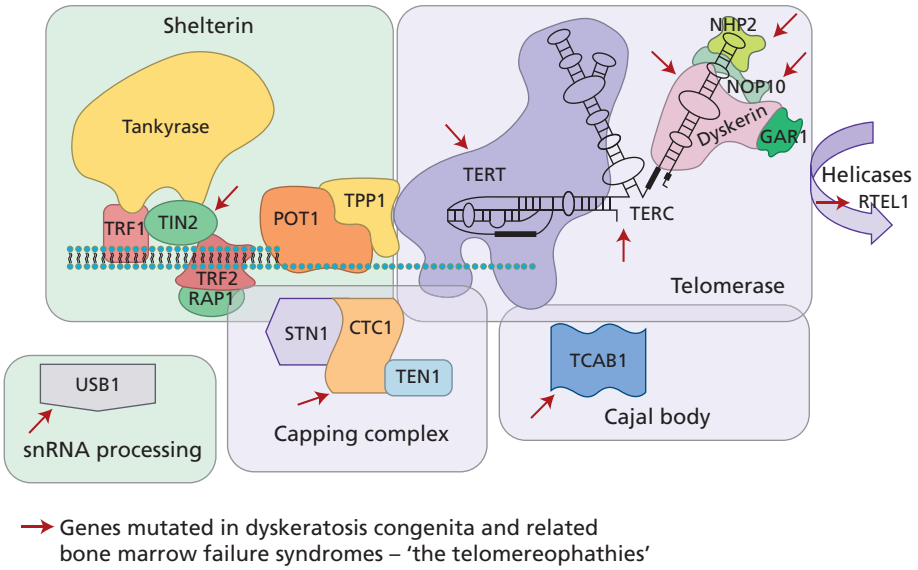


Figure 10.5 Schematic representation of complexes involved in telomere maintenance. The telomerase complex includes TERC, TERT, dyskerin, NOP10, NHP2 and GAR1. The shelterin complex includes the six proteins TRF1, TRF2, TPP1, POT1, RAP1 and TIN2. The telomere capping (CST) complex is composed of CTC1, STN1, and TEN1. Protein/RNA names indicated by red arrows are mutated in DC and related disorders. Hemizygous DKC1 (dyskerin) mutations are observed in X-linked DC and HH. Heterozygous TERC mutations are associated with DC, AA, MDS, AML and pulmonary fibrosis. Heterozygous TERT mutations are responsible for some cases of AA, DC, MDS, AML and pulmonary/liver fibrosis. Biallelic mutations in TERT can cause classic DC and HH. Heterozygous TIN2 mutations have been observed in DC, AA, HH and Revesz syndrome. Biallelic NOP10, NHP2, TCAB1, USB1 and CTC1 mutations have been seen in AR-DC. Biallelic RTEL1 mutations are observed in AR-HH.

telomerase is not present telomere shortening is part of the normal process of cellular ageing.

Mutations in patients with AD-DC were identified in *TERC* initially and it was through the identification of mutations within this molecule that led to significant expansion of the DC phenotype to include other haematological and non-haematological disorders. Firstly, heterozygous *TERC* mutations were identified in patients with AA and soon after in patients with myelodysplastic syndrome (MDS).

The identification of mutations in *DKC1* and *TERC* established the pathology of defective telomere maintenance (Figure 10.5) as being the principal underlying cause of DC; both dyskerin (encoded by *DKC1*) and *TERC* are now recognized to be core components of telomerase and patients with *DKC1* and *TERC* mutations have very short telomeres compared to age-matched controls. This led to further study of the telomerase complex to determine the genetic basis of the remaining uncharacterized patients. The next gene to have mutations identified was *TERT*, which encodes the enzymatic component of the telomerase complex. The clinical presentation in patients with *TERT* mutations is highly variable, ranging from near DC phenotype to just AA. Heterozygous mutations in *TERT* and *TERC* have also been identified in some patients with idiopathic pulmonary fibrosis, liver disease and leukaemia.

In 2008, mutations in a component (TIN2) of the shelterin complex were identified in one subtype of AD-DC. The shelterin complex (comprised of six proteins) has at least three effects on telomeres. It determines the structure of the telomeric terminus, it has been implicated in the generation of t-loops and it controls the synthesis of telomeric DNA by telomerase. The composition and protein interaction of the components of shelterin complex appears to be highly ordered, with TIN2 playing a pivotal role. In a subset of patients with DC, HH, AA and Revesz syndrome, heterozygous TIN2 mutations have been identified. This discovery extends the range of the DC spectrum of diseases even further. Revesz syndrome is characterized by bilateral exudative retinopathy, BMF, nail dystrophy, fine hair, cerebellar hypoplasia and growth retardation. Patients with TIN2 mutations tend to have severe disease and this is associated with very short telomeres. Interestingly, nearly all the patients have *de novo* TIN2 mutations, which give rise to a different mechanism causing the disease. In patients with heterozygous *TERC* and *TERT* mutations, studies have shown that the phenomenon of genetic anticipation is frequently involved; a parent of an affected child has the same telomerase mutation, but usually no overt signs of disease. However, in a child with the same heterozygous telomerase mutation the disease manifests itself at a much younger age and is usually more severe.

Autosomal recessive DC

Since 2007 progress has been made in the genetic basis of AR-DC. A large linkage study of 16 consanguineous families comprising 25 affected individuals did not identify a single common locus, suggesting there is genetic heterogeneity within AR-DC. Since this observation, mutations in seven genes have been identified as causing AR-DC. The first AR-DC gene to be identified was *NOP10*. The homozygous *NOP10* mutation identified in a large family affected a highly conserved residue. As a result of this mutation all the affected individuals had reduced telomere length and reduced TERC levels. In a subset of AR-DC biallelic mutations have been identified in *TERT*. These mutations give very different profiles regarding telomerase activity and telomere length, with both being greatly reduced compared with heterozygous *TERT* mutations. Biallelic mutations in *NHP2* have been identified in a third subset of AR-DC patients. Again telomere lengths and TERC levels are reduced in patients compared to normal controls. Both NOP10 and NHP2 are components of H/ACA ribonucleoprotein complex (H/ACA RNP). This complex is comprised of a RNA molecule and four proteins, dyskerin, GAR1 as well as NOP10 and NHP2. These four proteins are highly conserved and have been shown to be involved in ribosome biogenesis, pre mRNA splicing and telomere maintenance. Mutations have been identified in all components of this H/ACA RNP complex in patients with DC except for GAR1.

In 2011, biallelic mutations in the *TCAB1* were identified in two patients with AR-DC. TCAB1 is a telomerase holoenzyme protein that facilitates trafficking of telomerase to Cajal bodies, the nuclear sites of nucleoprotein complex modification and assembly. Compound heterozygous mutations in TCAB1 disrupt telomere localization to Cajal bodies, resulting in misdirection of telomerase RNA to nucleoli. This in turn prevents telomerase from elongating telomeres, thereby resulting in short telomeres.

In another subgroup of AR-DC, biallelic mutations have been found in *USB1*. It is noteworthy that mutations in this gene have unified this subgroup of DC with patients classified as having poikiloderma with neutropenia and Rothmund–Thomson syndrome. This subgroup of AR-DC patients appear to have normal length telomeres and therefore represent a biologically different subtype.

In 2012, biallelic mutations in *CTC1*, encoding conserved telomere maintenance component 1, were identified in a rare subgroup of DC patients. As *CTC1* mutations were initially identified in the pleotropic syndrome, Coats plus (characterized by retinopathy, intracranial calcifications and cysts, osteopenia and gastrointestinal abnormalities), this observation expands further the complexity of phenotypes associated with the 'telomereopathies'.

In 2013, biallelic mutations in *RTEL1* (regulator of telomere length 1) were identified in some patients with severe DC (close to HH) and these patients have very short telomeres as well as qualitative defects in telomeres. RTEL1 is a helicase with an important role in homologous recombination and telomere maintenance (dismantles telomere t-loops during DNA replication).

Treatment

BMF is the main cause of mortality in DC. Use of the anabolic steroids (oxymetholone and danazol) can produce improvement in the haemopoietic function. Approximately two-thirds of patients with DC will respond to oxymetholone; in some cases the response can last several years and involve all lineages. Patients with DC can respond to a dose as low as 0.25 mg oxymetholone/kg/day and this can be increased, if necessary, to 2–5 mg/kg/day. It is important to monitor for side-effects (e.g. liver toxicity). It is possible to maintain reasonable blood counts by this approach in many patients.

The only long-term cure for the haemopoietic abnormalities is allogeneic HSCT. There is significant mortality associated with BM transplants for DC patients than with other BMF syndromes. One of the main reasons for this is the high level of pulmonary/vascular complications that present in these patients, probably as a result of the underlying telomere defect. The conditioning regimen appears to have an impact on patient survival. The standard myeloablative conditioning regimes are associated with frequent and severe adverse effects, such as pulmonary complications and veno-occlusive disease. Recently, the adoption of non-myeloablative fludarabine-based protocols has allowed for successful engraftment in some patients with fewer complications and lower toxicity. The long-term survival, however, is unknown at present, but the initial response is encouraging as a more effective treatment for DC.

Shwachman–Diamond syndrome (SDS)

Clinical features

Shwachman and Bodian and their colleagues first reported this disease independently in 1964. SDS is now recognized as an AR disorder characterized by exocrine pancreatic insufficiency (100%), BMF (100%) and other somatic abnormalities (particularly involving the skeletal system). Signs of pancreatic insufficiency (malabsorption, failure to thrive) are apparent early in infancy (NB: pancreatic function can improve in a subset of SDS patients). Other somatic abnormalities include short stature (~70%), protuberant abdomen and an ichthyotic skin rash (~60%). Metaphyseal dysostosis is seen on radiographs in about 75% of patients. Other abnormalities include hepatomegaly, rib/thoracic cage abnormalities, hypertelorism, syndactyly, cleft palate, dental dysplasia, ptosis and skin pigmentation.

The spectrum of haematological abnormalities includes neutropenia (~60%), other cytopenias (~20% have pancytopenia), MDS and AML (~25%). The age at which leukaemia

Table 10.7 Genetic subtypes of Shwachman–Diamond syndrome.

SDS subtype	Approximate percentage of patients	Chromosome location	Protein product	Exons
Autosomal recessive	>90	7q11.22	SBDS	5
Uncharacterized	<10	?	?	?

develops varies widely, from 1 to 43 years. AML is the commonest category and there is an unexplained preponderance of cases of leukaemia in males.

Exocrine pancreatic insufficiency and haematological abnormalities are also seen in Pearson syndrome (PS), and this is therefore an important differential diagnosis. In PS, the anaemia is usually more prominent than neutropenia and the marrow usually shows ringed sideroblasts along with vacuolation of myeloid and erythroid precursors. In addition, acidosis, abnormalities of liver function and mitochondrial DNA rearrangements are seen in PS. PS has a worse prognosis than SDS, with many patients dying before the age of 5 years from liver or BMF. Other differential diagnoses to be excluded are cartilage hair syndrome and cystic fibrosis.

Cell and molecular biology

The SDS gene (*SBDS*) on 7q11.22 was identified in 2003. The majority (>90%) of SDS patients have biallelic mutations in this gene (Table 10.7). Data from yeast and human studies suggest that the SBDS protein has an important role in the maturation of the 60S ribosomal subunit (Figure 10.6). These observations suggest that SDS is principally a disorder of ribosome biogenesis.

Treatment

The malabsorption in SDS responds to treatment with oral pancreatic enzymes. For those with neutropenia, G-CSF may produce an improvement in the neutrophil count. Some patients

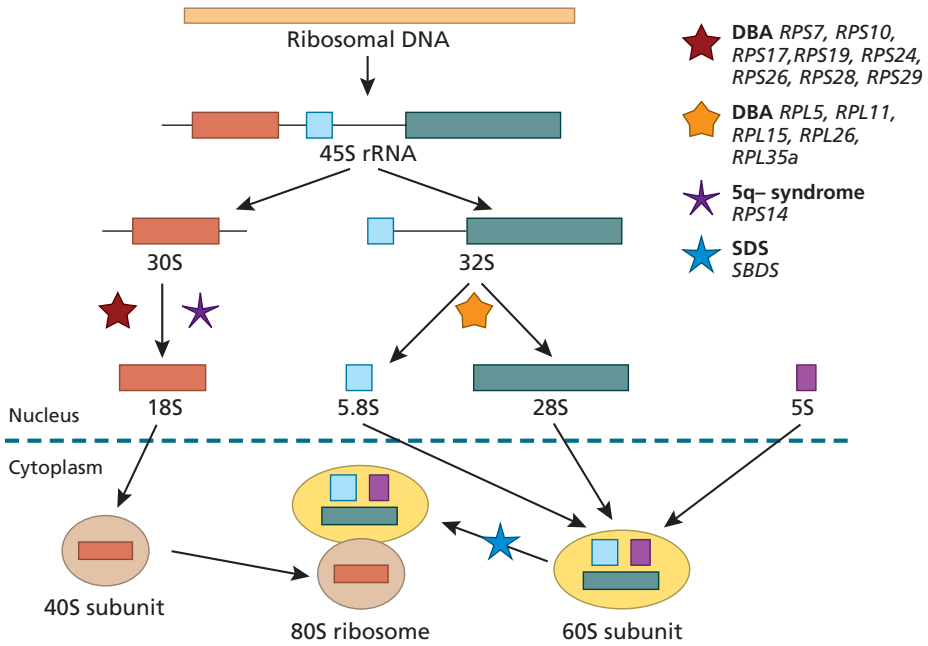


Figure 10.6 Schematic diagram showing scheme of ribosomal (r)RNA processing in human cells and the points at which this is possibly disrupted in the different BMF syndromes. The rRNAs are transcribed by RNA polymerase I as a single precursor transcript (45S rRNA). The 45S rRNA is then processed to 18S, 5.8S and 28S rRNAs. The 18S is a component of the 40S ribosomal subunit. The 5.8S and 28S together with 5S (synthesized independently) are components of the 60S ribosomal subunit. The 40S and 60S subunits are assembled to form the 80S ribosome. The processing steps affected in DBA (heterozygous mutations in *RPS7*, *RPS10*, *RPS17*, *RPS24*, *RPS26*, *RPS28*, *RPS29*, *RPL5*, *RPL11*, *RPL26* and *RPL35a*), 5q– syndrome (haploinsufficiency of *RPS14*) and SDS (biallelic mutations in *SBDS*) are indicated by the different coloured stars.

with anaemia and/or thrombocytopenia may achieve haematological responses with oxymetholone treatment. As in other types of BMF, supportive treatment with blood products and antibiotics is very important.

The development of leukaemia, often with features of myelodysplasia, usually has a poor prognosis. SDS patients with leukaemia treated with conventional courses of chemotherapy usually fail to regenerate normal haemopoiesis. As this is a constitutional disorder all somatic cells, including haemopoietic stem cells, are abnormal. In addition, the haemopoietic stem cells may have accumulated secondary abnormalities, as suggested by complex karyotypes (especially involving chromosome 7) often observed in the BM from these patients. Therefore, for those who develop leukaemia, the only approach likely to be successful is allogeneic HSCT, using conditioning regimens that include fludarabine. It is noteworthy that non-haematological malignancies have not been observed in SDS patients.

Diamond–Blackfan anaemia (DBA)

Clinical features

Red cell aplasia was first reported in 1936 by Josephs. In 1938, Diamond reported on four children with hypoplastic anaemia and this has now come to be recognized as DBA or congenital pure red cell aplasia. DBA usually presents in early infancy, with features of anaemia such as pallor or failure to thrive. The hallmark of classical DBA is a selective decrease in erythroid precursors and normochromic macrocytic anaemia associated with a variable number of somatic abnormalities such as craniofacial, thumb, cardiac and urogenital malformations. Hitherto the diagnostic criteria for DBA have comprised: (i) normochromic, usually macrocytic, but occasionally normocytic anaemia developing in early childhood; (ii) reticulocytopenia; (iii) normocellular BM with selective deficiency of erythroid precursors (erythroblasts <5%); (iv) normal or slightly decreased leucocyte counts and (v) normal or often increased platelet counts. More recently, elevated erythrocyte deaminase activity, macrocytosis and elevated fetal haemoglobin have been added to the list of supportive features of DBA. It has also been recognized that in a subset of cases the presentation may be in adulthood.

There is considerable heterogeneity in the associated somatic abnormalities and response to therapy. Analysis of cases recruited to the DBA Registry of North America (DBAR) has shown that the annual incidence of DBA is about 5 per million live births. The median age at presentation was 8 weeks and 93% of patients presented in the first year of life. In total, 79% were initially responsive to steroids, 17% were non-responsive and 4% were never treated with steroids; 31% of patients were receiving transfusions at analysis. The actuarial survival rates at older than 40 years were 100% for those in sustained remission, 87% for steroid-maintained patients and 57% for transfusion-dependent patients. Of the 36 deaths reported, 25 were treatment related: 5 from infections, 5 from complications of iron overload, 1 related to vascular access and 14 from transplant-related complications.

In the DBAR, 8.8% of families had more than one affected individual. Most of the familial cases displayed AD inheritance. Somatic anomalies, excluding short stature, were found in 47% of patients. Of these, 50% were craniofacial (high-arched palate, cleft lip, hypertelorism and flat nasal bridge), 38% were upper limb and hand (flat thenar eminence, triphalangeal thumb), 39% genitourinary and 30% cardiac. Height was below the third centile for age in about 30%.

MDS and AML have been reported in a few patients with DBA. There are also cases that have evolved into AA; neutropenia and thrombocytopenia are relatively common after the first decade. Giri and colleagues reported on moderate to severe BM hypocellularity in 21 of 28 (75%) patients with steroid-refractory DBA; marrow hypoplasia correlated with the development of neutropenia (9/21, 43%) and/or thrombocytopenia (6/21, 29%). Furthermore, using *in vitro* long-term culture-initiating cell assay, they provided evidence for a trilineage haemopoietic defect in patients with refractory DBA. Thus, although DBA has been regarded classically as a pure red cell aplasia, a more global haemopoietic defect is likely to be present, and this may be seen more frequently in the future as patients survive longer due to improved medical care.

Cell and molecular biology

A number of different defects of *in vitro* erythroid progenitor proliferation, differentiation and cytokine responsiveness have been reported. The establishment of DBA registries, and advances in genetics has led to genetic categorization of many DBA cases. The first gene to be identified was *RPS19* (1999). Subsequently, heterozygous mutations in other genes encoding ribosomal proteins of the small (RPS7, RPS10, RPS17, RPS24, RPS26, RPS28, RPS29) and large (RPL5, RPL11, RPL15, RPL26 and RPL35a) ribosomal subunits have also been reported (Table 10.8); collectively the genetic basis of approximately 65% of DBA patients can now be substantiated at the genetic level. These findings suggest that the primary defect in DBA is defective ribosome biogenesis, which then leads to other biological defects, including increased apoptosis and upregulation of p53. It will now be important to establish how precisely mutations in these ribosomal protein genes lead to altered cell growth (and anaemia), developmental anomalies and increased susceptibility to cancer. Recently, constitutional hemizygous *GATA1* mutations have been identified in rare patients with 'DBA-like' disease.

Review of clinical and genetic data has shown that patients with *RPL5* mutations tend to have multiple physical abnormalities, including craniofacial, thumb and heart anomalies, whereas isolated thumb malformations are predominantly seen in

Table 10.8 Genetic subtypes of Diamond–Blackfan anaemia.

DBA subtype	Approximate percentage of patients	Chromosome location	Protein product	Exons
Autosomal dominant	25	19q13.2	RPS19	6
	2	10q22–23	RPS24	7
	1	15q25.2	RPS17	5
	7	1p22.1	RPL5	8
	5	1p36.1–p35	RPL11	6
	3	3q29–qter	RPL35a	5
	1	2p25	RPS7	7
	7	6q21.31	RPS10	6
	3	12q13.2	RPS26	4
	1	17p13.1	RPL26	4
	<1	3p24.2	RPL15	4
	?	14q21.3	RPS29	3
	?	19p13.2	RPS28	4
X-linked recessive	<1	Xp11.23	GATA1	6
Uncharacterized	~30	?	?	?

patients with heterozygous *RPL11* mutations. Some genotype–phenotype correlations are therefore beginning to emerge.

It is noteworthy that recent advances have provided a very interesting connection between DBA, SDS, MDS and defective ribosome biogenesis (see Figure 10.6). Specifically: (i) DBA (as highlighted here) is caused by mutations in genes important in the biogenesis of the small and large ribosomal units; (ii) the *SBDS* gene mutated in SDS (see above) is required for the maturation of the large ribosomal unit; and (iii) the 5q– syndrome (a recognized subtype of acquired myelodysplasia) is associated with haploinsufficiency of the gene encoding the ribosomal protein RPS14. These advances highlight the importance of studies on rare syndromes, as often they can provide significant insights into the pathology of more common diseases.

Treatment

The first line of treatment for DBA remains corticosteroids. Once a maximal haemoglobin response has been achieved, the dose of prednisolone should be tapered slowly until the patient is on the lowest dose possible on an alternate-day regimen. For those patients who fail to respond or become refractory to steroids, blood transfusion is the mainstay of treatment. The major complication from transfusions is iron overload, and chelation of iron (avoiding deferiprone) should therefore be commenced as soon as patients have increased iron stores (Chapter 4). Splenectomy may be indicated in the event of an increased transfusion requirement secondary to hypersplenism. For patients who are transfusion dependent and who have a compatible sibling donor, HSCT may be appropriate.

Congenital dyserythropoietic anaemia (CDA)

CDA comprises a heterogeneous group of disorders of erythropoiesis characterized by anaemia, ineffective erythropoiesis and frequently morphological evidence of dyserythropoiesis. The first description of these disorders was in 1966 by Crookston and colleagues. In 1968 Wendt and Heimpel classified CDA into three types (I, II and III) (Table 10.9). Over the years many additional subtypes (IV, V, VI and VII) have been added to the list, often based on case reports.

CDA type I

The majority of patients present with splenomegaly and mild to moderate anaemia (~66–116 g/L); approximately 70% have macrocytosis. In some cases non-haematological features (e.g. skeletal abnormalities, abnormal skin pigmentation) have been observed. Ineffective erythropoiesis is evidenced by morphological abnormalities in the peripheral blood (anisocytosis) and in the marrow (megaloblastic erythroid precursors, internuclear chromatin bridging, binuclearity affecting 3–7% of erythroblasts; Figure 10.7a), as well as by increased markers of haemolysis (elevated lactate dehydrogenase and bilirubin). The defining ultrastructural feature is a spongy ('Swiss cheese') appearance of the heterochromatin in the majority of erythroblasts on electron microscopy (Figure 10.7b).

Recognized to be AR, the first gene responsible for CDA type I (*CDAN1,* called codanin) was identified in 2002. Recently a second gene, *C15ORF41*, was found to be responsible for some cases of CDA type I.

Table 10.9 Characteristics of common subtypes of congenital dyserythropoietic anaemia.

Feature	Type I	Type II	Type III
Inheritance	AR	AR	AD, AR
Erythrocytes	Macrocytic	Normocytic	Macrocytic
Erythroblasts			
Light microscopy	Megaloblastic, internuclear chromatin bridges	Normoblastic, binuclearity	Megaloblastic, up to 12 nuclei/cell
Electron microscopy	'Swiss-cheese' appearance	Peripheral double membranes	Non-specific
Serology			
Ham test	Negative	Usually positive	Negative
Anti-i-agglutinabiliy	Normal/strong	Strong	Normal/strong
SDS-PAGE	Normal	Band 3 thinner and migrates faster than normal	Band 3 migrates faster than normal
Genes identified	*CDAN1* (15q15.2), *C15ORF41* (15q14)	*SEC23B* (20p11.23)	*KIF23* (15q21)

Source: Wickramasinghe SN, Wood WG (2005) Advances in the understanding of congenital dyserythropoietic anaemia. *Br. J. Haematol.* **131**: 431–46. Reproduced and modified with permission of John Wiley & Sons.

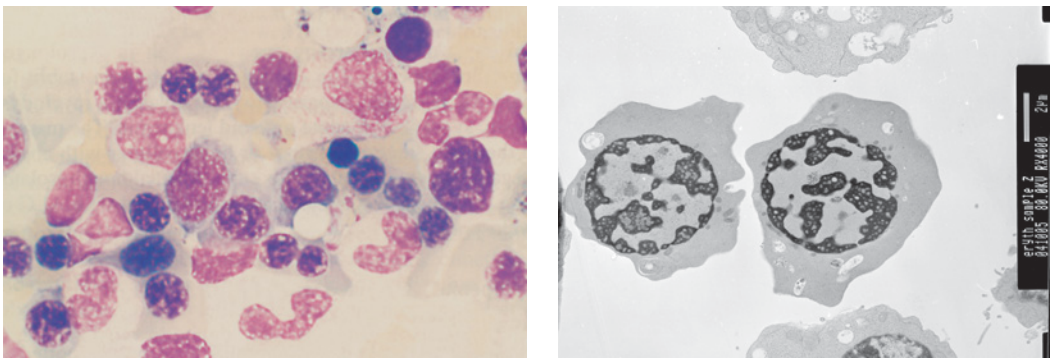
CDA type II

This is the most common subtype of CDA and was initially described as hereditary erythroblastic multinuclearity with a positive acidified serum lysis test (HEMPAS) in 1969. It is inherited as an AR trait. The anaemia is variable (Hb 80–110 g/L). Approximately 10% of cases require regular transfusions and some cases present with anaemia at birth. The clinical presentations include a variable degree of jaundice, hepatomegaly, splenomegaly and cirrhosis. Mental retardation has been reported in some cases.

Peripheral blood morphology shows moderate to marked red cell anisocytosis. BM features include normoblastic erythroid hyperplasia with usually more than 10% binucleate erythroblasts (Figure 10.8a). At the electron microscope level, the erythroid cells have a characteristic peripheral arrangement of the endoplasmic reticulum giving the appearance of a 'double membrane' (Figure 10.8b). Red cells from patients with CDA type II are haemolysed by some acidified sera, but not by the patient's own serum. In 2009 the gene encoding the secretory COPII component SEC23B was shown to be responsible for CDAII.

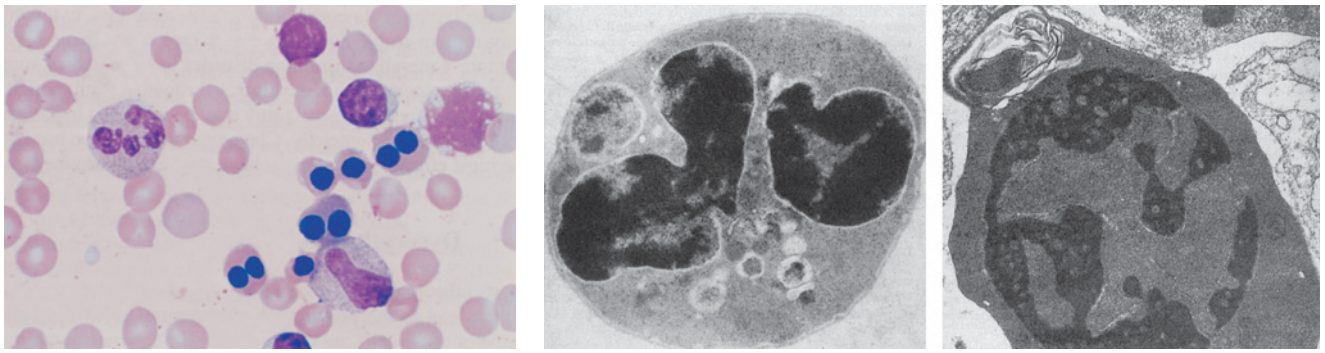
CDA type III

This subtype is rare. In one of the largest (Swedish) families investigated, the disease was characterized by giant multinucleated erythroblasts in the marrow (Figure 10.9). There appears to



(a) (b)

Figure 10.7 (a) CDA type I. BM aspirate showing internuclear bridging in normoblasts. (b) CDA type I. Electron micrograph of erythroblasts showing a spongy ('Swiss cheese') appearance of the heterochromatin (Source: David Fergusion and Noemi Roy. Reproduced with permission.)



(a) (b)

Figure 10.8 (a) CDA type II (HEMPAS). BM aspirate showing typical multinuclearity. (b) CDA type II. Electron micrographs of erythroblast showing peripheral arrangement of endoplasmic reticulum with 'double membrane' appearance.

be an increased prevalence of lymphoproliferative disorders in CDA type III. CDA III exhibits AD transmission and is caused by mutations in *KIF23*. *KIF23* encodes mitotic kinesin-like protein 1, which plays a critical role in cytokinesis during cell division.

Treatment

Those with mild anaemia require no major intervention. Folate supplementation should be given to prevent folate deficiency. If regular transfusions are necessary, early attention to iron chelation is essential. Splenectomy may be of benefit in some patients (CDA type II) and there are rare case reports of successful HSCT. In CDA type I there are also case reports of improvement in the haemoglobin after treatment with interferon-α. The mechanism of this therapeutic benefit remains unclear.

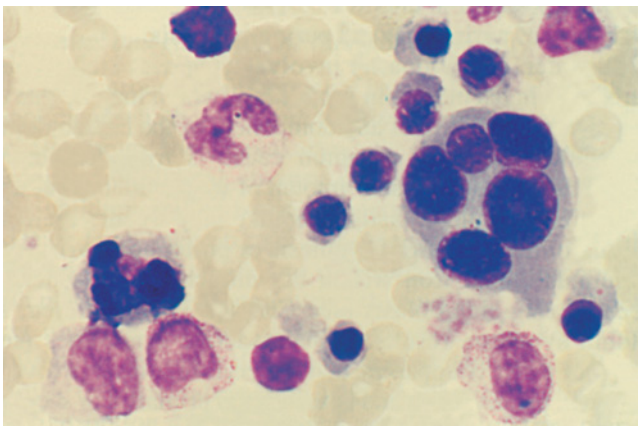


Figure 10.9 CDA type III. Giant multinucleated erythroblast from the marrow.

Congenital and cyclical neutropenias

Congenital neutropenia is a heterogeneous disorder (Table 10.10). It includes Kostmann syndrome, which was first described in 1954. Although the original description by Kostmann was of an AR disorder, other congenital neutropenia subtypes (both sporadic and AD) have subsequently been included in this category. The neutropenia is usually recognized at birth and the neutrophil count is often below 0.2×10^9/L. The haemoglobin and platelet count are usually normal and the BM shows maturation arrest of myelopoiesis at the level of the promyelocyte/myelocyte (with abundant promyelocytes, but with selective reduction in myelocytes, metamyelocytes and neutrophils).

The neutropenias are associated with severe infections and early death. No patient has developed AA, but myeloid leukaemias (~25% by 25 years) can occur. The availability of G-CSF has revolutionized the outcome for these children. These patients can progress to MDS and leukaemia, usually with acquisition of secondary mutations including in the G-CSF receptor. The precise contribution of G-CSF therapy to the development of G-CSF receptor mutations remains unclear. For patients who become refractory to G-CSF or who develop leukaemia, HSCT may be appropriate and curative.

Cyclical neutropenia is characterized by a neutrophil count that usually reaches a nadir with a 21-day periodicity. Around the nadir, patients may develop fever and mouth ulcers. Cyclical neutropenia is usually AD. Linkage analysis in affected families resulted in the localization of the disease gene to 19p13.3. Subsequent studies identified mutations in the gene (*ELANE*) encoding neutrophil elastase. An extraordinary twist to the story was the identification of *ELANE* mutations in many patients with congenital neutropenia as well. In cyclical neutropenia, the mutations are usually clustered around the active site of the molecule, whereas the opposite face of the molecule

Table 10.10 Genetic subtypes of neutropenia.

Subtype	Approximate percentage of patients	Chromosome location	Gene/protein product	Exons
Autosomal dominant	50–60	19p13.3	*ELANE*	6
	<2	1p22	*GFI1*	3
Autosomal recessive	10–15	1q21.3	*HAX1*	7
	~5	17q21.31	*G6PC3*	6
	Rare	1q21.2	*VPS45*	4

NB: Some patients initially presenting as isolated neutropenia may turn out to be cryptic presentations of other syndromes, such as Wiskott–Aldrich syndrome.

tends to be mutated in congenital neutropenia. Neutrophil elastase is a serine protease that is synthesized predominantly at the promyelocytic stage and can be expected to be important in neutrophil development. Recent studies suggest that *ELANE* mutations lead to accumulation of the non-functional protein, which in turn triggers an unfolded protein response leading to maturational arrest. The precise mechanism leading to maturation arrest of promyelocytes remains unclear. The original family described by Kostmann had AR severe congenital neutropenia, and has been shown to be associated with biallelic mutations in *HAX1*, predicted to lead to defects in cell death. Biallelic mutations in *HAX1* account for approximately 10% of congenital neutropenia. The HAX1 protein is a critical regulator of the mitochondrial membrane potential and cellular viability. While there are data which suggest that HAX1 is important in controlling apoptosis, it is unclear why premature death of neutrophils is consistently associated with HAX1 deficiency. Mutations in other genes (*GFI1, G6PC3* and *VPS45*) are also associated with SCN.

Thrombocytopenia with absent radii (TAR)

TAR is an AR disorder characterized by hypomegakaryocytic thrombocytopenia and bilateral radial aplasia. Babies with TAR often have haemorrhagic manifestations at birth, when the diagnosis is usually made because of the characteristic physical appearance combined with thrombocytopenia. Additional skeletal abnormalities (absent ulnae, absent humeri, clinodactyly) and other somatic abnormalities (microcephaly, hypertelorism, strabismus, heart defects) may be seen in some patients.

The platelet count is usually below 50×10^9/L. The leucocyte count can be normal or raised, sometimes up to 100×10^9/L ('leukaemoid reaction'). BM cellularity is normal and myeloid and erythroid lineages are normal or increased. Megakaryocytes are absent or decreased. Most patients bleed in infancy and then improve after the first year. The mainstay of management is prophylactic and therapeutic use of platelet transfusions. Patients with TAR have a very good prognosis after infancy. There have been no reports of AA or leukaemia.

In TAR patients thrombopoietin levels are usually elevated and thrombopoietin receptor expression on the surface of TAR platelets is normal. Therefore, defective megakaryocytopoiesis/thrombocytopoiesis does not appear to be caused by a defect in thrombopoietin production. Recently it has been determined that compound inheritance of low frequency regulatory SNP and a rare null mutation in *RBM8A* (which encodes a subunit of the exon-junction complex) causes TAR.

Congenital amegakaryocytic thrombocytopenia (CAMT)

CAMT is a rare disorder that usually presents in infancy and is characterized by isolated thrombocytopenia and reduction/absence of megakaryocytes in the BM, usually with no somatic abnormalities. It is genetically heterogeneous with AR and X-linked subtypes. Approximately 50% of patients will develop AA, usually by the age of 5 years. For patients with severe thrombocytopenia or AA, the treatment of choice is HSCT, if a compatible donor is available.

In a subgroup of patients with CAMT, mutations in the gene encoding the thrombopoietin receptor (*MPL*) have been identified (Table 10.11). As patients with *MPL* mutations can also have abnormalities in the leucocyte count and haemoglobin level and central nervous system (CNS) abnormalities (e.g. cerebral and cerebellar hypoplasia), this highlights the important role of the thrombopoietin receptor in haemopoiesis in general and in CNS development. It also substantiates the genetic heterogeneity of CAMT.

Conclusion

Since the identification of the first FA gene (*FANCC*) in 1992, there have been significant advances in our understanding

Table 10.11 Genetic subtypes of congenital amegakaryocytic thrombocytopenia.

CAMT subtype	Approximate percentage of patients	Chromosome location	Gene/protein product	Exons
Autosomal recessive	?	1p34	*MPL*	12
Uncharacterized	?	?	?	?

of FA, DC and other BMF syndromes. This has facilitated diagnosis, as highlighted in Table 10.12. It can be anticipated that further studies of the pathophysiology of these disorders is likely to lead to a better understanding of normal haemopoiesis and how this becomes defective in many patients presenting with the more common forms of AA and MDS. Indeed studies have already established a link between DC and AA and, in turn, to defective telomere maintenance. Equally, a link between DBA, SDS and MDS and, in turn, defective ribosome biogenesis has been recognized. These advances also suggest that new treatment strategies, based on targeting the primary defect in each syndrome, may now emerge.

Table 10.12 Laboratory tests useful in the investigation of patients with BMF.

Test	Diagnostic value
Peripheral blood	
Fetal haemoglobin	High level suggestive of generalized BMF
DEB/MMC chromosomal breakage	Increased in FA
Mutation analysis of specific genes	
FANCA–FANCQ	Mutated in FA
DKC1, *TERC*, *TERT*, *NOP10*, *NHP2*, *TINF2, TCAB1, USB1, CTC1, RTEL1*	Mutated in DC and related disorders
SBDS	Mutated in SDS
RPS19, *RPS24*, *RPS17*, *RPL5*, *RPL11*, *RPL35a*, *RPS7, RPS10, RPS26, RPL26, RPL15, GATA1, RPS29, RPS28*	Mutated in DBA
CDAN1, C15ORF41, SEC23B, KIF23	Mutated in CDA
ELANE, *HAX1*, *GFI1, G6PC3, VPS45*	Mutated in congenital and cyclic neutropenia
MPL	Mutated in CAMT
Mitochondrial DNA analysis	Deletions seen in PS
X-chromosome inactivation patterns	Skewed in carriers of X-linked DC
Ham test (CD59 analysis)	Abnormal in paroxysmal nocturnal haemoglobinuria
Telomere length	Short in AA, very short in DC
Constitutional karyotype	Abnormality suggestive of constitutional AA
Other investigations	
To identify somatic abnormalities	
Skeletal survey	Presence of somatic abnormalities in association with AA is suggestive of constitutional/inherited AA
Ultrasound of abdomen	
Pulmonary function tests	
Echocardiogram	
Exocrine pancreatic function	Abnormal in SDS and PS
Neutrophil chemotaxis	Abnormal in SDS
Fibroblast cultures	Abnormalities seen in DC

AA, aplastic anaemia; CAMT, congenital amegakaryocytic thrombocytopenia; CDA, congenital dyserythropoietic anaemia; DBA, Diamond–Blackfan anaemia; DC, dyskeratosis congenita; DEB/MMC, diepoxybutane/mitomycin-C; FA, Fanconi anaemia; PS, Pearson syndrome; SDS, Shwachman–Diamond syndrome.

Acknowledgements

I would like to thank my research colleagues past (Richard Beswick, Michael Kirwan, Stuart Knight, Anna Marrone, David Stevens and Philip Mason) and present (Shirleny Cardosa, Laura Collopy, Upal Hossain, Hemanth Tummala, Amanda Walne and Tom Vulliamy) and all the patients and clinicians on whom research in our laboratory depends. The research has been supported by funding from the Wellcome Trust and the Medical Research Council UK.

Selected bibliography

All inherited BMF syndromes

Alter BP, Young NS (1998). The bone marrow failure syndromes. In: *Haematology of Infancy and Childhood* (DG Nathan, HS Orkin, eds), pp 237–335. WB Saunders, Philadelphia, PA.

Dokal I (2014). Inherited bone marrow failure syndromes. *Hematology Education: the education program for the annual congress of the European Hematology Association* 2014 **8**: 299–308.

FA

Kottemann MC, Smogorzewska A (2013). Fanconi anaemia and the repair of Watson and Crick DNA crosslinks. *Nature* **493**: 356–63.

DC

Gramatges MM, Bertuch AA (2013). Short telomeres: from dyskeratosis congenita to sporadic aplastic anemia and malignancy. *Translational Research* **162**: 353–63.

SDS

Dror Y, Donadieu J, Koglmeier J *et al.* (2011). Draft consensus guidelines for diagnosis and treatment of Shwachman Diamond Syndrome. *Annals of the New York Academy of Science* **1242**: 40–55.

DBA

Ball S (2011). Diamond-Blackfan anemia. Hematology. American Society of Hematology Education Program **2011**: 487–91.

CDA

Iolascon A, Heimpel H, Wahlin A *et al.* (2013). Congenital dyserythropoietic anemias: molecular insights and diagnostic approach. *Blood* **122**: 2162–6.

Congenital and cyclical neutropenia

Hauck F, Klein C (2013). Pathogenic mechanisms and clinical implications of congenital neutropenia syndromes. *Current Opinion Allergy and Clinicl Immunology* **13**: 596–606.

CAMT

Ballmaier M, Germeshausen M (2011). Congenital amegakaryocytic thrombocytopenia: clinical presentation, diagnosis and management. *Seminars in Thrombosis and Hemostasis* **37**: 673–81.

CHAPTER 11

Acquired aplastic anaemia and paroxysmal nocturnal haemoglobinuria

11

Judith CW Marsh[1], Austin G Kulasekararaj[1], Neal S Young[3] and Peter Hillmen[2]

[1]King's College Hospital/King's College London, London, UK
[2]St James's University Hospital, Leeds, UK
[3]National Heart, Lung and Blood Institute, National Institutes of Health, Bethesda, Maryland, USA

Acquired aplastic anaemia

Characterization and definition

Aplastic anaemia (AA), a rare bone marrow failure syndrome (BMFS), is defined by pancytopenia with a hypocellular bone marrow in the absence of an abnormal infiltrate and with no increase in reticulin. The term 'aplastic anaemia' encompasses different entities, but here we discuss acquired AA, most often idiopathic, though sometimes a drug or chemical or a virus infection is implicated. 'Inevitable' AA/myelosuppression occurs after treatment with cytotoxic drugs or radiation; it is dose dependent and recovery is usually predictable, and is discussed no further here.

AA is a bone marrow failure disorder and shows considerable overlap with clonal disorders of bone marrow failure, including myelodysplastic syndrome (MDS), acute myeloid leukaemia (AML), paroxysmal nocturnal haemoglobinuria (PNH) and T-large granular lymphocyte leukaemia/lymphoproliferative disorder (T-LGL), and a tendency itself to later evolve to MDS/AML (Figure 11.1). The inherited forms of AA, such as Fanconi anaemia (FA), dyskeratosis congenita (DC) and Shwachmann–Diamond syndrome (SDS), are rarer than acquired AA and are discussed in detail in Chapter 10. However, the importance of excluding an inherited form of AA and the increasing proportion of adults with apparent acquired AA with related genetic lesions is highlighted in this chapter.

In AA there must be at least two of the following: (i) haemoglobin below 100 g/L, (ii) platelet count below 50×10^9/L and (iii) neutrophil count below 1.5×10^9/L. Patients with bilineage or trilineage cytopenias that are less severe than this are not classified as AA. However, they should have their blood counts monitored to determine whether they will develop AA later. The severity of the disease is graded into very severe, severe and non-severe AA, according to the blood count parameters and bone marrow findings, as summarized in Table 11.1. The assessment of disease severity is important in treatment decisions and is of prognostic significance.

Epidemiology

Aetiology and incidence

Studies indicate an incidence of AA in the West of 2 per million per year. There is a twofold to threefold higher incidence rate in Asia. In a large prospective study from Thailand, an incidence of 3.9 per million was reported from the metropolitan area of Bangkok compared with 5 per million in the northeast region of Khonkaen. An incidence of 7.4 per million was reported from a prospective study from China, although this may represent an overestimate, as a bone marrow trephine was not required for the diagnosis of AA. The reasons for the differences in incidence are not known, but may include environmental and genetic factors.

There is a biphasic age distribution, with peaks at 10–25 years and over 60 years; it is possible that some cases diagnosed

Postgraduate Haematology, Seventh Edition. Edited by A Victor Hoffbrand, Douglas R Higgs, David M Keeling and Atul B Mehta.

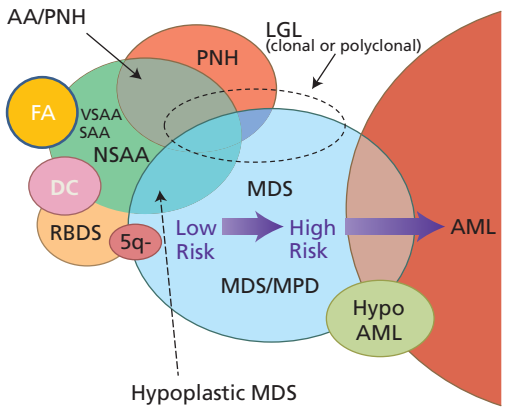


Figure 11.1 Overlapping bone marrow failure disorders. RBDS, ribosomal dysgenesis syndromes; DC, dyskeratosis congenital; FA, Fanconi anaemia; AA, aplastic anaemia; SAA, severe AA; NSAA, non-severe AA; VSAA, very severe AA; PNH, paroxysmal nocturnal haemoglobinuria; LGL, large granular lymphocytosis; AML, acute myeloid leukaemia.

in older patients represent hypocellular MDS rather than AA. There is no significant difference in incidence between males and females.

Posthepatitic AA accounts for around 5–10% of cases; invariably such cases are negative for the known hepatitic viruses. Patients, typically young males, present with jaundice and hepatitis symptoms, then, on average 6 weeks later, develop pancytopenia when the liver function has usually improved or normalized. Posthepatitic AA is usually severe, but responds well to immunosuppressive therapy (IST), and similar oligoclonal expansion of activated cytotoxic T cells is observed as in idiopathic AA, suggesting a common pathogenesis for the hepatitis and bone marrow failure. Rarely, AA may follow EBV-related infectious mononucleosis.

AA occurs in association with other systemic autoimmune disorders, especially eosinophilic fasciitis, but also systemic lupus erythematosus (SLE), Sjögren syndrome and coeliac disease. SLE can also produce pancytopenia with a cellular bone marrow or it may occasionally be associated with myelofibrosis. Thymoma may be associated with AA, although typically related to pure red cell aplasia (PRCA). Fatal AA is almost invariably the outcome of transfusion-associated graft-versus-host disease (GVHD); it is well documented clinically and can be reproduced *in vivo* in a mouse model.

The association between AA, other autoimmune disorders and TLGL, clonal proliferation of cytotoxic ($CD8^+$) T cells, has been known for decades and conceptualizes the theory of immune-mediated marrow suppression by a dominant T cell clone. Somatic *STAT3* mutations in autoimmune T cells in a proportion (7%) of AA possibly induces a persistently dysregulated autoimmunity, which is amenable to immunosuppression, akin to LGL leukaemia.

AA can rarely occur in pregnancy, although this may be due to chance and other possible causes should always be sought. A retrospective review from Leiden over a 24-year period found a similar frequency of AA in pregnancy compared to that in the general population. However, due to the rarity of AA, this may not exclude a possible association in a larger number of patients. The disease may remit spontaneously after termination, whether spontaneous or therapeutic, and after delivery, but not in all cases. There is a risk of relapse in pregnancy in patients who have previously responded to IST. In contrast, after successful allogeneic haemopoietic stem cell transplantation (HSCT), pregnancy does not trigger relapse.

Pathogenesis and its clinical relevance

The haemopoietic defect in AA

AA is characterized by both a quantitative and a qualitative defect in the haemopoietic stem cell compartment. The primitive long-term culture-initiating cells and more mature haemopoietic progenitors in the bone marrow (colony-forming cells) of all cell lineages are reduced or absent. In addition to the reduced numbers and impaired repopulating ability of bone marrow $CD34^+$ cells, these cells are more apoptotic than normal $CD34^+$ cells, as evidenced by transcriptome analysis. About 10–15% of patients with AA have shortened telomeres, as measured in blood leucocytes (see later) (Figure 11.2).

The bone marrow stromal cell microenvironment functions normally in most patients as assessed by an *in vitro* system, although this does not examine individual cellular components of the stroma and thus defects in particular cells cannot be excluded. The frequencies of fibroblast colonies (CFU-F) are normal in AA, and mesenchymal stem cells showed normal phenotype as defined by $CD34^-CD45^-CD44^+CD29^+CD90^+CD105^+CD106^+$. The differentiation capacity of AA mesenchymal stem cells has not yet been formally evaluated. Cultured AA mesenchymal stem cells support *in vitro* haemopoiesis after addition of normal bone marrow mononuclear cells.

AA is not due to a deficiency of any known haemopoietic growth factor (HGF). Long-term marrow culture studies have shown normal mRNA expression and/or secretion of granulocyte/macrophage colony-stimulating factor, granulocyte colony-stimulating factor (G-CSF), interleukin (IL)-6, stem cell factor and thrombopoietin (TPO) from stromal cells. Serum levels of most HGFs, including the above and recombinant human erythropoietin, are markedly elevated. These observations explain the lack of striking effects of HGFs in most AA patients, although recent data suggest a role for synthetic TPO-mimetic in refractory AA.

The immune-mediated nature of acquired AA

In acquired AA, it is proposed that an initiating event, such as a virus or drug, provokes an aberrant immune response,

Table 11.1 Laboratory assessment of a newly presenting patient.

1. Laboratory tests to confirm diagnosis	
FBC	Pancytopenia with reticulocytopenia. Single cytopenia especially thrombocytopenia may precede pancytopenia. Raised MCV common.
Blood film examination	Toxic granulation of neutrophils, anisopoikilocytosis, macrocytosis. Exclude blasts, dysplastic neutrophils and hairy cells.
Liver function tests, vitamin B12 and folate	To detect preceding hepatitis; exclude and correct vitamin B12 or folate deficiency.
Virology	Post-hepatitic AA occurs in 5–10% cases, serology usually negative for known hepatitis viruses; HIV and parvovirus are very rare causes of AA; check CMV exposure for potential BMT candidates.
Antinuclear antibody and dsDNA	SLE is very rare cause of AA, and other autoimmune syndromes associated with pancytopenia.
Radiology	Chest X-ray to exclude infection; HRCT scan of chest to exclude pulmonary fibrosis if DC or RUNX1 familial BMF suspected. Abdominal ultrasound or CT scan: splenomegaly or lymphadenothy indicate alternative diagnosis; abnormal or misplaced kidneys may occur in Fanconi anaemia.
BM examination	*Aspirate* fragments easily obtained, hypocellular fragments and trails. Reduced/absence of haemopoietic lineages. Dyserythropiesis common. Absence of dysplasia in granulocytic and megakaryocytic lineages. Lymphocytes, plasma cells, macrophages and mast cells often prominent. Iron stain to exclude ringed sideroblasts. CD34+ immunostain negative. *Trephine*: Good length (at least 1–2 cm) essential. Hypocellular. May be patchy cellularity. Absence of blasts. Lymphoid aggregates (reactive) common. Reticulin stain negative. CD34 immunocytochemistry negative. CD61 useful to exclude dysplastic megakaryocytes.
2. Tests to detect an associated abnormal clone	
PNH clone	Screen for glycosyl phosphatidyl inositol (GPI)-anchored proteins on red cells, granulocytes (and monocytes) using six-colour flow cytometry, including fluorescent aerolysin (FLAER). Size of PNH clone will help distinguish haemolytic PNH from AA/PNH and sub-clinical PNH. PNH clone detected in up to 40% of AA patients.
Cytogenetic clone using metaphase cytogenetics (MC) and cytoFISH	Detected in up to 12 % of AA patients in the absence of dysplastic changes. Monosomy 7 – typical of MDS and AA that has evolved to MDS. MC may fail due to hypocellular BM and small number of metaphases. Other cytogenetic abnormalities (del13q, del20q, +8) can be seen in morphologically typical AA.
T-large granular lymphocyte (T-LGL) clone	If blood/BM T-lymphocytosis, perform TCR gene rearrangement to detect clonal T-LGL; in this situation, exclude STAT3 mutation.
3. Tests to exclude an inherited bone marrow failure syndrome	
DEB test	PB chromosomal breakage (spontaneous and DEB/MMC-induced) increased in Fanconi anaemia. Screen patients up to 60 years old (FA has rarely been diagnosed in 5th decade) and all BMT candidates.
PB telomere length	Flow FISH or multiplex qPCR analysis. Very short telomeres (far below 1st percentile for age) are indicative of DC. Telomere gene mutation analysis then indicated (including TERC, TERT, TIN2, DKC1).
Other IBMFS	Specific gene mutation analysis when suspected, e.g. SBDS for Schwachmann–Diamond anaemia, GATA2 for Emberger's syndrome. See point number 6 for emerging technologies to detect other IBMFS.
4. Emerging diagnostic tests	
PB MDS gene mutation panel	Next generation sequencing technology to detect somatic mutations which may help distinguish AA from hypocellular MDS/AML.
Single nucleotide polymorphism (SNP)-array karyotyping	To detect abnormal genomic clones. Increases detection rate of chromosomal defects compared with MC. Use with MC.

Table 11.1 (*Continued*)

IBMFS custom array mutation analysis using exome/whole genome sequencing	To detect known (exome sequencing) or unknown (whole genome sequencing) IBMFS. See Chapter 19.
5. How severe is the aplastic anaemia?	
Severe AA (Camitta et al., 1976)	BM cellularity <25%, or 25–50% with <30% residual haemopoietic cells 2 out of 3 of the following: 1. neutrophils $<0.5 \times 10^9$/L 2. platelets $< 20 \times 10^9$/L 3. reticulocytes $< 60 \times 10^9$/L
Very severe AA (Bacigalupo *et al.*, 1988)	As for severe AA but neutrophils $<0.2 \times 10^9$/L.
Non-severe AA	Patients not fulfilling the criteria for severe or very severe aplastic anaemia.
6. Is there an alternative diagnosis?	
In the presence of pancytopenia and a hypocellular BM, exclude other possible causes:	
Hypocellular MDS/AML, and hypoALL in children	These features occur in hypoMDS/AML but not in AA: dysplastic neutrophils, BM dysgranulopoiesis and dysmegakayopoiesis, blasts in PB or BM, ringed sideroblasts, increased BM reticulin and ALIP.
Hodgkin or Non-Hodgkin lymphoma	Examine trephine carefully for foci of lymphoma and perform immunophenotyping and gene rearrangement analyses.
Myelofibrosis	Presence of splenomegaly, tear-drop poikilocytes and leucoerythroblastic changes on blood film, and increased BM reticulin and/or collagen are all features of myelofibrosis.
Mycobacterial infection, especially atypical forms	Other BM features of mycobacterial infection are granulomas, necrosis and increased reticulin.
Anorexia nervosa, starvation	Often presents with neutropenia before pancytopenia. BM features include stromal oedema and serous degeneration.
Thrombocytopenic disorders, e.g. ITP, amegakaryocytic thrombocytopenia (AMT)	Sometimes AA can present with isolated thrombocytopenia. BM in ITP usually shows normal or increased megakaryocytes, but occasionally reduced megakaryocytes. AMT can later progress to AA or MDS. Congenital AMT is due to c-MPL mutation.

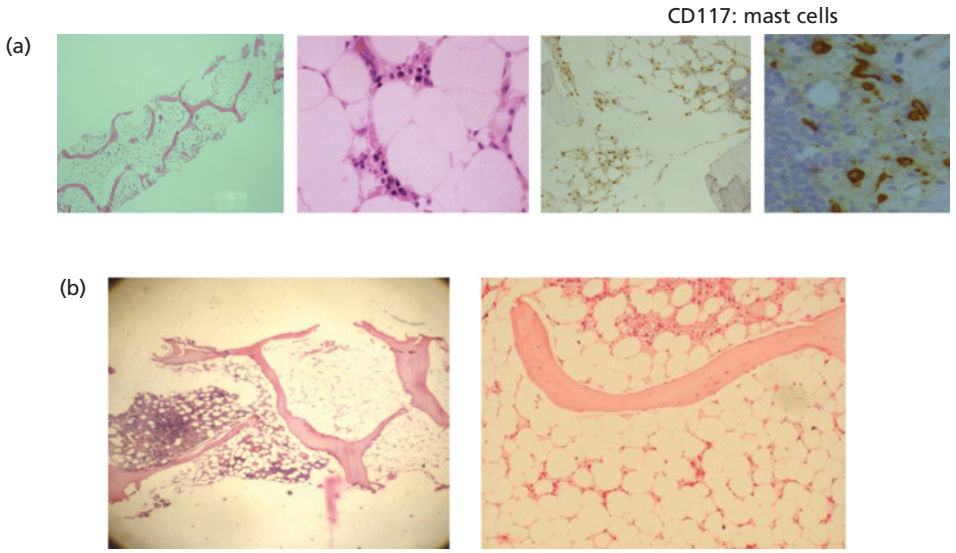


Figure 11.2 (a) BM trephine – severe AA; (b) Non-severe AA.

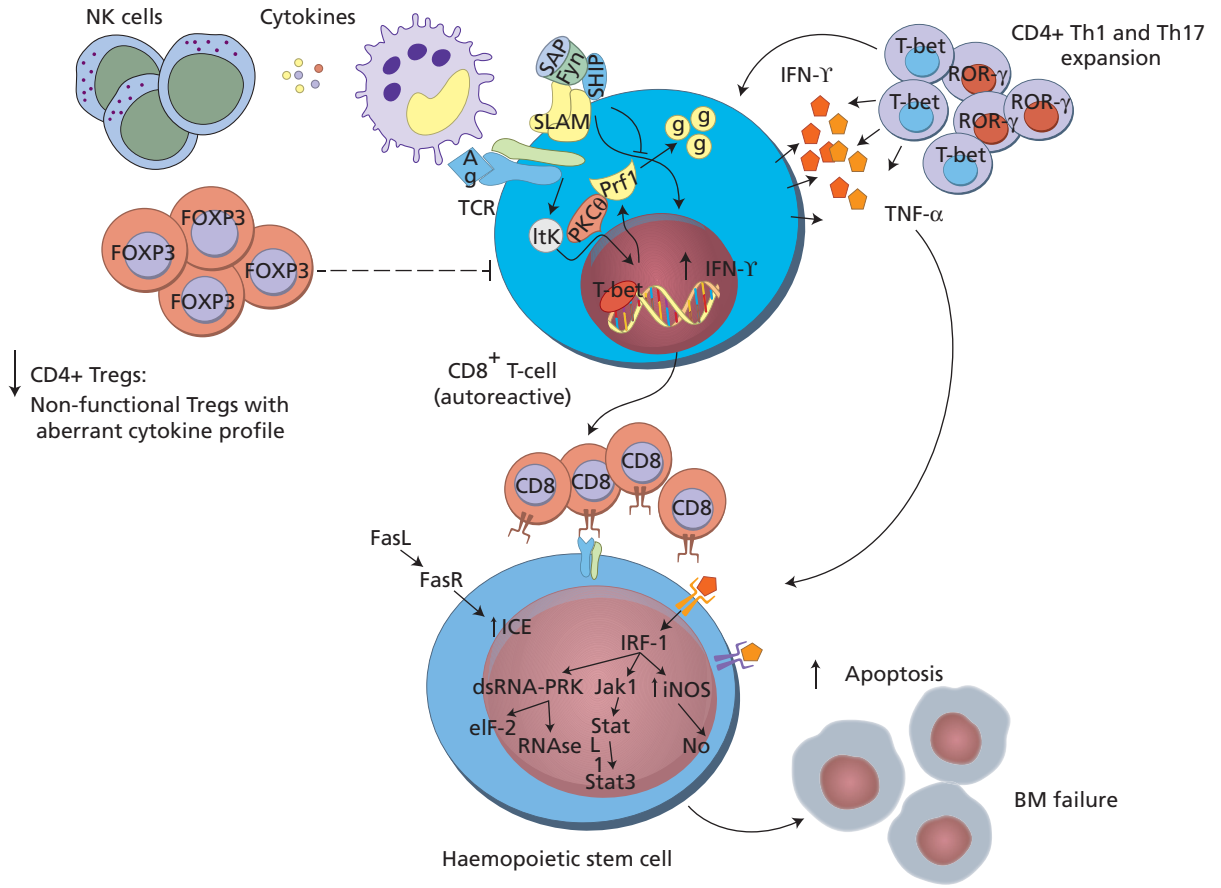


Figure 11.3 Immune mediated aplastic anaemia. Following an insult, most likely viral, there is immune-mediated damage to stem cells, resulting in a pro-inflammatory response. The most important component comprises the $CD4^+$ T-cells, with expansion of Th1 (clonal) and Th17 cells, and reduced and dysfunctional regulatory T-cells (Tregs). This leads to oligoclonal expansion of cytotoxic $CD8^+$ T-cells, resulting in apoptotic death of stem cells. (Source: Kordasti *et al.*, 2012 [*Blood* **119**(9):2033–43.]. Reproduced with permission of American Society of Hematology.)

triggering oligoclonal expansion of cytotoxic T cells that destroy haemopoietic stem cells (Figure 11.3). HSCT or IST leads to complete or partial response by eradicating or suppressing pathogenic T cell clones. Relapse occurs with recurrence of the immune response, and the immunologically stressed and depleted stem-cell compartment also allows selection of abnormal haemopoietic clones that manifest as MDS and occasionally AML.

Strong evidence that AA has an autoimmune nature is based on the following observations:

1 Haematological recovery after IST with antithymocyte globulin (ATG) and ciclosporin (CSA) occurs in the majority of patients.

2 There are activated autoreactive oligoclonal $CD8^+$ T cells present in blood and bone marrow that release interferon (IFN)-γ and tumour necrosis factor (TNF)-α, cytokines that inhibit haemopoiesis.

3 Intracellular IFN-γ levels in T cells correlate with response to IST.

4 Increased Fas expression on bone marrow $CD34^+$ cells indicates increased apoptosis.

5 T cell repertoire analysis shows oligoclonal expansion of $CD8^+$ T cells in AA, MDS and PNH.

6 Transcription factor T-bet is upregulated and binds to the IFN-γ promoter, resulting in increased expression of IFN-γ.

7 $CD4^+CD25^+FOXP3^+$ regulatory T cells are impaired and clonally restricted Th1 are expanded inducing an inflammatory microenvironment.

Gene expression profiling in AA $CD34^+$ cells has shown that more than half the upregulated genes are related to the immune response, including genes for cytokines and cytokine receptors and signal transduction genes, as well as other immune response genes. Many apoptosis and cell death genes are upregulated and some antiapoptotic genes downregulated.

It is unclear why T cells are activated in AA. HLA-DR2 and its split HLA-DR15 and DRB1*1501 and 1502 alleles are overrepresented in AA, as are class I HLA-B*4002 and HLA-A*0206, indicating a possible role for antigen recognition. There is correlation between ATG response and DRB1*1501, but most of these HLA data come from studies of Japanese patients. Studies of cytokine gene polymorphisms that may reflect a heightened immune response in AA are limited; polymorphisms in the TNF-α promoter, IL-6 and IFN-γ genes have been reported, but a systematic assessment of all potentially relevant cytokine genes has not been reported.

The specific cytotoxic T-cell targets on haemopoietic stem and progenitor cells in AA have not been identified. Potential candidates, identified by screening antibodies in patients' serum against a peptide library using leukaemia cell lines, include kinectin, DRS-1 (diazepam-binding inhibitor-related protein-1), PMS1 (postmeiotic segregation increased 1), moesin and hnRNPK (heterogeneous nuclear ribonucleoprotein K). However, the relevance of these findings is unclear, and they may represent epiphenomena rather than primary targets of cytotoxic T-cell attack.

Alternatively, a defect in the glycosylphosphatidylinositol (GPI) anchor may be the trigger for the immune response against normal progenitor cells through aberrant expression of intracellular GPI protein(s) while at the same time providing a protective mechanism for GPI-defective cells.

Detection of somatic mutations in AA

Using next generation sequencing, two large studies have recently reported similar incidence (20 to 25%) of somatic (acquired) mutations in AA in genes associated with myeloid malignancies (most frequently ASXL1, DNMT3A). These mutations are predictive for later evolution to MDS/AML especially in patients with history of AA for > 6 months, and also for worse response and survival following immunosuppressive therapy; in contrast, other mutations confer good prognosis (BCOR, BCORL1 and PIG-A genes). Their further significance in AA is discussed later (see 'Suggested further reading').

Genetic predisposition to AA and short telomeres

Inherited forms of AA have long been regarded as very rare causes of acquired AA, these including FA, DC and SDS (see Chapter 10 for detailed discussion). More recent awareness of these syndromes, with characterization of the genetic mutations in these and other bone marrow failure syndromes and the finding of very short telomeres in some patients with AA, indicate a more common occurrence than previously realized, and recognition that such cases may present later in adulthood.

The DC genes are involved in telomere maintenance, and all DC patients invariably have very short telomeres. About 10% of patients with apparent acquired AA have inherited mutations in *TERC*, *TERT*, with none or very mild mucocutaneous abnormalities. Families may have individuals with pulmonary fibrosis and hepatic cirrhosis, with or without cytopenia(s). SDS usually occurs in childhood with neutropenia, but can present in adult life.

In addition to mutations in the DC genes, other factors such as 'stressed' haemopoiesis as a result of reduced stem cell numbers in AA, and environmental factors such as smoking, contribute to the shortened telomeres. Telomeres are assumed to be short in all organs, but they are only measured in blood leucocytes (the tissue for which there are sufficient age-matched controls). The consequences of telomere shortening include genomic instability, resulting in increased risk of malignant transformation (MDS, AML), bone marrow failure and defects in repair and regeneration (pulmonary fibrosis, cirrhosis). The clinical consequences of missing late-onset inherited forms of AA, as they often lack somatic anomalies, include likely mortality after allogeneic HSCT using standard conditioning regimens in FA due to the defect in DNA repair, and in DC due to the higher risk of liver and/or respiratory failure, and failure to screen potential family bone marrow donors for inherited bone marrow failure.

That not all patients with acquired AA respond to IST could be explained in pathogenetic terms by: (i) complete stem cell depletion due to a severe autoimmune attack, (ii) insufficient degree of IST or (iii) a non-immune basis for the bone marrow failure, of which one mechanism may be a genetic predisposition resulting in genomic instability. The presence of shortened telomeres in idiopathic AA does not preclude initial response to IST, but predicts relapse after ATG and also late clonal disorders. There are insufficient data in patients with *TERC* or *TERT* mutations on response to IST (see also section 'Predictors of response to ATG').

Clinical features

Patients with AA most commonly present with symptoms of anaemia and skin or mucosal haemorrhage (ecchymoses or petechiae), including buccal haemorrhages, or visual disturbance due to retinal haemorrhage. Infection, particularly sore throat, may be a presenting feature, but is less common. There is no lymphadenopathy or hepatosplenomegaly in the absence of infection. An important careful history and clinical examination help to exclude an inherited form of AA, especially in children and young adults. This also applies to older patients as it has more recently been realized that typical features of inherited bone marrow failure syndromes may be absent and/or other features may be present in older patients. The findings of short stature, abnormal thumbs and forearms, café-au-lait spots and skeletal anomalies would raise the possibility of an inherited form of AA, specifically FA, although FA can present in adults without physical anomalies. Patients with FA most commonly present between the ages of 3 and 14 years, but can occasionally present later in their thirties or forties and very rarely early

fifties. The findings of leucoplakia, nail dystrophy and pigmentation of the skin are characteristic of DC, with a median age at presentation of 7 years (range 6 months to 26 years). Some affected patients may have none of these clinical features and the diagnosis is made later after failure to respond to IST. For patients with AA and an inherited *TERC*/*TERT* mutation, the family history is very important as the pedigree may reveal pulmonary fibrosis, cirrhosis, osteoporosis, avascular necrosis of bone, low blood counts or cancer. A previous history of malabsorption or neutropenia may underlie a diagnosis of SDS in children or young adults, especially as the malabsorption often resolves in later life. A preceding history of jaundice, usually 2–3 months, strongly suggest sero-negative, posthepatitic AA.

Diagnostic investigations and differential diagnosis (Table 11.1)

To diagnose AA with certainty, it is important to consider other possible causes of pancytopenia with a hypocellular bone marrow and to exclude an inherited form of AA, as this will have important implications for treatment options, the choice of conditioning regimen and donor for HSCT, genetic screening and counselling of family members. All patients should be screened by cytogenetics and for a PNH clone. See Table 11.1 for summary of laboratory assessment of a newly presenting patient.

Management

General comments relating to the management of AA

Because AA is a rare disease, the haematologist responsible for the patient should contact a centre/specialist with expertise in AA soon after presentation to discuss a management plan for the patient. Patients should be offered the opportunity to be reviewed in a specialist centre. Whenever possible, patients should be enrolled into prospective national or international trials.

Prior to administration of specific treatment for the disease, the patient should be stabilized clinically in terms of controlling bleeding and treating infection. The presence of infection is an adverse factor for outcome after HSCT, although it may sometimes be necessary to proceed with HSCT in the presence of active infection, particularly fungal infection. In this setting, transplantation may offer the best chance of early neutrophil recovery, and delaying the transplant may risk progression of the fungal infection.

Supportive care

Transfusions

Support with red cell and platelet transfusions is essential for patients with AA to maintain a safe blood count. Current practice is to give prophylactic platelets when the platelet count is below 10×10^9/L (or higher in the presence of fever, sepsis or bleeding). Fatal haemorrhage, usually cerebral, is more common in patients who have less than 10×10^9/L platelets, extensive retinal haemorrhages, buccal haemorrhages or rapidly spreading purpura. Other important practical measures to help prevent bleeding include good dental hygiene, and control of menorrhagia with appropriate hormone therapy.

A common problem in multitransfused AA patients is alloimmunization to leucocytes present in red cell and platelet transfusions by generating HLA or non-HLA (minor histocompatibility) antibodies. Universal leucodepletion has probably reduced, but not eliminated the incidence of alloimmunization; the incidence in AA was around 30–50% in the preleucodepletion era. Alloimmunization can result in platelet refractoriness, as well as an increased risk of graft rejection after allogeneic BMT. Patients who become refractory to platelet transfusions should be screened for HLA antibodies, once other non-immune causes have been excluded, if present HLA-matched platelets should be used. Novel approaches, including the use of HLA epitope-matched platelets, are currently being evaluated in clinical trials in view of the shrinking HLA-typed donor pool.

Platelet transfusions should be given to maintain a safe platelet count and blood transfusions to allow normal daily activities. Transfusions should not be withheld for fear of sensitizing the patient. Directed blood and platelet donations from family members should be avoided, as the recipient may become sensitized to minor histocompatibility antigens from the potential bone marrow donor, increasing the risk of graft rejection.

Even though there is no evidence base, the British Committee for Standards in Haematology now recommends empirically the use of irradiated blood components for all AA patients receiving ATG. The rationale for this is twofold. Firstly, animal data demonstrate that irradiation of all red cell and platelet transfusions before HSCT in addition to leucodepletion, further reduces the risk of sensitization to minor histocompatibility antigens (and hence reduced risk of graft rejection after allogeneic HSCT). Secondly, irradiated blood products would abolish the potential risk of transfusion-associated GVHD associated with ATG therapy. Transfusion-associated GVHD has very rarely been reported after ATG. Rabbit ATG is more immunosuppressive than horse ATG, results in a prolonged period of lymphopenia and has a longer half-life and higher-affinity IgG subtype to human lymphocytes than does horse ATG.

The use of granulocyte transfusions is being re-evaluated as supportive therapy in patients with life-threatening neutropenia. Adverse events, such as febrile reactions, HLA alloimmunization and transfusion-related acute lung injury, are well-recognized complications following granulocyte transfusions. Irradiated granulocyte transfusions are used to support patients with fungal/mould disease or bacterial sepsis who are severely neutropenic and not adequately responding to maximal antibiotics.

Iron chelation therapy

Many AA patients require multiple red cell transfusions leading to transfusion haemosiderosis. Iron overload impacts adversely on outcome after allogeneic BMT. Iron chelation therapy should commence when the serum ferritin is above 1000 μg/L, although the evidence base for this is lacking. It may be difficult to deliver subcutaneous desferrioxamine to AA patients on account of local haemorrhage and infection from subcutaneous injections. Alternative oral agents include deferiprone and deferasirox (Exjade). Because of a high incidence of agranulocytosis with deferiprone, it is not used routinely in patients with AA. Results of a recent prospective study of deferasirox in 166 iron-overloaded AA patients showed the safety and also the ability to reduce ferritin (median fall in serum ferritin of 946 μg/L), although in view of renal toxicity of the drug, extra vigilance should be used in those on CSA.

Infections

Patients with AA are at risk of bacterial and fungal infections, depending in part on the degree of neutropenia. Very severe neutropenia (neutrophils $< 0.2 \times 10^9$/L) carries a high risk. If there is severe neutropenia, either non-absorbable antibiotics or oral quinolone are administered to reduce the potential pathogenic load from the gastrointestinal tract, although neither is standard practice in the USA. Oral hygiene is important. Entry sites for venous access are potential sources of systemic infection. Patients with AA are at high risk of fungal infection, including *Aspergillus*. Fluconazole provides no protection against *Aspergillus* species, for which the drugs of choice are itraconazole, which has clinically significant, but manageable or avoidable interactions with other drugs, and posaconazole, which has not yet been shown to be superior in efficacy to itraconazole.

Fever should be treated with broad-spectrum antibiotics without waiting for laboratory isolation of organisms, and with early introduction of systemic antifungal therapy if fever fails to respond to antibiotics. Chest radiography, CT scan of chest and sinuses should be part of the investigation of new or persistent fever. G-CSF is usually ineffective in severe AA due to a marked reduction or absence of myeloid progenitor cells, although it is reasonable to consider giving a short course in life-threatening infections.

Indications for treatment of AA (Figure 11.4)

The two main treatment options for AA are allogeneic HSCT or IST with ATG and CSA, based on disease severity, age of the patient, availability of a matched sibling donor, patient comorbidities and patient choice. For patients with severe AA who are under 35–50 years of age, the first-line treatment is matched sibling donor HSCT. EBMT data show similar outcomes for patients aged 40–50 to those aged 30–40 years. However, comorbidities should be carefully assessed to determine fitness for up-front transplant instead of immunosuppressive therapy for patients aged 35–50 years. For all other patients who require treatment, namely those with non-severe AA or those older than 35–50 years, IST is the first option. Unrelated donor HSCT is indicated for adults after failure to respond to one course of IST, for those patients who lack a matched sibling donor. Upfront MUD HSCT may be considered in children who lack a MSD if a suitable donor is readily available.

The standard regimen for IST comprises the combination of horse ATG and CSA. Corticosteroids should not be given to treat AA, except to prevent and treat serum sickness, as they are ineffective and encourage fungal and bacterial infections.

Immunosuppressive therapy

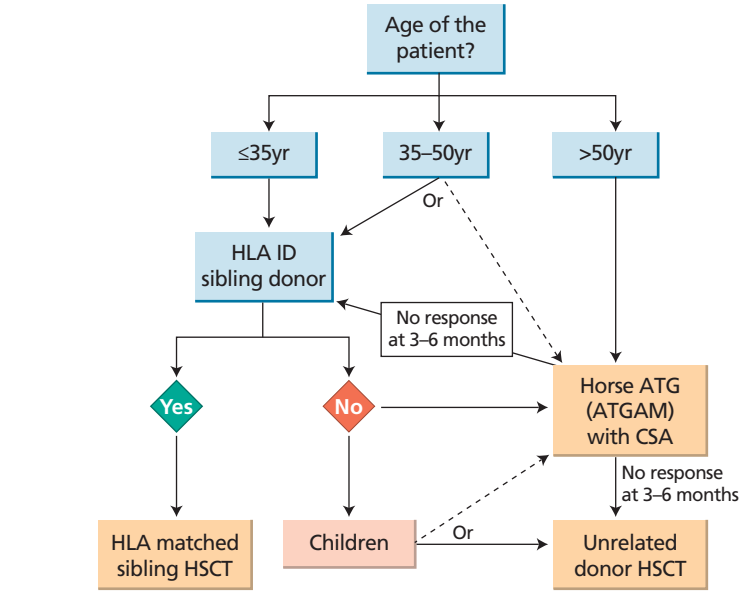
Antithymocyte globulin: properties, mechanism of action and administration

ATG is a polyclonal IgG antibody preparation produced by immunizing horses or rabbits with human thymocytes. Its mechanism of action in AA is not entirely clear. Possible mechanisms include the following:

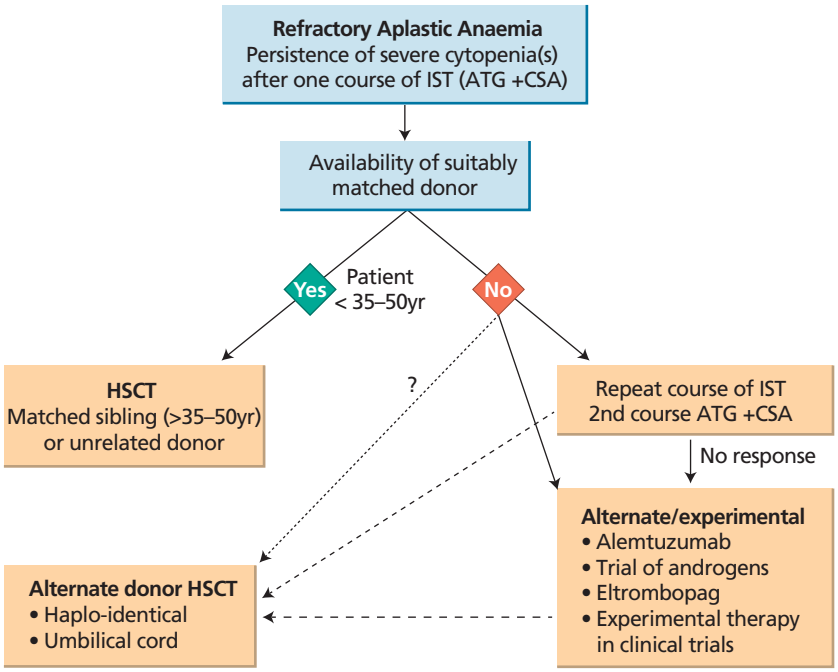
1 T cell depletion by complement-mediated lysis;
2 destruction of activated cytotoxic T lymphocytes by Fas-mediated apoptosis and antibody-dependent cellular cytotoxicity;
3 reduced apoptosis and Fas expression on AA $CD34^+$ bone marrow cells;
4 direct stimulation of normal and AA $CD34^+$ bone marrow cells and a mitogenic effect in the absence of complement resulting in release of HGFs;
5 direct stimulation of T regulatory cells has been demonstrated when normal bone marrow mononuclear cells are incubated with rabbit ATG.

In the UK, most of Europe and many other countries, the standard preparation of ATG has until recently been horse ATG (Lymphoglobuline, Genzyme). The rabbit preparation (Thymoglobuline, Genzyme) was usually reserved for second or subsequent courses. From June 2007, horse ATG (Lymphoglobuline) was withdrawn due to manufacturing difficulties in maintaining quality control. In the USA, horse ATG (ATGAM, Pfizer) is still widely used. Horse ATG (ATGAM, Pfizer) is preferred to rabbit ATG as it results in a higher response rate (68% at 6 months compared to 37% for rabbit ATG, Table 11.2) and better survival, despite rabbit ATG being more immunosuppressive than horse ATG in terms of the degree and duration of lymphodepletion, indicating that different mechanisms are important in the mode of action of the two agents.

ATG is a powerful immunosuppressive drug, given as an inpatient treatment and requires very careful monitoring, the prophylaxis and treatment of fevers and infections, as well as adequate (and sometimes intensive) platelet transfusion support. ATG is given as a daily intravenous infusion usually over 12–18 hours through a central venous catheter. Instead of giving a separate test dose, most centres give the first 100 mL of the first



(a)



(b)

Figure 11.4 (a) Treatment of severe aplastic anaemia; (b) treatment of refractory severe aplastic anaemia. For patients age 35–50 or >50 years, who fail to respond to first line IST, HSCT may be considered, using a matched sibling donor, or a suitably matched unrelated donor if no matched sibling is available.

infusion very slowly over 1 hour to assess for possible anaphylaxis. Allergists maintain that this is inadequate to prevent anaphylaxis as the amounts delivered are relatively large. The dosage of horse ATG (ATGAM) is 40 mg/kg daily for 4 days and rabbit ATG is 3.75 mg/kg for 5 days.

Immediate side-effects are allergic and occur commonly, including fever, rigors, rash, hypertension or hypotension and fluid retention. Each daily dose should be preceded by intravenous methylprednisolone and chlorpheniramine. In the USA, patients are pretreated with diphenhydramine and pethidine if fever and chills have occurred. Platelet transfusions should be given to maintain a safe platelet count (ideally $> 30 \times 10^9$/L; in the USA, a threshold of $> 20 \times 10^9$/L is used). Prior to starting ATG, patients should be assessed to ensure adequate platelet increments with random donor platelets. Poor platelet increments should be investigated beforehand. Patients are often in isolation with reverse barrier nursing, but this is not standard practice in the USA. Fevers are treated with

Table 11.2 Direct comparisons of horse and rabbit ATG with CSA as first line treatment for aplastic anaemia in the last 5 years.

		N, preparation		Response at 6 months		Overall survival		
	Study design	**Horse ATG**	**Rabbit ATG**	**Horse ATG**	**Rabbit ATG**	**Horse ATG**	**Rabbit ATG**	**Comments**
Scheinberg, 2011	Prospective, randomized	60, ATGAM	60, Thymoglobulin	68%	37%	96%	76% at 3yr	Significant difference between ATG preparations
Marsh, 2012	Prospective	105, Lymphoglobulin	35, Thymoglobulin	67%	40%	86%	68% at 2yr	Matched pair analysis between two preparations; Significant differences between ATGs. More infective deaths with rabbit ATG.
Jeong, 2014	Retrospective	297, Lymphoglobulin	158, Thymoglobulin	60%	55%	96%	87%	Study in children. Significant difference in survival but not response. More infections with rabbit ATG.
Afable, 2011	Retrospective	67, ATGAM	20, Thymoglobulin	58%	45%	76%	64% at 2yr	No significant difference between ATGs but study likely under-powered.
Atta, 2010	Retrospective	42, Lymphoglobulin	29, Thymoglobulin	60%	35%	78%	55% at 2yr	Significant difference between ATG preparations.
Chen, 2013	Retrospective	42, Lymphoglobulin	53, Thymoglobulin	39%	45%	84%	83%	No significant difference between ATGs

broad-spectrum antibiotics. Intravenous methylprednisolone (or oral prednisolone) and paracetamol are given at least 30 min before each daily dose of rabbit ATG 1–2 mg/kg, with reduction of the prednisolone dose by half every 5 days. Prednisolone is given to help prevent serum sickness. Serum sickness typically occurs between 7 and 14 days from the start of ATG treatment. The common symptoms of serum sickness include arthralgia, myalgia, rash, fever, mild proteinuria and platelet consumption, often necessitating increased platelet transfusion support.

There are no indications for routine use of daily G-CSF (administered for 3 months) with ATG and CSA as three prospective randomized studies comparing ATG, CSA and G-CSF with ATG and CSA alone demonstrated no difference in response and survival between the two groups, and there is concern about G-CSF stimulating low-level monosomy 7 clones. Attempts to augment responses and reduce relapse by incorporation of high-dose corticosteroids, mycophenolate mofetil (MMF), sirolimus or androgens to standard IST (ATG and CSA) have not met with any success.

Survival after ATG

Data from EBMT and large single centres confirm that over the last three decades overall survival after ATG therapy has improved steadily, with 5-year survival of around 80% reported (Table 11.2). The reasons for improved survival have been difficult to assess formally, but likely factors include: (i) improved supportive care for prevention and treatment of infections, especially superior antifungals and better availability and quality of blood products, (ii) the use of CSA with ATG and (iii) successful second-line therapies, including another round of ATG, unrelated donor HSCT (UDBMT) in younger patients and allogeneic HSCT in older adults (see also Chapter 35).

Predictors of response to ATG

Patients with non-severe AA are more likely to respond than are patients with severe or very severe AA. Other predictors are younger age, absolute reticulocyte count (ARC) $>25 \times 10^9$/L and absolute lymphocyte count (ALC) $>1 \times 10^9$/L, which correlate with response to ATG.

Several studies have examined whether the presence of a PNH clone is associated with response to ATG, with conflicting results. This may reflect differences in sensitivity of the test used to detect a PNH clone. Those using very sensitive tests to detect GPI-deficient clones of less than 0.003 cells show a strong correlation with response.

HLA type may also correlate with response to immunosuppressive therapy. A Japanese study showed that HLA-DRB1*1501 associated with response to CSA. Although HLA-DR15 or DRB1*1501 predicts response to ATG in white patients with MDS, no association was seen in AA patients. More recently, the same Japanese group demonstrated that of the 30 different DRB1 alleles, only DRB1*1501 and DRB1*1502 occurred more frequently in AA than in controls, and that response to ATG and CSA was only associated with DRB1*1501 in the presence of a PNH clone. Thus specific HLA haplotypes correlate with response to immunosuppressive therapy in Japanese patients, but similar studies in white Europeans are lacking.

Recent studies from NIH indicate that shortened telomere length does not preclude initial response to immunosuppressive therapy, but predicts relapse after ATG and is a risk factor for later cytogenetic abnormalities and evolution to MDS and AML. Thus telomere length at diagnosis may reflect depleted stem cell reserve so that prolonged stem cell division is not possible, resulting in later relapse; later clonal evolution reflects genomic instability of the critically shortened telomeres.

Repeat courses of ATG

A second course of immunosuppression may be administered (horse ATG, rabbit ATG, or alemtuzumab). The response rate following a second course of ATG (horse or rabbit) for relapse is around 60%, but only about 35% for patients who did not respond to the first course. A study from Japan examined prospectively the outcome of 52 children who failed one course of horse ATG, and who went on to receive either a second course of horse ATG or an unrelated donor HSCT. The response to a second course of ATG was only 11%, with a 5-year failure-free survival of only 9.5%. Another study has assessed the value of giving three courses of ATG. Among those patients who showed no response to the first or second courses, there were no sustained responses, but for those who had relapsed after two previous courses, all responded to a third. For patients receiving more than one course of ATG, the risk of anaphylaxis may be higher. Serum sickness is still unpredictable, though it may occur earlier following a second course of the same animal preparation of ATG.

Late complications after ATG

Relapse occurs in up a third of patients when CSA is withdrawn at 6 months. A more prolonged course of CSA with a later slow tapering of the drug reduces the relapse risk to around 13–16%, although recent data from NIH show that a long course of CSA only delays but does not prevent relapse. About one-third of patients are CSA dependent and require a small dose long term. CSA can be continued for at least 12 months after a maximal response before starting to taper the drug, followed by a very slow taper, for example by 25 mg every 3 months.

Patients treated with ATG are also at risk of developing clonal disorders, such as PNH, MDS and AML. Long-term follow-up from a large NIH cohort showed 35% relapse at 5 years, and 10% risk of PNH and 15% risk of clonal evolution, excluding PNH, at 10 years. A previous study from NIH showed 20% risk at 10 years of clonal cytogenetic abnormalities after ATG. Reported risk factors for developing MDS/AML include repeated courses of ATG, older age, and high doses, prolonged duration of G-CSF with

ATG and CSA, shortened telomeres and presence of acquired somatic mutations in genes known to be mutated in MDS/AML.

ATG treatment in older patients and children

For patients > 60 years old, the response rate and survival rate are lower compared with younger patients, with a higher risk of bacterial and fungal infections, bleeding and significant cardiac events after ATG. Although there is no upper age limit for ATG treatment, consideration for treatment should be preceded by medical assessment to exclude significant comorbidities and hypocellular MDS. For older patients who are not candidates for ATG, CSA may be considered, but there is increased risk of renal toxicity and hypertension.

Excellent response to ATG is seen in children. Current response rates are around 75%, with 90% long-term survival, but issues of relapse and clonal evolution persist. In view of excellent outcomes after MUD HSCT in children using alemtuzumab-based regimens, upfront MUD HSCT may be considered instead of ATG when an MSD is not available.

Other drugs used in AA

Cyclophosphamide

High-dose (200 mg/kg) and recently 'moderate-dose' (120 mg/kg) cyclophosphamide (Cy) has been used to treat AA in the absence of stem cell support, following the observation that complete autologous haematological recovery occurs in a small number of patients undergoing allogeneic sibling BMT using Cy. Although durable responses were seen, the predictable and markedly prolonged cytopenias exposed patients to a high risk of fatal fungal infections and a significant increase in use of blood and platelet transfusions, days of intravenous antibiotics and amphotericin, and inpatient days in hospital. In addition, HDC treatment does not eliminate the risk of clonal events or relapse. Consequently, the use of Cy in the treatment of AA is not recommended.

Alemtuzumab

The anti-CD52 monoclonal antibody alemtuzumab (Campath-1H) is effective in other autoimmune disorders such as multiple sclerosis, autoimmune cytopenias and the vasculitides. NIH investigated the use of alemtuzumab monotherapy (10 mg intravenous for 10 days) without CSA in AA and induced heterogeneous responses; 56% in relapsed setting, 37% in refractory disease and 19% in treatment-naïve AA. Alemtuzumab can alternatively be given subcutaneously. The ease of administration, its efficacy even without concurrent use of CSA and the relatively good safety profile, makes alemtuzumab an alternative choice of IST in relapsed/refractory AA for salvaging transplant-ineligible patients.

Eltrombopag

Eltrombopag, an oral thrombopoietin mimetic, licensed in chronic immune thrombocytopenic purpura, induces platelet maturation and release by binding to c-MPL receptors on megakaryocytes (Chapter 42). In a recent prospective study of 43 patients with refractory severe AA, eltrombopag induced responses in 40% patients; responses included tri/bilineage, were robust and led to normalization of cellularity. Trilineage responses, although surprising, might indicate stimulation of c-MPL receptors on remaining stem cells. The drug was well tolerated and there were no reports of increased reticulin or collagen fibrosis in the bone marrow, although new cytogenetic abnormalities, especially monosomy 7, were detected. The safety of eltrombopag needs to be evaluated further in prospective clinical trials, especially in view of the possible link to clonal evolution. The early and rapid response seen with eltrombopag has led to design of prospective trials with incorporation of this agent along with ATG+CSA for first-line treatment of AA. Eltrombopag has now had FDA approval for refractory SAA.

Androgens

Androgens, such as oxymetholone, have been used historically for the treatment of AA prior to availability of ATG/CSA, subsequently as an adjunct to ATG and are still used in certain developing countries for first-line treatment of AA. Androgens lead to increased telomerase activity via aromatization of oestradiol to steroids and hence induce responses in patients with telomeropathies who manifest as apparent acquired AA.

Haemopoietic stem cell transplantation

Outcomes following allogeneic HSCT for acquired AA, whether from a matched sibling donor (MSD) or a matched unrelated donor (MUD), have steadily improved over time. More recently, significant improvements have been observed for matched unrelated donor HSCT for AA patients failing immunosuppressive therapy, with 80–90% survival and the best results seen in younger patients (see also Chapter 35).

Allogeneic HSCT

Conditioning regimens

For MSD HSCT, the standard immunosuppressive, non-myeloablative conditioning regimen for patients < 30 years old is high-dose cyclophosphamide 200 mg/kg, which results in survival of at least 80% in younger patients For patients > 30 years old, a fludarabine-based regimen is used, with a lower dose of cyclophosphamide, with either ATG ('FCATG') or alemtuzumab ('FCC'), to reduce rejection and GVHD, and improve outcomes. Alemtuzumab is associated with a lower incidence of GVHD, especially chronic GVHD compared to ATG. Standard GVHD prophylaxis is with CSA and methotrexate, although CSA alone is sufficient in the setting of alemtuzumab. Because AA patients are at risk of late graft failure, CSA is usually continued for 12 months after HSCT.

For MUD HSCT, the conditioning regimen is identical to that used for older patients undergoing MSD HSCT using either FCATG, with addition of low dose (2Gy) TBI, or FCC

without TBI. Recent studies using alemtuzumab-based conditioning show 90% OS in children and around 80% in adults; and 70% in patients > 50 years old. High-resolution HLA tissue typing is now routine for selection of MUDs.

Dose and source of haemopoietic stem cells

It is important to give at least 3×10^8 nucleated marrow cells/kg and $> 2 \times 10^6$ $CD34^+$ cells/kg, because lower doses delay neutrophil engraftment and increase the risk of graft rejection. The strongly preferred stem cell source is BM for ATG-containing regimens, due to higher incidence of chronic GVHD and worse survival with PBSC, but either BM or PBSC can be used for alemtuzumab-based conditioning.

Chimerism

Using polymerase chain reaction for short tandem repeats on unfractionated bone marrow or peripheral blood mononuclear cells, the Dublin group, in collaboration with EBMT, reported mixed chimerism in 25% of HSCTs for AA. Patients with mixed chimerism may have either stable or progressive mixed chimerism. All the cases of graft failure occurred in the progressive mixed chimerism group, of whom 50% rejected their grafts. More often now, chimerism is examined in T-cells (CD3) and myeloid cells (CD15) and most patients have stable mixed T cell chimerism with full donor myeloid chimerism when using alemtuzumab-based conditioning, with excellent survival and very low incidence of chronic GVHD, suggesting a state of tolerance may be achieved.

Graft-versus-host disease

GVHD (especially chronic) is one of the most serious complications, as it adversely impacts on performance status, quality of life and survival after marrow transplantation, especially in a disease like AA where there is no advantage in having a graft-versus-disease effect (Table 11.3), but the risk is reduced when using alemtuzumab-based conditioning.

The risk factors for chronic GVHD include: (i) history of prior acute GVHD, (ii) infusion of non-irradiated donor buffy coat cells (this manoeuvre was used previously to overcome the problem of graft rejection, but was abandoned because of the high incidence of chronic GVHD), (iii) older patient age, (iv) the use of peripheral blood stem cells and (v) the use of ATG compared with alemtuzumab-based conditioning regimens.

Long-term complications of HSCT

Fertility is usually well preserved when irradiation is not used in the conditioning regimen, and the chances of pregnancy or fathering a child are much higher in patients who receive a marrow transplant for AA than for haematological malignancies. Pregnancies in patients transplanted for AA usually have a successful outcome. Less long-term data are available using fludarabine with lower-dose Cy regimens, although cases of successful pregnancy have been reported. For patients of child-bearing age, referral to an assisted conception unit for discussions on fertility should be offered. Men should be offered sperm storage.

There is an increased frequency of solid tumours after HSCT for AA, but the risk is lower than in patients transplanted for haematological malignancies when irradiation is avoided. The risk is also lower in the absence of chronic GVHD. Monitoring of late effects should follow international guidelines, including routine surveillance for secondary malignancy, endocrine, metabolic, bone (including avascular necrosis) and cardiovascular risks.

Table 11.3 Outcomes of HLA matched sibling donor HSCT from recent studies using high dose cyclophosphamide (Cy) ± antithymocyte globulin (ATG) conditioning.

Study	Number of patients	Conditioning	Follow-up (years)	Overall survival	Acute GVHD	Chronic GVHD	Graft failure
Bacigalupo, 2012 EBMT study	1886	Cy±ATG predominantly		BM: 92% PB: 80%	BM: 11% PB: 17% (Gd II-IV)	BM: 11% PB: 22%	BM: 5.5% PB: 7.2%
Konopacki, 2012 Paris study	61	Cy+ATG	6 (0.6–19)	85% at 6 yr	23% Gd II-IV	32% CI	3.2%
Sanders, 2011 Seattle study, children	118	Cy±ATG	22 (1–38)	82% at 30 yr	30% Gd II-III	30%	N/A; autologous recovery in 2
Gupta, 2010, CIBMTR study	1307	Cy±ATG predominantly		Age <20: 82% 20-40: 72% >40: 53% at 5 yr	11% 17% 27%	11% 25% 27%	17% (lack of 13% neutr. 12% recovery)

ATG, antithymosyte globulin; BM, bone marrow; CIBMTR, Center for International Blood and Marrow Transplant Research; Cy, ciclosporin; EBMT, European Group for Blood and Marrow Transplantation; gd, grade; N/A, not available; PB, peripheral blood.

Alternative donor HSCT

Umbilical cord blood transplantation

Experience in acquired AA is still limited. The largest study of unrelated cord blood transplantation (CBT) in acquired AA is reported by the combined EUROCORD/EBMT SAAWP. The main problem was engraftment failure in around 50% patients. A reduced intensity conditioning regimen is used. Because of the high risk of non-engraftment, a higher cell dose (EBMT recommendation is to use $> 4 \times 10^7$ TNC/kg) is required compared to doses used in CBT for leukaemia and with no more than 2/6 HLA mismatches in the cord unit(s). As for haploidentical HSCT (see below), all patients must be screened for HLA antibodies that may be directed against HLA antigens present on the cord units and thereby increase the risk of rejection, so that only cord units lacking that antigen(s) are used.

Haploidentical HSCT

The key advantages of haploidentical HSCT are that a graft is available for most patients, the time to procure the graft is short, and the cost is low. A novel approach is reduced intensity conditioning with high dose Cy given on day + 3 and + 4 post-transplant to prevent GVHD by depleting alloreactive donor T cells, but sparing quiescent, non-alloreactive T cells. This promising approach is undergoing further evaluation in AA throughout Europe.

Paroxysmal nocturnal haemoglobinuria

Introduction

Paroxysmal nocturnal haemoglobinuria (PNH) has fascinated haematologists since its first definitive description in 1882 by Paul Strübbing. PNH is unique as an acquired haemolytic disorder in which the defect is intrinsic to the red cell. Patients have a propensity to develop thromboses that are frequently life-threatening. PNH is also associated with bone marrow failure and indeed may provide a unique insight into the pathophysiology of a variety of bone marrow failure syndromes. Recently, the development of the terminal complement inhibitor eculizumab has revolutionized the treatment of PNH, but in turn has revealed further insights into the pathophysiology of the disease and into normal physiology. The treatment of PNH with eculizumab is not without problems in that a minority of patients have suboptimal responses and require modifications to maximize the benefits of therapy.

Pathophysiology

Glycosylphosphatidylinositol defect

PNH results from the expansion of an abnormal haemopoietic clone that arises following an acquired mutation in a gene critical for the biosynthesis of glycosylphosphatidylinositol (GPI) structures. These highly conserved GPI anchors are glycolipid structures through which a large number of cell-surface antigens are attached to the cell membrane. Almost all patients with PNH have the same biosynthetic defect in one of the early steps of the pathway, namely the transfer of *N*-acetylglucosamine from UDP-*N*-acetylglucosamine to phosphatidylinositol. In 1993, Miyata and colleagues reported the cloning of the phosphatidylinositol glycan complementation class A (*PIGA*) gene, which is part of the enzyme complex involved in this step of the pathway and which has subsequently been shown to be mutated in almost all cases of acquired PNH reported to date.

Mutations of the *PIGA* gene are different between patients as they are acquired not inherited. It appears that the mutation rate of *PIGA* is similar to that of other genes and in fact populations of PNH-like GPI-deficient cells can be observed at extremely low levels (1–10 per million cells) in most normal individuals, indicating that occasional haemopoietic stem cells by chance acquire *PIGA* mutations. Since the *PIGA* gene is located on the X chromosome (Xp22.1), each somatic cell, male or female, only has a single active copy and therefore a somatic mutation of this active gene leads to deficiency of GPI biosynthesis. GPI-deficient clones only expand if there is an additional factor that encourages their selection. The mechanism of this selective growth advantage of PNH clones, which could be a second genetic event, is of key importance in understanding the pathophysiology of PNH. A much more likely explanation is indicated by the association between PNH and other bone marrow failure syndromes.

Association with bone marrow failure

PNH clones are detectable in up to 50% of patients with aplastic anaemia and in a smaller proportion of patients with myelodysplastic syndrome. This suggests that such bone marrow failure syndromes are permissive for the expansion of PNH clones. There is overwhelming evidence that aplastic anaemia is an autoimmune disorder in which there is an aberrant immune attack, probably by CD8 cytotoxic T cells directed against the haemopoietic stem cell. It would therefore appear likely that PNH stem cells, presumably due to GPI deficiency, evade this immune attack. This implies that either the immune attack in bone marrow failure syndrome is, at least in part, directed through a GPI-linked structure or that the biology of the PNH stem cell is abnormal, protecting it from immune-mediated attack. Understanding PNH may well provide an insight into the mechanism of a variety of bone marrow failure syndromes and provide a unique therapeutic target for these conditions.

Epidemiology

If PNH is defined as the presence of a PNH clone identified during the investigation of bone marrow failure, haemolysis or thrombosis, then the incidence of PNH in a series from the UK was found to be 1.3 newly diagnosed patients per million per year and the prevalence 15.9 patients per million population. However, of these patients, 57% have less than 10% PNH neutrophils, indicating that they are unlikely to have clinically

significant haemolysis and only one-third reported macroscopic haemoglobinuria. However the identification of the PNH clone is important, even when it is small, as it suggests that the bone marrow failure is likely to be immune mediated and because the PNH clone can evolve over time so that patients may be at risk of haemolysis and/or thrombosis.

Clinical features

Clinical triad

The clinical picture in PNH depends on the balance between three components: intravascular haemolysis, bone marrow failure and thrombotic complications. The chronicity of the disease, with a median survival of 10 to 20 years even prior to targeted therapy, causes continuous, high-level intravascular haemolysis leading to other complications, including renal disease and cholelithiasis.

Haemolysis

PNH, the disorder with perhaps the most intense chronic intravascular haemolysis, is characterized by episodes of haemoglobinuria, during which the patient's urine is often black and which, for unknown reasons, is often worse in the morning. This episodic haemoglobinuria is frequently associated with disabling symptoms including profound lethargy out of keeping with the patient's anaemia, abdominal pain sometimes requiring opiates, dysphagia that can be temporarily absolute and erectile dysfunction. The degree of haemolysis is associated with the size of the PNH clone and particularly the proportion of PNH type III (completely deficient) red cells. Macroscopic haemoglobinuria is rarely seen unless there are at least 10% type III red cells. However, patients with apparently identical PNH clones can have a wide range of disease severity, from very occasional attacks of haemoglobinuria without anaemia to severe recurrent or even continuous haemoglobinuria and transfusion dependence, suggesting that other important factors influence the severity of the symptoms. It is possible that this variation is due to the severity of underlying bone marrow failure, to an inherent variation in complement activity or to some other unidentified factor.

Many of the symptoms and complications of PNH result directly from intravascular haemolysis. Free plasma haemoglobin is immediately bound to haptoglobin and removed, explaining why haptoglobin is depleted in every patient. Free haemoglobin binds to and removes other gases such as nitric oxide (NO). This depletion of NO in PNH results in smooth muscle dysfunction, vasoconstriction, and pulmonary and systemic hypertension. This leads to dysphagia, abdominal discomfort, possibly due to intestinal spasm, erectile dysfunction, severe lethargy, pulmonary hypertension and possibly thrombosis, all classic symptoms of PNH. This syndrome of NO consumption has only become clear since the development of eculizumab (see below), which stops intravascular haemolysis in almost all patients and abrogates the clinical features of PNH.

Bone marrow failure

The degree of anaemia and other cytopenias is a composite of the activity of intravascular haemolysis and the degree of underlying bone marrow failure. The platelet count in many patients is a suitable surrogate for marrow function, but in some patients, for example those with previous intra-abdominal thrombosis and subsequent hypersplenism, there may be other causes for a low platelet count. The presence of a PNH clone in patients with bone marrow failure probably indicates that there is a significant immune component to the marrow failure and may suggest that immune suppression is a reasonable therapeutic option. The degree of bone marrow failure will also impact on the efficacy of eculizumab, as this will only treat the component due to complement activity, such as intravascular haemolysis and thrombosis.

Thrombosis

Thrombosis occurs in 40–50% of patients with haemolytic PNH, with a predilection for certain veins, such as the hepatic veins (Budd–Chiari syndrome), other intra-abdominal veins or the cerebral veins (Table 11.1). Approximately 10% of patients present with a thrombosis as their initial clinical manifestation of PNH. Recent data indicate that the thrombotic tendency in PNH is in arterial as well as venous sites, including cerebrovascular accidents and myocardial infarctions at an early age. After the first thrombosis, patients have a 7.8-fold increased risk of dying compared with those patients who have not had a thrombosis. A classical clinical scenario is that of downward spiralling thrombotic events: after a first thrombosis, patients continue to experience further apparently discrete thromboses despite what would be considered adequate anticoagulation with warfarin and/or heparin until they eventually succumb. The cause of thrombosis in PNH is multifactorial including activation of PNH platelets by complement as well as intravascular haemolysis and NO consumption, which may lead to endothelial damage and thrombosis. It is now clear that there is a close relationship between the activation of both coagulation and complement with thrombin directly activating complement through the alternative pathway. This explains why frequently patients enter a spiral of multiple and increasingly severe thrombotic complications as the first thrombosis leads to complement activation leading to further thrombosis and more thrombin generation. Anticoagulants are relatively ineffective at preventing this circle of catastrophic thrombosis as they do not stop the activation of complement, whereas eculizumab stops the further activation of complement and usually interrupts this vicious circle of thrombosis.

Renal disease

All patients with haemolytic PNH develop heavy renal tubular iron loading due to the continuous filtration of haemoglobin. This is evident on both magnetic resonance imaging (MRI)

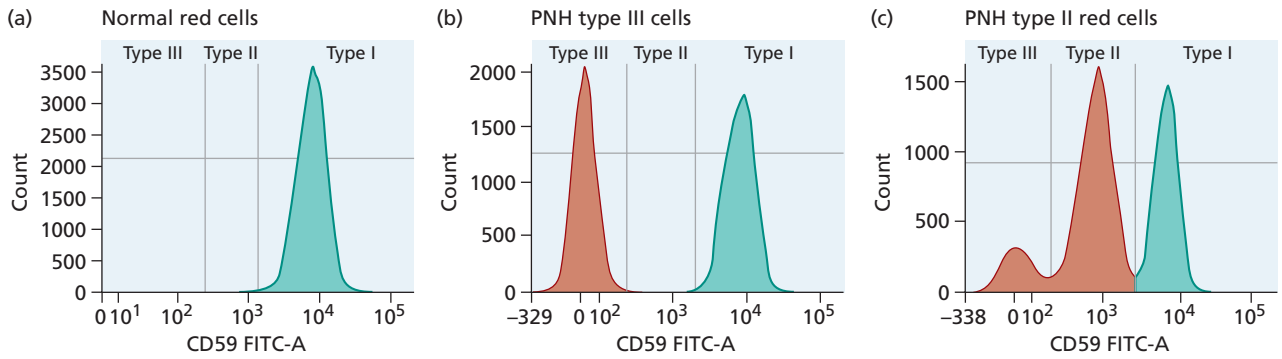


Figure 11.5 Identification of PNH red cells by flow cytometry: (a) GPI anchor protein (e.g. CD59) expression on normal red cells (type I); (b) A total of 46% of red cells are completely GPI anchor-deficient (type III); (c) a mixture of all three types of red cells are shown: type III, 10%; type II (partial GPI anchor expression), 54%; normal type I cells, 36%. (Source: S.J. Richards. Reproduced with permission.)

and post-mortem studies. The majority of patients will eventually develop chronic renal disease, with a minority progressing to established chronic renal failure requiring dialysis. In addition, during times of very intense intravascular haemolysis and haemoglobinuria, patients can rarely develop acute renal failure, which is potentially fully reversible, even though it may require dialysis. Some patients may develop repeated episodes of acute renal dysfunction. The use of nephrotoxic therapies, such as ciclosporin, adds to the renal insult, as do rare complications such as renal vein thrombosis.

Leukaemic transformation

There have been a number of case reports of myelodysplastic syndrome and acute myeloid leukaemia (AML) in patients with PNH, but only less than 5% of PNH patients in the larger series develop AML. This is similar to the incidence of AML in aplastic anaemia and since it seems the majority, if not all, patients with PNH have an underlying aplasia, then the rate of AML is no higher than would be expected, suggesting that the GPI deficiency has no impact on leukaemogenesis.

Spontaneous remission

PNH often develops in young adults between 20 and 40 years old. Historically, most patients with PNH have either died as a direct or indirect result of their disease or have suffered from the disease for the remainder of their life. In a series of 80 haemolytic PNH patients reported in 1995, initially diagnosed between 1940 and 1970, median survival was 10 years and 12 of the 35 patients surviving at least 10 years experienced spontaneous remission. Analysis of GPI-linked antigens on blood cells up to 20 years after spontaneous clinical remission demonstrated that the myeloid series was entirely normal with no residual PNH cells, whereas there was a small PNH clone in the lymphoid compartment (presumably these are long-lived memory cells). Early reports suggest that a proportion of patients treated with eculizumab have a progressive decrease in the size of their PNH clone, perhaps suggesting that remission may occur.

Investigation

The diagnosis of PNH was historically made by demonstrating the sensitivity of red cells to lysis by activated complement in the Ham test or similar. However, these tests only gave an indirect assessment of the proportion of PNH red cells and have been superseded by flow cytometry. Peripheral blood flow cytometric analysis of at least two cell lineages (e.g. neutrophils and red cells), with a transmembrane marker to positively identify the cell type and at least two separate GPI-linked antigens to clearly separate PNH cells from their normal counterparts (Figures 11.5 and 11.6). In most patients there is a population of red cells with complete deficiency of GPI-linked molecules (PNH type

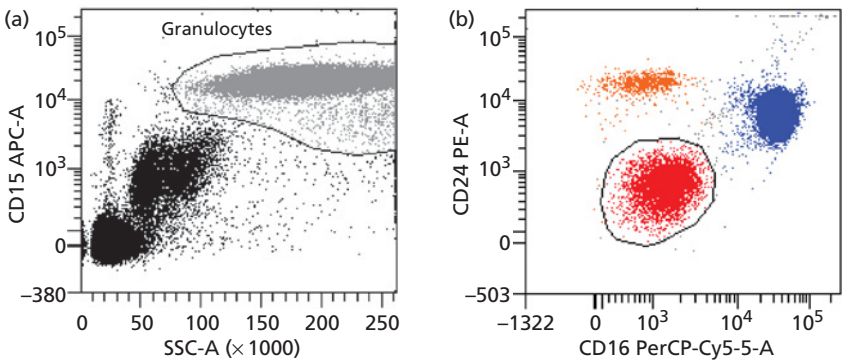


Figure 11.6 Identification of PNH granulocytes by multicolour flow cytometry: (a) granulocytes identified based on CD15 expression and side scatter; (b) analysis of CD16 and CD24 (both GPI-linked proteins) reveals a small population of PNH granulocytes (10%) that are deficient in both proteins. (Source: S.J. Richards. Reproduced with permission.)

III cells, the most complement-sensitive cells) and, in some, an additional population with partial GPI deficiency (PNH type II cells, intermediate complement sensitivity). Type II and III are usually not evident in the neutrophils. The size of the PNH clone correlates with the risk of complications, such as thrombosis, and the severity of haemolysis. In addition, the evolution of the clone (either its expansion, leading to more haemolytic disease, or its reduction, ultimately leading to spontaneous remission) can be tracked.

Treatment

Supportive care

Conventional management in PNH has been supportive. The severity of haemolysis varies greatly between patients, in part dependent on the size of the PNH clone and the degree of underlying bone marrow failure. However, even patients with no evidence of clinically apparent marrow failure and large PNH clones can have highly variable levels of haemolysis. Some patients experience recurrent, even continuous, haemoglobinuria and are transfusion dependent, whereas others experience haemoglobinuria only rarely, or not at all, do not require any transfusions and have a well-compensated haemolytic anaemia. Patients with a significant degree of haemolysis should receive folic acid supplementation. The constant haemosiderinuria means that patients, even those who require regular transfusions, have a tendency to become iron deficient. This iron deficiency can result in failure of the marrow to compensate for the increased red cell destruction and therefore increasing anaemia. Iron supplementation has been reported to precipitate episodes of intravascular haemolysis, but usually patients tolerate oral iron supplements well and should be treated when iron deficiency is present.

Patients should be transfused according to their symptoms. Incidental infections lead to an increase in haemolysis, presumably due to activation of the complement system, and frequent sudden drops in the level of haemoglobin. These episodes are often associated with severe symptoms of abdominal pain, dysphagia and debilitating lethargy.

Corticosteroids have been widely used in the treatment of haemolysis in PNH. At high doses, steroids appear to have an effect of reducing the activity of complement and some patients report an improvement in symptoms. However, high doses are required to have a clinically useful effect and as the haemolysis in PNH is chronic, the required dose is too high. Steroids are therefore not generally recommended in PNH as the long-term side-effects outweigh the potential benefits.

Thrombosis

The major cause of morbidity and mortality in PNH is thrombosis, which usually affects the venous system, particularly the intra-abdominal and cerebral vessels, but also appears to result in arterial thrombosis (Table 11.4). The chance of developing a thrombosis depends on the size of the PNH clone and the severity of haemolysis, two variables that are closely related; half of patients with over 50% PNH neutrophils will develop a thrombosis at some point in their disease course. In most series, approximately one-third of patients will eventually die as a result of thrombosis. It appears that primary prophylaxis with warfarin is effective in preventing thrombosis, but carries a significant risk in this patient group. Aspirin does not appear to be protective against thrombosis and has no effect on the symptoms of PNH such as abdominal pain (Peter Hillmen, unpublished observation). It seems very likely that primary prophylaxis with warfarin is unnecessary for patients receiving the anticomplement monoclonal antibody eculizumab, as this agent significantly reduces the risk of thrombosis in PNH (see further on). There are no published data for the use of novel anticoagulants in PNH.

Table 11.4 Sites of thrombotic events in haemolytic PNH.

Venous thrombosis	
Deep vein thrombosis	41 (33.1%)
Lower extremity	23 (18.5%)
Other	18 (14.5%)
Mesenteric/splenic vein thrombosis	23 (18.5%)
Hepatic/portal vein thrombosis	21 (16.9%)
Pulmonary embolus	8 (6.5%)
Cerebral/internal jugular thrombosis	7 (5.6%)
Superficial vein thrombosis	5 (4.0%)
Arterial thrombosis	
Cerebrovascular accident/TIA	17 (13.7%)
Myocardial infarction/unstable angina	2 (1.6%)
Total	124 (100%)

Source: Hillmen *et al*., 2007 [*Blood* 2007;**110**: 4123–8]. Reproduced with permission of The American Society of Hematology.

The treatment of established thrombosis is similar to the management of thrombosis in patients without PNH, except for the addition of immediate treatment with eculizumab (see further on). However, in view of the risk of recurrent thrombosis, patients should remain on lifelong anticoagulation after their first episode of thrombosis. Hepatic vein thrombosis (Budd–Chiari syndrome) is one of the more common thromboses seen in PNH and there have been reports of the successful lysis of such thromboses using tissue plasminogen activator; this should be considered even in patients who present with a relatively long history suggesting that their thrombosis occurred days or even weeks before. Current practice involves immediate use of eculizumab along with anticoagulation in acute Budd–Chiari syndrome, associated with PNH, to reduce mortality and long-term sequelae.

Allogeneic stem cell transplantation

The only curative strategy for PNH is allogeneic stem cell transplantation, but this carries a considerable risk of mortality. From

the reported series and in view of the fact that a proportion of patients will eventually experience a spontaneous remission of PNH and with the advent of potentially effective novel therapies such as eculizumab, transplantation should only be considered in selected cases, such as those with a syngeneic donor or with associated bone marrow failure. In these patients the indications for transplantation are similar to those for aplastic anaemia.

Complement blockade

The development of eculizumab (Soliris), a humanized monoclonal antibody that blocks the activation of terminal complement, has dramatically altered the management and prognosis in PNH. Eculizumab binds to C5 and stops it being cleaved and, as long as trough plasma levels remain adequate, prevents any activation of terminal complement. Individuals with inherited terminal complement deficiency are either asymptomatic or present with recurrent *Neisseria meningitidis* (meningococcus) infections and this highlights the main concern with eculizumab. The clinical features of PNH are due largely to the absence of CD59 from haemopoietic cells and therefore the uncontrolled activity of terminal complement on PNH cells, making eculizumab an ideal candidate for the targeted therapy of PNH. In the initial study, 11 patients with transfusion-dependent haemolytic PNH were treated with eculizumab using a schedule designed to maintain trough levels to ensure that complement was completely blocked. The responses were dramatic, with an immediate resolution of the symptoms of intravascular haemolysis. This study was followed by a randomized placebo-controlled Phase III trial (TRIUMPH) and a non-randomized trial (SHEPHERD). In total, 195 patients were included in these three trials, which led to the licensing of eculizumab in the USA and Europe in 2007. The results of these trials are summarized below.

Efficacy of eculizumab

Intravascular haemolysis

Eculizumab has a profound and immediate effect on intravascular haemolysis in PNH. The lactate dehydrogenase (LDH) concentration, which is typically as much as 10–20 times normal in PNH, falls immediately in almost all patients treated with eculizumab to normal or just above normal. In general this results in increasing haemoglobin level which, depending on the degree of coexistent bone marrow failure as well as the extent of extravascular haemolysis (see below), will reach a plateau usually between 90 and 120 g/L. The most dramatic effect of eculizumab is on the symptoms, with resolution of the abdominal pain and dysphagia, and improvement in the severe lethargy and the other features of haemolysis. In the vast majority of patients transfusion requirements improve, with over half of patients becoming transfusion independent. The dramatic improvement in PNH-related symptoms is due to a marked reduction in NO consumption during therapy as a result of the improvement in intravascular haemolysis and thereby the reduction in free haemoglobin. This probably explains many of these benefits of the drug and provides insights into the pathophysiology of the symptoms of PNH. There is a dramatic and clinically significant improvement in the quality of life of patients as measured by validated questionnaires. There is now evidence that the adverse consequences of intravascular haemolysis, such as renal damage and pulmonary hypertension, are ameliorated by eculizumab.

Thrombosis

Eculizumab also protects patients from thrombosis. Compared with thrombosis before patients commencing eculizumab (effectively using patients as their own controls), the thrombotic rate reduces by fivefold to tenfold. In patients who have had a previous thrombosis and who are on anticoagulation, there is still a high recurrent thrombosis rate prior to starting eculizumab. However, since eculizumab became available, such recurrent thromboses are extremely uncommon. Patients who commence eculizumab during a 'spiral' of thrombotic events stop having further thromboses, indicating that eculizumab specifically targets the mechanism of thrombosis in PNH and that this is a more effective strategy than conventional anticoagulation alone. There is now increasing confidence that the catastrophic thrombotic complications of PNH can be much more successfully managed with a combination of eculizumab and anticoagulation. This observation will have an impact on the decision to use warfarin prophylaxis: in cases where there are concerns over the safety or requirement for warfarin prophylaxis, such as those with low platelets or with borderline PNH clone sizes, it is clearly safer to withhold anticoagulation as long as the use of eculizumab is an option for the patient should a thrombosis occur.

Renal dysfunction

In the 195 patients entering the eculizumab trials, renal dysfunction or damage was observed in 65% of patients before they were treated with eculizumab. In this series, 27% of patients had developed major clinical kidney disease within 10 years from their initial diagnosis of PNH and 21% of patients developed late-stage chronic kidney disease (stage 3 or 4 as defined by the Kidney Disease Outcomes Quality Initiative) or kidney failure (stage 5). The early analysis of patients treated with eculizumab suggests that many of the patients with early renal dysfunction (stages 1 and 2) will improve, and the deterioration in renal function in patients with advanced renal dysfunction (stages 3–5) is frequently stabilized. Thus eculizumab appears to have a beneficial effect on renal function in PNH and this is presumably due to the marked reduction in intravascular haemolysis and therefore in haemoglobinaemia and haemoglobinuria.

Extravascular haemolysis with eculizumab

Despite the impressive impact on intravascular haemolysis, most patients remain somewhat anaemic, maintaining their haemoglobin between 90 and 120 g/L. Virtually all have a persisting reticulocytosis and many continue to have a raised bilirubin. These features are suggestive of ongoing extravascular haemolysis, which has previously been unreported in PNH. On further investigation it transpires that approximately two-thirds of PNH patients on eculizumab develop a positive direct antiglobulin test to complement only. Flow cytometry demonstrates that the PNH red cells are coated by early complement components (C3b and C3d), possibly because PNH red cells, as well as being deficient in the principal controller of terminal complement, namely CD59, do not express the inhibitor of C3 convertase decay-accelerating factor (DAF). It appears that preventing terminal complement activation leads to a build-up of the early complement components, which accumulate on PNH red cells due to their deficiency of DAF. If intense, this extravascular haemolysis results in a poor increase in haemoglobin and a minority of patients continuing to require transfusions. If the transfusions are due to poor marrow reserves and a lack of compensation, particularly in patients with evidence of renal dysfunction, then treatment with erythropoietin can lead to a clinically meaningful increase in haemoglobin.

Eculizumab administration and dosing

Eculizumab is given as a 30-min intravenous infusion. The aim is to rapidly block complement and to maintain complement blockade continuously. The standard dosage schedule for eculizumab comprises a loading dose of 600 mg every week for 4 weeks, followed by 900 mg the next week and then 900 mg every 2 weeks indefinitely. In the majority of patients this is adequate to maintain trough levels of eculizumab above 50 mg/L and therefore to block complement completely. In approximately 10% of patients this dose is inadequate and patients break through complement blockade. Patients appear well with no signs of haemolysis until immediately before a dose of eculizumab, but then develop dark urine often with abdominal pain, sometimes with dysphagia, and a sudden deterioration in the laboratory measures of haemolysis, such as LDH and bilirubin, and a fall in the level of haemoglobin. In this situation the maintenance dose of eculizumab is too low and an increase, either by reducing the interval or more conveniently by increasing the dose (from 900 mg every 2 weeks to 1200 mg every 2 weeks or even higher doses are usually effective), will raise trough levels of eculizumab above 50 mg/L and re-establish continuous control of haemolysis (Figures 11.7 and 11.8). Approximately 3% of Japanese individuals have a C5 polymorphism that prevents the binding of eculizumab and means that they do not respond to eculizumab. These individuals have no fall in their LDH after eculizumab dosing. Such a lack of response has not yet been reported in non-Japanese patients.

Infectious risk with eculizumab

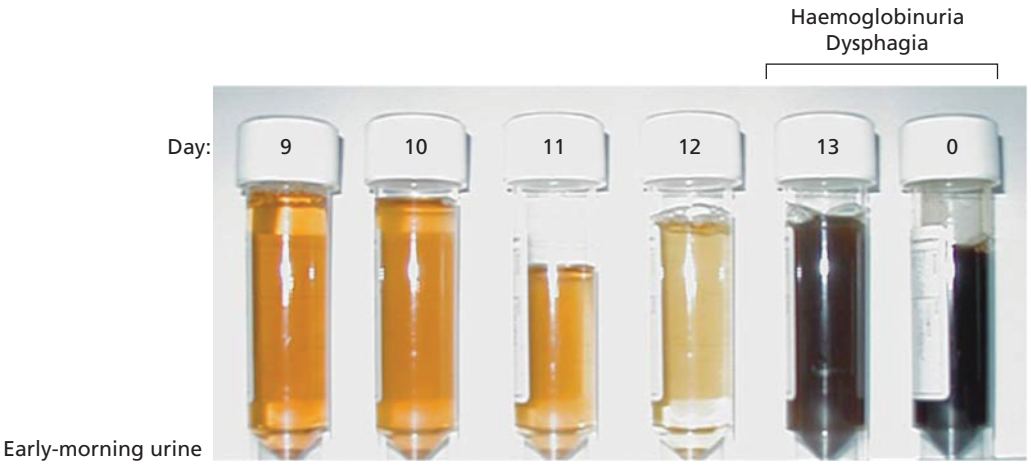
Eculizumab is generally very well tolerated with few infusion-related reactions. However, as noted above, blocking terminal complement activity would be expected to increase the risk of infection with *Neisseria meningitidis* (meningococcus). All patients commencing eculizumab should be vaccinated with a wide-spectrum meningococcal vaccine (ACWY Vax). The role of newly available vaccine (Bexsero) to cover serotype B, which is common in the UK, is being explored in the setting of complement deficiency. Although the risk of meningococcal infection is a concern, the observed risk is less than 0.5 cases of meningococcal infection per 100 patient-years on eculizumab. However, when these infections occur they can be life-threatening and it is vitally important to impress on patients that they should seek medical help if they suffer any symptoms suggestive of infection, and these are usually of septicaemia rather than meningitis. Prophylaxis with penicillin is also recommended. There is no convincing evidence as yet of an increased risk to any other organism except for *N. meningitidis*.

Pregnancy in PNH

It is difficult to estimate the true risk of pregnancy in PNH but there is undoubtedly an increased risk of maternal morbidity and mortality. The reported maternal mortality, mainly from thrombosis, is between 12 and 21%, although this is likely an overestimate due to biased reporting. In addition, an increased fetal loss rate has been reported, although again this is difficult to substantiate and is probably due to maternal factors. There have now been a number of successful pregnancies in women receiving eculizumab either later in pregnancy or throughout the pregnancy from conception to delivery. The reports to date are positive with little or no eculizumab crossing the placenta into the fetus. It does appear that the metabolism of eculizumab may be altered in pregnancy and women seem more likely to break through complement blockade and may need higher doses in the latter part of pregnancy.

Prognosis

A series of 80 patients with haemolytic PNH diagnosed between 1940 and 1970 were reported in 1995 to reveal a median survival of approximately 10 years from initial diagnosis of PNH. A more recent series of 465 patients reported from France, including patients with small PNH clones as well as haemolytic PNH, reported a median survival of 22 years. The major causes of death in haemolytic PNH are thrombotic and since these are virtually abolished by eculizumab therapy, a recent single centre report demonstrated that the survival of patients with PNH receiving eculizumab is very similar to an age-matched control population.



Days		0*	1–8	9	10	11	12	13	0*	1	2
Urine scale	1 = Clear; 10 = Black	10	2–3	3	3	3	3	9	10	3	3
LDH	160–480 IU/L	2624	–	784	–	–	697	1687	2917	–	–
AST	0–40 IU/L	119	–	38	–	–	31	87	–	–	–
PK	C5 blocked >35 μg/mL	27	–	45	–	–	35	29	26	–	–
PD	No lysis <20%	71	–	1	–	–	2	56	56	–	–

*Eculizumab 900mg

Figure 11.7 Breakthrough from complement blockade by eculizumab due to inadequate dose level. The urine is clear with lactate dehydrogenase (LDH) just above normal and no haemolytic activity in the patient's serum at day 12 after a 900-mg dose of eculizumab. On day 13, the patient suddenly begins to haemolyse as the level of eculizumab falls below 35 mg/L (the level at which complement is blocked). Immediately after the next dose (day 0), the haemolysis stops. Urine scale: patient assesses urine colour first thing in the morning (red or black urine at 6+). AST, aspartate aminotransferase; PK, pharmacokinetics (serum level of eculizumab); PD, pharmacodynamics (haemolytic potential of the patient's serum *in vitro*).

Future challenges and developments

The development of targeted therapy for PNH appears to alter the natural history of the disease and other newer anticomplement therapies are being evaluated in preclinical/Phase I trials. A global PNH registry has been established to document the changes in the natural history of the disease, as well as to record any unexpected complications of the newer therapeutic interventions. The current therapy of PNH controls rather than cures the disease. It is hoped that with more detailed understanding of the pathophysiology of PNH, a curative strategy may be developed.

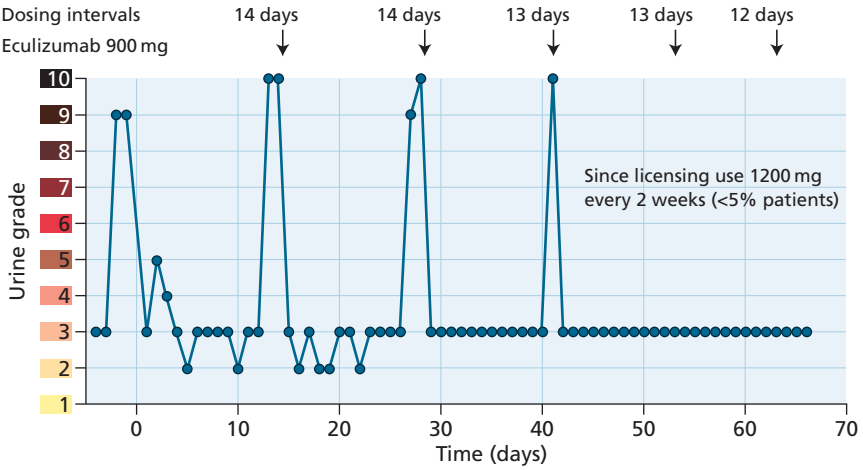


Figure 11.8 Urine colour with increasing eculizumab dose. Patient records early-morning urine colour according to scale on *y*-axis. He is breaking through from complement blockade for 2 days when receiving 900 mg eculizumab every 14 days. When the frequency of eculizumab 900 mg is increased to every 12 days, his breakthroughs cease. He is now maintained on 1200 mg eculizumab every 14 days.

Suggested further reading

Clinical significance of acquired somatic mutations in AA

The group at King's College Hospital, London, recently reported the first large study of acquired somatic mutations in AA patients using next generation sequencing. Almost 20% of patients had mutations in myeloid specific genes, most commonly ASXL1, DNMT3A and BCOR. Subsequently, a larger study from NIH, Cleveland Clinic USA and Japan has confirmed similar results with ASXL1, DNMT3A, BCOR and BCORL1 in 36% of AA patients. In both studies, the size of these mutated clones in AA was small in most patients and lower than found in MDS. The presence of somatic mutations was highly predictive of later evolution to MDS/AML (in 40% patients) compared to only 4% who lacked these mutations, for patients with a history of AA of > 6 months. Mutations in ASXL1 and DNMT3A were associated with worse response to immunosuppressive therapy and worse survival; in contrast BCOR/BCORL1 and PIG-A mutations conferred better response to IST and better survival. When all these mutations are combined with single nucleotide polymorphism (SNP-A) karyotyping, evidence of clonal haemopoiesis was demonstrated in 47% of AA patients. However, the interpretation of clonal haemopoiesis in AA is not straightforward. Age-related clonal haemopoiesis with mutations in especially ASXL1 and DNMT3A, and others including TET2, JAK2 but not BCOR/BCORL1 or PIG-A, has been reported in 10% normal individuals older than 65 years, and increasing further with subsequent decades of life. In AA, analogous to PNH and abnormal cytogenetic clones, the level of these clones may fluctuate over time and may be selected in the context of immune dysregulation in the setting of bone marrow failure through increased proliferative pressure on the residual haemopoietic stem cell/progenitors. A possible unifying explanation is that the low level 'age-related' clones represent a predilection (founder) stage that requires subsequent co-operating mutations for clonal expansion and disease, alongside the increased telomere attrition seen in AA.

Selected bibliography

Brodsky RA (2014) Paroxysmal nocturnal hemoglobinuria. *Blood* **124**: 2804–11.

Desmond R, Townsley DM, Dumitriu B *et al*. (2014) Eltrombopag restores trilineage hematopoiesis in refractory severe aplastic anemia that can be sustained on discontinuation of drug. *Blood* **123**(12): 1818–25.

Hill A, Kelly RJ, Hillmen P (2013) Thrombosis in paroxysmal nocturnal hemoglobinuria. *Blood* **121**(25): 4985–96.

Hillmen P, Young NS, Schubert J *et al*. (2006) The complement inhibitor eculizumab in paroxysmal nocturnal hemoglobinuria. *New England Journal of Medicine* **355**(12): 1233–43.

Jaiswal S, Fontanillas P, Flannick J *et al*. (2014) Age-Related Clonal Hematopoiesis Associated with Adverse Outcomes. *New England Journal of Medicine* **371**: 2488–98.

Kulasekararaj AG, Jiang J, Smith AE *et al*. (2014) Somatic muttaions identify a sub-group of aplastic anemia patients that progress to myelodysplastic syndrome. *Blood* **124**: 2698–2704.

Marsh JC, Pearce RM, Koh MB *et al*. (2014) Retrospective study of alemtuzumab versus ATG based conditioning without irradiation for unrelated and matched sibling donor transplants in acquired severe aplastic anaemia: a study from the British Society for Blood and Marrow Transplantation (BSBMT). *Bone marrow Transplantation* **49**: 42–48.

Marsh JC, Kulasekararaj AG (2013) Management of the refractory aplastic anemia patient: what are the options? *Blood* **122**(22): 3561–7.

Scheinberg P, Young NS. (2012) How I treat acquired aplastic anemia. *Blood* **120**(6): 1185–96.

Townsley DM, Dumitriu B, Young NS. (2014) Bone marrow failure and the telomeropathies. *Blood* **124**: 2775–83.

Yoshizato T, Dumitriu B, Hosokawa K et al. (2015) Somatic mutation and clonal hematopoiesis in aplastic anemia. *New England Journal of Medicine* **373**(1): 35–47.

CHAPTER 12

Red cell immunohaematology

12

Geoff Daniels[1], Marcela Contreras[2] and Shubha Allard[3]

[1]International Blood Group Reference Laboratory, NHS Blood and Transplant, Bristol, UK
[2]University College London and Blood Transfusion International, London, UK
[3]Barts Health NHS Trust & NHS Blood and Transplant, London, UK

Introduction

Red cell immunohaematology is the study of the interactions between blood group antigens and their corresponding antibodies, both *in vivo* and *in vitro*, and of the presence or absence of clinical consequences following those interactions. It encompasses: (i) determination of the phenotype of red cells from donors and patients, defined by antibodies, (ii) the search for and identification of antibodies in patients' sera, with red cells of known phenotype, (iii) in the presence of clinically significant antibodies, compatibility testing of patients' sera against selected cell samples from ABO compatible donor units and (iv) investigation of adverse effects due to incompatibility.

The aim of this chapter is to provide an introduction to blood group serology, to include aspects of immunology, biochemistry and molecular genetics that contribute to our understanding of blood group antigens, antibodies and antigen–antibody reactions, and to provide some details on the blood group systems of most relevance to blood transfusion.

Blood group systems

The International Society of Blood Transfusion recognizes 347 red cell surface antigens, 308 of which belong to one of 36 blood group systems (Table 12.1). Each system represents either a single gene or two or three very closely linked homologous genes. Each system is genetically discrete from all others. In addition, there are 39 antigens that have not been included in systems, owing to inadequate genetic evidence. Most blood groups are inherited as Mendelian characters, although environmental factors may occasionally affect blood group expression. The 36 systems represent a total of 41 genes: MNS contains 3 loci, Rh, Xg and Ch/Rg 2 loci each, and the other 32 systems each contain a single gene. All of the genes are autosomal, except for *XG* and *XK*, which are on the X chromosome, and *CD99*, which is located on both the X and Y chromosomes. All 41 genes have been identified and sequenced.

The red cell membrane and chemistry of blood group antigens

The red cell membrane is composed of about 40% (w/w) lipids and up to 10% carbohydrates, the remainder being protein. Blood group antigens may be integral proteins or glycoproteins of the red cell membrane, or they may be membrane glycolipids. In proteins and glycoproteins, blood group polymorphisms and variants may represent differences in the amino acid sequence (e.g. Rh and Kell antigens). In glycoproteins and glycolipids, the blood group activity may reside in the carbohydrate moiety and polymorphism is associated with differences in the oligosaccharide sequence (e.g. ABO). In some glycoproteins, blood group polymorphism may be caused by amino acid substitutions, but antigen expression is also dependent on glycosylation of the polypeptide.

The integral proteins of the red cell membrane penetrate the lipid bilayer and, in some cases, are bound to the membrane skeleton or cytoskeleton. Cell-surface proteins are not necessarily discrete units within the cell membrane; they may congregate together in lipid rafts or might be part of complexes of proteins in which the components contribute to the function

Postgraduate Haematology, Seventh Edition. Edited by A Victor Hoffbrand, Douglas R Higgs, David M Keeling and Atul B Mehta.

Table 12.1 Human blood group systems.

No.	Name	Symbol*	No. of antigens	HGNC gene symbol(s)	Chromosome	Structure	CD number
001	ABO	ABO	4	*ABO*	9	Carbohydrate	
002	MNS	MNS	48	*GYPA*, *GYPB*, *GYPE*	4	Glycoprotein	CD235a/b
003	P1PK	P1PK	3	*A4GALT*	22	Carbohydrate	
004	Rh	RH	54	*RHD*, *RHCE*	1	Protein	CD240D/CE
005	Lutheran	LU	22	*BCAM*	19	Glycoprotein	CD239
006	Kell	KEL	35	*KEL*	7	Glycoprotein	CD238
007	Lewis	LE	6	*FUT3*	19	Carbohydrate	
008	Duffy	FY	5	*DARC*	1	Glycoprotein	CD234
009	Kidd	JK	3	*SLC14A1*	18	Glycoprotein	
010	Diego	DI	22	*SLC4A1*	17	Glycoprotein	CD233
011	Yt	YT	2	*ACHE*	7	Glycoprotein	
012	Xg	XG	2	*XG*, *CD99*	X/Y	Glycoprotein	CD99**
013	Scianna	SC	7	*ERMAP*	1	Glycoprotein	
014	Dombrock	DO	10	*ART4*	12	Glycoprotein	CD297
015	Colton	CO	4	*AQP1*	7	Glycoprotein	
016	Landsteiner–Wiener	LW	3	*ICAM4*	19	Glycoprotein	CD242
017	Chido–Rodgers	CH/RG	9	*C4A*, *C4B*	6	Glycoprotein	
018	H	H	1	*FUT1*	19	Carbohydrate	
019	Kx	XK	1	*XK*	X	Protein	
020	Gerbich	GE	11	*GYPC*	2	Glycoprotein	CD236
021	Cromer	CROM	18	*CD55*	1	Glycoprotein	CD55
022	Knops	KN	9	*CR1*	1	Glycoprotein	CD35
023	Indian	IN	4	*CD44*	11	Glycoprotein	CD44
024	Ok	OK	3	*BSG*	19	Glycoprotein	CD147
025	Raph	RAPH	1	*CD151*	11	Glycoprotein	CD151
026	John Milton Hagen	JMH	6	*SEMA7A*	15	Glycoprotein	CD108
027	I	I	1	*GCNT2*	6	Carbohydrate	
028	Globoside	GLOB	2	*B3GALT3*	3	Carbohydrate	
029	Gill	GIL	1	*AQP3*	9	Glycoprotein	
030	RHAG	RHAG	4	*RHAG*	6	Glycoprotein	CD241
031	Forssman	FORS	1	*GBGT1*	9	Carbohydrate	
032	JR	JR	1	*ABCG2*	4	Glycoprotein	
033	Lan	LAN	1	*ABCB6*	2	Glycoprotein	
034	Vel	VEL	1	*SMIM1*	1	Glycoprotein	
035	CD59	CD59	1	*CD59*	11	Glycoprotein	CD59
036	Augustine	AUG	2	*SLC29A1*	6	Glycoprotein	

HGNC, Human Genome Organisation Gene Nomenclature Committee.
*ISBT gene name in italics.
**Does not include Xg glycoprotein.

of the whole (Figure 12.1). There are four types of integral proteins in the red cell membrane. Type 1 glycoproteins, typified by the glycophorins, carrying the MNS antigens, cross the membrane once, with the amino-terminal domain on the outside and the carboxy-terminal domain in the cytosol. The Kell glycoprotein is a type 2 structure. It also crosses the membrane once, but has the carboxy-terminal domain on the outside and the amino-terminal domain in the cytosol. Type 3 glycoproteins are polytopic: they cross the membrane several times giving rise to a series of extracellular loops, one of which is usually glycosylated. These N-glycans express ABO, H and Ii blood group activity. In most of these type 3 proteins, both terminal domains are cytosolic (e.g. Rh proteins, band 3, glucose transporter, Kidd glycoprotein), but the Duffy glycoprotein has an extracellular amino-terminal domain. Type 5 membrane glycoproteins, which include the Dombrock

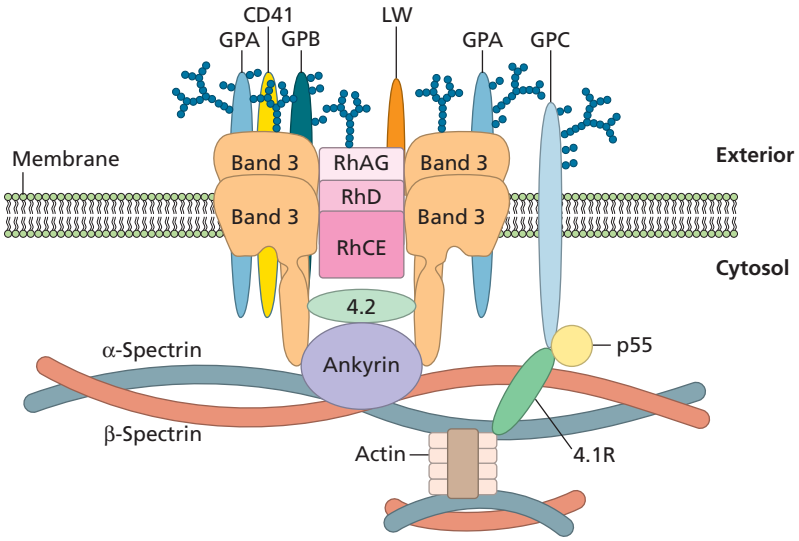


Figure 12.1 Diagrammatic representation of some integral red cell membrane proteins, including the band 3/Rh macrocomplex and components of the membrane cytoskeleton (see also Chapter 8).

and Cromer glycoproteins, reside in membrane microdomains rich in cholesterol known as lipid rafts and are attached to the lipid bilayer by a glycolipid, the glycosylphosphatidylinositol (GPI) anchor. There are no Type 4 proteins in the red cell membrane.

On the inside of the lipid bilayer is a network of glycoproteins that comprise the cytoskeleton, which is responsible for maintenance of the shape and integrity of the red cell as it squeezes through the tiniest capillaries. The main components of the membrane skeleton are the α- and β-subunits of spectrin, which form long flexible rods, plus actin, ankyrin, proteins 4.1R and 4.2, and several other glycoproteins. Band 3, the anion exchanger and Diego blood group antigen, has an extended N-terminal domain attached to the membrane skeleton through ankyrin and proteins 4.2 and 4.1R. Glycophorin (GP)C and GPD, the Gerbich glycoproteins, have C-terminal domains attached to the membrane skeleton through protein 4.1R and p55. In addition there is evidence that several other integral membrane proteins interact with components of the membrane skeleton. Absence or reduced levels of these integral membrane proteins that interact with the membrane skeleton often result in some degree of abnormal red cell morphology (see Chapter 8).

ABH and Ii antigens are found predominantly on the carbohydrate moieties of the major red cell glycoproteins band 3 (anion exchanger) and on the glucose transporter, although they are also present on type 3 and some other minor glycoproteins and on the carbohydrate portions of membrane glycolipids. P, P1 and P^K antigens are expressed on the carbohydrate of glycolipids. The M and N antigens arise from interactions between the carbohydrate and polypeptide in glycophorin A. In addition to the antigens on integral membrane components, the Lewis and Chido/Rogers antigens are not of erythroid origin, but are adsorbed from the plasma.

Blood group antibodies

Several terms have been used in the past and are still sometimes used to describe different types of blood group antibodies, including 'naturally occurring' and 'immune' antibodies, and 'cold' and 'warm' antibodies. They are described below and an attempt is made to correlate these terms with the class of immunoglobulin involved.

Naturally occurring and immune antibodies

Antibodies are naturally occurring when they are produced without any obvious immunizing stimulus, such as pregnancy or transfusion. These antibodies are not present at birth and, in the case of anti-A,B, anti-A and anti-B, start to appear in the serum at about 3–6 months of age. ABO antibodies are probably produced in response to antigens of bacteria, viruses and other substances that are inhaled or ingested. Despite this probable antigenic stimulus, the term 'naturally occurring' is retained for these 'non-red-cell-induced' antibodies. 'Immune' red cell or alloantibodies are only produced after pregnancy or following transfusion or injection of blood or blood group substances.

Cold and warm antibodies

Cold antibodies give higher agglutination titres at low temperatures (0–4 °C), and many of them will not agglutinate red cells at 37 °C. Most naturally occurring antibodies are cold reacting. Some, such as naturally occurring anti-A,B, have a wide thermal range and will react at 37 °C, activating complement and leading to red cell lysis. However, the titre will be much higher at 0–4 °C. Cold antibodies that fail to react above 30 °C are of no clinical significance and can be ignored for blood transfusion purposes.

Immune antibodies have a thermal optimum of 37 °C. Any red cell antibody reacting above 30 °C should be considered potentially capable of destroying red cells *in vivo*.

IgM and IgG

IgM antibodies agglutinate red cells suspended in saline. They are often called saline or directly agglutinating antibodies. Conversely, IgG antibodies do not usually agglutinate saline-suspended red cells. However, lack of agglutination does not mean that the antibodies have not bound to their antigen, and it can be shown that they have reacted by using antiglobulin reagents, which facilitate agglutination of antibody-coated cells (see below). Most naturally occurring antibodies are IgM and preferentially cold-reacting, whereas most immune antibodies are IgG and warm-reacting, although some may be IgM. Exceptionally, very potent IgG antibodies may be directly agglutinating.

Monoclonal antibodies

Monoclonal antibodies have increased the repertoire of well-standardized blood grouping reagents of guaranteed quantity, quality and potency. By fusing the spleen cells of immunized mice or rats with drug-sensitive myeloma cells and selecting for drug-resistant hybrids, it has been possible to establish permanently growing cloned cell lines in tissue culture that secrete antibodies of desired specificities. Several murine hybrids secreting potent human blood-group-specific monoclonal antibodies have now been established, many of which were raised using immunogens other than intact red cells (e.g. anti-A, -B, -Le^a, -Le^b, -M and -N). Unfortunately, rodents have proved remarkably resistant to efforts to produce antibodies to the human D antigen.

Human monoclonal antibodies specific for RhD, other Rh antigens, Jk^a, Jk^b, and several other antigens, have been obtained by transformation of isolated peripheral blood lymphocytes from immunized individuals with Epstein–Barr virus or by fusion of human lymphocytes with mouse myelomas to form heteromyelomas.

Monoclonal antibodies of murine and human origin have widely replaced polyclonal blood grouping reagents in everyday ABO and D grouping. The future of antibody reagent production may lie in recombinant DNA technology, but a useful blood grouping reagent is still to be produced by these methods.

Lectins

Lectins are sugar-binding proteins, mostly extracted from plants and lower vertebrate animals. They are useful tools for routine and experimental blood group serology. They combine with simple sugars (e.g. fucose, galactose, *N*-acetylgalactosamine) present on the glycolipids and glycoproteins of cell membranes and body fluids. The three most commonly used blood-group-specific lectins are extracts from *Dolichos biflorus*, *Vicia graminea* and *Ulex europaeus*, which have anti-A_1, -N and -H specificities, respectively. Several other lectins have proven valuable in investigating red cell polyagglutinability. Lectins are also used for determining ABH secretor status and for partially purifying and identifying blood-group-active membrane glycoproteins.

Clinical significance of red cell antibodies

The clinical significance of red cell antibodies depends partly on their destructive capacity *in vivo* and partly on their frequency. ABO and D antibodies are by far the most significant.

Several factors influence immune red cell destruction *in vivo*:

1 *Plasma concentration and avidity of the antibody.*

2 *Thermal amplitude of the antibody.* Only antibodies reacting at 37 °C can cause haemolysis *in vivo* because complement can only be fixed at such a temperature and because antibody binding to Fc receptors can only take place at 37 °C.

3 *Immunoglobulin class and subclass.* The complement-fixing ability of most warm-reacting IgM antibodies, such as anti-A,B, anti-A and anti-B, makes them clinically significant. Of the IgG subclasses, IgG1 and IgG3 have clinical importance because of their capacity to bind to the Fc receptors of mononuclear phagocytic cells, the effector cells of extravascular immune red cell destruction. In addition, some IgG1 and IgG3 antibodies can fix complement up to C3b.

4 *Antibody specificity.* Several warm-reacting antibodies are incapable of causing *in vivo* red cell destruction (e.g. anti-Ch, -Rg, -Cs^a, -Kn^a, -Xg^a and most examples of anti-Yt^a).

5 *Antigen density on the red cell membrane.* The likelihood and degree of sensitization of a red cell with antibody and/or complement increases with the number of antigen sites on the surface. There are many more A sites than D sites on red cells.

6 *Volume of incompatible red cells transfused.* A small volume of incompatible red cells will be destroyed more rapidly than a large volume from the same donor. If antibodies are not very potent, larger volumes of cells may exhaust the circulating antibody available and saturate the mononuclear phagocytic system.

7 *Presence of antigen in donor plasma.* Lewis antigens (Le^a and Le^b) and the Ch/Rg antigens are primarily in plasma and are only secondarily adsorbed onto red cells. ABH antigens are present on red cells and plasma. The free antigen in plasma of transfused red cells can react with part of the recipient's antibody and inhibit its binding to red cells.

8 *Activity of cells of the mononuclear phagocyte system.* The ability of macrophages to remove sensitized red cells varies between individuals. Splenectomy and drugs such as corticosteroids and immunosuppressants will decrease the clearance of IgG-sensitized cells.

9 *Sensitivity of red cells to complement.*
10 *Extent of complement activation.* Some antibodies regularly bind complement and others do so rarely or not at all. Of the complement-binding antibodies, IgM (e.g. anti-A,B, -A, -B, -Le^a, -PP1P^K) will activate the complement cascade through to C9, resulting in intravascular haemolysis, but with IgG antibodies, such as anti-Fy^a, -Jk^a and -K, the cascade is interrupted at the C3 stage. Red cells coated with IgG and C3b will be destroyed extravascularly in the liver. As a rule, Rh antibodies do not fix complement.

Detection of red cell antigen–antibody reactions

Agglutination techniques

There are various ways of detecting antigen–antibody reactions *in vitro*, either manually or with automated methods. In manual methods, tubes, microplates or gels can be used. The most widely used methods employ the following techniques.

Direct agglutination

Most IgM antibodies will directly agglutinate saline-suspended red cells of the appropriate phenotype. This method is used routinely for ABO grouping with polyclonal or monoclonal antibodies and RhD typing with monoclonal antibodies.

Indirect agglutination

Apart from ABO, antibodies against most blood group antigens are IgG and generally will not produce direct agglutination of red cells. Such antibodies can be detected by the antiglobulin technique or with the aid of agents that enhance agglutination, for example proteases, such as papain, albumin and other colloids, aggregating agents such as polybrene and polyethylene glycol, and low ionic strength solutions (LISS).

Antiglobulin or Coombs test

The antiglobulin test is used to detect IgG antibodies that do not cause direct agglutination of saline-suspended red cells. The technique can be used to test directly, with an antiglobulin reagent, for the presence of antibodies or complement components that are bound to the red cells *in vivo*, as in autoimmune haemolytic anaemia, haemolytic transfusion reactions (HTRs) or haemolytic disease of the fetus and newborn (HDFN); this is the so-called direct antiglobulin test (DAT). Alternatively, the test can be used before transfusion to detect IgG antibodies in a patient's serum by adding the appropriate screening test red cells and then, after incubation and thorough washing (except in microcolumn tests), adding an antiglobulin reagent that will agglutinate cells coated *in vitro* with antibody or complement components; this is the indirect antiglobulin test (IAT). If IgG antibodies are found, these will be identified by IAT with an extended red cell panel. The IAT will then be used with the appropriate antigen-negative cells in the cross-match, to ensure compatibility between the patient's serum and donors' red cells. IATs are also used with many reagent antibodies, such as anti-K, -Fy^a, and -Jk^a for determining blood group phenotypes (Figure 12.2).

The first step in the antiglobulin test is the antibody uptake during the incubation of antibody with red cells carrying the antigen (Figure 12.2). When an IAT is undertaken with red cells suspended in normal-ionic-strength saline (NISS) solution, maximum antibody uptake, and hence maximum sensitivity, is achieved within 60–90 min incubation. The incubation phase can be reduced to 10–15 min by using low-ionic-strength saline (LISS) solutions. When the DAT and IAT are performed in tubes or microplates, it is essential that, after the antibody uptake, the red cells are washed three or four times with a large volume of saline before adding the antiglobulin reagent, as any free IgG or complement will neutralize the anti-IgG or anticomplement reagent and lead to false-negative reactions. With microcolumns containing a matrix

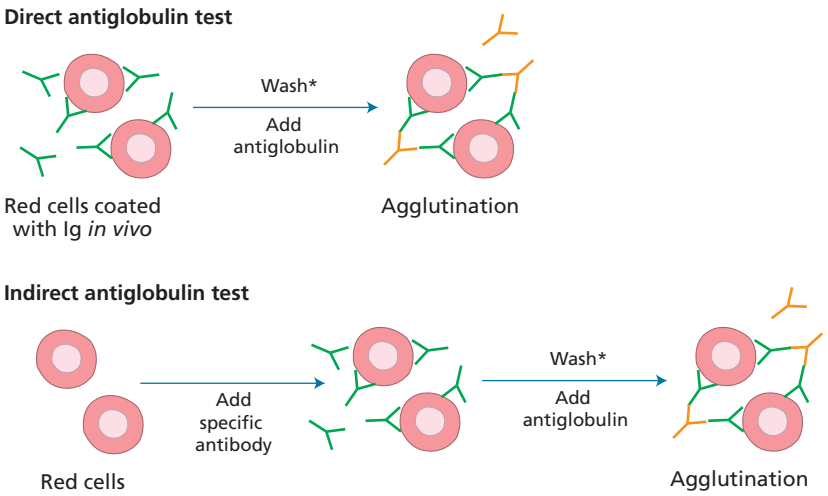


Figure 12.2 Direct and indirect antiglobulin tests. Antiglobulin antibodies are shown in red. *Washing is not required in microcolumn technology.

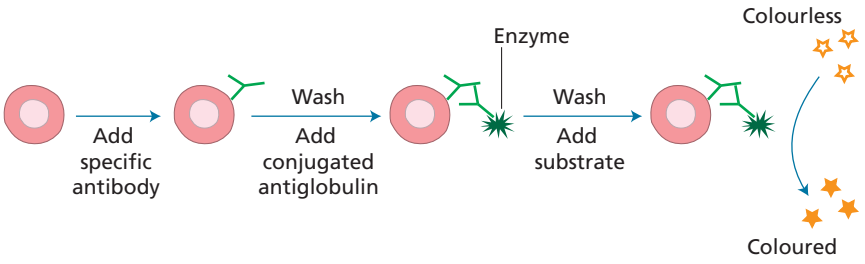


Figure 12.3 Enzyme-linked immunosorbent assay (ELISA) for cell-bound antibodies. Enzyme substrate is represented by open and closed stars, for colourless and coloured derivatives respectively.

of either a gel or glass beads, the wash phase has been eliminated.

Antibodies against IgG and complement should always be present in polyspecific antiglobulin reagents. The anti-IgG component is essential for pretransfusion antibody screening and cross-matching, as the vast majority of clinically significant antibodies, apart from anti-A,B, -A and -B, are IgG. The anticomplement component is needed for the detection of occasional examples of weak complement-binding antibodies (e.g. some anti-Jk^a) and for the detection of *in vivo* complement coating of red cells (i.e. in the DAT). Antiglobulin reagents are generally produced from polyclonal antibodies to IgG, made in experimental animals, and with monoclonal antibodies to C3d.

Inhibition of agglutination

Expected agglutination reactions with known antigens and antibodies can be neutralized by soluble antigens of the appropriate specificity. For example, the saliva of group A secretors inhibits the agglutination of group A cells by anti-A. Hydatid cyst fluid with P1P^K activity can be used to confirm the presence of anti-P1 in sera. Soluble antigens produced by recombinant DNA technology are useful in reference laboratories as aids to sorting out complex serological problems.

Haemolysis

After the incubation of red cells with fresh serum, red cell lysis indicates a positive antigen–antibody reaction mediated by IgM complement-fixing antibodies. A pink- or red-coloured supernatant after settling or centrifugation of red-cell–antibody mixtures is an indication of haemolysis.

Adsorption and elution tests

Specific antibodies can be removed from serum by adsorption on red cells carrying the corresponding antigen. Bound antibodies can be subsequently recovered from the washed, sensitized red cells by elution with heat treatment, freeze-thawing, low pH or chloroquine treatment.

Specialized antiglobulin techniques

Flow cytometry and enzyme-linked immunosorbent assays (ELISAs) are used for the estimation of the number of antigen sites on red cells, for the quantitation of anti-D in pregnant women and when more sensitive antiglobulin techniques are required. The essence of these approaches is the use of a labelled antiglobulin: in flow cytometry, fluorescent anti-IgG is used (or antibodies can be directly coupled with a fluorescent dye); in the case of ELISA, enzymes are attached covalently (i.e. conjugated) to the antiglobulin (Figure 12.3).

In these assays antibody-coated cells or particles are incubated with the labelled antiglobulin reagent. After incubation, excess unbound antiglobulin is washed away and bound labelled antiglobulin is then measured. For flow cytometry, a cytofluorimeter is used to measure fluorescence of each cell. For ELISA, the bound enzyme conjugate is detected through the enzyme's ability to modify its substrate, usually effecting a colour change. The colour intensity of the modified substrate is then measured with a spectrophotometer.

Microcolumn tests (gel and beads)

The principle of microcolumn tests is the separation of agglutinated from non-agglutinated red cells by centrifugation through a miniature filtration column. For blood grouping, red cells are layered on microcolumns impregnated with blood grouping sera; for antibody screening and identification, phenotyped panel red cells are mixed with patients' sera within the incubation chamber of the microcolumn. After centrifugation, agglutinated red cells are retained towards the top of the microcolumn because the agglutinates are trapped by the column matrix, whereas unagglutinated red cells form a button at the bottom of the column (Figure 12.4).

Antiglobulin tests can be performed by centrifuging the preincubated mixture of red cells and patient's serum through a column impregnated with anti-IgG or polyspecific antiglobulin. Any red cells coated with IgG will be agglutinated by the free anti-IgG in the matrix of the column, giving a positive result. No washing of sensitized cells is required, as IgG in the serum does not penetrate the column, as it is less dense than the matrix, and so does not neutralize the anti-IgG in the column. The advantages of microcolumn techniques are the ease of reading and reproducibility, and the fact that the tests can be stored for later examination, checking, photocopying or photographing. These systems can be automated and the agglutination results evaluated by image analysis.

Figure 12.4 Results of a gel microcolumn test. The subject is group B D-negative. Red cells remaining at the top of the gel represent a positive result; red cells collected at the bottom of the tube represent a negative result.

Microplate techniques

Semi-automated blood grouping and antibody screening can be performed in microtitre plates, which can also be used for extended phenotyping of red cells, antibody identification and large-scale screening for rare red cells and antibodies. A single microplate is equivalent to 96 short test tubes and the same basic principles of discrete analysis of agglutination apply. Several commercial microplate-based blood grouping systems are available, using bar codes to identify samples, automated liquid handling of samples and reagents, and plate readers linked to computers for easy and accurate record-keeping.

Solid-phase microplate systems, with antigen or antibody adhered to the surface of the plastic, reduce the variability between tests, inherent in liquid-phase systems when undertaken by different operators, and easily lend themselves to automated reading of results. Blood grouping (e.g. for ABO and D groups) can be accomplished with U-shaped microplate wells coated with the relevant antibody (e.g. anti-A, -B, -D); suspensions of patients' or donors' red cells are added to the wells and then centrifuged. Positive results appear as a carpet of cells coating the bottom of the well. Negative results appear as a tight pinhead of unattached cells in the centre of the well.

In antibody screening and identification by solid-phase, the microplates are coated with a panel of phenotyped red cell ghosts. The wells are incubated with patient serum and LISS, washed and then anti-IgG is added. Anti-IgG will bind to those wells where the patient's IgG antibody has bound to the relevant red cell ghost; bound anti-IgG is then easily detected by adding indicator red cells coated with IgG (e.g. D-positive cells coated with anti-D), followed by brief centrifugation, as described above.

Automated techniques

Fully automated blood grouping and antibody screening, using microcolumn techniques or microplates, is carried out in transfusion centres, where large numbers of donor samples are tested daily, and, increasingly, in hospital transfusion laboratories. In some automated systems test samples are mixed with typing sera or screening cells in individual wells of a special microplate with a terraced surface at the bottom of the well. After incubation and settling of the red cells, agglutination patterns are distinguished either on the basis of light transmission or by image analysis with the aid of a computer-controlled camera.

There are fully automated walk-away systems, based on microcolumns or solid-phase microplate systems. One system applies magnetized red cells for blood grouping and antibody detection, in order to avoid centrifugation.

Blood grouping reagents

In pretransfusion testing, it is essential to determine, as a minimum, the correct ABO and RhD groups of donors and recipients. To avoid potential fatalities resulting from errors, it is essential that the ABO and RhD typing reagents have suitable potency and comply with the European Directive on *in vitro* diagnostic devices and the associated Common Technical Specifications and carry the 'CE' mark to show they are in conformance.

IgM monoclonal reagents are generally used for ABO, Rh and K typing. Such reagents are extensively tested prior to release and are generally free of unwanted contaminating antibodies. Blood grouping reagents from polyclonal antisera are more likely to have unwanted antibodies and should have been exhaustively tested with an extensive panel of cells to exclude common and rare specificities before they are issued for routine use.

Antibody screening and identification

With plenty of anticipation, for all patients with any possibility of needing a transfusion, sera should be screened against unpooled group O cells from selected individuals known to carry the following antigens between them: D, C, E, c, e, M, N, S, s, P1, Le^a, Le^b, K, k, Fy^a, Fy^b, Jk^a and Jk^b. Ideally, one cell sample should be R_1R_1 (DCe/DCe) and the other R_2R_2 (DcE/DcE), so that the Rh antigens are all in double dose. A minimum homozygous expression (i.e. double dose) of Fy^a and Jk^a should be present on one of the red cell samples. It is generally possible to meet these requirements with two screening cell samples, but if more antigens with homozygous expression are required, three cell samples might be needed. Zygosity can be checked by molecular genetic methods.

Antibody screening of patients' sera need only consist of a well-controlled sensitive IAT, commonly with microcolumns. If the antibody screening is positive, antibody identification against a panel of 8–12 fully phenotyped red cells should be performed. For identification, in addition to the IAT, a second

sensitive technique (e.g. enzyme-treated cells or the polyethylene glycol or manual polybrene test), is recommended. Saline tests are not essential for antibody screening or identification and all tests should be performed at 37 °C, as antibodies reacting at lower temperatures are of no clinical importance.

Molecular techniques for blood grouping

It is possible to predict blood group phenotypes from DNA with a high degree of accuracy. This is usually performed when a blood group phenotype is required but a suitable red cell sample is not available. The most important application is the determination of fetal blood groups. When a pregnant woman has a blood group antibody with the potential to cause severe HDFN, it is beneficial to determine whether her fetus has the corresponding antigen and consequently whether it is at risk from HDFN. Hence, molecular RhD typing is usually required, but typing for Rhc and K antigens may sometimes be needed. The usual source is the small quantity of cell-free fetal DNA present in the maternal plasma. In some countries fetal RhD typing is applied to all RhD-negative pregnant women in order to avoid giving antenatal Rh immunoglobulin prophylaxis unnecessarily when the fetus is RhD-negative.

Molecular methods for blood grouping are useful in transfusion-dependent patients, where serological methods are not possible because of the presence of transfused red cells in the patient's blood. Tests are carried out on DNA isolated from whole blood of the transfused patient. If genotypes can be determined for all clinically important blood group polymorphisms, then matched blood can be provided to prevent the patient from making multiple red cell alloantibodies. Another application is blood grouping of patients with autoimmune haemolytic anaemia, whose red cells are coated with immunoglobulin, making serological typing difficult.

Several types of tests are commonly used in blood group genotyping, most of which are available commercially:

1 Amplification of a portion of a blood group gene to determine its presence. This only applies to *RHD*.

2 Amplification of a portion of a gene followed by detection of the polymorphism with restriction endonucleases.

3 Selective amplification of a specific allele by the use of an allele-specific primer.

4 Amplification of a portion of a gene followed by direct sequencing of the amplified product.

5 Allelic discrimination by polymerase chain reaction (PCR) incorporating allele-specific fluorescently labelled probes, in which relative quantities of a pair of alleles are measured.

6 Application of microarray technology, usually involving coloured microbeads, in which numerous polymorphisms can be tested from a relatively small quantity of DNA. Testing of large numbers of blood donors for multiple blood groups is beginning to be introduced, in order to establish a database of donors typed for all clinically significant groups. This would be valuable for transfusion-dependent patients. High-throughput methods are now being introduced for such testing and may involve the application of high-throughput sequencing technologies in the near future.

The ABO system

ABO was the first system to be recognized and remains the most important in transfusion and transplantation. Almost everybody over the age of about 6 months has clinically significant anti-A and/or anti-B in their serum if they lack the corresponding antigens on their red cells. Thus, if we consider the incidence of ABO blood groups in the UK (Table 12.2), transfusions given without regard to ABO would result in a major incompatibility (patient has the antibody and the antigen is on the transfused red cells) about once in every three cases.

Antigens of the ABO system

The vast majority of human bloods can be grouped into six main ABO phenotypes (Table 12.3), although several rare weak variants can be distinguished serologically. The incidence of ABO groups varies very markedly in different parts of the world and in different races.

A_1 and A_2 subgroups

The distinction between the A_1 and A_2 subgroups is usually made by using anti-A_1, which will agglutinate A_1, but not A_2, red cells. There are several sources of anti-A_1: (i) by absorbing anti-A (from group B people) with A_2 red cells, (ii) in the serum of some A_2 and A_2B persons (Table 12.3), (iii) from a saline extract of the seeds of the hyacinth bean *Dolichos biflorus* and (iv) mouse monoclonal anti-A_1. Anti-A_1 is not used routinely as it is not necessary to distinguish A_1 from A_2 red cells for most transfusion recipients. There is no specific antibody for A_2 red cells; if anti-A is absorbed with A_1 cells, all the antibody is removed. Group B serum can therefore be thought of as containing two antibodies: anti-A, which

Table 12.2 Incidence of ABO groups in southern England.

Phenotype	Frequency (%)	
O		44.9
A_1	30.8	41.1
A_2	10.3	
B		10.1
A_1B	2.7	3.9
A_2B	1.2	

Table 12.3 ABO blood grouping.

Agglutination of test cells with				Agglutination of test serum of			ABO group of test sample	Possible genotype
Anti-A	Anti-A_1	Anti-B	Anti-A,B*	A cells	B cells	O cells		
−	−	−	−	+	+	−	O	*O/O*
+	+	−	+†	−	+	−†	A_1	A^1/A^1, A^1/O, A^1/A^2
+	−	−	+	−/+‡	+	−	A_2	A^2/A^2, A^2/O
−	−	+	+	+	−	−	B	*B/B*, *B/O*
+	+	+	+	−	−	−†	A_1B	A^1/B
+	−	+	+	−/+‡	−	−	A_2B	A^2/B

*Anti-A,B (group O serum) is not generally used in routine laboratories.
†Some group A_1 and A_1B individuals may have weak anti-H in their plasma.
‡Serum from a proportion of A_2 (1–8%) and A_2B (22–35%) individuals contains anti-A_1.

agglutinates both A_1 and A_2 red cells, and anti-A_1, which agglutinates only A_1 red cells. The anti-A component of group O serum also has both antibodies.

The difference between the A_1 and A_2 subgroups is partly quantitative: the red cells of A_1 and A_1B subjects have more A antigen sites than A_2 and A_2B subjects, respectively. For practical purposes, A_2 can be regarded as a weaker form of A. When both A and B antigens are present, there are less sites for each than when either is present alone. The practical importance of this lies in the fact that the A antigen of A_2B red cells may give an extremely weak reaction with anti-A, which could be missed in routine grouping tests if reagents of inadequate potency are used. Moreover, if the same person's serum contains anti-A_1 and is tested in the reverse grouping only with A_1 and not A_2 red cells, he/she will be grouped as B. For this reason, potent anti-A reagents reacting with A_2B cells must be used in routine blood grouping. This is now possible with monoclonal antibodies.

There is also a qualitative difference between A_1 and A_2, but this must be very subtle because A_2 red cells can absorb all the anti-A from group B serum if the absorption is carried out at 0–4 °C for sufficient time. The chemical basis is unresolved, but A-active oligosaccharide structures called type 4A are only present on A_1 cells.

H antigen

Group O cells have no antigens of the ABO system but do possess H antigen, the precursor upon which the products of the *ABO* genes act. The *H* gene (called *FUT1*) is on chromosome 19, whereas *ABO* is on chromosome 9. The H antigen is present to some extent on almost all red cells, regardless of the ABO group, but the amount of H antigen varies with the ABO group as follows: O > A_2 > A_2B > B > A_1 > A_1B.

Individuals with the rare Bombay phenotype are homozygous for inactive *FUT1* alleles (*h/h*). Their red cells are not agglutinated by anti-A or anti-B, regardless of *ABO* genotype, but are not group O as they are also not agglutinated by anti-H. The serum of Bombay subjects contains potent anti-H, anti-A and anti-B that will only allow transfusion with red cells of the scarce Bombay phenotype.

Distribution of the A, B and H antigens

ABH antigens are often referred to as histo-blood group antigens because they are widely distributed in the body. They are therefore very important in transplantation. They are present on white cells, platelets and epidermal and other tissue cells. They are also present in the plasma, regardless of ABH secretor status, and in the saliva and other secretions of ABH secretors (see later).

Rare ABO variants

Rare ABO variants are usually disclosed because an expected ABO antibody is missing. A sample typed as group O that has anti-B but no anti-A will usually prove to be a weak A variant. The presence of weak A or B antigens can be demonstrated either by using potent antisera or by adsorption and elution. A variety of A and B variant phenotypes exist with the symbols A_3, A_x, A_{end}, A_m A_{el}, B_3, B_x, B_m and B_{el}. All are extremely rare, are usually recognized by their variable reactions with anti-A and/or anti-A,B sera, and arise from mutations in the coding or regulatory regions of the *ABO* gene.

Weakening of the A antigen can occur in acute myeloid leukaemia. The A antigen may revert to almost normal in remission. Similar weakening of B, H and I has been described.

B-like antigens may be acquired by group A individuals who are suffering from bowel infections, usually associated with carcinoma or strictures of the large bowel. Red cells with an acquired B antigen are agglutinated by some anti-B, including some monoclonal anti-B, but not by the patient's own anti-B. Bacterial deacetylases convert *N*-acetylgalactosamine, the immunodominant sugar of the A antigen, into galactosamine,

a structure similar to galactose, the immunodominant sugar of the B antigen.

Antibodies of the ABO system

Anti-A, anti-B and anti-A,B

It is likely that 'naturally occurring' ABO antibodies arise in response to A- and B-like antigens present on bacterial, viral or animal molecules. Titres of ABO antibodies vary considerably with age, reaching a peak in young adults and then declining in old age.

Sera taken from people over the age of about 6 months that do not contain the expected A and B antibodies (Table 12.3) are very rare. They should always be investigated thoroughly; often, some interesting explanation will be found, for example a rare subgroup of A, a blood group chimera or congenital absence of IgM.

Anti-A and anti-B have a wide thermal range. Although they are active at 37 °C, they react better at lower temperatures. ABO antibodies always have some IgM component and, in group A and B persons, they are almost entirely IgM. Antibodies from group O individuals, even before immunization, usually have some IgG anti-A,B, an antibody that cross-reacts with both A and B structures.

Following immunization with red cells or blood group substances, the thermal characteristics of the antibodies change, but group A and B subjects continue to produce antibodies that are mainly IgM. Most group O persons, however, will produce IgG as readily as IgM anti-A,B. Consequently, mothers of children with ABO HDFN are almost always group O. Immune anti-A,B are mainly IgG2, which does not cause HDFN because there are no Fc receptors for IgG2 on the cells of the mononuclear phagocyte system. When the maternal serum contains potent IgG1 and/or IgG3 ABO antibodies, HDFN may occur, although this is comparatively milder than Rh HDFN (see Chapter 13). Some IgA anti-A or anti-B may be produced following immunization with A or B substances.

Dangerous 'universal' donors

Group O red cells can be transfused to A, B or AB recipients and group O donors were formerly, and inappropriately, called 'universal donors'. Group O donors have anti-A, anti-B and anti-A,B in their plasma, which will react with the recipient's A or B cells. Normally, if group A, B or AB recipients are transfused with a relatively small number of group O units of whole blood, the anti-A or anti-B that is transfused will be diluted out and neutralized by the plasma of the adult recipient, especially if plasma-reduced blood, or red cells in additive solution are used. However, the transfused units of red cells may contain potent ABO haemolytic antibodies, which may cause marked destruction of the recipient's A or B red cells, leading to a severe acute HTR. For this reason, the practice of transfusing group O whole blood, plasma platelets, or even plasma-reduced red cells, to non-O recipients should be strongly discouraged.

There is usually a shortage of group O blood, and not infrequently a surplus of group A blood. In the vast majority of cases, including emergencies, there is enough time to perform a rapid ABO group on the patient's cells, which will allow the transfusion of group-specific blood. If there is no time to do an ABO group before transfusion, group O red cells in optimal additive solution, devoid of plasma, should be given until the patient's blood group is known. The practice of transfusing group O platelets to non-O patients should be discouraged as the dose of adult platelets will contain at least 300 mL of plasma, unless part of it has been replaced by platelet-additive solution. Group O fresh-frozen plasma and cryoprecipitate should only be given to group O recipients.

Anti-A_1

Anti-A_1, reactive at room temperature (18–22 °C), can be found in the serum of 1–8% of group A_2 and 22–35% of group A_2B persons. Most of these antibodies are not of clinical importance because they do not agglutinate A_1 red cells at 30 °C and above. Very rarely, anti-A_1 able to react at 37 °C may lead to significant destruction of A_1 red cells *in vivo*. The appropriate group A_2 or A_2B red cells should be provided in these rare instances.

Anti-H

Several forms of anti-H exist:

1 Clinically significant 'true' anti-H occurs in the serum of the very rare person with Bombay phenotype. It is active at 37 °C and only Bombay phenotype blood may be transfused.

2 Anti-H and anti-HI, commonly found in the serum of group A_1, B and A_1B persons, react much more strongly with adult than with cord red cells. Anti-H is inhibited by secretor saliva; anti-HI is not. These antibodies do not usually agglutinate O cells above 30 °C. Very occasionally, anti-H/anti-HI may cause rapid destruction of at least some of the transfused O red cells *in vivo*. However, these antibodies will not interfere with red cell survival if ABO identical units, i.e. A_1, B or A_1B units, are transfused to A_1, B or A_1B recipients, respectively.

ABH secretor status

About 80% of Caucasian populations, ABH secretors, have H antigen, plus A or B according to their ABO genotype, in a water-soluble form in their body secretions. The remaining 20% are non-secretors.

Biochemistry and biosynthesis of ABH antigens

ABH antigens

A, B and H antigens on red cell membranes (Figure 12.1) are predominantly glycoproteins, the majority being on the N-glycans of the anion exchanger (band 3) and the glucose transporter

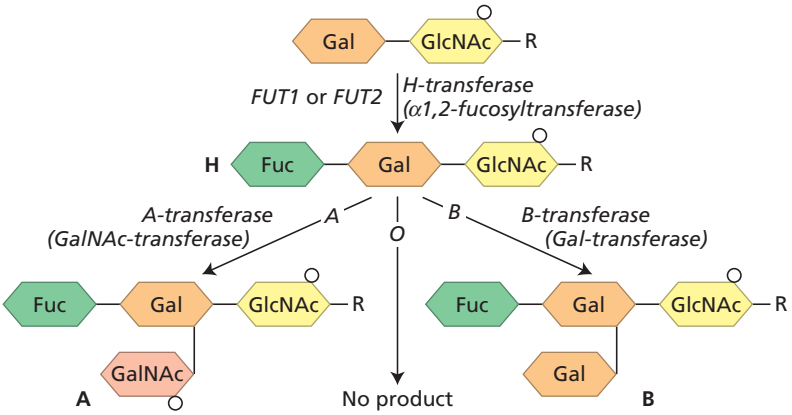


Figure 12.5 Biosynthetic pathway of H antigen from its precursor, and of A and B antigens from H. In group O subjects, H remains unconverted in the absence of *A* or *B* gene products. R, remainder of molecule.

(GLUT1). ABH antigens on red cells are also expressed on glycosphingolipids. Soluble ABH antigens are glycoproteins. Differences in the terminal sugars of the glycoproteins and glycolipids determine the specificity of these antigens: L-fucose (Fuc) for H, L-fucose plus *N*-acetyl-D-galactosamine (GalNAc) for A and L-fucose plus D-galactose (Gal) for B.

Two major types of carbohydrate chain endings serve as acceptors for the fucosyltransferases that synthesize H antigen: type 1 and type 2 chains have Gal joined to *N*-acetylglucosamine (GlcNAc) through 1→3 and 1→4 linkages, respectively. A- and B-transferases transfer GalNAc and Gal, respectively, from their donor substrates UDP-GalNAc and UDP-Gal to the terminal galactosyl residue of type 1 H and type 2 H, creating A and B epitopes and masking H specificity (Figure 12.5).

Secretory glycoproteins possess both type 1 and type 2 linkages, whereas red cells synthesize type 2 chains only. Other chains, called type 3 and 4, are also present in low numbers on red cells, but probably only on glycolipids.

ABO genes

ABO is located on the long arm of chromosome 9, comprises seven exons and encodes the A and B glycosyltransferases. Products of the *A* and *B* alleles differ by four amino acids encoded by exon 7, two of which determine whether the enzyme has GalNAc-transferase (A) or Gal-transferase (B) activity.

The majority of *O* alleles (called O^1) resemble *A*, but have a single-base deletion in exon 6, which creates a shift in the reading frame, truncating any putative polypeptide. About 3% of *O* alleles (called O^2) have a single-nucleotide polymorphism (SNP) that changes one of the vital amino acids in the catalytic site, inactivating the enzyme.

The A^2 allele has a single-base deletion immediately before the usual termination codon, creating a reading frameshift and abolition of this stop codon. This produces an A-transferase with 21 extraneous amino acids on its C-terminus, reducing its efficiency as a GalNAc-transferase.

H genes

At least two genes, *FUT1* and *FUT2*, on chromosome 19, are responsible for production of H antigen. Both encode α1,2-fucosyltransferases that catalyse the transfer of fucose to the terminal galactose of the H precursor chain (Figure 12.5). *FUT1* is active in mesodermally derived tissues, including haemopoietic tissues, and is responsible for H expression on red cells. Homozygosity for inactivating mutations in *FUT1* gives rise to Bombay and related phenotypes. *FUT2* is responsible for the expression of H antigen in endodermally derived tissues, including secretions, and hence is the gene responsible for ABH secretion. Secretors are homozygous or heterozygous for an active *FUT2*; non-secretors are homozygous for an inactive allele usually containing a nonsense mutation.

The Lewis system

Lewis antigens and their biosynthesis

Lewis is primarily a system of soluble antigens present in secretions and in plasma: Lewis antigens on red cells are adsorbed passively from the plasma, and plasma is needed to maintain Lewis antigens on the red cells. There are two Lewis antigens: Le^a and Le^b. Expression of either requires the presence of an active Lewis gene, but Lewis phenotypes are also governed by the gene controlling H secretion (*FUT2*). Lewis antigens in saliva and plasma are glycoproteins and glycolipids, respectively.

The Lewis gene, *FUT3*, encodes an α1,4-fucosyltransferase that catalyses the addition of L-fucose in 1→4 linkage to the subterminal GlcNAc of type 1 chains (Figure 12.6). Since red cells have only type 2 chains, they cannot make Lewis antigens. If the type 1 chain has been unmodified, the Lewis-transferase makes Le^a antigen, whereas if the secretor α1,2-fucosyltransferase has converted the type 1 chains to type 1 H, Le^b is produced. In a white population, 75% have active Lewis and secretor genes and hence Le(a−b+) red cells; 20% have an active Lewis gene but are H non-secretors and have Le(a+b−) red cells; 5% are

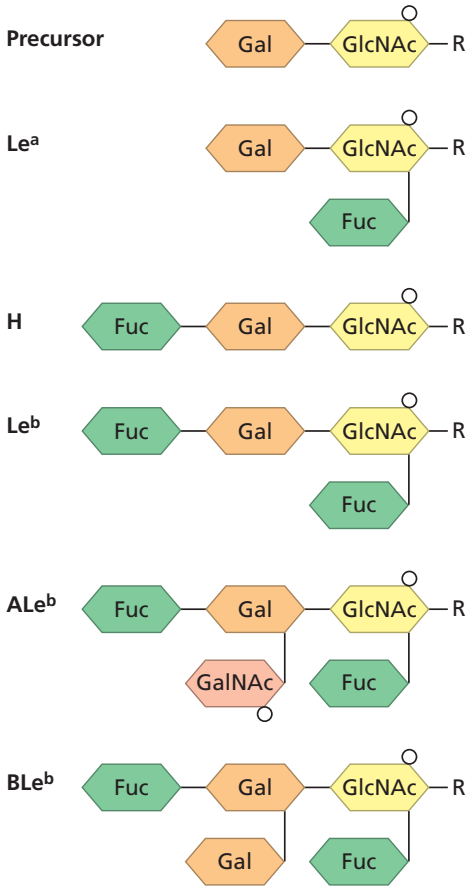


Figure 12.6 Diagrammatic representation of H and Lewis antigens in plasma and secretions. Le^a requires the action of the Lewis α1,4-fucosyltransferase (encoded by *FUT3*), H the action of the H α1,2-fucosyltransferase (encoded by *FUT2*), Le^b the action of both Lewis and H fucosyltransferases, and ALe^b and BLe^b the action of Lewis and H fucosyltransferases and the A or B glycosyltransferases.

homozygous for Lewis genes with inactivating mutations and are Le(a−b−) (Table 12.4). Le(a+b−) and Le(a−b+) red cells incubated in Le(a−b−) plasma shed their Lewis antigens into the plasma. Similarly, if Le(a+b−) or Le(a−b+) red cells are transfused into an Le(a−b−) recipient, the transfused cells will gradually lose their Lewis antigens and will group as Le(a−b−) within 1 week of transfusion.

Development of Lewis antigens on red cells

The Lewis antigens are poorly developed at birth and red cells from cord blood are usually Le(a−b−). Thereafter, Le^a develops first, followed by Le^b if the relevant Lewis and secretor genes are present. Red cells of children between the ages of 6 months and 4 years often type as Le(a+b+) if they are destined to become Le(a−b+). The definitive adult Lewis phenotype may not be reached until the age of 4–5 years.

Lewis antibodies

Lewis antibodies are generally made only by Le(a–b–) individuals. Anti-Le^a occurs fairly frequently. The incidence of Le(a−b−) is much higher in people of African and Southeast Asian origin than in Europeans, and Lewis antibodies may be found in up to 10% of random serum samples from black people.

Anti-Le^b commonly accompanies anti-Le^a. Pure anti-Le^b is uncommon.

Lewis antibodies are predominantly IgM. They usually agglutinate the appropriate cells at 20 °C. All Lewis antibodies that react at 37 °C will bind complement and may lyse antigen-positive cells. Lewis antibodies do not agglutinate red cells in saline at 37 °C, but can be detected with anticomplement in an indirect antiglobulin test (IAT).

Anti-Le^a active at 37 °C is usually more haemolytic than anti-Le^b, and some anti-Le^{a+b} can be very potent, leading to intravascular red cell destruction in the initial stages of transfusion. All Lewis antibodies active at 37 °C lead to two-component survival curves of transfused incompatible red cells, i.e. the first cells are destroyed at an accelerated rate and the remainder, which have

Table 12.4 The Lewis system and secretion of ABH.

Genotype		Antigens in saliva			Plasma/red cells*	
FUT2 (secretor)	**FUT3 (Lewis)**	**Le^a**	**Le^b**	**ABH**	**Le^a**	**Le^b**
Se/Se or *Se/se*	*Le/Le* or *Le/le*	+	++	++	−	++
Se/Se or *Se/se*	*le/le*	−	−	++	−	−
se/se	*Le/Le* or *Le/le*	+++	−	−	++	−
se/se	*le/le*	−	−	−	−	−

*Le^a and Le^b on red cells are passively adsorbed from plasma.
Le, active allele at *FUT3* locus; *le*, inactive allele; *Se*, active allele at *FUT2* locus; *se*, inactive allele.

lost their adsorbed Lewis antigens into the recipient's Le(a–b–) plasma, will have a normal survival.

For patients with Lewis antibodies reacting at 37 °C, ABO-identical red cells, compatible in an IAT cross-match at 37 °C, should be transfused. It is important to choose ABO-identical cells because group A, B and AB cells carry fewer Lewis antigens than group O cells, which are the cells routinely used in antibody screening. The provision of pre-typed Le(a−b−) blood for patients with Lewis antibodies is not necessary, as it is always easy to find ABO identical, cross-match-compatible red cells at 37 °C.

Lewis antibodies do not cause HDFN as they are IgM and, moreover, newborn infants have Le(a−b−) red cells.

P blood groups

P1 antigen is present on red cells of about 75% of Caucasians though P1 frequency varies in different populations. P1 is weakly expressed at birth and its strength varies considerably in adults. For this reason, identification of anti-P1 can be difficult, as panel cells will have varying expression of the antigen.

Anti-P1 is naturally occurring and commonly found in the serum of P1-negative individuals. Unlike anti-A and anti-B, anti-P1 rarely causes transfusion reactions because it is usually not reactive above 30 °C.

Anti-PP1P^K and anti-P invariably occur in the sera of the very rare individuals with p and P^K phenotypes, respectively. Anti-PP1P^K reacts with all red cells except those of the p phenotype; anti-P reacts with all red cells except those of the P^K and p phenotypes. They are usually strong IgM antibodies, often lytic at 37 °C, and can cause severe haemolytic transfusion reactions. If p or P^K patients need transfusion, autologous blood or red cells from the international panel of rare cells should be given. Occasionally, IgG anti-P or anti-PP1P^K have been associated with spontaneous early abortion.

The biphasic Donath–Landsteiner autoantibody, found in the sera of patients suffering from paroxysmal cold haemoglobinuria, is always IgG and usually has anti-P specificity.

I and i antigens and antibodies

I and i antigens are carbohydrates and are on the interior structures of the complex oligosaccharides that carry ABO, H and Lewis antigens; i is the biosynthetic precursor of I. The i antigen represents linear structures that are converted to I-active branched structures by the transferase coded by the *I* gene. This enzyme is not active in neonates. Consequently, red cells of most adults are agglutinated strongly by anti-I and only weakly by anti-i, whereas red cells from cord blood give the opposite result. Adults who are homozygous for rare inactivating mutations in the *I* gene have the adult i phenotype; their red cells react weakly with anti-I and strongly with anti-i, and their serum contains anti-I.

Autoanti-I active at 20 °C, but not at 30 °C, occurs in a variety of disorders and after blood transfusions. These patients' red cells give a negative direct antiglobulin test. A transient increase in strength, titre and thermal range of anti-I regularly occurs after infections with *Mycoplasma pneumoniae* and occasionally leads to acute haemolysis; red cells of such patients give a positive DAT, as do those of patients suffering from chronic cold haemagglutinin disease (CHAD). In CHAD, the autoantibody is nearly always monoclonal and usually has anti-I specificity.

Autoanti-i is found transitorily in many patients suffering from infectious mononucleosis. Very occasionally, the titre and thermal range of this antibody may lead to acute haemolysis. Autoanti-i may occasionally be the antibody specificity in chronic CHAD. Such patients often have an underlying lymphoma.

The Rh system

Rh is the second most important blood group system in blood transfusion. Anti-D is formed readily when D-positive blood is transfused to a D-negative person. The Rh system contains a total of 54 antigens, but D is the most important, owing to its high immunogenicity and ability to cause severe HDFN and HTRs.

Rh antigens

Alloanti-D was initially considered to be identical to antibodies produced in rabbits and guinea pigs immunized with rhesus monkey red cells. Later animal anti-Rh was shown to detect a different determinant renamed LW. LW is now known to be intercellular adhesion molecule-4 (ICAM-4). A remnant of the original name, rhesus, remains in the name of the blood group system, Rh. For clinical purposes, individuals are often called Rh-positive if they have the D antigen and Rh-negative if they lack D. Approximately 85% of Caucasians, 95% of Africans and over 99% of eastern Asians are D-positive.

Two pairs of antithetical antigens, C (RH2) and c (RH4), and E (RH3) and e (RH4), are closely associated with D genetically. There is no antithetical antigen to D, so d represents the absence of D. From serological analyses, D, C/c and E/e behave like the products of three closely linked genes and the eight Rh haplotypes are often written as *DCe*, *DcE*, *dce*, etc (shorthand, R^1, R^2, *r*, respectively) (Table 12.5).

Molecular genetics has shown that there are only two Rh genes, one encoding D, the other encoding the Cc and Ee antigens, but as the Cc and Ee polymorphisms are determined by separate regions of a single gene, the DCE terminology is still

Table 12.5 Eight Rh haplotypes and their frequencies in English, Nigerian and Hong Kong Chinese populations.

Haplotype		Frequencies (%)		
CDE	Rh-Hr	English	Nigerian	Chinese
DCe	R^1	42	6	73
dce	*r*	39	20	2
DcE	R^2	14	12	19
Dce	R^0	3	59	3
dcE	r''	1	Very rare	Very rare
dCe	r'	1	3	2
DCE	R^z	Rare	Very rare	Rare
dCE	r^y	Very rare	Very rare	Rare

Results of testing, with anti-D, -C, -c, -E and –e, red cells from 2000 English donors, 274 Yoruba of Nigeria and 4648 Cantonese from Hong Kong.

suitable for understanding Rh at most levels. The approximate frequencies of the Rh gene complexes in three populations are shown in Table 12.5. Genotype frequencies vary considerably in different parts of the world: e.g. *dce*/*dce* (Rh-negative) varies from about 35% in Basques to 0.3% in Japanese and Chinese.

Probable Rh genotype

When a person's Rh phenotype is known, the probable genotype can be discerned and its likelihood calculated from known genotype frequencies within the same population. When probable genotype determinations are carried out, it is very important that the ethnic origin of the person is known; figures for one population will not apply to people of other populations. For example, in white populations, *dce* is 15 times more common than *Dce*, whereas in African populations *Dce* has a slightly higher frequency than *dce*. Consequently, the phenotype D+ C+ c+ E– e+ represents a probable genotype of *DCe*/*dce* in a white person, but of *DCe*/*Dce* in a black person.

Molecular genetics of Rh

Rh genes and proteins

Rh antigens are encoded by two closely linked genes with 92% sequence homology. *RHD* encodes the D antigen and *RHCE* the Cc and Ee antigens. Each consists of 10 exons and, unusually for homologous genes, the two genes are in opposite orientations on the chromosome (Figure 12.7). Each gene encodes a 416-amino-acid polypeptide of 30–32 kDa that is palmitoylated but not glycosylated. The polypeptides encoded by *RHD* and *RHCE* differ by 31–35 amino acids, depending on the *RHCE* genotype. The Rh polypeptides traverse the lipid bilayer 12 times, with both

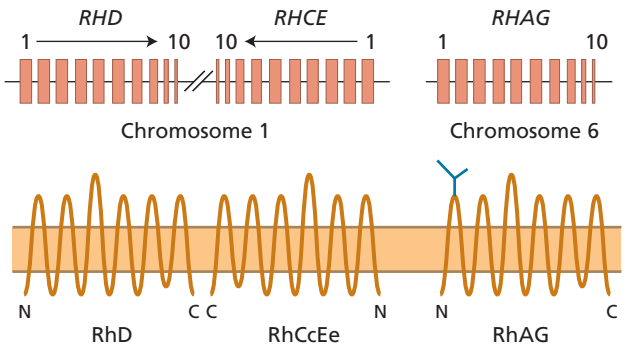


Figure 12.7 Rh and related genes (in pink) and the polypeptides they encode (in orange), showing the 10 exons of *RHD* and *RHCE* in reverse orientation on chromosome 1 and of *RHAG* on chromosome 6, and the RhD and RhCcEe polypeptides and RhAG glycoprotein crossing the membrane 12 times.

termini in the cytoplasm and six extracellular loops that provide the putative sites for antigenic activity (Figure 12.7). Rh antigen activity is very dependent on the conformation of the proteins within the membrane and may involve interactions between two or more of the extracellular loops.

D, C/c and E/e polymorphisms

In white people, the D-negative phenotype almost always results from homozygosity for a complete deletion of the *RHD* gene. D-negative represents the absence of the RhD protein from the membrane. D-positive people may be homozygous or hemizygous for *RHD*. However, most D-negative black Africans have an inactive *RHD*, called the *RHD* pseudogene (*RHDψ*).

The Cc polymorphism is associated with three or four amino acid substitutions encoded by exons 1 and 2 of *RHCE*, although the definitive change is Ser103 for C and Pro103 for c in the second extracellular loop of the RhCcEe protein. E and e are associated with Pro226 and Ala226, respectively, in the fourth extracellular loop.

The Band 3/Rh molecular macrocomplexes

The two Rh proteins are closely associated in the red cell membrane with the Rh-associated glycoprotein (RhAG). RhAG has a similar conformation, spanning the membrane 12 times, but is glycosylated (Figure 12.7). The complex of Rh proteins and RhAG is part of a membrane macrocomplex, with band 3 (the anion exchanger and Diego blood group antigen) at its core, and also containing the LW glycoprotein (ICAM-4), CD47 and glycophorins (GP)A and B. This macrocomplex is attached to the cytoskeleton, primarily through ankyrin and protein 4.2 (Figure 12.1). Another macrocomplex containing band 3 and the Rh proteins also contains the Kell, Kx and Duffy proteins, plus GPC (Gerbich), which links the complex to the cytoskeleton.

The Rh proteins are only present on erythroid cells, but RhAG is more widely distributed.

Variants of D

Most D-negative individuals lack the whole RhD protein from their red cells and, when immunized by D-positive red cells, can make antibodies to an array of epitopes on the external loops of the RhD protein. About 30 D epitopes have been defined with monoclonal antibodies. There are numerous variants of D. These used to be divided into two types, weak D (formerly D^u) and partial D, though these terms are not adequately defined and of little value clinically. The usual understanding is weak D red cells express all epitopes of D at a low level and individuals with weak D phenotype cannot make anti-D, whereas partial D red cells lack some epitopes of D and so individuals with a partial D phenotype, if immunized by a normal complete D antigen, can make antibodies to the D epitopes they lack. These antibodies behave as anti-D in most circumstances. However, things are not that simple and some phenotypes named weak D have subsequently been found in individuals who have made anti-D and, in addition, some so-called partial D antigens are expressed very weakly. So, the terms weak D and partial D should be abolished and the abnormal D antigens all included in the term D variant.

The most important D variant, from a transfusion perspective, is DVI ('D six'). It is not the most common, but lacks most of the D epitopes and is therefore the D variant most commonly associated with the production of anti-D. In the UK it is recommended that anti-D reagents for typing patients should not detect DVI. Consequently, DVI patients will be typed as D-negative and transfused with D-negative red cells and DVI pregnant women will be given anti-D prophylaxis. Ideally, one of the anti-D reagents used for typing donor red cells should detect DVI; hence DVI donors will be typed as D-positive. This strategy means that DVI individuals should be considered D-negative patients, but D-positive donors. Weak D types 1, 2, and 3 are three of the most common D variants and are almost never associated with anti-D production.

In view of the high quality of monoclonal anti-D in the UK, the recommended method for D typing of patients requires direct agglutination tests, in duplicate, with potent IgM monoclonal anti-D reagents. An antiglobulin test is no longer required. Patients whose red cells appear D-positive, but give weak or inconsistent reactions, suggestive of a D variant, are treated as D-positive and transfused with D-positive blood unless they are females of childbearing potential or transfusion dependent. In those cases, they are treated as D-negative and transfused with D-negative blood unless they are shown to have the weak D type 1, 2, or 3. This strategy is designed to conserve valuable stocks of D-negative red cells for patients who really need them.

Similarly, donors are not typed any longer for D by an antiglobulin test because it is unlikely that transfusion of very weak D red cells to a D-negative patient will result in immunization of the patient. In some countries, however, the antiglobulin test for D typing of donors is still required and in Switzerland a molecular test is mandatory for confirming D-negative status of donors.

Molecular basis of D variants

D variants result either from missense mutations in *RHD*, encoding single amino acid changes, or from an *RHD* in which a segment has been replaced by the equivalent segment of an *RHCE* gene due to unequal crossing over or gene conversion. For example, the most common form of DVI is produced by an *RHD–CE–D* hybrid gene in which exons 4–6 have the nucleotide sequence of *RHCE*, and external loops 3 and 4 of the encoded hybrid protein have the amino acid sequence of an RhCe protein. Hence, a protein is produced with a sequence similar enough to RhD to be stable in the membrane, but different enough to lack many epitopes of D. A variety of names have been used to denote the different types of D variants, for example DIII, DVI, DFR, DAR, DBT, which can be distinguished by sophisticated serological techniques, and weak D types 1–76, which can only be distinguished by molecular techniques.

Antibodies of the Rh system

Naturally occurring antibodies

Generally, Rh antibodies are only produced following immunization with red cells. However, anti-E is often naturally occurring; about half may occur without a history of pregnancy or transfusion. Rarely, naturally occurring anti-D and anti-C^w are found. All such naturally occurring Rh antibodies react optimally with enzyme-treated cells; if they react at 37 °C by IAT, they are clinically significant.

Immune antibodies

The clinical importance of the Rh system lies in the readiness with which anti-D arises in D-negative subjects after stimulation with D-positive red cells by pregnancy or transfusion. Prophylaxis against D immunization with anti-D immunoglobulin in pregnancy (see below) has led to a significant decrease in the incidence of anti-D, but it still remains the most common immune antibody of clinical relevance detected in a routine blood transfusion laboratory. D is considerably more immunogenic than the other Rh antigens, which have the following order of immunogenicity: c > E > e > C.

About 20–30% of anti-D sera also appear to contain anti-C. Usually, this anti-C is not a separable antibody and is probably more correctly called anti-G. About 1–2% of anti-D sera also contain anti-E. Anti-C (and anti-G) in the absence of anti-D is very uncommon.

The incidence of other Rh antibodies is much lower, but together they are more common than anti-K (Kell), which is the most immunogenic antigen after D. In routine screening, pure

anti-E is the most common, followed by anti-c, although anti-c is a more common cause of HDFN, which can be severe. This is probably because many anti-E are weak, naturally occurring antibodies. Anti-e, like anti-C, is very rare.

The vast majority of Rh antibodies are IgG and do not fix complement. Anti-D may occasionally be partly IgA. IgM anti-D is very rare.

D immunization and immunoprophylaxis

RhD is very immunogenic; over 80% of D-negative recipients of a whole unit of D-positive red cells make anti-D and 17% of D-negative women will be immunized by one pregnancy with an RhD-positive offspring.

Prediction of fetal Rh genotype by molecular methods

Knowledge of the molecular basis for D-negative phenotypes has made it possible to devise tests for predicting fetal D type from fetal DNA. This is valuable in determining whether the fetus of a woman with anti-D is at risk from HDFN. Most methods involve PCR-based tests that detect the presence or absence of *RHD*, but it is important to test for more than one region of *RHD*, so that hybrid genes responsible for partial D antigens do not give a false result, and to test for the *RHD* pseudogene (*RHDψ*), so that this does not give rise to a false-positive result.

Although fetal DNA can be obtained by amniocentesis, this procedure is associated with a significant risk of fetal loss and feto-maternal haemorrhage. The best source is the small quantities of free fetal DNA present in maternal plasma. This non-invasive form of fetal D typing is now provided for pregnant women with significant levels of anti-D as a routine service in the UK and other parts of Europe, and should be considered the standard of care in these patients.

It is also common practice to offer antenatal Rh prophylaxis to all D-negative pregnant women, yet in a predominantly white population about 40% of these women will have a D-negative fetus and receive the treatment unnecessarily. Successful trials have been carried out on high-throughput methods for determining fetal D type from DNA in maternal plasma, showing that fetal D typing is reliable from about 11 weeks gestation. Fetal D testing of all D-negative pregnant women is provided as a routine service in several European countries.

The MNS system

Antigens of the MNS system

M and N are inherited as codominant Mendelian traits, giving rise to three common genotypes *M*/*M*, *M*/*N* and *N*/*N*. The *Ss* locus, which is closely linked to *MN*, also consists of two codominant alleles in Europeans and Asians, producing S or s antigens. In northern Europeans, haplotype frequencies are as follows: *MS*, 25%; *Ms*, 28%; *NS*, 8%; *Ns*, 39%. The system also contains many variants. About 2% of black West Africans and 1.5% of African-Americans are S−s− and most of these lack the U antigen that is present when either S or s is expressed.

The MN antigens are carried on glycophorin A (GPA), which is encoded by the *GYPA* gene on chromosome 4. There are about 1 million molecules of GPA per cell, yet their absence in the rare En(a−) phenotype does not affect red cell function or survival. The negative charge of the red cells is mainly due to the ionized COOH groups of sialic acid (neuraminic acid), which is mostly carried on the oligosaccharides of GPA and can be removed with neuraminidase. En(a−) cells therefore have a reduced negative charge and behave serologically as if they have been treated with neuraminidase. GPB carries the S and s determinants. GPB is encoded by *GYPB*, which is closely linked and homologous to *GYPA*. S−s−U− red cells lack GPB. GPA and GPB are exploited as receptors by the malaria parasite *Plasmodium falciparum*.

Antibodies of the MNS system

Anti-M is uncommon and is usually naturally occurring, but can be immune and, very rarely, causes HDFN. Anti-N is also rare and is nearly always cold-reactive IgM. Useful anti-N lectin can be prepared from the seeds of *Vicia graminea*. Anti-S, the rarer anti-s and anti-U are usually immune, IgG and can cause HDFN. They have also been implicated in HTRs. Anti-U only occurs in S−s− black people and reacts with all cells that have the S or s antigens and up to 50% of cells that are S−s− and have a variant U antigen. Finding compatible blood for a patient with anti-U can prove difficult and it is often necessary to resort to the panel of rare cell donors.

The Kell blood group system

The Kell system consists of 35 antigens, including one triplet and four pairs of allelic antigens – K and k; Kp^a, Kp^b and Kp^c; Js^a and Js^b; K11 and K17; K14 and K24 – all of which represent amino acid substitutions in the Kell glycoprotein. Three sets of antigens are clinically important: K/k; Kp^a/Kp^b /$Kp^c$and Js^a/Js^b.

In European whites, the incidence of the K/k phenotypes, the most important clinically, is as follows: K+k−, 0.2%; K+k+, 8.7%; K−k+, 91.1%. K is rare in populations other than whites. Js^a is present in about 20% of black people, but is extremely rare in other ethnic groups. K_o, a very rare null phenotype of the Kell system in which no Kell antigens are expressed, results from homozygosity for a variety of inactivating mutations in the *KEL* gene.

The Kell antigens are located on a glycoprotein that crosses the cell membrane once, with a large glycosylated C-terminal extracellular domain, maintained in folded conformation by

multiple disulfide bonds. Reduction of these bonds by 2-aminoethylisothiouronium bromide (AET) results in loss of expression of all Kell antigens, leading to artificial K_o cells with increased expression of Kx.

Anti-K is an important antibody in white populations; it is nearly always immune, IgG and complement-binding. It causes severe HTRs and HDFN. K stimulates the formation of anti-K in about 10% of K-negative people given one unit of K-positive blood. About 0.1% of all cases of HDFN are caused by anti-K; most of the mothers will have had previous blood transfusions. HDFN caused by anti-K differs from Rh HDFN in that anti-K causes fetal anaemia by suppression of erythropoiesis, rather than immune destruction of mature fetal erythrocytes. Kell antigen is expressed by erythroid cells at a very early stage of erythropoiesis and anti-K would facilitate destruction of early erythroid progenitors, before they become haemoglobinized. Anti-k is very rare, reacting with 99.8% of random blood samples. It is always immune and has been incriminated in HDFN. Most other Kell system antibodies are rare, best detected by the IAT and are usually clinically significant.

The Kell glycoprotein is linked by a single disulfide bond to the Kx protein, produced by an X-linked gene, *XK*. Absence of Kx, resulting from hemizygosity in males of a deletion of *XK* or of *XK* containing inactivating mutations, gives rise to the McLeod syndrome, in which there is no expression of Kx antigen, weak expression of all Kell system antigens and a neuroacanthocytosis that is usually characterized by late-onset muscular, neurological and psychiatric disorders, plus about 30% red cells with acanthocytosis and associated haemolytic anaemia. In some cases, an X chromosome deletion is large enough to encompass *XK* and *CYBB*, a gene for a subunit of flavocytochrome b_{558}, which leads to McLeod syndrome and chronic granulomatous disease (CGD). However, the majority of CGD patients have normal Kell antigens and Kx. If boys with McLeod syndrome and CGD are transfused, they are liable to make anti-Kx, which can cause severe HTRs, needing to resort to the International Rare Donor Panel for compatible blood.

Some other blood group systems

Fy^a and Fy^b antigens of the Duffy system represent a single amino acid substitution in the extracellular N-terminal domain of the Duffy glycoprotein. Their incidence in the UK is as follows: Fy(a+b−), 20%; Fy(a+b+), 46%; Fy(a−b+), 34%. About 70% of African-Americans and close to 100% of West Africans are Fy(a−b−). Anti-Fy^a is not infrequent. It is IgG, often complement-fixing and can cause HTRs, but seldom HDFN. It is best detected by IAT and does not react with red cells treated with the proteases papain and ficin. Anti-Fy^b is very rare and is always immune.

Kidd has two alleles, Jk^a and Jk^b, which represent a single amino acid change in the Kidd glycoprotein. Phenotype frequencies in the UK population are as follows: Jk(a+b−), 25%; Jk(a+b+), 50%; Jk(a−b+), 25%. The Kidd null phenotype, Jk(a−b−) is very rare in most populations, but reaches an incidence of greater than 1% in Polynesians. Anti-Jk^a is uncommon and anti-Jk^b is even rarer, but they may both cause severe transfusion reactions and, to a lesser extent, HDFN. Kidd antibodies have often been implicated in delayed HTRs; they are IgG and predominantly complement-fixing up to C3b, but may be difficult to detect because they tend to disappear and then reappear promptly in an anamnestic response. The antiglobulin test is the best method for detection; reactions are enhanced if the cells are protease-treated or if fresh serum is added as a source of complement.

Lutheran is a complex system comprising 22 antigens, including Lu^a and Lu^b, which represent a single amino acid substitution in the Lutheran glycoprotein. Phenotype incidence in the UK is as follows: Lu(a+b−), 0.1%; Lu(a+b+), 7.5%; Lu(a−b+), 92.4%. The extremely rare Lu_{null} phenotype, in which no Lutheran antigens are expressed, results from homozygosity for an inactive Lutheran gene; anti-Lu3 may be produced. Lutheran antibodies are uncommon and are not generally clinically significant, although anti-Lu^b may have caused mild delayed HTRs. Lu^a may be omitted from antibody screening cells.

The Diego system antigens represent single amino acid substitutions in band 3, the red cell anion exchanger. Di^a, is very rare in white and black people, but relatively common in Mongoloid people, with frequencies varying between 1% in Japanese and 50% in some native South Americans. Anti-Di^a and anti-Di^b are immune and rare, but can cause HDFN. Wr^a has a frequency of about 0.1%. Naturally occurring anti-Wr^a is present in approximately 1% of blood donors. Anti-Wr^a is often found in the serum of patients who have made other antibodies or who are suffering from autoimmune haemolytic anaemia. Very rarely, anti-Wr^a causes HDFN.

Do^a and Do^b of the Dombrock system represent a single amino acid substitution on the Dombrock glycoprotein. Approximately 67% of northern Europeans are Do(a+) and 82% Do(b+). Dombrock antibodies are extremely rare, immune and are best detected by IAT; they have been implicated in severe acute and delayed HTRs. Dombrock reagents are rare and unreliable, so typing is best achieved by molecular methods.

Co^a and Co^b of the Colton system represent a single amino acid substitution in the water channel aquaporin-1. Colton phenotype frequencies in white people are as follows: Co(a+b−), 90.5%; Co(a+b+), 9.0%; Co(a−b+), 0.5%. Colton antibodies are very uncommon. Anti-Co^a has caused severe HDFN, and has been implicated in acute and delayed HTRs.

Yt^a and Yt^b of the Yt system represent a single amino acid change in red cell membrane acetylcholinesterase. Yt^a and Yt^b have frequencies of about 99.8% and 8% respectively. Anti-Yt^a and anti-Yt^b are exceptional and most are of no clinical significance.

JR, Lan and At^a are clinically significant antigens of very high frequency. Anti-Vel, an antibody to an antigen of high frequency causes severe HTRs.

There are many other antigens, of either very high or very low incidence, not assigned to blood group systems. Some have caused HDFN and/or HTRs. For those rare individuals who have formed antibodies to high-frequency antigens, the provision of compatible blood can be a problem and it is often necessary to approach the national or international panels of rare donors for compatible units. Antibodies to low-frequency antigens are usually naturally occurring; they may occasionally give rise to unexpected incompatible cross-matches. Some antibodies to low-frequency antigens have caused HDFN.

Polyagglutinable red cells

Erythrocyte polyagglutination is the agglutination of red cells irrespective of blood group by many sera from normal adults. Polyagglutinable red cells are not agglutinated by the patient's own serum. The abnormality is a property of the red cells, not of the sera, in contrast to panagglutination, which is the agglutination of most red cells by one serum.

There are two main categories of polyagglutination, acquired and inherited. The acquired forms can be subdivided into: (i) microbial polyagglutination, which results from the action of microbial enzymes on red cell surface oligosaccharides (T, Tk, acquired B); and (ii) non-microbial polyagglutination, which is caused by somatic mutation (Tn). There are four types of inherited polyagglutination: Cad (strong expression of Sd^a), congenital dyserythropoietic anaemia type II (CDAII or HEMPAS), NOR and Hyde Park. Lectins are required for the identification of the different types of polyagglutination.

The biological significance of blood group antigens

The functions of many red cell membrane proteins bearing blood group antigenic determinants are known, or can be deduced from their structure. Some are membrane transporters, facilitating the transport of biologically important molecules through the lipid bilayer: band 3 glycoprotein, the Diego antigen, provides an anion exchange channel for HCO_3^- and Cl^- ions; the Kidd glycoprotein is a urea transporter in red cells and renal endothelial cells; the Colton and GIL glycoproteins are aquaporin water and glycerol channels; JR and Lan are ATP-binding cassette transporters, which translocate various substrates across biological membranes. RhAG probably functions as a channel for rapid CO_2 transfer in and out of the cell. CO_2 in the red cell is converted to bicarbonate ions by carbonic anhydrase II and then transported rapidly out of the cell to the plasma by band 3, part of the same macrocomplex as RhAG, where it is carried to the lungs.

The Lutheran, LW and Indian (CD44) glycoproteins are adhesion molecules, possibly serving their functions during erythropoiesis. The Duffy glycoprotein is a chemokine receptor and might function as a 'sink' or scavenger for unwanted chemokines. The Cromer and Knops antigens are markers for decay-accelerating factor and complement receptor 1, respectively, which protect the cells from destruction by autologous complement. Some blood group glycoproteins have enzyme activity: the Yt antigen is acetylcholinesterase and the Kell antigen is an endopeptidase that synthesizes endothelin-3, a vasoconstrictor. The C-terminal domains of the Gerbich antigens, GPC and GPD, and the N-terminal domain of the Diego glycoprotein, band 3, anchor the membrane to its skeleton. The carbohydrate moieties of the membrane glycoproteins and glycolipids, especially those of the most abundant glycoproteins, band 3 and GPA, constitute the glycocalyx, an extracellular coat that protects the cell from mechanical damage and microbial attack.

The difference between red cell antigens that represent the products of alleles (e.g. A and B, K and k, Fy^a and Fy^b) is small, often being just one monosaccharide or one amino acid. The biological importance of these differences is unknown and there is little evidence to suggest that one antigen confers any significant advantage over another. Most blood group systems have a null phenotype in which the whole blood group protein is absent from the red cells or any other cells. These usually result from homozygosity for gene deletions or inactivating mutations within the genes. In most cases, individuals with these null phenotypes are apparently healthy, suggesting that whatever the precise function of the missing structure may be, some other structure must be able to substitute in its absence. However, there are exceptions; rare Rh_{null} individuals, who lack D and CcEe Rh proteins, as well as RhAG, have chronic haemolysis, which may be compensated by increased red cell production, but may require splenectomy for stabilization. Absence of the Kx protein causes weakness of expression of all Kell antigens, as a result of linkage between the two proteins in the membrane, but is also associated with neuroacanthocytosis. Very rare patients lacking the RAPH glycoprotein, CD151, have disruption of basement membranes causing hereditary nephritis, epidermolysis bullosa and neurosensory deafness. A patient lacking the Diego antigen, the anion transporter, only survived with extreme medical intervention. On the other hand, people with the rare Bombay phenotype lack ABH antigens from all cells and tissues, with no apparent ill effect or red cell abnormality.

Some blood group antigens are exploited by pathological microorganisms as receptors for attaching and entering cells. Consequently, in some cases absence of antigens can be beneficial. The Duffy glycoprotein is used by *Plasmodium vivax* to penetrate red cells. Hence, the Fy(a−b−) phenotype, common in Africans, confers resistance to *P. vivax* malaria. *Plasmodium falciparum* malaria appears to have played a part in the

establishment of the ABO polymorphism. The geographic distribution of group O is consistent with selection pressure by *P. falciparum* in favour of group O individuals in malaria endemic regions.

Selected bibliography

Anstee DJ (2009) Red cell genotyping and the future of pretransfusion testing. *Blood* **114**: 248–56.

Daniels G (2013) *Human Blood Groups*, 3rd edn. Wiley-Blackwell, Oxford.

Daniels G, Bromilow I (2013) *Essential Guide to Blood Groups*, 3rd edn. Wiley-Blackwell, Oxford.

Daniels G, Finning K, Martin P, Massey E (2009) Non-invasive prenatal diagnosis of fetal blood group phenotypes: current practice and future prospects. *Prenatal Diagnosis* **29**: 101–7.

Fung MK, Grossman BJ, Hillyer C, Westhoff CM (2014) *AABB Technical Manual*, 18th edn. American Association of Blood Banks, Bethesda, MD.

International Society of Blood Transfusion Working Party on Red Cell Immunogenetics and Blood Group Terminology. http://www.isbtweb.org/nc/working-parties/red-cell-immunogenetics-and-blood-group-terminology/ (accessed May 2015).

Klein HG, Anstee DJ (2014) *Mollison's Blood Transfusion in Clinical Medicine*, 12th edn. Wiley-Blackwell, Oxford.

Kumpel BM (2007) Efficacy of Rh monoclonal antibodies in clinical trials as replacement therapy for prophylactic anti-D immunoglobulin: more questions than answers. *Vox Sang* **93**: 99–111.

National Blood Service (2013) *Guidelines for the Blood Transfusion Services in the United Kingdom*, 8th edn. The Stationery Office, London (available at http://www.tsoshop.co.uk/bookstore.asp; accessed May 2015).

Poole J, Daniels G (2007) Blood group antibodies and their significance in transfusion medicine. *Transfusion Medicine Reviews* **21**: 58–71.

CHAPTER 13

Clinical blood transfusion

13

Shubha Allard and Marcela Contreras

Barts Health NHS Trust and NHS Blood and Transplant
Blood Transfusion International, London, UK

Introduction

This chapter covers key aspects of blood donation and collection, and the preparation and storage of blood components; it describes pretransfusion testing of donors' and recipients' blood, complications and adverse effects of blood transfusion, as well as the appropriate use of components and transfusion alternatives available in different clinical scenarios. The features of haemolytic disease of the fetus and newborn are also described.

Blood transfusion and regulatory aspects

Concerns about blood safety and in particular transfusion-transmitted infection have resulted in an increasingly stringent regulatory framework for transfusion medicine. The World Health Organisation (WHO) and the Council of Europe (CoE) have defined a code of ethical principles, which are not legally binding, but widely serve as guidance for countries, on blood donation and transfusion practice. The European Union (EU) Blood Directives set standards of quality and safety for the collection, testing, processing, storage and distribution of blood and components. In the UK these Directives have been transposed into law as the Blood Safety and Quality Regulations 2005.

The blood donor (Tables 13.1 and 13.2)

> Blood donation shall in all circumstances be voluntary. Financial profit must never be a motive for the donor or for those collecting the donation.

These statements sum up the attitude of the WHO and the International Society of Blood Transfusion towards the principle of blood donation. Repeated voluntary donations are associated with the lowest risk of transfusion-transmitted infection. However, in a number of countries worldwide, whole blood donation and, especially, apheresis plasma donation, are still remunerated.

Selection criteria and blood donation

Donor selection criteria are aimed at ensuring the safety of the donor and also minimizing risk to recipients needing transfusion, in particular risk of the transmission of infection. Blood donors should be healthy adults within a permitted age range with minimum requirements for weight, haemoglobin, donation volume and donation frequency (see Table 13.1).

The minimum age for blood donation in many countries is 17 years, essentially aimed at avoiding donation in adolescence when iron requirements are high. Similarly pregnant and lactating women are not accepted. In the UK the upper age limit for first-time donors is 65 years, with no upper age limit for regular donors, but with annual health review after their 66th birthday. A full donor medical history (Table 13.2) will help identify those not suitable for donation (e.g. cancer, cardiovascular or renal disease). A full travel history is essential together with any activities known to be associated with increased risk of acquiring infections with temporary or permanent deferral as needed (see below). Due to the risk of delayed faints, donors in hazardous occupations or hobbies (e.g. scaffolding, firemen, climbers) should not be allowed to return to work or their hobbies until the following day.

The haemoglobin level is assessed prior to each donation generally by a semi-quantitative, gravimetric method using a drop

Postgraduate Haematology, Seventh Edition. Edited by A Victor Hoffbrand, Douglas R Higgs, David M Keeling and Atul B Mehta.

Table 13.1 Measures to protect the donor from adverse effects of large volume blood withdrawal.

- Age 17–70 years (60 at first donation). In the UK, the upper age limit for first-time donors is 65 years, with no upper age limit for regular donors, but with annual health review after their 66th birthday.
- Weight above 50 kg
- Haemoglobin > 135 g/L for men, 125 g/L for women
- Minimum donation interval of 12 weeks (16 weeks for females)
- Pregnant and lactating women excluded because of high iron requirements
- Exclusion of those with:
  - Known cardiovascular disease e.g. ischaemic heart disease
  - Significant respiratory disorders
  - Epilepsy and other CNS disorders
  - Inflammatory bowel disease
  - Insulin-dependent diabetes
  - Chronic renal disease (CKD ≥3)
  - Ongoing medical investigation or clinical trials
- Exclusion of any donor returning to occupations such as driving bus, plane or train, heavy machine or crane operator, mining, scaffolding, etc. because delayed faint would be hazardous

of capillary blood in a copper sulfate solution, but this may be supplemented by use of portable haemoglobinometers. The EU Blood Directive states a haemoglobin standard of 134 g/L for a male donor and 124 g/L for a female donor. No more than 15% of the estimated blood volume should be taken during any one donation and, in general, 450 mL+10% of blood is collected with an interval of 12 to 16 weeks between donations. Individual components such as platelets and plasma can also be collected by apheresis using a cell separator with a maximum of 24 procedures in 12 months.

Donors with minor red cell abnormalities, such as thalassaemia trait and hereditary spherocytosis, are acceptable, providing the haemoglobin screening test excludes anaemia. Red cells from donors with sickle cell trait containing HbS may block leucodepletion filters and have limited survival under conditions of reduced oxygen tension and so should not be transfused to newborn infants and patients with sickle cell disease. Red cells from donors with glucose-6-phosphate dehydrogenase deficiency survive normally, unless the recipient is given oxidant drugs. Individuals with hereditary haemochromatosis are accepted for donation provided they meet other donor selection criteria and are under ongoing medical review.

Maintaining donor numbers and attendance is a key challenge in most countries. Much time, effort and investment, including information technology and social media, are needed to promote donor recruitment and retention to maintain adequate national blood stocks.

Table 13.2 Conditions in the donor, that lead to exclusion/deferral, in order to protect recipients (based on UK guidelines).

- All potential donors provided with information, so those at risk of HIV through lifestyle will refrain from donation (sexual practices, piercing, tattooing)
- Donors with history of hepatitis deferred until 12 months after recovery
- All potential donors who have themselves received a blood component transfusion since 1980 (due to risk of third-party vCJD transmission)
- Those who have received pituitary-derived hormones or cadaveric dura mater or corneal grafts, and those with family history of CJD
- Those whose travel history or country of origin places them at risk of malaria, Chagas disease (unless antibody test available)
- Permanent exclusion of any donor who has had filariasis, bilharzia, yaws or Q fever
- Exclusion for varying time periods following vaccinations
- Temporary exclusion after known exposure to infectious illnesses such as varicella
- Exclusion of anyone with a malignant condition except fully excised BCC of skin or carcinoma *in situ*
- Exclusion of those with diseases of unknown origin, e.g. Crohn's disease
- Donor deferral for certain drugs based on the underlying illness, e.g. insulin, malignancy
- Exclusion of those taking teratogenic drugs or those that accumulate in the tissues

BCC, basal cell carcinoma; vCJD, variant Creutzfeldt–Jakob disease

Transfusion-transmitted infection (TTI)

A number of diseases have the potential to be transmitted by transfusion of blood or its components (see Table 13.3). The epidemiology of infection in the population of a particular country can help guide the microbiological testing required. Careful donor selection criteria, with detailed questionnaires, aimed at excluding individuals at high risk of carrying infection, are supported by stringent microbiological testing (Table 13.4) with additional use of measures where appropriate to protect the recipient (see Table 13.5). Despite these measures, infection may still be transmitted, although rarely, due to the donor being in the incubation or 'window period' before testing becomes positive for infective markers or due to transmission of infections of unknown aetiology, or where no screening test available to date (e.g. prion disease).

The viruses that pose the greatest potential risk for transmission by transfusion are those that have long incubation periods (often causing subclinical infection), and especially those that

Table 13.3 Transfusion-transmissible agents.

Agents		Characteristics related to transfusion
Viruses		
Hepatotropic	HAV	Very rarely transfusion transmitted during incubation; no carrier state; faecal–oral transmission
	HEV	As above but person-to-person spread is rare
	HBV	2–6 month incubation period; carrier state; readily transmissible by blood
	HCV	Majority of cases asymptomatic; carrier state; readily transmissible by blood
Retroviruses	HIV-1 and HIV-2	Carrier state and latent in WBCs; readily transmissible by blood
	HTLV-I and HTLV-II	Latent in WBCs
Herpesviruses	CMV	50% of UK adults have been infected; latent in WBCs
	EBV	Most UK adults have been infected (therefore already exposed pretransfusion); latent in WBCs
Others	Parvovirus B19	Generally mild or asymptomatic, posing no transfusion risk except for non-immune aplastic anaemia patients and fetuses Approximately two-thirds of UK adults have been infected Seasonal variation (and epidemic years) in incidence rate
	West Nile virus	Recently exhibiting epidemic rates of transmission in summer months in North America
	Dengue virus	Transfusion transmission has been reported in Asia
Bacteria		
Endogenous	*Treponema pallidum*	Inactivated by storage at 4 °C No transfusion transmissions reported in the past 20 years
	Yersinia enterocolitica	Very occasional transmissions, usually contaminated red cells transfused late in the storage period
Exogenous	e.g. *Staphylococcus epidermidis*, *Micrococcus*, *Sarcina*	Mainly skin commensals or contaminants Most common cause of platelet contamination
Parasites		
	Malaria	Frequently transmitted in endemic countries in Africa, Asia and the Caribbean. Only five verified transfusion-transmission cases reported in UK in 25 years (all *Plasmodium falciparum*)
	Chagas disease	No transmission of *Trypanosoma cruzi* by transfusion has been reported in UK; many cases in Latin America
Prions		
	Abnormal PrP	Transfusion risk from vCJD. Only five possible cases of transmission in 18 years (three with disease)

WBCs, white blood cells

may be carried by asymptomatic individuals for many years, or even lifelong. Alternatively, some viruses that are transfusion transmissible exhibit cell-associated latency. If the virus is latent in white blood cells, recrudescence of that virus, stimulated by allogeneic transfusion, can cause infection of the recipient.

More rarely, a few viruses causing acute infection can be transmitted in the short presymptomatic viraemic phase, if this coincides with the time of blood donation. Such agents pose a risk only if recipients have not been previously infected and donors have become infected during an epidemic period (e.g. parvovirus B19) or after visiting endemic countries (e.g. hepatitis A). Bacteria, especially skin contaminants and blood-borne parasites, are also potentially transmissible by blood transfusion.

Surveillance schemes used by blood services with regular analyses of the number of infections detected in blood donors can help provide figures for the calculation of the risk of undetected infection (see Table 13.6). This also helps provide information on where and how the infection was acquired and can be used to improve donor selection guidelines. In the UK, in the period between 1995 and 2011, HIV infections detected in blood

Table 13.4 Microbial testing in England and North Wales.

HIV	ELISA (combined HIV-1 antigen plus anti-HIV-1 and anti-HIV-2); NAT for RNA
HBV	HBsAg ELISA; NAT for DNA
HCV	Anti-HCV ELISA plus NAT for RNA
HTLV	Anti-HTLV ELISA (on pools of samples)
CMV	Anti-CMV for selected immunosuppressed recipients only
Malaria	Antibody screening of potentially exposed donors
Chagas disease	Antibody screening of potentially exposed donors
Bacteria	All donations are tested for antibody to syphilis; all platelet preparations are tested by culture methods

donors were most often sexually transmitted while HBV infection was predominantly related to country of birth or parent's country of birth (with high prevalence of infection) and HCV infection was mainly associated with intravenous drug abuse.

Some groups of people may be at particular risk of blood-borne infections, e.g. intravenous drug abusers and individuals paid to have sex, who are accordingly asked never to give blood. Other individuals may be asked to wait 12 months before donating e.g. after tattooing or after sex with a high risk partner. In the UK, a major review by the Advisory Committee for the Safety of Blood, Tissues and Organs (SaBTO) resulted in a change in donor selection guidelines. While previously men who had ever had sex with other men were excluded from blood donation permanently, the revised guidance states that such individuals should be deferred from donation for 12 months after sex with another man.

Table 13.5 Measures to protect the recipient.

Measure	UK currently
Donor selection	Yes
Donor deferral/exclusion	Yes
Stringent arm cleansing	Yes
Diversion of the first 20–30 mL of blood collected	Yes
Microbiological testing of donations	Yes
Immunohaematological testing of donations	Yes
Leucodepletion of cellular blood components	Yes
Postcollection viral inactivation of FFP (e.g. MB or SD treatment)	For at-risk patients
Safest possible sources of donors for plasma-based products	For at-risk patients
Monitoring and testing for bacterial contamination	Yes
Pathogen inactivation of cellular components (e.g. psoralens)	No
Prion filtration	No

MB, methylene blue; SD, solvent-detergent

Table 13.6 Estimated risk per million blood donations of hepatitis B virus, hepatitis C virus and HIV entering the blood supply in the UK 2010–2012 (data and information collected by the NHSBT/Public Health England Epidemiology Unit).

	Hepatitis B virus	Hepatitis C virus	HIV
All donations	0.79	0.035	0.14
Donation from repeat donors	0.65	0.025	0.14
Donations from new donors	2.23	0.133	0.18

The risk estimates in the UK, 2010–2011 per unit of donated blood are:
Hepatitis B: 1 in 1.3 million; Hepatitis C: 1 in 28 million; HIV: 1 in 6.7 million
https://www.gov.uk/government/collections/bloodborne-infections-in-blood-and-tissue-donors-bibd-guidance-data-and-analysis

Hepatitis viruses

Donors with a history of hepatitis are deferred for 12 months. The hepatitis viruses can be transmitted in cellular and plasma components, as well as plasma products. The hepatitis B virus (HBV) is readily transmitted by blood or body fluids, including sexual intercourse and IV drug abuse. In endemic areas, such as the Far East and China, vertical transmission mother to child is common. The majority of patients will recover after initial acute hepatitis, but a proportion may develop a chronic carrier state, estimated at 350 million individuals worldwide with a long-term risk of cirrhosis and hepatocellular cancer.

Hepatitis B virus (HBV), when examined with the electron microscope, has three types of particle; the large (42 nm diameter) Dane particle is the actual virus with its central nucleocapsid core, which has its own antigenic constituent HBc. The core contains partially double-stranded DNA and DNA polymerase, and is surrounded by a lipoprotein coat carrying the hepatitis B surface antigen (HBsAg). The other two types of particles are 18–22 nm rods and spheres and represent overproduction of surface antigen material. HBsAg can be detected for a few months in cases of acute HBV infection, or for years in chronic carriers. The HBe antigen, a marker of high infectivity, is in soluble form and is present in the incubation period, during acute infection

and during the first years of the carrier phase. Dane particles are very rare in the plasma of low-infectivity carriers. When HBc disappears, anti-HBc appears, persisting for a long time, in the presence or absence of HBsAg. All donations are tested for the presence of HBsAg by sensitive ELISAs that can detect at least 0.2 IU/mL of HBsAg. More sensitive screening tests, such as HBV NAT, have been introduced in the UK and many countries, since they are increasingly effective in shortening the window period of infectivity. Hepatitis B remains the most commonly reported viral TTI in the UK because of window-period transmissions either during the incubation period or at the tail end of the carrier state.

Hepatitis C (HCV) is the cause of the majority of cases of what was previously known as non-A non-B (NANB) hepatitis. There are around 170 million affected individuals worldwide. Initial infection is often asymptomatic, but around 80% of patients develop a chronic carrier state with long-term risk of cirrhosis, liver failure and liver cancer. The risk of transmission by blood transfusion has fallen dramatically since the introduction of HCV antibody screening and progressively more sensitive nucleic acid testing (NAT).

Hepatitis A virus is rarely transmitted by transfusion, as there is no carrier state. Any donor who has been in close contact with a case (e.g. household contact) or developed hepatitis A is deferred for 12 months, since HAV is only transmissible during the incubation period or acute phase.

Hepatitis E virus (HEV) is spread by the oral–faecal route and can cause a self-limiting illness. It can lead to chronic infection, especially in immunocompromised patients. Although HEV is present in some UK donors, its potential transmissibility through transfusions, as well as its clinical relevance, are unknown. At present, donor screening for HEV is not advocated.

Human immunodeficiency virus (HIV-l and HIV-2)

The vast majority of the literature on AIDS refer to HIV-1; a second retrovirus capable of causing AIDS, HIV-2, mainly occurs in West Africa. HIV can be transmitted in both cellular and plasma components, Most of the patients infected by the transfusion of blood components were transfused before the introduction of screening of blood donations for HIV antibodies. Moreover, the majority of recipients of fractionated plasma products infected in the past were transfused before 1985 with unheated non-pasteurized pooled plasma products, mainly factor VIII and factor IX concentrates. Thus, HIV infection became an important sequel to transfusion of factor VIII concentrates to haemophiliacs in the early 1980s. HIV is more heat labile than HBV, especially when in solution and is also susceptible to other viral inactivation methods such as solvent-detergent treatment. On the other hand, albumin solutions, which have always been pasteurized, and intramuscular immunoglobulin (Ig) preparations, including anti-RhD Ig, which undergo a different manufacturing process, have never been associated with HIV transmission.

In large areas of sub-Saharan Africa and Southeast Asia there is a particularly high prevalence of HIV seropositivity in the general population. Inhabitants of these areas and their partners are considered to be at greater risk of HIV infection than heterosexual non-drug users in other areas of the world. Donor education and encouragement of those whose behaviour may have exposed them to HIV to exclude themselves from blood donation are highly cost-effective methods for the prevention of transmission of HIV infection by blood transfusion.

Routine screening, combined with the well-established donor education and self-deferral schemes, can reduce the already small risk of HIV transmission by transfusion in those countries where blood donors are voluntary and mostly regular.

Human T-cell leukaemia viruses

Human T-cell leukaemia viruses (HTLV-I and HTLV-II) are related retroviruses. The importance of HTLV-II is not clear; it appears to be associated with intravenous drug use in the Western world and has no known association with any clinical condition. However, HTLV-I is endemic in the Caribbean, parts of Africa and in Japan, where 3–6% of the population are seropositive. The virus is transmitted from mother to child through breast-feeding. Infection with HTLV-I is associated with at least two distinct clinical conditions; it can occasionally lead to adult T-cell leukaemia, with an incubation period of approximately 20 years, and, on even rarer occasions, to tropical spastic paraparesis (also known as HTLV-I-associated myelopathy), with a shorter incubation period. Only about 1% of patients who are seropositive develop T-cell leukaemia. Both HTLV-I and HTLV-II are cell associated and not transmitted in plasma. In highly endemic areas, transmission of HTLV-I by transfusion was relatively common before mandatory screening was introduced. Both tropical spastic paraparesis and adult T-cell leukaemia have been associated with transfusion-transmitted HTLV-I. The prevalence of anti-HTLV in previously untested UK blood donors is roughly 1 in 50,000. Screening using a sensitive ELISA for anti-HTLV-I/II on pooled plasma samples became mandatory in 2002. In the UK, the residual risk of transmission of HTLV by transfusion is of the order of 1 in 24 million donations.

Cytomegalovirus

Although most cases of post-transfusion cytomegalovirus (CMV) infection are subclinical, the syndrome of post-transfusion infectious mononucleosis-like illness is well recognized, especially after the transfusion of large amounts of blood. The infection is characterized by fever, splenomegaly and atypical lymphoid cells in the peripheral blood, with a negative Paul Bunnell test. The usually benign course of CMV infection in

most recipients and the high prevalence of CMV antibodies in the general population have meant that there has been no necessity to screen all donors for evidence of past infection. Immunosuppressed individuals are at greatest risk from potentially fatal pneumonitis or disseminated CMV infection.

In 2011, the Specialist Advisory Committee on the Safety of Blood, Tissues and Organs in the UK (SaBTO), following a review of available evidence, recommended that leucodepletion of all blood components (other than granulocytes) provided adequate CMV risk reduction for most patients requiring transfusion (haemopoietic stem cell transplant recipients, organ transplant patients and immune deficient patients, including those with HIV) without the requirement to provide them, in addition, with CMV seronegative cellular blood components. However CMV seronegative red cell and platelet components should continue to be provided for intrauterine transfusions and for neonates, as well as for pregnant women requiring repeat elective transfusions during the course of pregnancy.

Syphilis

According to WHO data, the global incidence of syphilis has increased in recent years. Each donation is tested by a serological test for syphilis. Any donation from an individual giving a positive result is discarded, and subjects with positive tests are permanently debarred from donation, even after effective therapy. In the past, syphilis testing was believed valuable as a surrogate marker for lifestyles known to be associated with high risk of HIV infection.

Malaria

Plasmodia remain viable in red cells stored at 4 °C and are readily transmissible by blood transfusion. In some endemic areas, all recipients are treated with antimalarial drugs. In non-endemic areas, there is a real risk of failure to recognize post-transfusion malaria owing to the rarity and unexpectedness of the infection. This fact, combined with increasing travel to tropical areas, necessitates the careful vetting of blood donors by direct questioning and, in some centres, by tests for malarial antibodies. In the UK, donors with a history of possible malaria exposure (provided it was more than 6 months previously and they are free of symptoms) can be accepted if they are negative by ELISA for malarial antibodies.

Other infections

Acute viral infections, although associated with short phases of viraemia, can pose transmission risk by transfusion especially if their incidence is high. This has been amply demonstrated by West Nile virus (WNV) in North America. WNV is an arthropod-borne infection of birds, with humans and horses as incidental hosts. The annual outbreaks in recent years in the USA and Canada have resulted in several transmissions by transfusion, with severe morbidity and mortality in immunosuppressed and elderly patients. The EU Blood Safety Directive (and the Blood Safety and Quality Regulations) requires that travellers from an area with ongoing transmission of WNV in humans should be deferred for 28 days. Outbreaks of WNV infection in a number of areas within Europe in 2010/11 led to the introduction of WNV NAT testing of donations in affected areas.

Arboviruses, such as dengue, though rarely, can be transmitted by transfusion. Chinkungunya virus might be transmissible.

Diseases such as Chagas disease cause significant problems for the blood transfusion services in Latin America and, due to immigration, in the USA and Spain. Potential exposure of UK donors necessitates stringent donor selection criteria and specific serological testing of those individuals to exclude the possibility of infection. *Babesia microti*, the agent of Nantucket fever, still poses a risk in certain areas of the USA. There are very few reports of possible transmission of *Leishmania donovani*, all outside Europe. Subjects giving a history of brucellosis are not accepted as donors in the UK due to the risk of transmission of this agent.

Bacteria

Bacteria from the donor arm can contaminate blood components. Rarely, low-grade or asymptomatic infections in the donor, with *Salmonella* or *Yersinia*, may be an endogenous source of contamination of blood components. Coagulase-negative staphylococci form normal skin flora are rarely pathogenic, whereas *Staphylococcus aureus*, and Gram-negative organisms, such as *E. coli*, *Klebsiella* spp. and *Pseudomonas* spp., may result in life-threatening infections. Platelet components are at particular risk due to their storage at 20–24 °C, the risk increasing with length of storage. It is estimated that up to 1 in 2000 platelet packs contain detectable bacteria 5 days after donation and in the past, fatal reactions have been reported in 1 in 25,000–80,000 transfusions. Most pathogenic bacteria grow poorly in refrigerated temperatures needed for red cell storage, but some Gram-negative organisms, such as *Yersinia enterocolitica* and *Pseudomonas* spp., can grow under these conditions.

Because haemovigilance reports showed that the risk of bacterial contamination by transfusion was greater than the risk of viral infection, UK blood services introduced enhanced methods of donor arm cleansing, and 'diversion' of the first 20 mL of the donation to reduce such risk. Routine bacterial screening of platelet preparations is also now in place in the UK, with extension of the shelf-life from 5 to 7 days.

Pathogen inactivation (PI) technologies for platelets and red cells, such as the use of light-activated psoralens that kill organisms by damaging their DNA or RNA have the potential to eliminate both bacterial and viral transfusion-transmitted infections.

While PI methods are now available, mostly for platelets, the cost-effectiveness of these techniques, together with the long-term risks, are uncertain, with the loss of viable platelets being a concern. Accordingly, although some parts of Europe have introduced PI, the potential for wide implementation in the UK is still under review.

Prions

Variant Creutzfeldt–Jakob disease (vCJD), the human form of bovine spongiform encephalopathy (BSE), is considered a potential threat to blood safety, particularly applicable to the UK, where 179 cases of vCJD have so far been reported. The number of cases of clinical disease has been falling since it peaked in the year 2000. To date, since 1996, when vCJD was first reported, there have been only four cases in the UK where blood transfusion of non-leucodepleted red cells may have been implicated in transmission of vCJD. All four cases occurred before 2007. A further possible transmission of vCJD prions was reported in February 2009 in a patient with haemophilia who had received batches of factor VIII to which a donor who subsequently developed vCJD had contributed plasma. The patient died of other causes but was found to have evidence of abnormal prion accumulation in his spleen.

There is no blood test at present for detecting prions in blood donations. The full risk of vCJD in the UK population remains uncertain and the UK blood services have taken a number of precautionary measures to reduce the potential risk of transmission of prions by blood, plasma and blood products. These include: (i) barring various individuals from blood donation, such as those who have received a blood transfusion or tissue/organ transplant since 1980 or anyone who has received human pituitary-derived hormones or grafts of human dura or ocular tissue, (ii) universal leucocyte depletion of all blood donations since 1998, (iii) discarding UK plasma and importation of plasma from countries other than the UK for fractionation into plasma products and (iv) importation of fresh frozen plasma for use in patients born after January 1996.

Laboratory tests on blood donations

Samples for laboratory testing are taken at the time of donation in order to avoid later entry into the sterile blood pack. The routine tests are automated if large numbers of donor samples are tested daily, as is the case in the UK.

All blood donors in the UK are tested at each donation for syphilis, HBsAg, anti-HIV-1 and anti-HIV-2 and HIV antigen, anti-HCV, HCV and HIV RNA, HBV DNA and anti-HTLV. Some donations are screened for CMV antibodies for patients in need of CMV-negative blood (as discussed previously). Additional testing may be needed of donors at particular risk based on travel history e.g. malaria antibodies, *T. cruzi* (Chagas) antibodies.

ABO and RhD grouping are determined routinely on each occasion. Typing for other Rh antigens (C, E, c and e) and K is now routinely performed on all blood donations in the UK. Phenotyping for other red cell antigens such as Duffy, Kidd, MNSs is performed on a restricted number of units in order to provide antigen-negative blood for alloimmunized patients and for transfusion-dependent patients likely to make multiple antibodies, such as those with sickle cell disease. Ideally, girls and women of childbearing age should be matched for c and K, as anti-c and anti-K are, after anti-D, the major causes of severe haemolytic disease of the newborn (HDN). In the UK it is now accepted practice to give K-negative blood to females of childbearing potential to minimize the risk of forming anti-K.

All donations are also screened for the presence of atypical red cell alloantibodies. The incidence of clinically significant red cell alloantibodies in blood donors is very low (<0.3%) compared with the incidence in potential recipients (1–2%). Donations with potent clinically significant alloantibodies are not issued to hospitals. There are additional requirements for blood provision for neonatal and intrauterine use, including the use of repeat, rather than new, donors, HbS-negative and exclusion of high titre anti-A and -B. Group O blood for emergency or 'flying squad' use, which may be given in an emergency to ungrouped recipients, should be given as red cells in additive solution such as saline-adenine-glucose-mannitol (SAGM) in order to avoid problems that might be caused by high-titre anti-A,B in donor plasma. As soon as the patient's group is known, group-specific blood should be given. If group O platelets have to be given to non-O recipients, donors with high-titre anti-A,B should be excluded as the providers of the plasma used to suspend the platelets; alternatively, platelet-additive solutions can be used to replace most of the plasma.

Molecular techniques are becoming increasingly available for blood group testing and extended red cell genotyping; they are likely to come into more routine use within blood services with the application of rapid automated technology.

Storage and Processing of Blood

Donor blood is collected into plastic packs containing anticoagulant and then transported without delay to the blood centre for processing. In the UK, all units undergo leucodepletion at different stages of processing, depending on the final components required (see below). Processing (see Figures 13.1 and 13.2) is undertaken to produce red cells, platelets and plasma or cryoprecipitate under stringent standards of quality control, as mandated by the EU Directives 2005.

Storage changes

Loss of red cell and platelet viability are the most important practical considerations. Progressive loss of viability varies according to the combination of anticoagulant used, how blood is stored

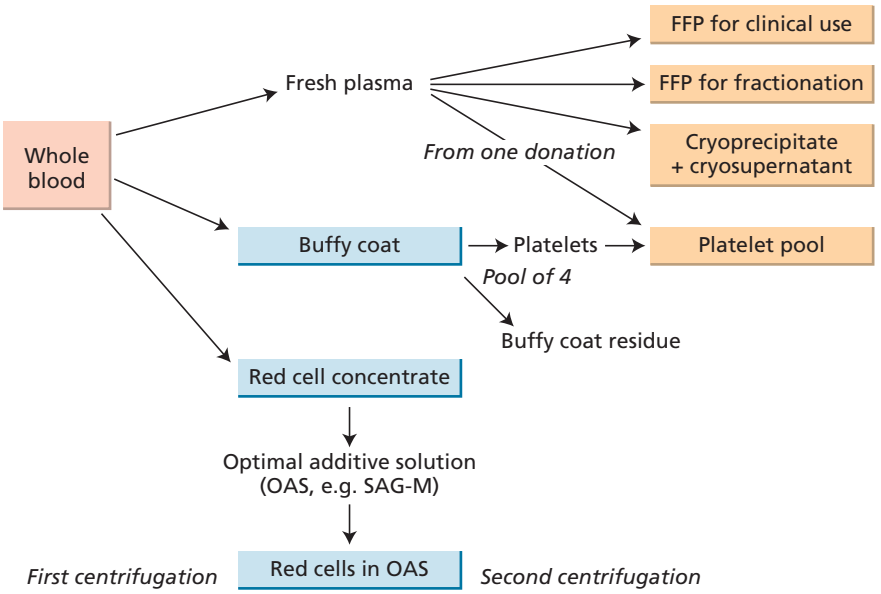


Figure 13.1 Diagrammatic representation of the preparation of components from whole blood by the 'top and bottom' or 'buffy coat' method. Items in boxes represent final components. Ideally, 200 mL of FFP to resuspend the platelet pool should be from one of the male donors of the 4 buffy coats. FFP, fresh-frozen plasma.

prior to component separation, what storage medium the components are stored in, and the pack systems used. The time limit for storage of blood or its component parts is set, taking all these into consideration. After transfusion of stored red cells or platelets, a proportion is removed from the circulation within the first 24 hours. The remainder appear to survive normally. There are various changes in the *in vitro* characteristics of red cells during storage (the so-called storage lesion), including depletion of metabolic substrates such as adenosine triphosphate (ATP) and 2,3-diphosphoglyerate (2,3-DPG), leakage of potassium and

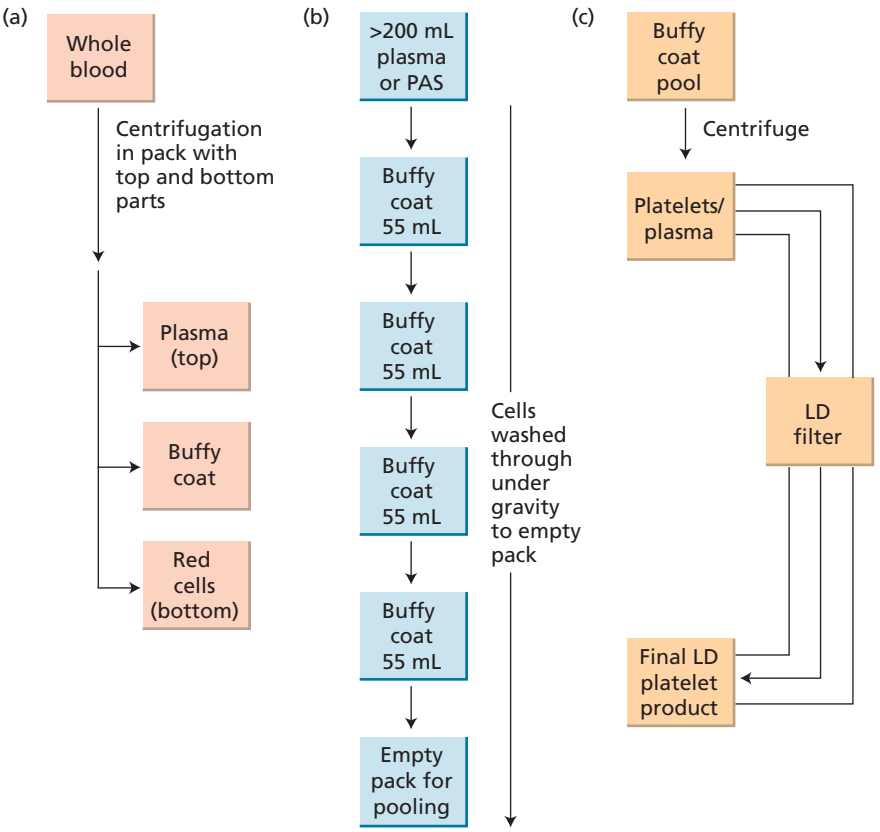


Figure 13.2 Preparation of leucodepleted pooled, buffy-coat-derived platelet preparations. After gentle centrifugation of the buffy coat pool, the platelets remain in the plasma supernatant and the buffy coat residue is discarded. PAS, platelet additive solution; LD, leucodepletion.

changes in red cell morphology. Some of these changes can be partially reversed *ex vivo* by incubation with purine nucleosides, or *in vivo*, following transfusion. ATP seems to be an important determinant of red cell viability, although not the only one. 2,3-DPG is almost completely depleted in red cells after 14–21 days storage. Red cell 2,3-DPG is restored to normal by approximately 24–48 hours after transfusion. The clinical significance of the low 2,3-DPG level of stored red cells is unknown, but only likely to be an important consideration in recipients with severe anaemia or coronary artery insufficiency.

During storage of red cells, potassium gradually leaks out through the cell membrane once active transport has been halted by the cooling of blood to 4 °C. There is rapid restoration of electrolyte levels after transfusion. The pH of blood decreases with storage, but most recipients can handle the acid load during transfusion without ill effect. Likewise, there are a number of changes in platelet function that occur during platelet storage, including increased platelet activation and decreased responsiveness to agonists such as ADP. For both red cells and platelets, the relationship between these changes *in vitro* and the function of the cells following transfusion is complex, and it is notoriously difficult to predict post-transfusion viability from the results of laboratory testing.

Anticoagulants and optimal additive solutions

When blood is stored in a liquid state there is a progressive loss of viability of the red cells, and of red cell ATP, and depletion of 2,3-DPG. The purpose of modern anticoagulants used for the collection of blood, and additive solutions, to store red cells and platelets, is to reduce these changes to a minimum.

The vast majority of whole blood donations are collected into CPD (citrate-phosphate-dextrose) anticoagulant. Platelets from apheresis are collected into acid citrate dextrose. Optimal additive solutions such as SAGM containing saline, adenine, glucose and mannitol have been developed to improve the viability of plasma-depleted red cells on storage, and this is now the most usual preservative solution for red cells in the UK. Due to theoretical concerns of adenine and mannitol toxicity, CPD continues to be used for the relatively large volume of red cells used for neonatal exchange transfusion and for intrauterine transfusion, with SAGM red cells used for top-up neonatal transfusions.

Leucodepletion

Many developed countries have introduced universal leucocyte depletion of blood components. In the UK, leucodepletion of all components, to a level below 5×10^6 white cells per unit, has been mandatory since 1999, as a measure which might decrease the risk of vCJD transmission by transfusion. Added benefits of leucodepletion include a significant decrease in the incidence of febrile non-haemolytic transfusion reactions, in the risk of human leucocyte antigen (HLA) alloimmunization and in the risk of transmission of intracellular viruses such as CMV and HTLV. Leucodepletion filters can be used at the initial stage on whole blood, leaving leucodepleted red cells and plasma, or later on in the processing, for the preparation of platelet concentrates from pooled buffy coats, when the final components, red cells, platelets and plasma must be filtered individually (see Figures 13.1 and 13.2). These two alternative methods are necessary because whole blood filters remove a significant number of platelets, limiting production of platelet concentrates from filtered whole blood.

Red cells

There are no specific clinical indications for the transfusion of whole blood; therefore, the vast majority of donor blood is separated into its components. A multiple 'bottom and top' blood collection pack is used (Figure 13.1) if the donation will be used to make platelets; otherwise, a simpler 'top and top' system is used to make red cells and plasma. In the bottom and top system, after centrifugation, whole blood collected in CPD is separated into plasma at the top, a buffy coat in the middle and red cells at the bottom (Figure 13.1). Fresh plasma is expressed from the top and red cells from the bottom of the pack, leaving the buffy coat in the original pack. The red cells are transferred to a pack containing an optimal additive solution to preserve red cell function during storage, and are also filtered. The resulting red cells have a shelf-life of 35–42 days (35 days in the UK). Adult red cells in SAGM in the UK have a volume of 220–340 mL with haematocrit of 0.5–0.7 and a residual plasma volume of 5–30 mL.

The specification for red cells for intrauterine (IUT) neonatal transfusion is shown in Table 13.7.

Washed red cells

Patients with recurrent or severe allergic or febrile reactions to red cells and especially those with severe IgA-deficiency with anti-IgA may need washed red cells. These are produced either manually with a 24-hour shelf life or in a closed, automated system in which the red cells are washed to remove most of the plasma and then re-suspended in 100 mL SAG-M with a shelf life of 14 days after washing.

Frozen red cells

Red cells can be stored for a prolonged period (up to 10 years) without damage if glycerol is added before freezing. Thawed red cells must be washed free of glycerol before transfusion. This method of storage is expensive and time-consuming, but is invaluable as a means of storing red cells with rare phenotypes (e.g. Bombay blood) and national banks of frozen rare cells have been established for this purpose (see Chapter 12).

Platelet preparations

Platelets are produced from whole blood (Figure 13.1) or by apheresis. Following the processing of whole blood, the buffy

Table 13.7 Specifications of red cells for intra uterine transfusion and neonatal transfusion (adapted from UK guidelines).

Red cell components for IUT
Plasma reduced (haematocrit 0.7–0.85)
In citrate phosphate dextrose (CPD) anticoagulant (theoretical risk of toxicity from additive solutions)
Leucocyte-depleted
Less than 5 days old (to avoid hyperkalaemia)
Cytomegalovirus (CMV)-antibody negative
Sickle screen negative
Irradiated to prevent TA-GvHD (shelf life 24 hours)
Usually group O with low-titre haemolysins
RhD- and K-negative and red cell antigen-negative for maternal alloantibodies
Compatible with the mother's plasma by an IAT cross-match
Red cells for neonatal exchange transfusion
Plasma reduced with haematocrit of 0.5–0.6 to reduce the risk of post-exchange polycythaemia
In CPD anticoagulant
Less than 5 days old
Irradiated (essential if previous IUT)
CMV-antibody negative
Sickle screen negative
Usually produced as group O (with low-titre haemolysins)
RhD negative (or RhD identical with neonate) and Kell negative
Red cell antigen negative for maternal alloantibodies
Compatible with the maternal plasma by IAT cross-match
Red cells for small-volume transfusion of neonates and infants
Haematocrit 0.5–0.7
In SAG-M anticoagulant/additive solution (approximately 20 mL residual plasma)
Up to 35 days from donation
Group O (or ABO-compatible with baby and mother) and RhD negative (or RhD compatible with the neonate)
CMV negative

coat layer contains most of the platelets, over 80% of the white cells and 5–10% of the red cells. This is pooled with three other buffy coats and more than 200 mL of plasma from only one of the donations, preferably from a male donor (in order to avoid the risk of transfusion of donor white cell antibodies that might cause transfusion-related acute lung injury), using a sterile connecting device (Figure 13.2). The pool undergoes a second gentle centrifugation step and the supernatant is used to produce a platelet preparation, containing, on average, 3×10^{11} platelets. The residue from the buffy coat pool, containing mainly white cells and red cells, is then discarded. The systems that enable the separation of buffy coats in a semi-automated manner are called blood separators, e.g. Optipress.

Unlike red cells, platelets must be stored at 20–22 °C, since storage at 4 °C results in poor survival following transfusion. Platelets have a shelf-life of 5 days, and are stored in permeable bags that allow the diffusion of oxygen into the pack, which, with constant gentle agitation, maintains aerobic metabolism and reduces the rate of fall of pH. The shelf-life of platelet preparations can be extended to 7 days provided that systems are in place to monitor bacterial contamination (now used by the UK blood services) or to inactivate pathogens (not widely used at present).

The equivalent of two or even three adult doses of platelets (minimum 2.4×10^{11} each) may be obtained from one donor, with adequate platelet counts, by an apheresis procedure lasting approximately 90 min. A large HLA-typed donor panel is available in England to provide HLA-matched platelets for immunologically refractory patients. Most apheresis platelet donors are also typed for human platelet antigens (HPA), in order to provide HPA-1a-negative platelets for newborns affected with neonatal alloimmune thrombocytopenia and also for a few HPA-1a-negative patients who have become immunologically refractory to platelet transfusions.

Platelets in additive solution

This component is indicated for patients with recurrent severe allergic or febrile reactions to standard platelet transfusions. After 'washing' to remove most of the plasma, the platelets are re-suspended in 200 mL of platelet additive solution (PAS) but may still contain around 10 mL residual plasma.

Granulocyte concentrates

With the increase in intensive chemotherapy regimens, demand for granulocytes has risen in order to prevent or treat fungal or other serious, antibiotic-resistant infections in severely neutropenic patients.

Granulocytes are extremely labile; they must be separated from whole blood immediately after collection and transfused within a maximum of 24 hours of preparation. Granulocytes prepared from routine blood donations (buffy coats) are heavily contaminated with red cells and platelets. Buffy coats from at least 10 donors are required to produce a therapeutic dose for an adult (at least 1×10^{10} granulocytes). Ten buffy coats also contain the equivalent of two units of red cells and 2.5 pools of platelets. A new pooled granulocyte component produced from whole blood donations that has reduced volume and red cell contamination is now available in the UK. Due to the overall high red cell content, granulocyte preparations should be ABO and RhD compatible with the recipient and cross-matched by indirect antiglobulin test (IAT) against the recipient's serum. Prospective recipients of granulocyte concentrates should be screened for HLA antibodies; if present, they may cause severe transfusion reactions, with respiratory distress; All granulocyte concentrates must be irradiated prior to transfusion.

Administration of granulocyte colony-stimulating factor (GCSF) or steroids can produce a higher yield of granulocytes collected by apheresis, but this is not permitted for non-directed volunteer donors in the UK. A directed donation of granulocytes by apheresis from a suitable relative may sometimes be possible.

Fresh-frozen plasma (FFP)

Plasma that has been separated from whole blood (Figure 13.1) or obtained by apheresis is frozen within 24 hours to a temperature that will maintain the activity of the labile factors V and VIII. FFP contains all coagulation factors and should be stored at −30 °C or below for up to 24 months or even longer. When needed, the plasma is thawed rapidly at 37 °C and then transfused without delay at a dose of 15 mL/kg in adults (see below).

Single units of FFP can be treated with photosensitizing chemicals such as methylene blue (MB), riboflavin or amotosalen and exposed to visible or ultraviolet light to inactivate pathogens that may be present. If required, the residual chemicals can be removed before the plasma is rapidly frozen to −30 °C. Alternatively, FFP can be pooled with around 1500 other units and treated with solvent-detergent (SD) in order to inactivate pathogens. All methods of viral inactivation reduce the levels of labile clotting factors in FFP. In the UK, FFP for use in patients born after 1996 is sourced from countries with a low risk of vCJD and is pathogen inactivated (methylene blue or solvent detergent). Similarly, as a precautionary measure, UK plasma is discarded and not used for fractionation for the manufacture of immunoglobulins, albumin and coagulation factor concentrates.

In recent years in the UK, every effort is made to prepare FFP only from male donors in order to minimize the risk of passive transfer of donor white cell antibodies that can cause transfusion related acute lung injury (TRALI). Similarly, pooled platelets are, whenever possible, suspended in male donor plasma.

Cryoprecipitate

Cryoprecipitate is prepared from FFP that is allowed to thaw slowly (classically at 4 °C overnight). After removal of the supernatant, factor VIII: C, von Willebrand factor (VWF), fibrinogen, fibronectin and factor XIII are left as a precipitate, which is then refrozen in approximately 30 mL of plasma and stored at −30 °C or below for up to 24 months. Each unit should contain a minimum of 70 IU of factor VIII: C and 140 mg of fibrinogen. Cryoprecipitate is now used mainly as a source of fibrinogen in cases of disseminated intravascular coagulation (DIC), hepatic failure and hypofibrinogenaemia. A standard adult dose of cryoprecipitate is 10 units, which are thawed at 37 °C in about 10 min and should be used immediately. Pools of five single units are available in the UK. Fibrinogen concentrate is now available in the UK but only licensed for use in congenital hypofibrinogenaemia.

Irradiated blood components

Patients at risk of transfusion-associated graft-versus-host disease (TA-GvHD, see further on) require irradiation of cellular blood components (red cells, platelets and granulocytes). For adults, red cells can be irradiated using gamma or X-rays within 14 days of donation, with a shelf-life of 14 days from irradiation. Red cell components for exchange neonatal transfusion and intrauterine transfusion (IUT) should be irradiated within 5 days of donation with a postirradiation shelf-life of 24 hours. Platelets should also be irradiated to prevent TA-GvHD in susceptible patients and they retain their normal shelf life post irradiation. Frozen plasma components such as FFP and cryoprecipitate do not need irradiation.

Clinical and laboratory transfusion practice

In the UK, a series of three 'Better Blood Transfusion' Health Service circulars published in 1998, 2002 and 2007 promote safe transfusion practices within hospitals, with emphasis on the appropriate use of blood and components in all clinical areas. Many of the principles of 'Better Blood Transfusion' are now encompassed in Patient Blood Management (PBM), an evidence-based, comprehensive multidisciplinary approach to the management of patients needing transfusion (see below).

All hospitals should have hospital transfusion committees (HTCs), with multidisciplinary representation from the clinical, laboratory and management areas. These committees are responsible for overseeing implementation of guidelines, and for the audit and training of all staff involved in transfusion. The HTC is essential within the hospital clinical governance and must be accountable to the chief executive. The hospital transfusion team (HTT), composed of transfusion specialists with a nursing or biomedical background, the transfusion laboratory manager and the consultant haematologist specialist in transfusion undertakes various activities on a day-to-day basis to achieve the objectives of the HTC.

EU Blood Directives (UK Blood Safety & Quality Regulations 2005): impact on hospital transfusion practice

The EU Blood Directives not only set standards for quality within the blood services, but also have had a far-reaching impact on hospital transfusion practice. In the UK, the chief executive of each hospital with a transfusion laboratory needs to submit a formal annual statement of compliance to the Medicines and Healthcare products Regulatory Agency (MHRA) as the competent authority implementing the Blood

Safety and Quality Regulations 2005. Hospital transfusion laboratories can be inspected by the MHRA and in the event of significant deficiencies be given the order to 'cease and desist from activities'. The key requirements for hospital transfusion laboratories include:

- a comprehensive quality management system based on the principles of 'good practice', including stringent requirements for storage and distribution of blood and components, with emphasis on 'cold chain' management;
- traceability, requiring all hospitals to trace the fate of each unit of blood/blood component (including name and patient ID) with records kept for 30 years;
- education and training of staff involved in blood transfusion, with maintenance of all training records;
- haemovigilance, with reporting of all adverse events of transfusion.

Hospital transfusion laboratories undertaking any processing activities, such as irradiation, must have a license from the MHRA indicating they have a blood establishment status. NHSBT blood centres have such licenses.

Laboratory tests in patients

Pretransfusion group and screen

The ABO and RhD groups of all potential recipients should be determined before transfusion, due to the universal presence of ABO antibodies in all subjects lacking the corresponding antigens and to the high immunogenicity of RhD (see Chapter 12).

The patient's red cells are grouped for ABO using monoclonal anti-A and anti-B. A 'reverse group' is carried out on the patient's plasma to demonstrate the presence of the corresponding antibody, using A and B reagent cells. RhD typing is undertaken using 2 IgM monoclonal antiD reagents (see Chapter 12).

As described in Chapter 12, the patient's serum should also be screened for the presence of atypical red cell antibodies using indirect antiglobulin test (IAT). If a positive result is obtained, further investigation using a red cell panel of 8–10 cells is required to identify the antibody or antibodies and eventually provide antigen-negative cells for cross-matching.

The majority of hospital laboratories in the UK carry out blood grouping and antibody screening using automated analysers with computer control of specimen identification and electronic transfer of results. This is safer than manual grouping techniques with manual transcription of results.

Since errors related to having the 'wrong blood in tube' are relatively common, with the potential risk of ABO mismatched transfusions, the current British Committee for Standards in Haematology (BCSH) guideline recommends that a second sample should be requested for confirmation of the blood group of any new patient, provided this does not impede the delivery of urgent red cells or components.

For most patients without clinically significant antibodies in their plasma, ABO- and RhD-compatible red cells are provided if transfusion is required. If the patient has clinically significant antibodies, the red cells provided must also be negative for the relevant red cell antigen(s), and cross-matched before transfusion (see Chapter 12). Women of childbearing age should also receive K-negative units to avoid the risk of forming anti-K.

The BCSH Compatibility Guidelines 2012 provide pragmatic recommendations on the timing of pretransfusion samples; for the majority of patients this is 7 days, but testing should be performed using blood collected no more than 3 days in advance of the transfusion when the patient has been transfused or pregnant within the preceding 3 months. An extension to 7 days may be considered for regularly transfused patients and pregnant women with no significant alloantibodies who need blood on standby for a potential obstetric emergency such as placenta praevia.

Compatibility testing (cross-match)

Traditionally, a cross-match, by IAT at 37 °C, between the patient's plasma and red cells from the donor unit has been the final step before blood is issued for transfusion. This still applies in specific cases (see below), but now there is no need for serological cross-match in many patients with no red cell alloantibodies, where the relevant laboratory requirements for safe electronic issue are in place.

Electronic cross-match

ABO- and RhD-compatible red cells can be provided using electronic issue (or 'computer cross-match') provided the patient does not have any alloantibodies and there are robust automated systems in place for blood grouping and antibody testing, interfaced with laboratory information systems to minimize risk of human error. If a patient has clinically significant red cell antibodies electronic issue should not be used and an IAT cross-match is required with antigen-negative red cell units. Electronic issue is also not suitable in certain patient groups, e.g. those with a history of an ABO-incompatible stem cell transplant or those who have had an ABO-incompatible solid organ transplant in the preceding 3 months.

In hospitals where electronic issue is not routinely used, a maximum (or standard) surgical blood ordering schedule (MSBOS, SBOS) can reduce the number of compatibility tests performed and avoid the reservation of red cell units unlikely to be transfused. The MSBOS is based on a retrospective comparison of the number of units of blood cross-matched and the number actually transfused, for each elective surgical procedure, at a given hospital. It should be agreed with the transfusion laboratory, surgeons and anaesthetists and ratified by the hospital transfusion committee. Procedures that are likely to require blood transfusion have a ratio of crossmatched:transfused below

2:1. Regardless of the MSBOS, if atypical antibodies are detected, antigen-negative blood should be cross-matched and reserved before surgery.

In minor incompatibility, passenger lymphocytes from the donor, may engraft in the recipient and produce anti-AB capable of causing severe haemolysis in the recipient (see further on).

Special requirements for the selection of blood

Certain patient groups have special requirements regarding the selection of blood and components for transfusion. Clinical teams should be made aware of these requirements through the use of guidelines supported by education. Appropriate alerts for special requirements should be used within the transfusion laboratory. Serious Hazards of Transfusion (SHOT) haemovigilance has highlighted the relative frequency of errors in this area of transfusion practice. The special requirements for intrauterine and neonatal transfusion are as highlighted in Table 13.7.

Haemoglobinopathy

Patients with haemoglobinopathy, in particular those with sickle cell disease (SCD), are at a high risk of developing red cell alloantibodies, with reports of 20–30% in the literature (see also Chapters 6 and 7). Accordingly, from the start, patients with haemoglobinopathy should receive blood that has 'extended matching' for K and Rh antigens in addition to D, i.e. Cc Ee, in order to reduce the risk of alloimunization. Blood for transfusion to patients with SCD should also be HbS negative. Blood for top-up transfusion in haemoglobinopathy patients should be <2 weeks old and <7 days old if used for exchange transfusion in patients with SCD.

In addition, the red cells of patients with SCD should be phenotyped extensively prior to transfusion for antigens apart from Rh and K such as for Kidd, Duffy and Ss antigens. Where patients have already been transfused, the genotype can be determined by molecular techniques. This information can in particular assist with investigation of haemolytic transfusion reactions and expedite issue of blood in cases with complex serology.

Haemato-oncology

Irradiation of cellular blood components (red cells, platelets, granulocytes) is required for certain groups of haematology patients, as highlighted in Table 13.8. Patients should be informed of the need for irradiated blood, supported by written information and cards flagging up their special requirements. While CMV infection can be transmitted by cellular blood components and may produce fatal infection in immunocompromised patients, the risk can be reduced by leucodepletion (see previously). The need for further selection of CMV-negative components is discussed above.

Patients undergoing stem cell transplantation may have a mismatch of ABO groups with their donor (e.g. up to 25% of HLA-identical sibling transplants) with a risk of haemolysis. There can be major ABO incompatibility where the recipient's plasma contains anti-A, anti-B or anti-A,B incompatible with donor red cells (e.g. group A donor and group O recipient), minor ABO incompatibility where the donor's plasma contains anti-A, anti-B or anti-A,B that can react with the recipient's red cells (e.g. donor group O and recipient group A) or bidirectional ABO incompatibility where both the donor and recipient's plasma contain anti-A, anti-B or anti-A,B reactive with recipient and donor red cells respectively (e.g. donor group A and recipient group B). Clear communication is essential between the clinical and laboratory teams, with an agreed post-transplant transfusion policy (see Table 13.9).

Antenatal testing

All women should have blood samples taken at antenatal booking for blood grouping and antibody screening. If no clinically significant antibody is detected, further samples should be tested at 28 weeks' gestation, before routine antenatal prophylaxis is given to RhD-negative women (see below).

Neonatal 'top-up' transfusion

Premature infants are among the most widely transfused patients, with 'top-up' transfusions being very frequent (Chapter 50). During the first 4 months of life, ABO antigens may be poorly expressed on red cells and the corresponding ABO antibodies may not have yet developed so that results of 'reverse grouping' will be unreliable. Passive transfer of maternal IgG ABO antibodies may result in the detection of these antibodies in neonatal plasma. Samples from both the mother and infant should be tested for ABO and RhD grouping wherever possible, with an antibody screen performed on the maternal sample. If there are no atypical maternal antibodies, top-up transfusions can be given without further testing during the first 4 months of life.

Ideally, the unit of red cells used for the first transfusion should be aliquoted into several (six to eight) satellite (or 'paedi') packs and used for the same infant until expiry in order to decrease exposure to multiple donors. Measures should be in place in all neonatal units and their supporting laboratories to minimize the quantities of blood required for testing, by use of microsampling techniques and near-patient testing, since frequent blood sampling significantly contributes to anaemia in these infants. Erythropoietin has been extensively studied as an alternative to transfusion in the anaemia of prematurity, but it does not consistently reduce the need for transfusion in these patients.

Table 13.8 Indications for irradiated cellular blood components[a] in haemato-oncology patients. (Adapted with permission from the Handbook of Transfusion Medicine, based on BCSH guidelines 2010.)

Patient group	Irradiated blood components
Adults or children with acute leukaemia	Not required (except for HLA-matched platelets or donations from first- or second-degree relatives)
Recipients of allogeneic (donor) HSC transplantation	From the start of conditioning chemo-radiotherapy, continue while receiving GvHD prophylaxis If chronic GvHD or on immunosuppressive treatment, continue irradiated blood components
Bone marrow and peripheral blood stem cell donors	Provide irradiated cellular components during and for 7 days before the harvest
Bone marrow or peripheral blood HSC harvesting for future autologous re-infusion	Provide irradiated cellular components during and for 7 days before the harvest
Autologous HSC transplant patients	From start of conditioning chemoradiotherapy until 3 months post-transplant (6 months if total body irradiation was used)
Adults and children with Hodgkin lymphoma at any stage of the disease	Irradiated cellular components indefinitely
Patients treated with purine analogues (fludarabine, cladribine and deoxycoformicin)[b]	Irradiated cellular components indefinitely
Patients treated with alemtuzumab (anti-CD52) therapy[c]	Irradiated cellular components indefinitely

[a]Red cells, platelets and granulocytes
[b]Irradiated components are recommended for newer purine analogues and related compounds, such as bendamustine, until further data are available
[c]Irradiated components are also recommended for solid organ transplant patients receiving alemtuzumab

Transfusion in autoimmune haemolytic anaemia

Patients with autoimmune haemolytic anaemia pose a significant challenge (see also Chapter 9). While unnecessary transfusion should be avoided, this must not be withheld in the context of brisk haemolysis, in particular where there is risk of cardiac compromise. Close liaison between clinical teams, the transfusion laboratory and the specialist immunohaematology team within blood services is essential to assist with prompt

Table 13.9 Selection of ABO blood group of components transfused in the early* post-transplant period. (Adapted with permission from Handbook of Transfusion Medicine 2014.)

	Donor	Recipient	Red cells	Platelets	FFP
Major ABO incompatibility	A	O	O	A	A
	B	O	O	B	B
	AB	O	O	A	AB
	AB	A	A/O	A	AB
	AB	B	B/O	B	AB
Minor ABO incompatibility	O	A	O	A	A
	O	B	O	B	B
	O	AB	O	A	AB
	A	AB	A/O	A	AB
	B	AB	B/O	B	AB
Bidirectional ABO incompatibility	A	B	O	B	AB
	B	A	O	A	AB

*Once conversion to donor blood group is complete, components of the new ABO group can be given

investigation and issue of blood. These patients often have panreacting autoantibodies that can mask an underlying alloantibody with the need for investigation with serological techniques such as elution and absorption. Blood should then be selected that is ABO compatible, matched with the patient's own full Rh and K type and negative for the antigens to which alloantibodies have been identified, if any. Avoidance of the specificity of the autoantibody (e.g. auto-anti-e in RhD-negative female patients) is not routinely recommended since this may expose the patients to antigens provoking the formation of clinically significant alloantibodies. Transfused red cells are likely to be destroyed by the autoantibody at the same rate as the patient's own blood. Hence, transfusion, if needed in AIHA, will serve only as a temporary 'holding measure' and should be accompanied with effective immunosuppressive therapy.

Safe administration of blood

The principles underpinning the safe bedside transfusion practice should be encapsulated in a hospital transfusion policy covering aspects of requesting, collecting and administration of blood components, as well as the monitoring of the patient during transfusion.

Patient identification

Positive patient identification is essential at all stages of the transfusion process, including blood sampling, collection of blood from storage, delivery to the clinical area and administration. Patients should have an identification band with the minimum identifiers, including their last name, first name, date of birth and unique identification number (e.g. hospital number).

'Wrong blood in tube' (WBIT) errors can result in death. These occur due to human error either when the blood in the tube is from the wrong patient and is labelled with the intended patient's details, or blood is taken from the intended patient, but the tube is labelled with the wrong details. It is estimated that this occurs in 1 in every 2000 samples taken. The Serious Hazards of Transfusion (SHOT) UK Haemovigilance scheme, continues to report an increasing number of WBIT errors with 643 cases in 2013; these are mostly 'near miss' events which could have resulted in the wrong blood being transfused. SHOT estimates that approximately 1 incorrect blood component is transfused for every 100 near miss events.

Detection and approach to transfusion reactions

All patients should be transfused in clinical areas where they can be directly observed to allow prompt detection and immediate management of acute reactions. Baseline observations and regular monitoring of patients undergoing transfusion, are essential so that reactions are detected and promptly managed. Patients with severe transfusion reactions can present suddenly with cardiovascular collapse, but the underlying cause may not be immediately apparent; these can include bacterial contamination, acute haemolytic transfusion reactions (AHTRs, usually due to ABO-incompatible transfusion), anaphylaxis, TRALI and transfusion-associated circulatory overload (TACO). If the presumed transfusion reaction is severe or life-threatening the transfusion must be discontinued, with immediate assessment and the following steps taken:

1 stop and disconnect the blood pack and giving set immediately (but do not discard);
2 maintain venous access with physiological saline, commence resuscitation, if needed;
3 take samples for examination of the blood for colour and smell, full blood count, renal and liver function tests, blood cultures, coagulation screen, repeat compatibility testing, DAT, LDH and assessment of urine for haemoglobin;
4 check identification details for the patient, their identity band and the blood component/s concerned;
5 consider key/additional features, e.g. fever and shock without anaphylaxis – consider ABO-mismatched transfusion or bacterial sepsis; dyspnoea – consider TRALI or TACO – check O_2 sats or blood gases and CXR; laryngeal oedema with dyspnoea/stridor – consider anaphylaxis or severe allergic reaction;
6 seek early support and advice from the critical care team.

Less severe acute transfusion reactions include haemolytic transfusion reactions due to other red cell antibodies in the recipient, febrile non-haemolytic transfusion reactions (FNHTR) and allergic reactions. While the latter two can often be managed with an antipyretic or antihistamine, respectively, ongoing close monitoring of the patient is essential, since these could be the early signs of a more severe transfusion reaction.

Complications of blood transfusion

The complications of blood transfusion can be conveniently divided into acute and delayed, immunological and non-immunological categories (Table 13.10).

Immunological complications

Sensitization to red cell antigens

As only the ABO and RhD antigens are routinely matched for the selection of blood, sensitization can occur to other red cell antigens following transfusion. Athough the D antigen is the most immunogenic, other antibodies can form with alloimmunization being more likely in multitransfused patients. Red cell antibodies can also develop following pregnancy. The consequences of sensitization may be negligible, but can lead to difficulty with compatibility testing, haemolytic disease of the fetus and

Table 13.10 Hazards of transfusion.

Immediate (hours)
Non-immune complications
Bacterial: acute sepsis or endotoxic shock
Hypothermia
Hypocalcaemia (↓ Ca^{2+}) in infants and massive transfusion
Transfusion-associated circulatory overload (TACO)
Immune complications
Febrile non-haemolytic transfusion reactions
Acute haemolytic transfusion reactions: intravascular (IgM), extravascular (IgG)
Allergic reactions (urticarial)
Anaphylactic reactions (anti-IgA)
Transfusion-related acute lung injury (TRALI)
Delayed (days to years)
Non-immune complications
HIV, HCV, HBV, CMV
Others: parvovirus B19, HAV, HEV, WNV, dengue, malaria, Chagas disease, brucellosis, syphilis, vCJD
Immune complications
Delayed haemolytic transfusion reactions (due to anamnestic immune responses with red cell alloantibodies)
Post-transfusion purpura
Transfusion-associated graft-versus-host disease

CMV, cytomegalovirus; HAV, hepatitis A virus; HEV, hepatitis E virus; WNV, West Nile virus.

Table 13.11 Antibodies associated with haemolytic transfusion reactions.

Blood group system	Antibodies implicated in intravascular haemolysis	Antibodies implicated in extravascular haemolysis
ABO	-A,B; -A; -B	
Hh	-H (Bombay)	
Rh		All
Kell		K, k, Kp^a, Kp^b, Js^a, Js^b
Kidd		Jk^a, Jk^b, Jk^3
Duffy		Fy^a, Fy^b, Fy^3
MNS		M, S, s, U (some)
Lutheran		Lu^b (some)
Lewis	Le^a, Le^b, Le^{a+b}	
Cartwright		Yt^a (some)
Vel	Vel	Vel (some)
Colton		Co^a, Co^b
Dombrock		Do^a, Do^b

newborn (HDFN) and haemolytic transfusion reactions (HTRs) in the future.

Haemolytic transfusion reactions

This is premature destruction of transfused red cells reacting with antibodies in the recipient. Naturally occurring antibodies, such as ABO antibodies, are mainly IgM and, if warm-reacting, can destroy red cells *in vivo* by complement fixation. Red cell alloantibodies in response to exposure through previous transfusions or pregnancies are IgG antibodies. HTRs may occur immediately after the transfusion or may be delayed for anything up to 2–3 weeks.

Immediate or acute haemolytic transfusion reactions

These may be intravascular or extravascular (Table 13.11).

Immediate intravascular red cell destruction is the most dangerous type of HTR; it is associated with activation of the full complement cascade by IgM antibodies and is practically always due to ABO-incompatible blood transfusions (haemolytic anti-A,B, anti-A or anti-B present mainly in the recipient or, rarely, in the donor plasma). Most of these ABO-incompatible transfusions are due to errors, and occur with an approximate frequency of 1 in 180,000 patients transfused. Major morbidity with patients needing intensive care or renal dialysis occurs in 20–25% of all cases with a mortality rate of 5–10%. In recipients (mostly group O) of ABO-incompatible blood with serious reactions, the symptoms are usually dramatic and severe; most are due to the anaphylatoxins C3a and C5a that are liberated during complement activation, releasing vasoactive amines and hydrolases from mast cells and granulocytes (see Chapter 12). The cytokines interleukin (IL)-1, IL-8 and tumour necrosis factor (TNF) also play an important role, causing inflammation, smooth muscle contraction, platelet aggregation, increased capillary permeability, and hypotension. Typically, within less than 1 hour of the start of the transfusion, when the reaction is symptomatic, the patient complains of heat or pain in the cannulated vein, throbbing in the head, flushing of the face, chest tightness, nausea and lumbar pain. These symptoms are usually accompanied by tachycardia and hypotension. In severe cases, there is profound hypotension, leading to shock and renal failure. Rigors and pyrexia usually follow.

Intravascular destruction of red cells also brings about liberation of thromboplastin-like substances that activate coagulation and lead to DIC. The bleeding diathesis and increased destruction of red cells (which may eventually involve the recipient's cells) further exacerbates the problem. Intravascular destruction of red cells liberates haemoglobin into the circulation. Once haptoglobins are saturated, haemoglobin will also appear in the urine. If haemoglobinuria is very severe, haemosiderinuria may be seen. Renal complications consist of acute renal failure with oliguria and anuria, possibly the result of hypotension and/or

the action of activated complement. The initial symptoms may of course be modified or abolished in anaesthetized or heavily sedated patients, in whom evidence of DIC, hypotension or the presence of haemoglobinuria may be the first signs.

Immediate intravascular destruction of recipient red cells should be avoidable. In practice, the main cause is error, when the incorrect blood component is transfused. The most severe reactions occur in major incompatibility, when a group O recipient with high-titre anti-A,B is transfused with group A, B or AB red cells. Less severe intravascular haemolysis occurs when group A red cells are transfused to a group B recipient, or vice versa, because group B and A subjects have less potent ABO antibodies than group O subjects. More rarely, intravascular red cell destruction may occur when group O plasma is transfused by mistake to A, B or AB recipients. For this reason, group O red cells, even in additive solution, should not routinely be used for non-O recipients; furthermore, this practice leads to unnecessary shortages of group O blood. If unavoidable, the group O blood must first be screened for the presence of high-titre haemolysins or should have very small volumes of plasma. ABO-compatible cryoprecipitate, FFP and platelet transfusions should be selected for all recipients, especially for children because of their smaller blood volume. Group A, B or AB plasma components are safe for group O recipients. There is also the possibility of a laboratory error when tubes are transposed or there are transcription errors or misrecording of results, or there is insufficient time to complete an antibody screen or compatibility test.

In the UK haemovigilance (SHOT) system, approximately 30% of cases of incorrect blood component transfusion reported in the last 12 years have been primarily due to clerical or technical errors that originated in the hospital laboratory. Obviously, not all the errors led to ABO-incompatible transfusions, in fact 10% of errors led to ABO incompatibilities. The remainder of reports (70%) relate to clerical or administrative errors in the ward, collection of the blood from the blood bank, failure to confirm the identity of the patient when taking samples, mislabelling of the sample of blood or failure to perform proper checks before removing the units from the refrigerator or transfusing the blood. The potentially serious consequences of failures resulting in ABO-incompatible transfusions emphasize the need for set protocols with meticulous checking at all stages. If an identification mistake has been made, it is important to check, as a matter of urgency, that the units intended for the patient under investigation have not also been misdirected to another recipient. Such events are results of errors that need further investigation with a root cause analysis so that corrective action can be taken to minimize recurrence.

Extravascular red cell destruction is mediated by IgG antibodies (Figure 13.3). Mononuclear phagocytic cells have receptors for the Fc fragment of IgG1 and IgG3; the binding of IgG-coated cells to these receptors is inhibited by free IgG in plasma. There are no receptors for IgM on macrophages. Red cells sensitized with IgG1 and/or IgG3 antibodies may or may not activate complement up to C3b only. If they do not, they are removed extravascularly (by phagocytosis or cytotoxicity) by mononuclear phagocytic cells, predominantly in the red pulp of the spleen, where the plasma is largely excluded and the IgG on the red cells can compete with free IgG in the plasma. However, cells coated with IgG antibodies, which activate complement up to C3b, adhere to the C3b receptor on macrophages and monocytes. The presence of C3b on red cells greatly enhances the extravascular destruction of IgG-coated cells. This is because the binding to C3b receptors is not inhibited as there is no native C3b in plasma and consequently IgG/C3b-coated cells are destroyed by phagocytosis or cytotoxicity, predominantly in the liver, where there are abundant macrophages (Kupffer cells) and a generous blood flow. As C3b is rapidly inactivated and converted into C3dg by the action of factors H and I and proteases, a proportion of the cells re-enter the circulation, coated with C3dg, and are resistant to further lysis (Figure 13.3). Red cells coated with potent IgG antibodies, especially if they are C3b-binding, are destroyed mainly by cytotoxicity. Red cells coated with less potent antibodies are mostly destroyed by phagocytosis. Very rarely, red cell alloantibodies too weak to be detectable by routine pretransfusion testing may destroy donor red blood cells carrying the corresponding antigen.

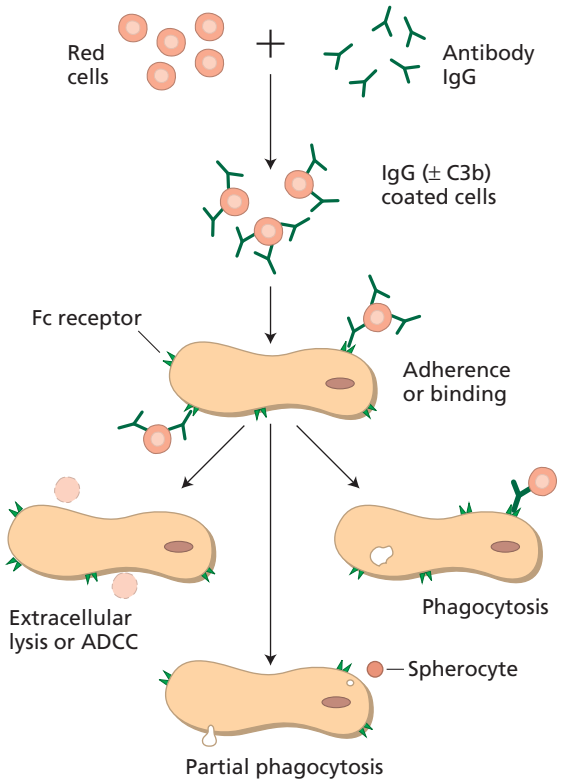


Figure 13.3 Mechanisms of extravascular destruction of red cells coated with IgG1 or IgG3 (± C3b). ADCC, antibody-dependent cell-mediated cytotoxicity.

The features of an immediate HTR vary according to a number of factors: whether the red cells are destroyed within the circulation or in the mononuclear phagocytic system, the strength, class and subclass of the antibody, the nature of the antigen, the number of incompatible red cells transfused and the clinical state of the patient. When antibodies are present in the circulation in low titres and a large volume of incompatible blood is given, all circulating antibody will bind to the incompatible red cells, coating them weakly, without destroying them. There will then be no antibody detectable in the serum for a number of days until secondary antibody production is stimulated by the immune challenge. On the other hand, in the presence of an overloaded or poorly functioning mononuclear phagocytic system, large volumes of IgG-sensitized incompatible red cells can be present in the circulation with minimal or no premature removal, so the haemoglobin level may be stable with little evidence of haemolysis. The DAT will be positive, but as there is no free antibody, elution techniques will be necessary for antibody identification.

Immediate extravascular destruction of red cells may be accompanied by hyperbilirubinaemia, occasionally haemoglobinaemia due to antibody-dependent cytotoxicity (in severe cases), fever and failure to achieve the expected rise in haemoglobin level. The signs and symptoms are less severe and dramatic than in intravascular haemolysis and usually appear more than 1 hour after the start of transfusion. There may be no signs or symptoms at all. Renal failure is very rare, even when the antibody binds the earlier components of the complement cascade. Symptoms are attributed in large degree to liberation of cytokines from mononuclear phagocytic cells after binding to IgG-coated red cells and to liberation of C3a when complement is bound up to C3b. The mortality is extremely low, but in an already sick patient the added complication of destruction of transfused red cells may contribute to death.

The management of immediate HTRs should be to terminate the transfusion immediately the patient develops the appropriate signs or symptoms. The identity of the patient and the units transfused should be checked against the appropriate documentation. Blood samples must be taken for investigation (Table 13.12). The transfusion laboratory should be immediately informed and all packs of transfused units should be returned. Pretransfusion samples should be tested in parallel. All urine passed during the first 24 hours should be measured and examined for haemoglobin. The circulating blood volume should be restored and blood pressure and urinary flow maintained using fluid challenges and furosemide (frusemide). Monitoring on a high-dependency unit may be required. The renal team should be involved early if urine output is poor (<1 mL/kg per hour) and haemofiltration may be necessary. Appropriate blood component therapy will be required if there is DIC.

Systems for correct patient identification can be strengthened with electronic transfusion management systems using barcodes on identity bands and blood components, with hand-held scanners linked to the laboratory information systems; however, these are not yet widely available.

Table 13.12 Investigation of suspected acute haemolytic transfusion reaction.

Blood test	Rationale/findings
Full blood count	Baseline parameters, red cell agglutinates on film
Plasma/urinary Hb, haptoglobin, bilirubin	Evidence of intravascular or extravascular haemolysis
Blood group of patient and units transfused	Compare with re-tested pretransfusion sample, to detect ABO error Unexpected ABO antibodies may arise from transfusion of incompatible plasma. Re-checking labels is often sufficient
DAT	Positive in majority. Compare with retested pretransfusion sample. May be negative if all incompatible cells destroyed
Compatibility testing	Repeat antibody screen and compatibility testing on pre- and post-transfusion samples. Elution of antibody from post-transfusion cells may aid antibody identification or confirm specificities in plasma in cases of non-ABO incompatibility
Urea, electrolytes and creatinine	Baseline renal function
Coagulation screen	Detection and monitoring of DIC
Blood cultures of patient and units	In event of possible septic reaction caused by bacterial contamination of unit

Delayed haemolytic transfusion reactions (DHTRs)

Such reactions may not be predictable or preventable. They are always caused by IgG antibodies, hence the haemolysis is always extravascular. In the majority of cases, an individual has been previously sensitized to one (or more) red cell antigen(s) by previous transfusion or pregnancy. Antibody is not detectable in routine pretransfusion testing, but the transfusion of blood containing the antigen/s to which the recipient has been sensitized previously, provokes a brisk anamnestic response that is

characteristic of the secondary immune response. Within days, the antibody level rises and the transfused incompatible cells are removed from the circulation. The effects of the secondary immunization are usually seen about 5–10 days after the transfusion, when the antibody level has been boosted and the recipient may already have left hospital. The clinical features are the same, though less severe than in immediate extravascular HTR.

The possibility of delayed HTRs underlines the importance of always taking fresh serum samples for antibody screening, direct antiglobulin test (DAT) and compatibility testing if a transfusion has been given more than 72 hours previously. Awareness of this complication may avoid unnecessary investigations to exclude infection when fever develops a few days after a transfusion. Most importantly, it will detect any alloantibody that will have been boosted by the transfusion, thus enabling the provision of compatible blood. Similarly it is important to have access to historical transfusion information since antibodies previously formed (e.g. Kidd antibodies) may become undetectable with time and may not be apparent in the current antibody screen. This is particularly important if patients are managed at more than one hospital.

Laboratory investigations of suspected DHTRs include full blood count and reticulocytes, examination of the blood film, plasma bilirubin, renal function tests and lactate dehydrogenase test (LDH). Serological investigations should include repeat blood group and antibody screen (on pre- and post-transfusion patient samples), DAT and elution of antibodies from the post-transfusion red cells for identification.

Treatment of DHTRs is usually supportive, sometimes requiring further transfusion. The presence of any antibodies should be documented within the laboratory and clinical records, with patients issued with an 'Antibody Card'.

Reactions related to white cell and platelet antibodies

Febrile non-haemolytic transfusion reactions (FNHTRs)

Febrile reactions are most frequently due to the transfusion of blood components containing white cells to patients sensitized to white cell antigens. Antibodies are usually against HLA antigens, or sometimes against granulocyte and platelet-specific antigens; they are stimulated by previous transfusions or pregnancies. Cytokines released from white cells during storage may also be pyrogenic. These reactions are characterized by fever, sometimes accompanied by shivering, muscle pain and nausea, but are much less common since the introduction of leucodepletion of blood components. They can occur up to 2 hours after completion of the transfusion and are more common in multi-transfused patients receiving red cells.

Mild FNHTRs (pyrexia >38 °C, but <2 °C rise from baseline) can often be managed simply by slowing (or temporarily stopping) the transfusion and giving an antipyretic, such as paracetamol. However the patient should be monitored closely, in case these are the early signs of a more severe transfusion reaction.

Transfusion-related acute lung injury (TRALI)

TRALI is non-cardiogenic pulmonary oedema, presenting during or within 6 hours of transfusion, with severe breathlessness accompanied by cough, fever and rigors. It is often associated with hypotension (due to loss of plasma volume) and there may be transient peripheral blood neutropenia or monocytopenia. Chest X-ray shows bilateral nodular shadowing in the lung fields with normal heart size, low oxygen saturation and low or normal central venous pressure. TRALI may be confused with acute respiratory distress syndrome (ARDS) of other causes, or with acute heart failure due to circulatory overload (see Table 13.13), but treatment with diuretics may worsen the outcome. Management is essentially supportive, requiring care in a high-dependency unit with ventilatory support, if needed. Steroid therapy is not effective. Managed appropriately, there is a high rate of survival and most patients recover within 1 to 3 days without long-term sequelae.

The reaction is due in most cases to passive transfer of leucoagglutinins (mostly anti-HLA class I or class II or, more rarely, granulocyte antibodies, i.e. anti-HNA) in donor plasma, reacting with granulocytes in the recipient's lung, leading to complement activation, endothelial and epithelial injury, alveolar damage and inflammatory changes, mediated by anaphylatoxins, cytokines and other inflammatory mediators. This complication is more likely after infusion of a plasma-containing component and the implicated donors are usually multiparous women. In the UK, the incidence of TRALI has declined to approximately 1 in 150,000 units transfused, since the introduction of the policy to produce FFP and plasma to suspend platelet pools mainly from male donors.

Transfusion-associated graft-versus-host disease (TA-GVHD)

TA-GVHD is a very rare, but usually fatal, complication of blood transfusion. It is caused by transfusion of viable donor lymphocytes with engraftment and clonal expansion of HLA-compatible donor lymphocytes in the recipient. It is characterized by fever, skin rash, diarrhoea, impaired liver function and pancytopenia 7 to 14 days after transfusion. Patients at risk include fetuses receiving intrauterine transfusion, newborn infants who have received IUTs, patients with inherited immunodeficiency disorders affecting T cell function, patients treated with allogeneic stem cell transplantation or specific chemotherapy drugs such as purine analogues. TA-GvHD has occasionally been reported in non-immunosuppressed patients receiving a blood transfusion from an HLA-matched donor or a close relative with HLA types in common. Patients receiving conventional combination chemotherapy for cancer are not at increased risk.

The diagnosis is made on the basis of typical features of acute GvHD in biopsies of affected organs, with detection of donor-derived cells or DNA in the patient's blood or tissues. The first SHOT reports in 1996 included 13 cases

Table 13.13 Comparison of transfusion related acute lung injury (TRALI) and transfusion associated circulatory overload (TACO) (adapted from BCSH Guideline on the Investigation and Management of Acute Transfusion Reactions, 2012, with permission of British Committee for Standards in Haematology).

	TRALI	TACO
Patient characteristics	? More common in haematology and surgical patients. Any age	Most common in age >70, but can occur at any age
Implicated blood components	Usually plasma or platelets	Any
Onset	Up to 6 hours from transfusion (usually within 2 hours)	Within 6 hours of transfusion
Oxygen saturation	Reduced	Reduced
Blood pressure	Often low	Often high
Jugular venous pressure	Normal or low	Elevated
Temperature	Often raised	Normal
Chest X-ray	Bilateral peri-hilar and nodular shadowing or 'white out', heart size normal	Enlarged heart and characteristics of pulmonary oedema
Echocardiogram	Normal	Abnormal
Pulmonary artery wedge pressure	Normal	Elevated
Blood count	Fall in neutrophils and monocytes followed by neutrophil leucocytosis	No specific changes
Fluid challenge	Improves	Worsens
Response to diuretics	Worsens	Improves

of transfusion-associated graft-versus-host disease, all fatal, in patients for whom gamma-irradiated blood components were not indicated at the time. Only one case of TA-GvHD has been reported in the UK since 2000 (an IUT of non-irradiated maternal blood). Given the potential fatal outcome of this complication of transfusion, it is essential to ensure that all at-risk patients receive irradiated red cells or platelet components.

Post-transfusion purpura (PTP)

PTP is a rare complication of blood transfusion, characterized by sudden onset of severe thrombocytopenia 7–10 days following transfusion of, usually, red cells. The patients are mostly female with always a history of previous blood transfusions or pregnancies. The most frequent cause is the presence in the recipient of an antibody (anti-HPA-1a) against the platelet-specific antigen HPA-1a (PI^{A1}); next in frequency is anti-HPA-5b. It appears that the antigen–antibody reaction causes both transfused and autologous platelets to be prematurely destroyed, either by the formation of immune complexes (in a manner similar to the 'innocent bystander' mechanism) or by cross-reaction of the causative antibody with the patient's own platelets. The disease is self-limiting, but in severe cases, prompt therapy with intravenous immunoglobulin is associated with a good response. Platelet transfusion is not recommended and should only be considered in life-threatening bleeding. This severe, and potentially fatal, complication has become even rarer since the introduction of leucodepleted blood components.

Reactions due to plasma protein antibodies

Mild allergic reactions

Mild urticarial reactions without other symptoms are not uncommon during blood transfusion; they occur with an approximate incidence of 1% and are mediated by IgE antibodies, usually against plasma proteins or other allergens present in donor plasma. Symptoms include urticaria (hives) and pruritis, but with no change in vital signs. They are most common in patients receiving plasma-rich components such as FFP or platelets.

Mild urticarial reactions may be treated effectively with antihistamines, and do not always recur. Standard components can be given unless symptoms are recurrent and severe.

Severe anaphylactic reactions

Anaphylaxis is an acute, life-threatening emergency associated with shock or severe hypotension with associated features including bronchospasm, stridor from laryngeal oedema and swelling of face, limbs or mucous membranes (angioedema). Components with a high plasma component such as platelets or FFP are most likely to be implicated, but such reactions may occur with all blood components, as they all contain some plasma. Staff in clinical areas carrying out blood transfusions must be trained in the emergency management of anaphylaxis, and epinephrine must be available for emergency use. The UK Resuscitation Council Guidelines recommend the urgent administration of intramuscular (IM) epinephrine to treat anaphylaxis (adult dose 0.5 mL of 1:1000 (500 μg)). The IM route is

rapidly effective (and life-saving); it avoids delay in attempting to obtain venous access in a shocked patient and is not contraindicated in patients with coagulopathy or low platelet count. After initial resuscitation, parenteral steroids may be given.

Although IgA deficiency is relatively common and may occur in up to 1 in 700 individuals tested, only a small minority of IgA-deficient patients are at risk of developing severe anaphylactic reactions to blood components. Those with severe IgA deficiency (< 0.07 g/L), with anti-IgA are at greatest risk. Patients with less severe IgA deficiency, as part of a more generalized (e.g. common variable immunodeficiency or secondary to a lymphoproliferative disorder) antibody deficiency disorder and mild cases frequently detected when screening for IgA coeliac antibodies, are not at risk. Patients with no history of severe reactions to blood transfusion should be transfused with standard blood components. Accordingly, it is the small minority of patients with severe IgA deficiency and a clear history of serious anaphylactic reaction to blood components that require special precaution and, if possible, transfusion of blood components from IgA-deficient donors. However, this might be feasible only in the elective setting, in countries with well-established blood services. In an urgent situation, washed red cells and platelets resuspended in platelet additive solution may be available, but in an extreme emergency, transfusion with standard blood components should not be withheld, but given with careful monitoring to allow timely detection and treatment of the anaphylactic reaction, as above.

Non-immunological complications of blood

Transfusion-transmitted infection

Bacterial contamination of blood components (see also section above under TTI)

This complication can rapidly be fatal and may occur in particular with platelet components, which are stored at 22–24 °C, rather than with red cells, which are refrigerated at 2–6 °C. These reactions can either be due to the septicaemia itself, to endotoxins, or both. The patient can present dramatically with collapse, high fever, shock and DIC.

Various measures to reduce bacterial contamination from the donor arm as highlighted above have significantly reduced this risk, but awareness and prompt detection and management remain essential. The immediate management and investigation follows the principles as outlined above. In particular, the pack should be inspected for abnormal discoloration or aggregates, but may appear normal. Blood cultures should be taken from the patient and immediate treatment started for shock, as appropriate and with intravenous broad-spectrum antibiotics, covering Gram-negative and Gram-positive bacteria. Any packs of implicated components must be returned to the transfusion laboratory for instigation of further testing including Giemsa staining for direct microscopic examination, as well as culture. The blood transfusion centre should be informed without delay so that any associated components from the implicated donation can be identified and withdrawn before they can be transfused to other patients. In the UK, the implicated component packs should be returned to the Blood Service for investigation, where comprehensive bacterial testing is undertaken together with typing of strains to confirm the identity of contaminating bacteria.

Transfusion-associated circulatory overload (TACO)

TACO is defined as acute or worsening pulmonary oedema within 6 hours of transfusion with acute respiratory distress, tachycardia, raised blood pressure and evidence of fluid overload (see Table 13.13 for differentiating features from TRALI). It causes significant morbidity and mortality; in 2013 SHOT received 96 reports of TACO contributing to the death of 12 of these patients with major morbidity in a further 34 patients. The risk factors for TACO include age > 70 years and comorbidity, including cardiac failure, renal impairment and hypoalbuminaemia. Low weight patients, such as the frail elderly and children, are at increased risk of receiving inappropriately high-volume and rapid blood transfusions, predisposing to TACO. Comprehensive pretransfusion assessment, with particular attention to the speed and volume of transfusion, fluid balance monitoring and use of diuretics, if indicated, may help prevent this complication. The common assumption that one unit of red cells produces a rise in Hb of 10 g/L only applies to patients of 70–80 kg. A red cell dose of 4 mL/kg will produce a rise of about 10 g/L. Single-unit transfusions in small, frail adults or prescription in millilitres (as in paediatric practice) should be considered. The treatment of TACO involves stopping the transfusion and administering oxygen and diuretic therapy, with critical care support, as required.

Other adverse effects of transfusion

Haemosiderosis is a very real complication of repeated blood transfusions, seen more commonly as long-term blood transfusion therapy improves the survival of patients suffering from some chronic anaemias. It is most commonly seen in thalassaemic patients, who commence transfusions in early childhood. Each unit of blood contains approximately 200 mg of iron, whereas the daily excretion rate is about 1 mg, and the body has no way of excreting the excess. Unless a patient is actively bleeding, hence losing iron, iron accumulation is inevitable. Significant iron overload is generally present after approximately 50 units of blood have been transfused to an average-sized adult. It is routine practice to give chelation therapy to transfusion-dependent thalassaemic patients (see Chapters 4 and 6).

In addition to the more tangible adverse events of transfusion described above, there are increasing concerns about the impact of transfusion on overall patient outcome in relation to mortality

and morbidity in critically ill patients, for example with potential impact on postoperative infection rates and length of hospital stay after surgery. This has led to more restrictive transfusion practices as described below. There is also an ongoing focus on the age of blood, with questions around the clinical impact of the age of blood on patient outcomes being tackled by randomized controlled studies.

Haemovigilance

Haemovigilance is defined by the International Haemovigilance Network as 'a set of surveillance procedures organized from the collection of blood and its components to the follow-up of its recipients, with the purpose of collecting and evaluating information on the undesirable and unexpected effects resulting from the therapeutic use of labile blood components and of preventing their occurrence or recurrence'.

The UK's professionally led haemovigilance system, SHOT, was introduced in 1996 with voluntary reporting of adverse transfusion reactions, errors and also near-miss events. Now, under the Blood Safety and Quality Regulations 2005 there is a legal requirement to report serious adverse reactions and events (SABRE) to the Medicines and Healthcare Products Regulatory Agency (MHRA). SHOT and MHRA now work closely together and have a joint reporting system through the SABRE website.

There is a high level of participation by hospitals in reporting to SHOT, with a steady increase in the level of reporting, but a reduction in the frequency of the most serious events, such as ABO-incompatible transfusions and cases of TRALI. These trends follow changes in national policies on training and competency for staff involved in blood transfusion, and a switch to male-only FFP. Overall mortality directly attributable to transfusion has decreased, which highlights the benefits of an effective haemovigilance scheme. However, errors related to transfusion remain the largest category of serious events, emphasizing the significant potential for further improving safety (see Figures 13.4 and 13.5). Transfusion-associated circulatory overload (TACO) is an important preventable cause of death or major morbidity.

SHOT also collects data on 'near-miss' events where the error was detected prior to administration of the transfusion, highlighting that further improvements are needed, particularly to ensure accurate patient identification. Whilst SHOT does not comprehensively collect data on fractionated plasma products, reports on events related to the inappropriate administration of anti-D Ig- are collated with feedback of key learning points.

Overall the SHOT data demonstrate that in high-resource countries, virological safety of the blood supply is advanced. Efforts should now be concentrated in other areas of transfusion medicine, such as the encouragement of appropriate use of blood, safe administration of blood components, and accurate patient and sample identification.

Appropriate use of blood and alternatives to allogeneic blood transfusion

In view of the inherent risks of blood transfusion and of the dependency of the blood supply on volunteer donors, blood and components should only be transfused when the benefits outweigh the risks. In addition to the clearly defined adverse events of transfusion, there are several reasons for aiming to reduce unnecessary allogeneic blood transfusions:

1 safety of the patient, by avoiding errors, as well as microbiological and immunological risks;
2 increasing difficulties in the recruitment of blood donors;
3 cost containment, with the ever-increasing demands by the regulatory agencies and technological advances;
4 high anxiety levels in patients, disproportionate to the real residual infectious risk of transfusion.

Patient blood management (PBM)

PBM may be defined as an evidence-based multidisciplinary approach to optimizing the care of patients who might need transfusion, considering alternatives, where feasible, and giving transfusion only when appropriate. PBM covers all aspects of decision-making in transfusion therapy, including initial patient evaluation with use of appropriate indications and transfusion triggers, minimizing blood loss and optimization of the patient's haemoglobin wherever possible. The application of the principles of PBM can reduce the need for transfusion and therefore health-care costs, while ensuring that limited donor blood supplies are available for the patients who need them most. In the surgical setting, the principles of PBM have been described as the 'three pillars', but it is important to note that the overall PBM strategy applies to all patients who might need transfusion, whether adults or children, elective or emergency, in any clinical setting, i.e. medical, surgical, obstetrics, intensive care, etc. There are now active PBM initiatives in many countries with ready web-based access to national guidance and resources to promote implementation.

There should be clear guidelines for safe and appropriate use of red cells and components in all clinical specialties. These should be based on national and international data, with appropriately graded recommendations reflecting the quality of evidence available within the published literature.

Education and training are fundamental to transfusion safety and good clinical practice, helping ensure that clinicians make the right decisions for their patients and use limited resources appropriately. Yet various studies show significant gaps in essential knowledge in transfusion medicine amongst undergraduates and doctors.

There should be regular multidisciplinary audits of transfusion practice, with feedback of results to all relevant staff and with re-audit to ensure that changes in practice have been

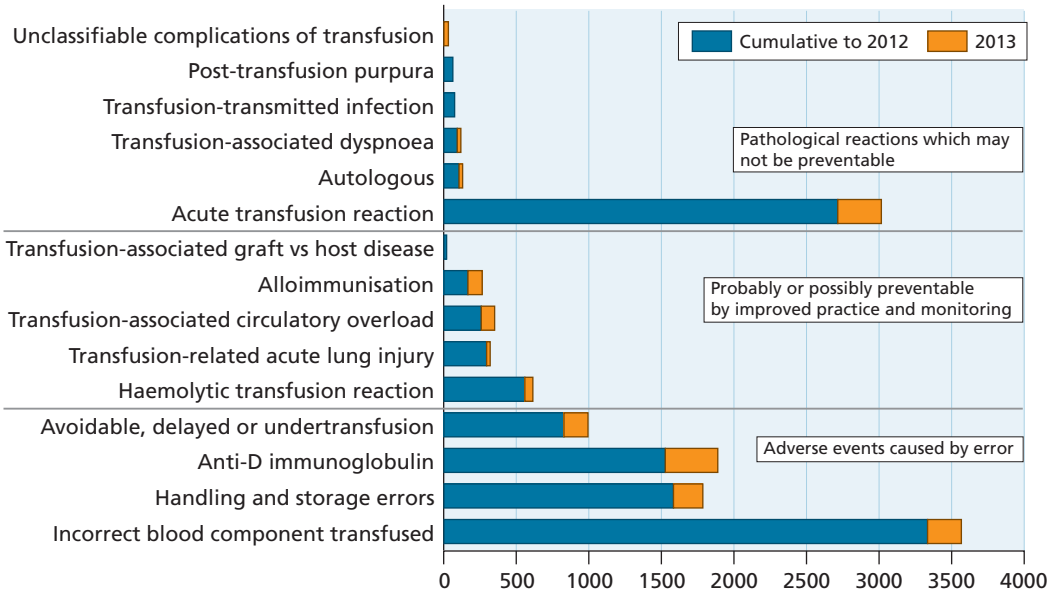


Figure 13.4 Cumulative data for SHOT categories 1996–2013. The figure subdivides reports into categories of transfusion-related adverse events that are due to errors and hence largely avoidable, compared to others that may be preventable in some cases and those that are unlikely to be preventable. 'Acute transfusion reactions' includes febrile, allergic and anaphylactic reactions. The category 'Autologous' includes reactions reported in relation to cell salvage. Further details at www.shotuk.org (Source: Serious Hazards in Transfusion (SHOT), www.shotuk.org. Reproduced with permission.)

undertaken, where needed. The National Comparative Audit of Transfusion is a collaborative initiative between the Royal College of Physicians in England and NHS Blood and Transplant, with active participation from hospitals across the country. Greater awareness of transfusion with emphasis on appropriate use of blood have resulted in a significant decrease in overall blood usage by more than 500,000 units in the UK in the last 15 years. However, several large regional and national audits of the use of red cells, platelets and plasma components continue to show ~15–20% inappropriate use, with much variation in practice, highlighting the ongoing need to implement effective patient blood management.

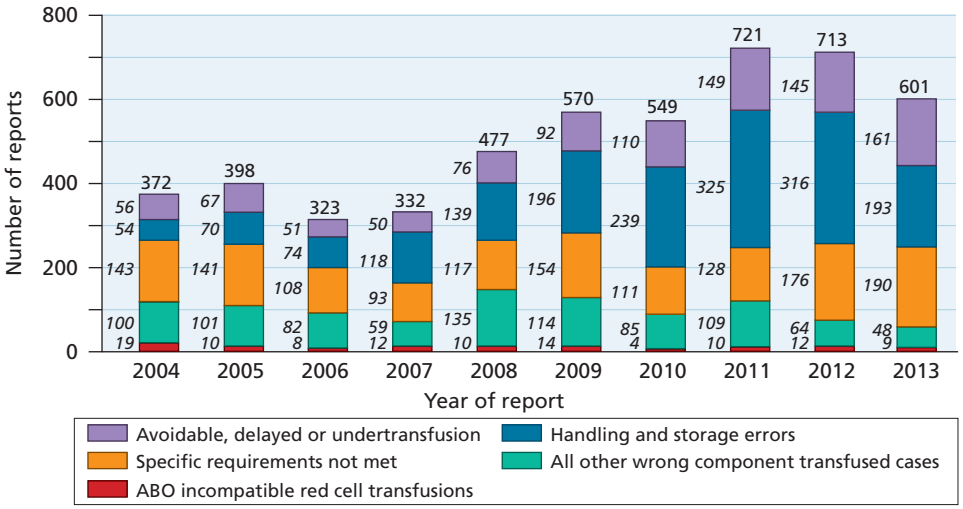


Figure 13.5 Cumulative data on causes of errors reported to SHOT 2004–2013. (Source: Serious Hazards in Transfusion (SHOT), www.shotuk.org. Reproduced with permission.)

Clinical decision to transfuse

The decision to transfuse is not just dependent on laboratory parameters, such as the haemoglobin level, platelet count or clotting screen, but requires a careful assessment of the patient's overall clinical status. Timely information should be made available to patients, ideally in written leaflets, about the indication for transfusion, the risks and benefits, and any alternatives available. These principles encompass 'valid consent', which does not also need a written signature from the patient. The indication for transfusion and the patient's consent should be documented in the clinical records. The initial assessment should include an evaluation of the patient's age, body weight and any comorbidity that can predispose TACO. Any special requirements, e.g. extended Rh- and K-matched blood for haemoglobinopathy patients or irradiated components for haemato-oncology patients should be considered when prescribing and requesting blood and components from the transfusion laboratory. Single unit red cell transfusions are recommended where possible, especially in non-bleeding adult patients, and further units should not be prescribed without further clinical assessment and monitoring of the patient's haemoglobin. Computer-assisted clinical decision-making may help with appropriate transfusion therapy, but such systems need significant resources and are not widely available.

Strategies to minimize transfusion of blood and components

Early preoperative identification of anaemia (see Table 13.14), followed by appropriate management with haematinics to optimize the haemoglobin level, can reduce the need for transfusion. The key aims of preoperative assessment are to assess a patient's fitness to undergo surgery and anaesthesia, anticipate complications, arrange for supportive therapy perioperatively and liaise with the appropriate specialists regarding non-surgical management. This assessment needs to take place at a presurgical clinic at least 1 month before the planned date of surgery. After the clinic, it is imperative that the results are evaluated so that the necessary action, such as the treatment of iron deficiency anaemia, can be undertaken.

Iron deficiency anaemia is also common in the obstetric setting, with often inadequate systems in place for timely intervention; oral or intravenous iron, as required, should be prescribed. Active management of anaemia is also indicated in the medical setting, with targeted management of true or functional iron deficiency, other haematinic deficiency (i.e. B12 and folate) or anaemia of chronic disease.

The volume of blood samples taken at phlebotomy should be minimized since this can exacerbate anaemia, particularly in premature babies and the intensive care setting. Patients on anticoagulants and antiplatelet drugs need careful preoperative assessment with a plan for discontinuation, where appropriate, before surgery. Surgical and anaesthetic techniques, use of fibrin glues and sealants provides additional scope for avoidance of donor transfusion in the context of surgery. Intraoperative cell salvage should be used where feasible but preoperative autologous blood deposit is no longer favoured, as evidence of efficacy is patchy and contradictory and the value of preoperative haemodilution needs further assessment. Pharmacological agents, in particular tranexamic acid, can reduce blood loss and the need for transfusion in many surgical settings with very low drug costs. Desmopressin (DDAVP) is effective in mild haemophiliacs. Erythropoietin can help reduce red cell requirements, but is expensive and therefore of limited application. Point of care testing, where available, may aid decision-making around use of plasma components and platelets, but further clinical studies are still needed in this area. Validated protocols supported by training and appropriate quality control are essential where these techniques are used.

Table 13.14 Preoperative assessment of patients.

- Take a full history and examination, including previous surgical episodes and bleeding history, transfusion and obstetric history
- Arrange full blood count (FBC), group and antibody screen, routine chemistry, coagulation screen (if indicated) and tube for haematinics assessment (ferritin level for iron stores, vitamin B_{12} and folate), which can be put on hold pending FBC results
- Consider use of tranexamic acid to prevent surgical blood loss
- Make cell salvage available if there is a likelihood of significant blood loss requiring transfusion and a 'clean' surgical procedure
- If iron deficiency give oral iron; use intravenous iron if intolerant or unresponsive to oral iron or if short timeframe prior to surgery (< 2 weeks)
- Prescribe folic acid supplement or B12 if evidence of deficiency
- If patient on antiplatelet therapy or oral anticoagulant drugs then assess and implement appropriate management plan to minimize risk of bleeding in the perioperative period

Red cell transfusion triggers

The seminal Transfusion Requirements in Critical Care (TRICC) study demonstrated no mortality benefit for patients with a higher transfusion threshold when compared to a restrictive haemoglobin trigger. The results of the TRICC study are strongly supported by more recent studies in high-risk patients needing hip surgery, cardiac surgery and acute upper gastrointestinal bleeding. This evidence from prospective randomized controlled studies demonstrates that a restrictive transfusion strategy results in equivalent or possibly better clinical outcomes than more liberal transfusion and should now inform clinical decision-making. Accordingly a threshold for

red cell transfusion of 70–80 g/L haemoglobin is appropriate in the majority of patients. A higher threshold of 80–90 g/L is indicated in patients with acute coronary syndromes, pending further studies.

Platelet transfusions

Platelet transfusions are indicated for the prevention and treatment of haemorrhage in patients with thrombocytopenia or platelet function defects. The largest group of patients, receiving up to 67% of all platelet concentrates, are haematology patients. Most of these platelet transfusions are prophylactic. The decision to transfuse platelets should be based on clinical assessment, taking into account clinical risk factors for bleeding and the extent and site of bleeding, together with the platelet count. The lack of evidence base supporting platelet transfusion practice has prompted increasing research in this area. Randomized controlled studies have attempted to tackle key questions around decision-making on prophylactic platelet transfusions in haemato-oncology patients. In such patients, the platelet count threshold for prophylaxis is 10×10^9/L, but a higher threshold may be indicated, depending on additional patient risk factors e.g. sepsis. The dose of prophylactic platelets has been shown to have no effect on the incidence of bleeding – accordingly, only single doses ($2.4–3 \times 10^{11}$) of platelets should be administered prophylactically, with little justification for routine use of higher, or 'double', doses of platelets. There is a lack of evidence for platelet thresholds, indicating the need for transfusion, in other clinical settings and, in particular, prior to procedures or surgery, with the use of empirical guidance to support practice e.g. 50×10^9/L for minor surgery and 80×10^9/L for major surgery.

Platelet transfusion response and refractoriness

The efficacy of platelet transfusions can be assessed by monitoring clinical response and by checking the platelet count increment. In patients requiring routine prophylactic platelet transfusions, a repeated failure to obtain a satisfactory response to two consecutive ABO-matched, relatively fresh platelet transfusions signifies refractoriness to platelet transfusions. Formulas to derive platelet recovery or corrected count increment are of limited value in the clinical setting and simpler, more pragmatic indicators include a failure of the post-transfusion (10 minutes to an hour) platelet increment to exceed the pretransfusion count or a rise of less than 10×10^9/L after 24 hours post transfusion.

Refractoriness to platelet transfusions is commonly due to non-immunological causes associated with increased platelet consumption or losses such as infections, antibiotics (amphotericin B and fluoroquinolones), splenomegaly/hypersplenism, DIC and major haemorrhage. Immunological refractoriness to platelet transfusions is due mainly to antibodies against antigens on platelets (HLA, HPA, ABO) and, more rarely, to platelet autoantibodies and drug-dependent antibodies or immune complexes. The most common cause of immunological refractoriness is HLA antibodies in the recipient. Although HLA alloimmunization is less common following universal leucodepletion of blood components, it is still a major problem in haemato-oncology patients needing prophylactic platelet transfusions.

The refractory patient should be assessed for the presence of non-immunological causes and if these have been excluded, and ABO-matched platelets have caused no increment, screening for HLA antibodies should be undertaken. If these are present, then platelet transfusions, collected by apheresis, from donors matched, as closely as possible, with the patient's HLA-A and HLA-B antigens should be given. The blood services in the UK have a large panel of HLA-typed donors, but this is not so in most countries, where the only alternative is to give cross-match-compatible platelets from the available stock. Immediate (10 to 60 minutes) and 24-hour post-transfusion platelet increments should be measured for evidence of effectiveness of therapy. If there is no response, then screening for HPA and other less common antibodies should be considered after advice from a transfusion medicine specialist. HLA-matched platelet preparations should be irradiated to prevent TA-GvHD.

Use of fresh frozen plasma and cryoprecipitate

There is a relative lack of good evidence to support the clinical use of FFP and therefore transfusions are given empirically in many situations. FFP is indicated in patients with acute DIC in the presence of bleeding and abnormal coagulation results (Chapter 40). The indication for FFP use in major haemorrhage is described below. FFP can also be used to treat inherited single coagulation factor deficiencies if the appropriate factor concentrates are not available. FFP should not be used for reversal of oral anticoagulation where the use of prothrombin complex concentrate is indicated. The coagulopathy of liver disease is also complex, with many studies showing the lack of evidence of clinical benefit of prophylactic FFP. A large proportion of the FFP used in the UK and in many other countries is misused, exposing patients to unnecessary risks.

Cryoprecipitate is indicated for fibrinogen replacement largely in the major haemorrhage setting (see below). Cryoprecipitate must not be used for replacement of coagulation factors in inherited conditions such as haemophilia or von Willebrand's disease, since specific factor concentrates are available (Chapter 38).

Major haemorrhage

Major haemorrhage can occur in many settings, e.g. trauma, obstetric, GI bleeds, with significant challenges in management.

Red cell transfusion is usually required when 30–40% of blood volume is lost (1500 mL in 70 kg male); more than 40% blood volume loss (1500–2000 mL) is life threatening and requires immediate transfusion. Red cells are necessary for their oxygen-carrying capacity and also because they contribute to improved haemostasis through a rheological effect leading to axial flow and therefore margination of platelets.

Metabolic changes in stored blood include low pH, hypocalcaemia and hyperkalaemia. The reduced oxygen-carrying capacity of stored blood becomes significant only after 21 days' storage (for CPD-AI blood) and is due to low 2,3-DPG levels (see Chapter 20). Although, theoretically, excess citrate in transfused blood could cause toxicity, its metabolism in the liver is usually rapid. In practice, the only situations when citrate toxicity is a real problem is with extremely rapid transfusion (one unit every 5 min), or in infants, especially if premature, having exchange transfusion with blood stored in citrate for longer than 5 days. Hypocalcaemia and hyperkalaemia are usually transient and rapidly corrected once the transfused blood is circulating. Acidosis is not generally significant, as citrate metabolism leads to an alkalosis. However, if a patient is severely shocked and under-transfused, acidosis may be a clinical problem. All these changes due to stored blood are exacerbated by hypothermia. Cardiac irregularities, in particular ventricular fibrillation, may result from transfusion of large quantities of cold blood. The optimal functioning of coagulation factors and of platelets is also temperature dependent and effectiveness is reduced by hypothermia. Thus, the use of a blood warmer and keeping the patient warm are important measures in the management of patients with major haemorrhage.

Approximately one-quarter of patients bleeding with severe trauma present with clotting abnormalities, indicating that the coagulopathy is not merely dilutional related to infusion of fluids or red cells. Indeed, coagulopathy associated with massive haemorrhage is likely to be multifactorial with contributory factors, including activation of fibrinolysis and consumptive coagulopathy, exacerbated by hypothermia and hypocalcaemia. The use of anticoagulant/antiplatelet drugs prior to surgery may further contribute to bleeding post surgery together with the heparin needed for bypass for cardiac surgery.

The practical management of major haemorrhage in any setting needs a coordinated multidisciplinary approach. This should incorporate significant advances in techniques for resuscitation as well as surgical, radiological and endoscopy interventions to control bleeding. The CRASH-2 study has demonstrated the benefits of early use of antifibrinolytics in trauma and therefore tranexamic acid should be incorporated into trauma protocols. Tranexamic acid may also have benefits in major obstetric haemorrhage and acute upper gastrointestinal haemorrhage, with multicentre clinical trials ongoing in these areas.

Baseline blood samples should be taken for full blood count, chemistry, coagulation screen, and group and screen. In patients with severe haemorrhage the initial use of group O red cells is indicated before the blood group is available; RhD negative units should be given to females of childbearing age. A switch to group-specific blood should be made as soon as possible.

While there is increasing focus on early and liberal administration of FFP, it is worth noting that many of the studies investigating the use of FFP in traumatic massive haemorrhage in a ratio of 1:1 with red cells have been retrospective, often in the military situation and hampered by the effect of survivor bias. There may be significant delays in obtaining laboratory clotting tests to initiate therapy. Therefore an empirical approach includes the early administration of FFP at a dose of 15–20 mL/kg, with further treatment guided by laboratory test results (aiming to maintain the prothrombin time (PT) and activated partial thromboplastin time (APTT) at a ratio of $<1.5 \times$ normal) or near patient tests of coagulation, if available.

Low fibrinogen levels commonly occur in massive haemorrhage. FFP may help improve fibrinogen levels, but specific therapy for fibrinogen replacement is needed if bleeding continues and the fibrinogen levels fall to < 1.5 g/L. In the UK, cryoprecipitate is most commonly used for this indication with the adult dose being two pools containing five donor units each. Fibrinogen concentrate, while available for replacement in congenital hypofibrinogenaemia, is not licensed for massive haemorrhage, but is used extensively in Europe as an alternative to cryoprecipitate at a dose of 3–4 g. Platelet transfusion should be given if the count falls below 75×10^9/L in major haemorrhage.

Haemolytic disease of the fetus and newborn (HDFN)

Haemolytic disease of the fetus and newborn is a condition in which the lifespan of the fetal/neonatal red cells is shortened due to maternal alloantibodies against red cell antigens inherited from the father. Maternal IgG can cross the placenta, and thus IgG1 and IgG3 red cell alloantibodies can gain access to the fetus. If the fetal red cells contain the corresponding antigen, then binding of antibody to red cells will occur. When the antibody is of clinical significance (e.g. anti-D, -c, -K), and of sufficient potency, the coated cells will be prematurely removed by the fetal mononuclear phagocytic system. The effects on the fetus/newborn infant may vary according to the characteristics of the maternal alloantibody. The blood film of a fetus affected by HDN shows polychromasia and increased numbers of nucleated red cells (Figure 13.6). In most cases (except a few due to ABO antibodies), the DAT (Coombs test) on the infant's cells is positive owing to IgG coating.

The antibodies giving rise to HDFN most commonly belong to the Rh or ABO blood group systems. The morbidity of Rh HDN is explained by the great immunogenicity of the D antigen; HDFN due to anti-c is also important and its incidence comes second among the cases of severe HDFN, closely followed by anti-K. (The disease caused by anti-K is more properly called

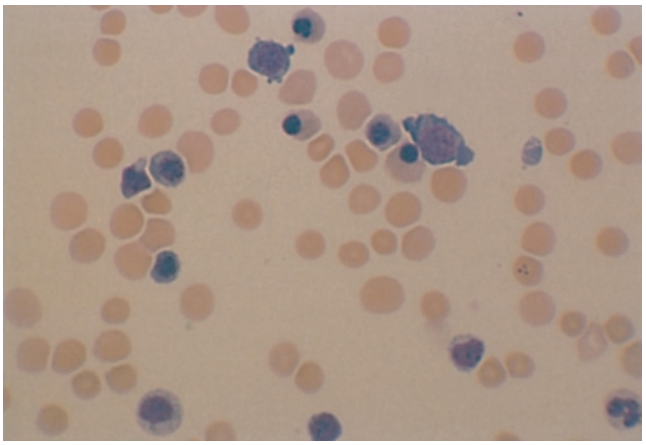


Figure 13.6 Blood film of a fetus affected by HDFN showing polychromasia and increased numbers of normoblasts.

alloimmune anaemia of the fetus and newborn as it is due to direct inhibition of erythropoiesis by the antibody and haemolysis is not a feature.) Antibodies against antigens in almost all the blood group systems (e.g. Duffy, Kidd), and against the so-called 'public' and 'private' antigens, have also been occasionally responsible for HDFN. However, IgM cannot cross the placenta and Lewis and P_1 antibodies, which occur frequently during pregnancy, are usually IgM and do not lead to HDN. Furthermore, the Lewis antigens are not red cell antigens per se and are not fully developed at birth.

Some antibodies (particularly anti-D, -K and -c) are associated with significant fetal and neonatal risks, such as anaemia requiring intrauterine or neonatal transfusion, jaundice or perinatal loss due to hydrops fetalis or kernicterus. There are many antibodies that are unlikely to significantly affect the fetus, but can cause neonatal anaemia and hyperbilirubinaemia), while others may cause problems for the screening and timely provision of appropriate blood for the mother or fetus/neonate. ABO antibodies can also cause mild to moderate anaemia and jaundice in the neonate and very occasionally in the fetus. In particular, IgG anti-A,B, anti-A and anti-B in group O mothers can cross the placenta and frequently cause HDFN in group A or B offspring (see also Chapter 50).

The blood group and antibody status of the mother should be tested at booking and at 28 weeks of gestation to identify the ABO group and D status and to detect clinically significant red cell antibodies (see Figure 13.7). If red cell antibodies are detected in the booking sample, further testing of maternal blood should be undertaken to determine the specificity and level of antibody or antibodies, and to assess the likelihood of them causing HDFN (Figure 13.8). The levels of anti-D and anti-c are quantified using automated analysers, whereas the titre of all other antibodies is determined by doubling dilution. Once detected, anti-D, anti-c and anti-K levels should be measured every 4 weeks up to 28 weeks' gestation and then every 2 weeks, until delivery. All other antibodies should be re-tested at 28 weeks gestation. Referral to a fetal medicine specialist should occur when there are rising antibody levels, but referral should also take place if there is a history of previous HDFN or of pregnancy with unexplained severe neonatal jaundice or anaemia requiring transfusion or exchange transfusion.

An anti-D level of > 4 IU mL but < 15 IU/mL correlates with a moderate risk of HDFN and an anti-D level of > 15 IU/mL can cause severe HDFN. An anti-c level of > 7.5 IU/mL but < 20 IU/mL correlates with a moderate risk of HDFN, whereas a level of > 20 IU/mL correlates with a high risk of HDFN. For anti-K, severe fetal anaemia can occur, even with low titres, so early referral to a fetal medicine specialist is advisable.

Where clinically significant maternal red cell antibodies are detected, the father's sample should also be tested for paternal phenotype to determine the risk of the fetus carrying the relevant antigen. If the partner is heterozygote or not available, free fetal DNA (fDNA) extracted from maternal plasma can be tested for RhD and Rhc genotype from 16 weeks onward and for K from 20 weeks (see Chapter 12).

If the fetus is antigen positive, then the pregnancy must be assessed and monitored for the development of HDFN by a fetal medicine specialist. Non-invasive techniques have largely replaced invasive testing, in particular with Doppler flow velocity; the use of middle cerebral artery (MCA) peak systolic velocity is a highly sensitive method for detecting fetal anaemia (see Figure 13.9). Ultrasound monitoring may help in detecting features suggesting severe fetal anaemia including polyhydramnios, skin oedema and cardiomegaly. The MCA peak systolic velocity related to gestation helps guide the need for further intervention including fetal blood sampling (see Figure 13.10) and intrauterine transfusion. The specifications for blood used for IUT are shown in Table 13.7.

Where the mother is known to have immune red cell antibodies, at delivery, cord blood should be taken for ABO and D typing and also for DAT, Hb and bilirubin levels. A positive DAT indicates that the infant's red cells are coated with antibody, but in itself cannot predict the severity of haemolysis, if any. Notably the DAT may be negative in ABO HDFN. It is therefore essential to also determine the Hb and bilirubin level to ascertain the degree of anaemia and haemolysis at birth, as this helps guide management, either with phototherapy or transfusion (see also Chapter 50). The specifications for blood for exchange transfusion are shown in Table 13.7. Some infants may have anaemia persisting for a few weeks following birth. While this may be due to passively acquired maternal antibodies causing continued haemolysis, it may also be due to suppression of erythropoiesis with extremely low reticulocytes in the affected infants and further top-up transfusions may be needed.

Prevention of RhD HDFN

The introduction of prophylaxis with anti-D immunoglobulin (anti-D Ig) for D-negative mothers has greatly reduced the risk

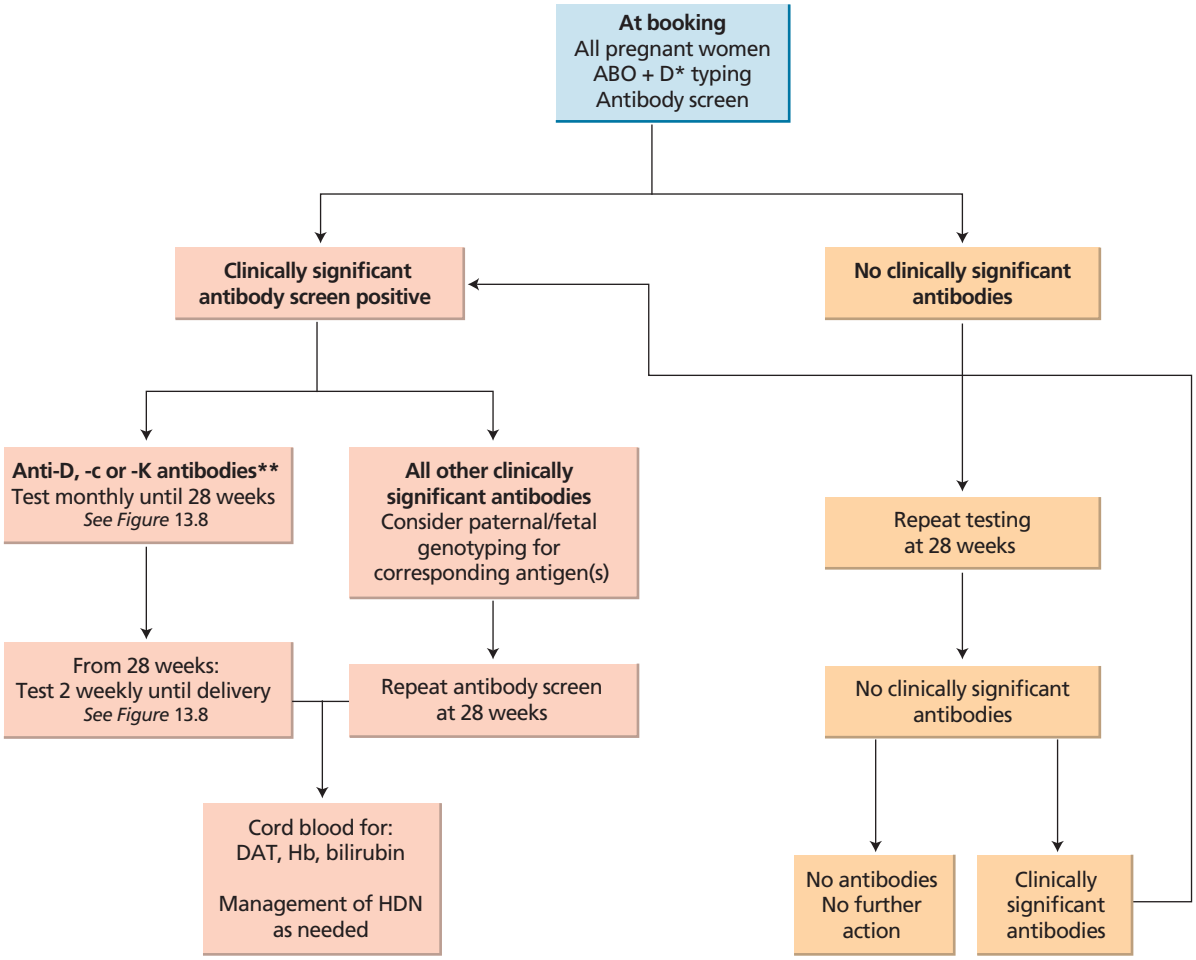


Figure 13.7 Blood group and antibody testing in pregnancy. If D-negative mother with no immune anti-D, advise anti-D Ig prophylaxis for any potentially sensitizing events in pregnancy and give routine antenatal anti-D Ig prophylaxis (RAADP), either as a single dose or as two doses (see RCOG anti-D guidelines); after delivery check cord sample for D type and maternal sample for fetomaternal haemorrhage (e.g. Kleihauer) testing to check if further anti-D Ig needed, in addition to the standard dose, which should be given in the first instance after delivery. DAT, direct antiglobulin test; Hb, haemoglobin; RAADP, routine antenatal anti-D Ig prophylaxis. (Source: RCOG Greentop guidelines 65, 2014. Reproduced with permission of The Royal College of Obstetricians and Gynaecologists.)
**Pregnancies with immune anti-D, -K or -c are at particular risk of severe fetal HDFN, so further early assessment and referral to fetal medicine specialist is indicated.

of HDFN, a previously significant cause of morbidity and mortality. Prior to the availability of anti-D prophylaxis, the first ABO-compatible, RhD-positive offspring resulted in the primary immunization of about 17% of D-negative women. The rate of alloimmunization fell considerably after administration of anti-D Ig postpartum and following other potentially sensitizing events (PSEs) during pregnancy, as listed in Table 13.15. It was, however, recognized that 'silent' transplacental bleeds could still occur in the later stages of pregnancy and the introduction of routine antenatal anti -D prophylaxis further reduced the sensitization rate to very low levels, ranging from 0.17 to 0.28%. Associated with this reduction in sensitization is a reduction in mortality due to HDFN, from 46/100,000 births to 1.6/100,000 births.

Following sensitising events, anti-D Ig should be administered as soon as possible and ideally within 72 hours of the event. If, exceptionally, this deadline has not been met, some protection may be offered if anti-D Ig is given up to 10 days after the sensitizing event. In pregnancies less than 12 weeks' gestation, anti-D prophylaxis is only indicated following ectopic pregnancy and therapeutic termination of pregnancy, but may also be indicated in some cases of uterine bleeding, where this is repeated, heavy or associated with abdominal pain. For PSEs between 12 and 20 weeks' gestation, a minimum dose of 250 IU (50 μg) should be

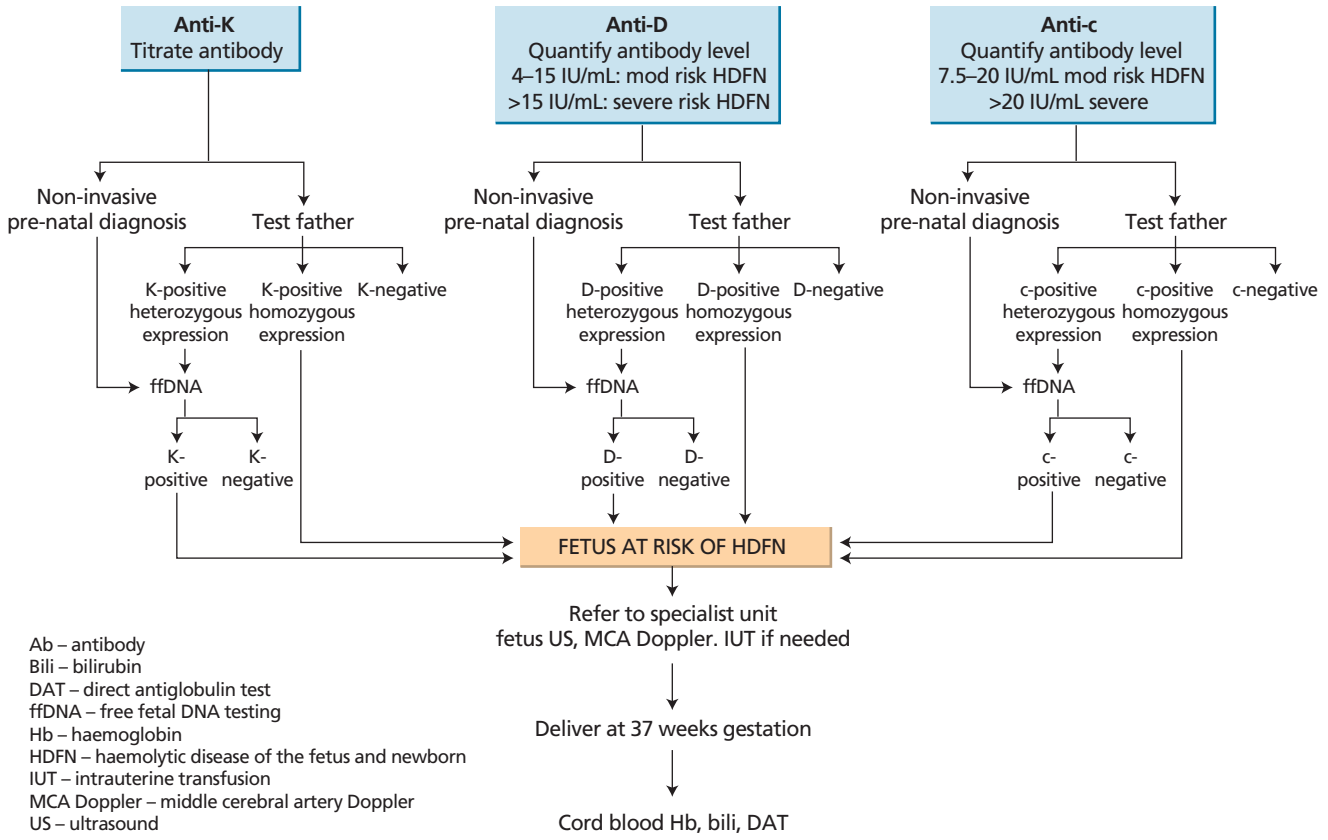


Figure 13.8 Management algorithm for pregnancies complicated with anti-D, anti-K or anti-c alloimmunization from RCOG Greentop guideline 2014. (Source: RCOG Greentop guidelines 65, 2014. Reproduced with permission of The Royal College of Obstetricians and Gynaecologists.)

given. For PSEs after 20 weeks' gestation a minimum anti-D Ig dose of 500 IU (100 μg) is indicated and testing for fetomaternal haemorrhage (FMH) should be performed to determine if additional anti-D Ig is needed.

The National Institute for Clinical Excellence (NICE) has recommended that D-negative pregnant women who do not have immune anti-D, should be offered routine antenatal anti-D prophylaxis (RAADP) during the third trimester of pregnancy. This can be with a single dose regimen at 28 weeks, or two dose regimen given at 28 and 34 weeks. It is important that the 28-week sample for blood group and antibody screen is taken prior to the first routine prophylactic anti-D Ig injection. RAADP should be

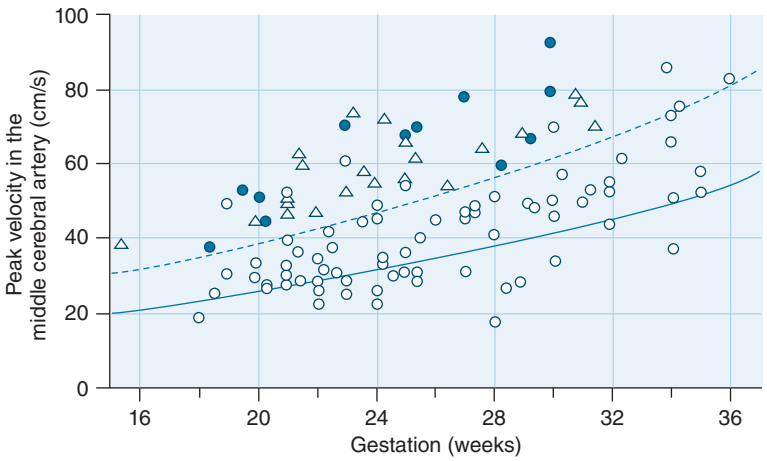


Figure 13.9 Middle cerebral artery Doppler. Peak velocity of systolic blood flow in the middle cerebral artery in 111 fetuses at risk for anaemia due to maternal red cell alloimmunization. Open circles indicate fetuses with either no anaemia or mild anaemia (≥0.65 multiples of the median Hb concentration). Triangles indicate fetuses with moderate or severe anaemia (<0.65 multiples of the median Hb concentration). Solid circles indicate the fetuses with hydrops. Solid curve indicates the median peak systolic velocity in the middle cerebral artery and the dotted curve indicates 1.5 multiples of the median. (Source: Professor Charles Rodeck. Reproduced with permission.)

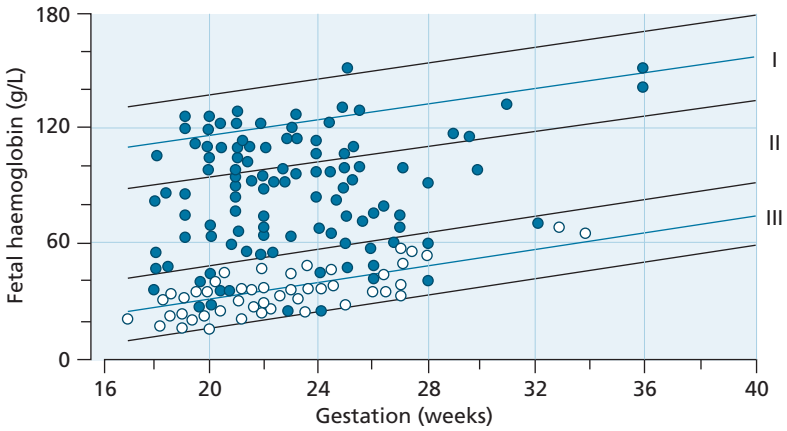


Figure 13.10 Fetal haemoglobin (Hb) concentration of 48 hydropic (open circles) and 106 non-hydropic (closed circles) fetuses from red cell isoimmunized pregnancies at the time of first blood sampling. Values are plotted on the reference range of fetal Hb for gestation. The individual 95% confidence intervals of the normal Hb for gestation define zone I and the individual 95% confidence intervals of the Hb for gestation of the hydropic fetuses define zone III. Zone II indicates moderate anaemia. (Source: Professor Charles Rodeck. Reproduced with permission.)

regarded as a separate entity and administered regardless of, and in addition to, any anti-D Ig that may have been given for a PSE.

Following birth, the cord blood should be tested and if the baby is D-positive, all D-negative, previously non-sensitised, women should be offered at least 500 IU (50 μg) anti-D Ig within 72 hours following delivery. Maternal samples should be tested for FMH and additional dose(s) given as guided by FMH tests (see below). If there is an intrauterine death (IUD) and no sample can be obtained from the baby, an appropriate dose of prophylactic anti-D Ig should be administered. Anti-D prophylaxis is also indicated if intraoperative cell salvage is used during caesarean section in D-negative women, where the cord blood group is confirmed as D-positive (or unknown) with estimation of FMH on a maternal sample taken 30–45 minutes after reinfusion to check if additional doses of anti-D Ig are needed.

A dose of 500 IU (50 μg), intramuscularly (IM) is considered sufficient to treat an FMH of up to 4 mL fetal red cells. Where it is necessary to give additional doses of anti-D Ig, as guided by tests for FMH, the dose calculation is traditionally based on 125 IU anti-D Ig/mL fetal red cells for IM administration. However the manufacturer's guidance should be referred to, depending on which product is used. In cases of large FMH, and particularly if FMH is in excess of 100 mL, a suitable preparation of intravenous anti-D Ig should be considered. Good communication between the antenatal teams and the transfusion laboratory is essential, with clear documentation of issue and administration of anti-D immunoglobulin. In the UK, the SHOT haemovigilance scheme continues to report on many errors with delayed or missed administration of anti-D prophylaxis.

Table 13.15 Potentially sensitizing events in pregnancy (from BCSH guidelines 2014).

- Amniocentesis, chorionic villus biopsy and cordocentesis
- Antepartum haemorrhage/PV bleeding in pregnancy
- External cephalic version
- Fall or abdominal trauma (sharp/blunt, open/closed)
- Ectopic pregnancy
- Evacuation of molar pregnancy
- Intrauterine death and stillbirth
- *In-utero* therapeutic interventions (transfusion, surgery, insertion of shunts, laser)
- Miscarriage, threatened miscarriage
- Therapeutic termination of pregnancy
- Delivery – normal, instrumental or caesarean section
- Intraoperative cell salvage

Testing for fetomaternal haemorrhage (FMH)

Commonly used techniques for quantification of FMH include the acid elution method, a modification of the Kleihauer–Betke test, and flow cytometry.

The principle of the Kleihauer test relies on the different properties of fetal haemoglobin (HbF) and adult haemoglobin (HbA), whereby HbF is more resistant than HbA to both acid elution and alkaline denaturation. When a blood film from the maternal sample is fixed and immersed in an acid buffer solution, HbA is denatured and eluted, leaving red cell ghosts. Fetal red cells containing HbF that is resistant to elution stand out after staining in a background of maternal ghost cells. The relative percentages can then be counted using light microscopy to calculate an estimated volume of FMH. It is essential to have positive and negative controls. This is a manual test with a high coefficient of variation. Accordingly current BCSH guidelines recommend that more accurate testing with flow cytometry should be undertaken if the estimated FMH volume is >2 mL.

Flow cytometry is used to quantify the population of RhD-positive fetal cells in the RhD-negative maternal sample, using a fluorochrome-conjugated IgG monoclonal anti-D to label the D-positive cells.

A maternal sample should be taken for FMH testing after any sensitizing event occurring after 20 weeks' gestation. When the

sensitizing event is before 20 weeks, the fetal blood volume is considered insufficient to exceed that covered by the minimum anti-D Ig dose, as recommended above. FMH testing should also then be undertaken after delivery, following which a standard dose of anti-D should be given after the baby is confirmed to be D-positive on cord blood testing. The result of the FMH testing is then used to decide if additional doses of anti-D Ig are needed. If additional doses of anti-D Ig are indeed needed, then follow-up FMH testing is indicated to confirm clearance of D-positive cells from the maternal circulation. Repeat FMH testing should be undertaken after 48 hours if anti-D is given intravenously or 72 hours after an intramuscular dose.

Rh D-positive platelet transfusions

Whenever possible, D-negative platelets should be transfused to D-negative girls or women of childbearing potential, but on occasions, depending on availability and clinical need, it may be necessary to transfuse D-positive platelets. In these circumstances, prophylaxis against possible Rh alloimmunization by red cells contaminating the platelet product may be needed. In the UK, a dose of 250 IU (25 μg) anti-D Ig is considered sufficient to cover up to five adult therapeutic doses of D-positive platelets given within a 6-week period. In severely thrombocytopenic patients with platelet counts of $\leq 30 \times 10^9$/L, anti-D Ig should be given subcutaneously, or if an appropriate preparation is available, intravenously, to avoid the risk of intramuscular bleed following IM injection.

Inadvertent transfusion of RhD-positive blood to RhD-negative women of childbearing potential

When less than one unit of red cells has been transfused, the appropriate dose of anti-D Ig may be given (see above). If more than 15 mL have been transfused, the larger anti-D immunoglobulin preparations (1500 or 2500 IU), if available, should be considered with intravenous anti-D immunoglobulin being preferable, achieving adequate plasma levels immediately. The quantitation of D-positive red cells should be performed by flow cytometry after 48 hours if an intravenous dose of anti-D has been given, or 72 hours if an intramuscular dose has been given and further anti-D Ig given until there are no detectable D-positive red cells in circulation.

When more than one unit of D-positive blood has been transfused, a red cell exchange transfusion should be considered to reduce the load of D-positive red cells in the circulation and the dose of anti-D Ig required to suppress immunization. A single blood-volume red cell exchange transfusion will achieve a 65–70% reduction in D-positive red cells; a double volume exchange will achieve an 85–90% reduction. Shortly after the exchange transfusion, the residual volume of D-positive red cells should be estimated using flow cytometry to guide anti-D Ig dosage needed.

Advice should be sought from a transfusion medicine specialist. The patient should be counselled regarding the implications of both non-intervention (for future pregnancies) and of treatment, including any hazards from receiving donated blood, the exchange procedure itself and of larger doses of anti-D Ig, including intravenous anti-D. Passive anti-D Ig given in large doses may remain detectable and tests for immune anti-D may be inconclusive for several months.

Neonatal alloimmune thrombocytopenia

This is a condition in which the platelets of the fetus and newborn are destroyed by maternal platelet-specific alloantibodies against platelet antigens inherited from the father (see also Chapter 50). It is analogous to HDFN, but for platelets. Although more than 15 HPA systems have been described, most cases are due to anti-HPA-1a in alloimmunized HPA-1b mothers. The difference with neonatal alloimmune thrombocytopenia is that the offspring of the first pregnancy can be affected and that the potency of the platelet antibodies is often not correlated with the severity of the fetal or neonatal thrombocytopenia. The most serious complication is intracranial haemorrhage, which may lead to death or severe neurological sequelae. Treatment consists of intravenous immunoglobulin and/or transfusion of HPA-1anegative platelets, available from the national Blood Services' to 'Pending confirmatory testing, transfusion of HPA-1a negative platelets is indicated and these are available from the national Blood Services. Specialist transfusion advice should be sought.

Selected bibliography

Barbara JAJ, Regan FAM, Contreras MC (eds.) (2008) *Transfusion Microbiology*. Cambridge University Press, Cambridge.

Blood Safety and Quality Regulations (2005) Statutory Instrument 2005 No. 50. www.opsi.gov.uk/si/si2005/20050050.htm (accessed May 2015).

British Committee for Standards in Haematology (BCSH) Guideline on the Administration of Blood Components (2009) http://www.bcshguidelines.com (accessed May 2015).

British Committee for Standards in Haematology (BCSH) Guidelines for the use of Prophylactic Anti-D Immunoglobulin http://www.bcshguidelines.com/documents/Anti-D_bcsh_07062006.pdf (accessed May 2015).

British Committee of Standards of Haematology. Available at: http://www.bcshguidelines.com/documents/ATR_final_version_to_pdf. Accessed 11 August 2015.

Contreras M (ed.) (2009) *ABC of Transfusion*, 4th edn. Wiley-Blackwell/BMJ Books, Oxford/London.

Hunt BJ, Allard S, Keeling DM *et al*. (2015) A practical guideline for the haematological management of major haemorrhage. *British Journal of Haematology*. doi: 10.1111/bjh.13580. [Epub ahead of print]

Norfolk D (ed.) *Handbook of Transfusion Medicine*, 5th edn. HMSO, London.

Milkins C, Berryman J, Cantwell C *et al.* (2013) BCSH Guidelines for pre-transfusion compatibility procedures in blood transfusion laboratories. *Transfusion Medicine* **23**(1): 3–35.

Murphy MF, Waters JH, Wood EM, Yazer MH (2013) Transfusing blood safely and appropriately. *British Medical Journal* **16**: 347.

RCOG Greentop guidelines 65 May 2014 The management of women with red cell antibodies in pregnancy. https://www.rcog.org.uk/globalassets/documents/guidelines/rbc_gtg65.pdf (accessed May 2015).

Serious Hazards of Transfusion. http://www.shotuk.org (accessed May 2015).

Specialist Advisory Committee on the Safety of Blood, Tissues and Organs (SaBTO) www.dh.gov.uk/en/Publicationsandstatistics/Publications/PublicationsPolicyAndGuidance/DH_132965 (accessed May 2015).

Tinegate H, Birchall J, Gray A *et al.* (2012) Guideline on the investigation and management of acute transfusion reactions Prepared by the BCSH Blood Transfusion Task Force. *British Journal of Haematology* **159**(2): 143–53.

UK Blood Transfusion and Tissue Transplantation Services. (2005) Guidelines for the Blood Transfusion Services in the UK (Red-Book), 7th edn. HMSO, London, 2005. Also available at: http://www.transfusionguidelines.org.uk/index.aspx?Publication=RB (accessed May 2015).

CHAPTER 14

Phagocytes

14

John Mascarenhas, Marina Kremyanskaya and Ronald Hoffman

Tisch Cancer Institute, Icahn School of Medicine at Mount Sinai, New York, New York, USA

Introduction

White blood cells have fundamental roles in defence against invading microorganisms and the recognition and destruction of neoplastic cells, as well as their role in acute inflammatory reactions. Furthermore, through their phagocytic function, white blood cells are influential in clearing senescent and apoptotic cells, hence allowing tissue repair and remodelling. Production of various cytokines by white blood cells influences the functions of other cells and affects processes such as cellular and humoral immunity, and allergic phenomena. The phagocytic actions of white blood cells can cause damage to the host tissue, leading to inflammation. This occurs either as a by-product of their microbial killing actions or as a direct attack on the host in autoimmune disorders.

Normal haemopoiesis, including generation of appropriate numbers of white blood cells and types, is dependent upon intricately regulated signalling cascades that are mediated by cytokines and their receptors. Orderly function of these pathways leads to the generation of a normal constellation of haemopoietic cells, and their abnormal activation results in impaired apoptosis, uncontrolled proliferation and neoplastic transformation. Cytokines function in a redundant and pleiotropic manner; different cytokines can exert similar effects on the same cell type and any particular cytokine can have several differing biological functions. This complexity of function is a result of shared receptor subunits, as well as overlapping downstream pathways, culminating in transcription of similar genes. Increased understanding of the role of cytokines and other growth factors in the control of normal haemopoiesis has led to better delineation of the pathogenetic events that affect the function and number of these cells.

In this chapter, we consider the normal production and function of white blood cells involved in phagocytosis and describe various disorders causing their altered number and activity.

Mechanisms of phagocyte function

Locomotion

Phagocytes are an important part of the innate host defence system, performing their function either as resident cells in tissues (e.g. macrophages) or as circulating defenders (e.g. neutrophils, eosinophils and monocytes). Phagocytosis of invading microorganisms by both types of defenders involves the synthesis of highly toxic derivatives of molecular oxygen by the respiratory burst NADPH oxidases and the delivery of stored antimicrobial proteases into the vacuoles containing microbes.

Circulating phagocytes such as neutrophils respond to spatial gradients of chemotaxins. The signals generated by such gradients activate the cytoplasm of the cells leading to propulsive and retractive events. The locomotion of neutrophils is achieved by the contraction of an actin filamentous network in the cortical gel at the leading front. Phagocytic cells possess a number of cell–cell adhesion receptors and ligands, which mediate their recruitment, migration and interaction with other immune cells (Table 14.1). These include integrins, members of the immunoglobulin superfamily and the selectins. Migration of macrophages is mediated by cytokine-regulated expression of intercellular adhesion molecules (ICAMs) on the surface of both phagocytes and endothelial cell. ICAMs share a similar structure with the immunoglobulin family and other immunoglobulin-like adhesion molecules, such as vascular cell

Postgraduate Haematology, Seventh Edition. Edited by A Victor Hoffbrand, Douglas R Higgs, David M Keeling and Atul B Mehta.

Table 14.1 Phagocytic cell adhesion molecules.

Adhesion molecule	CD number	Cellular distribution	Ligand	Function
Integrin family				
Very late-acting antigens				
$\alpha_1\beta_1$ (VLA-1)	CD49a/29	Mo, EC	Collagen I, IV, laminin	Cell adherence to ECM
$\alpha_2\beta_1$ (VLA-2)	CD49b/29	Mo, EC, platelets	Collagen I, IV, laminin	
$\alpha_3\beta_1$ (VLA-3)	CD49c/29	Mo	Collagen I, laminin, fibronectin	Cell adherence to ECM
$\alpha_4\beta_1$ (VLA-4)	CD49d/29	Mo, eos, bas	Fibronectin, VCAM-1	Cell adherence to ECM and cell–cell adhesion matrix
$\alpha_5\beta_1$ (VLA-5)	CD49e/29	Mo, neut, EC	Fibronectin	Cell adherence to ECM
$\alpha_6\beta_1$ (VLA-6)	CD49f/29	Mo	Laminin	Cell adherence to ECM
Leucocyte integrins (LFA-1 family)				
$\alpha_D\beta_2$	–/18	Ma	?	
$\alpha_L\beta_2$ (LFA-1)	CD11a/18	Mo, Ma, granulocytes	ICAM-1, ICAM-2, ICAM-3	Cell–cell adhesion and cell–matrix adhesion
$\alpha_M\beta_2$ (CR3, Mac-1)	CD11b/18	Mo, Ma, granulocytes	ICAM-1, C3bi, fibronectin, factor X, microbial antigens	Endothelium adherence/ extravasation
$\alpha_X\beta_2$ (p150,95)	CD11c/18	Mo, Ma, granulocytes	C3bi, fibronectin	Adhesion during inflammatory response
Cytoadhesins				
$\alpha_V\beta_3$ (vitronectin receptor)	CD51/61	Mo, EC	Vitronectin, fibronectin, collagen, thrombospondin, vWF	Cell adherence to ECM
$\alpha_R\beta_3$ (leucocyte response integrin)		Mo, granulocytes	Vitronectin, fibronectin, collagen, thrombospondin, vWF	Cell adherence to ECM
$\alpha_V\beta_5$	CD51/–	Mo	Vitronectin, fibronectin	Cell adherence to ECM
$\alpha_V\beta_7$	CD51/–	Ma	?	
Immunoglobulin superfamily				
ICAM-1	CD54	Mo, EC	$\alpha_L\beta_2$, $\alpha_M\beta_2$	Cell–cell adhesion
ICAM-2	CD102	Mo, EC	$\alpha_L\beta_2$	Cell–cell adhesion
ICAM-3	CD50	Mo, granulocytes	$\alpha_L\beta_2$	Cell–cell adhesion
VCAM-1	CD106	Ma, EC, dendritic cells	$\alpha_4\beta_1$	Recruitment
PECAM-1	CD31	Mo, EC, platelets	CD31, $\alpha_V\beta_3$	Transmigration
HCAM	CD44	Ubiquitous	Collagen I, IV, fibronectin	Extravasation
Selectin family				
L-selectin	CD62L	Mo, granulocytes	Carbohydrate determinants on EC	Migration, rolling on vessel wall
E-selectin	CD62E	Neutrophil, EC	Mo, neut, eos	Migration, rolling on vessel wall
P-selectin	CD62P	EC, platelets	Mo, neut, eos	Adhesion to activated platelets and EC

Bas, basophil; CD, cluster of differentiation; EC, endothelial cell; eos: eosinophil; ICAM, intercellular adhesion molecule; Mo, monocyte; Ma, macrophage; neut, neutrophil.

adhesion molecule (VCAM)-1, and serve as ligands for the β_2 integrins. Another member of the endothelial immunoglobulin superfamily, PECAM-1 (CD31), plays an important role in the transmigration of neutrophils into mucosa or other body tissues.

The β_2 integrin family consists of three leucocyte-restricted integrins, LFA-1 (CD11a/CD18), CR3 (MAC-1, CD11b/CD18) and p150/95 (CD11c/CD18). LFA-1 and ICAM-1 are both present on monocytes and mediate their attachment to endothelial cells and to lymphocytes bearing the corresponding receptor/ligand, thereby facilitating antigen presentation. Recently, serum response factor (SRF) has been shown to be critical to neutrophil adhesion and migration at sites of inflammation via regulation of CD11b.

Selectins are expressed on all leucocytes (L-selectins) as well as postcapillary endothelial surfaces (E-selectins) and in platelet α-granules and endothelial cell Weibel–Palade bodies (P-selectins). The interaction of E- and P-selectins on cytokine-activated endothelial cells, and L-selectins on macrophages with their appropriate ligands, targets phagocytic cells to the endothelium at sites of vascular injury and initiates the rolling movement of leucocytes along the vessel wall. Paracellular diapedesis is the purposeful coordinated movement of leucocytes though the endothelial lining towards sites of inflammation. Selectin-mediated leucocyte tethering and rolling across endothelial surfaces is followed by integrin-mediated firm adhesion and leucocyte polarization at endothelial cell junctions. CD157, a glycosylphosphatidylinositol (GPI)-anchored surface protein, is crucial to this process of locomotion and is expressed on the surface of neutrophils and vascular interendothelial junctions.

Phagocyte receptors

Phagocytes express a number of surface receptors that recognize microbial surfaces, as well as altered tissue components and apoptotic bodies (Table 14.2). Furthermore, non-specific components of the innate immune response, such as the components of the complement cascade, can tag and thereby identify invading microorganisms, thus allowing their opsonization via another family of receptors, leading to the uptake of complement-coated microorganisms. Similarly, other molecules such as matrix proteins (i.e. fibronectin and vitronectin) can act as opsonins allowing recognition and uptake by the phagocytic cells.

Phagocytes also express a number of receptors for the Fc portion of immunoglobulin molecules IgA, IgE and IgG. They mediate opsonization of particles and microorganisms, enhancing their phagocytic uptake. The antigen-binding site (Fab) of IgG binds to bacteria exposing the Fc-binding site, which is in turn recognized by one of three classes of receptors, FcγRI, FcγRII or FcγRIII. Polymeric IgA antibody can also function as an opsonin and is recognized by a specific IgA receptor (FcαR), which is widely distributed on the circulating haemopoietic cells, but is more highly expressed in phagocytes in the secretions of the gut and the lung.

A number of miscellaneous receptors involved in phagocytosis have been described, such as the macrophage mannose receptors (MMR), lectin-like molecules that bind mannose and fucose residues on the surfaces of yeast, bacteria and parasites. These receptors are present only on the surface of macrophages and their activation mediates endocytosis, phagocytosis and cytotoxicity by reactive oxygen intermediates. In addition to recognizing the haemoglobin–haptoglobin complex, the macrophage scavenger receptor CD163 has also been shown to bind Gram-positive and Gram-negative bacteria and elicit an inflammatory cytokine response as part of the innate immune response. CD14 is the receptor for complexes of lipopolysaccharide (LPS)-binding protein with LPS, which coat Gram-negative bacteria and enhance phagocytosis. CD14 is constitutively expressed on monocytes and macrophages and TNF, IL-1 and IL-6 increase its expression, whereas IFN-γ and IL-4 decrease it.

Different subsets of dendritic cells and macrophages likely play distinct roles in the clearance of cells at different stages of apoptosis. IL-10-producing anti-inflammatory macrophages are the professional phagocytes involved in the recognition and clearance of early apoptotic cells. The membrane of apoptotic cells binds increased amounts of thrombospondin, a macrophage secretory product that is recognized by both CD36 (thrombospondin receptor) and integrin αvβ3 (vitronectin receptor). Similarly, phagocytic lectin receptors bind carbohydrate determinants exposed on the surface of apoptotic cells. A specific receptor for phosphatidylserine has also been described, which acts alone or in association with CD36 to recognize the exposed phosphatidylserine. CD14 has also been reported to be involved in tethering of apoptotic lymphocytes via interaction with ICAMs. Apoptotic cell-associated molecular patterns (ACAMPs) are highly conserved molecular changes expressed in cells undergoing apoptosis that are recognized by innate immune system receptors such as CD14, CD91, C1q, C3bi, collectins and pentraxins. Neutrophils also recognize chemoattractant signals through receptors expressed on their cell surface. N-formyl-methionyl tripeptide receptors such as f-Met-Leu-Phe (FMLP) are similar to naturally occurring bacteria-derived factors. Each demonstrates time-dependent saturable binding kinetics and a high-affinity dissociation constant (KD) for the specific chemoattractant.

Phagocytic signalling

The complex process of phagocytosis is regulated by events related to the activation of various receptors such as the FcγRs. Such receptor activation results in initiation of downstream signalling events through immunoreceptor tyrosine-based activation motifs (ITAMs). As a result of cross-linking of these receptors under appropriate conditions, downstream effector functions are activated, resulting in phagocytosis, stimulation of

Table 14.2 Opsonic receptors mediating phagocytosis.

Receptor	Marker	Opsonic ligand	Binding affinity (K_a)	Cell type	Function
FcγRI	CD64	IgG1	High (50 nmol/L)	Monocytes, macrophages, neutrophils (after IFN-γ exposure)	Phagocytosis, respiratory burst
FcγRII	CD32	IgG1 = IgG3 ≥ IgG4 = IgG2	Low (1 μmol/L)	Neutrophils, monocytes, macrophages	Phagocytosis, respiratory burst
FcγRIII		IgG1 = IgG3	Low (110 nmol/L)	Neutrophils, monocytes, macrophages	IIIB: phagocytosis (requires CR1 or FcγRII)
IIIA	CD16a, 1 allotype NA1		Low (470 nmol/L)		
IIIB	CD16b, two allotypes NA1 and NA2				
FcαR	CD89 My43 IgM	IgA1, IgA2, secretory IgA1 and IgA2		Neutrophils, monocytes, macrophages, T and B cell subsets, NK cells, erythrocytes	Phagocytosis, respiratory burst, bacterial killing
CR1	CD35, four alleles	C3b and C4b dimers	High (0.5 nmol/L)	All phagocytes, some T lymphocytes	Phagocytosis
CR3	CD11b/CD18 Mac1	C3bi	High (0.5 nmol/L)	All phagocytes, NK cells, γδ T cells	Phagocytosis, respiratory burst

CR, complement receptor; NK, natural killer.

the respiratory burst, degranulation of bactericidal proteins and activation of transcription factors, in turn leading to enhanced expression of genes encoding cytokines and other inducible proteins.

Members of the Src family of tyrosine kinases (e.g. Lyn, Fgr and Hk) associate with FcγRs and are likely responsible for their tyrosine phosphorylation. The tyrosine-phosphorylated ITAMs then serve as binding sites for downstream kinases such as Syk, which propagate signals important for phagocytosis.

Degranulation and secretion

Degranulation and secretion are processes whereby the contents of phagocytic storage granules are released into the phagocytic vacuoles (degranulation) or into the extracellular space (secretion). Degranulation and secretion can be triggered by invading microorganisms, immune complexes, cytokines, chemotactic factors, and adhesion to tissue surfaces and activation of ICAMs. The processes of degranulation and secretion begin with the onset of phagocytosis. A number of morphological changes occur within the granules and they translocate and fuse their membranes with those of the phagocytic vacuoles formed by the invagination of the plasma membrane.

Cytoskeletal proteins are essential for this process, facilitating granule transfer to plasma membrane or phagolysosomes. A number of soluble molecules capable of provoking the release of granule content into the extracellular space have been described. These include many chemotactic factors such as C5a, FMLP, platelet-activating factor (PAF) and LTB_4, as well as phorbol myristate acetate and several non-chemotactic interleukins, cytokines and various growth factors. Importantly, phagocytes are capable of rapidly replenishing cellular stores of proteins and forming new granules, thus allowing repeated degranulation and secretion.

Phagocytic killing: the respiratory burst

Activation of phagocytes is associated with a rapid and dramatic increase in oxygen consumption described as the respiratory burst, forming the highly reactive hydroxyl radical ($OH^{\bullet}$), which is highly microbicidal. This and other reactive oxidative products not only contribute to microbial killing, but also activate metalloproteinases such as elastase and collagenase, leading to the surrounding tissue injury that often accompanies phagocyte activation.

The NADPH oxidase is a multicomponent enzyme, with several subunits located in various regions of the quiescent phagocytes (Figure 14.1). Activation of NADPH oxidase results from PKC-mediated phosphorylation of p47 phox and the subsequent binding of the p47 phox phosphoprotein and the p67-phox–p40-phox complex to the membrane flavocytochrome b_{558} located primarily in the specific granules, gelatinase granules and secretory vesicles of neutrophils. The cytochrome b_{558} is the terminal component of the superoxide-generating system, allowing the efficient transfer of electrons from the cytoplasmic NADPH to the surface of phagolysosome or the extracellular surface, where oxygen is reduced to the superoxide anion. Other GTP-binding proteins such as Rap1A and Rac2 associate with the above molecules and regulate phagocyte activation and oxidase activity through their preferred substrate GTP (Figure 14.1).

Phagocytic killing: nitric oxide

Generation of reactive nitrogen intermediates through nitric oxide synthase (NOS) is another mechanism for phagocytic killing of microbes. Nitric oxide (NO) is the highly reactive free radical product of the oxidation of L-arginine in phagocytes during inflammation. Three distinct isoforms of NOS have been described in blood cells. Two are found in endothelial and neuronal cells and the third, inducible NOS, is induced by cytokines, such as IFN-γ, in a number of cell types. NO has an important role in the antimicrobial activity of both neutrophils and mononuclear phagocytes. NOS is located in the cytoplasm and mediates defence against facultative intracellular pathogens, as well as other prokaryotic and eukaryotic pathogens.

Phagocytic killing: antimicrobial proteins

Phagocytic cells such as neutrophils can kill microorganisms using proteins present in various granules. The contents of each phagocytic cell type are specific. Neutrophil antimicrobial proteins include defensins, serpocidins (including cathepsin G and azurocidin) and bacterial permeability-increasing protein (BPI). The antimicrobial proteins within neutrophils are summarized in Table 14.3.

These microbicidal proteins exert their killing effect through enzymatic processes such as proteolysis or by non-catalytic mechanisms, and the combination of such events potentiates bactericidal activity. Cathepsin G is a serine protease that exerts its bactericidal action by binding to penicillin-binding proteins of bacteria and interfering with the synthesis of peptidoglycans. Azurocidin is a serine protease effective against a number of bacteria and fungi. Defensins constitute as much as 50% of neutrophil granule protein content and exert their cidal activity by inserting into hydrophobic channels, forming voltage-dependent ion channels in the lipid bilayer. BPI kills Gram-negative bacteria by binding to their LPS capsule and

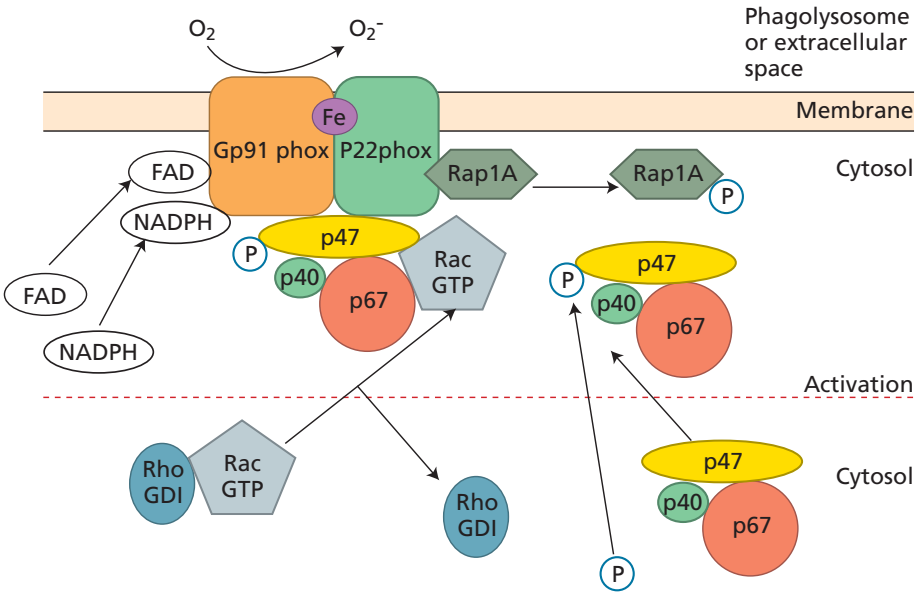


Figure 14.1 Components of the NADPH oxidase system. The components include a 47-kDa cytosolic protein (p47), a 67-kDa cytosolic protein (p67), a 40-kDa cytosolic protein (p40), cytosolic G-proteins (Rac and Rap1) and a membrane-bound cytochrome (b_{558}). The cytochrome consists of haem-containing p22-phox and gp91-phox. The gp91 subunit is an FAD-dependent flavoprotein shuttling electrons to molecular oxygen, forming O_2^-. The p47 component can be phosphorylated to various extents. In activated cells, the p40, p47 and p67 proteins translocate to the membrane to form an activation complex with cytochrome b_{558}. Similarly, the Rac and Rap1 proteins also translocate. The activated oxidase passes electrons from NADPH via FAD to oxygen, thereby generating superoxide.

Table 14.3 Neutrophil microbicidal proteins.

Protein	Characteristics	Target organisms	Effects on target
Non-enzymatic proteins			
BPI	Highly cationic, neutralizes LPS, most potent cidal protein, not released from granule	Gram-negative bacteria	Binds lipid A region of LPS, increases bacterial membrane permeability, activates bacterial degradative enzymes
Defensins	Comprise 30–50% of azurophil granule protein	Gram-positive > Gram-negative bacteria, fungi, viruses, mammalian cells	Increases membrane permeability
Lactoferrin	Cationic, stimulates hydroxyl radical formation	Gram-positive and Gram-negative bacteria, fungi	Oxidative damage
Catalytic proteins and analogues			
Proteinase 3	Serine proteinase	*Escherichia coli, Streptococcus faecalis, Candida albicans*	Growth inhibition
Cathepsin G	Serine proteinase	Gram-positive and Gram-negative bacteria, fungi	Inhibition of peptidoglycan synthesis
Azurocidin	Serine proteinase	Gram-negative bacteria	Non-catalytic mechanisms
Lysozyme	Cationic	Gram-negative bacteria, few Gram-positive bacteria	Potentiation of complement and H_2O_2 killing, cleavage of cell wall peptidoglycans
Elastase	No direct cidal activity		Coactive with lysozyme, potentiation of MPO–halide–H_2O_2 system

BPI, bacterial permeability-inducing factor; LPS, lipopolysaccharide; MPO, myeloperoxidase.

altering their bacterial membrane permeability to extracellular solutes. Other neutrophil granule proteins include lactoferrin which binds iron and kills some Gram-negative bacteria by generating free radicals, and lysozyme, which is involved in the digestion of killed bacteria in phagolysosomes of neutrophils.

Production, structure and dysfunction of phagocytes

White blood cells are produced from pluripotent haemopoietic stem cells located within the bone marrow. Development of white blood cells along different lineages is governed by external stimuli, including cytokines, matrix proteins and other cellular products within the marrow environment. The combination of specific cytokines and growth factors influences the maturation of white blood cell progeny along specific lineages.

Neutrophils (Figure 14.2a)

Development and function

Neutrophils are the predominant white blood cells involved in phagocytic killing of bacteria and certain fungi. They are also referred to as polymorphonuclear or segmented leucocytes, owing to their characteristic lobulated nucleus (their nucleus is segmented into two to five lobes, connected by thin chromatin strands). They represent the terminal stage of maturation (Figures 14.3 and 14.4) and are generally uniform in size (13 μm in diameter), with pink cytoplasm and fine azurophilic granules. The production of neutrophils involves the action of a variety of growth factors, including granulocyte colony stimulating factor (G-CSF), granulocyte macrophage CSF (GM-CSF), interleukin 3 (IL-3) and macrophage colony-stimulating factor (M-CSF).

Granulopoiesis is the process of terminal differentiation from a pluripotent haemopoietic stem cell via a multipotent common myeloid progenitor and bipotent granulocyte–macrophage progenitor stage to a committed mature neutrophil. Granulocyte differentiation depends on the coordinated expression of certain transcription factors, including Pu.1, CCAAT enhancer-binding protein (C/EBPα), C/EBPε and GFI1.

Neutrophils contain four types of granules that can be identified by marker enzymes or proteins (Table 14.4). The lysozyme-like azurophilic granules, otherwise known as primary granules, are present in promyelocytes and all further stages of neutrophil differentiation and contain microbicidal proteins and acid hydrolases (e.g. myeloperoxidase, defensins,

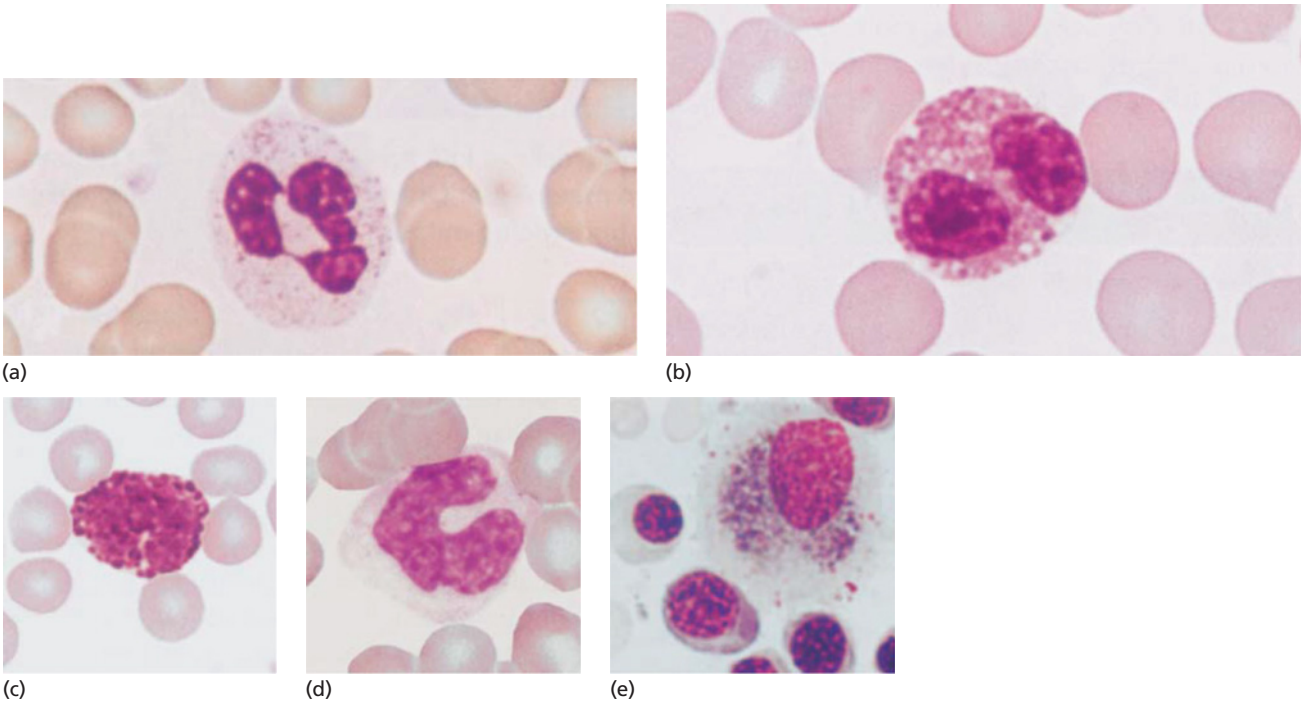


Figure 14.2 Morphology of phagocyte cell types: (a) neutrophil; (b) eosinophil; (c) basophil; (d) monocyte; (e) macrophage.

lysozyme) involved in oxidative and non-oxidative killing of bacteria and fungi. Specific or secondary granules are smaller than azurophilic granules and contain other distinct hydrolases, as well as chemotactic, opsonic and adhesion protein receptors. They release their contents into both phagocytic vesicles and the extracellular medium. Haptoglobin has been identified in specific granules of mature neutrophils. The release of haptoglobin from specific granules after neutrophil activation in sites of inflammation would act to reduce local tissue injury and bacterial growth. Other granules, collectively known as tertiary granules, include secretory vesicles, which contain alkaline phosphatase, and gelatinase granules rich in gelatinase. Degranulation of neutrophils begins with the onset of phagocytosis and involves their translocation and fusion with phagocytic vacuoles created by invagination of the plasma membrane. Degranulation may also occur by reverse endocytosis as a result of the action of complement, aggregated immunoglobulin or certain cytokines.

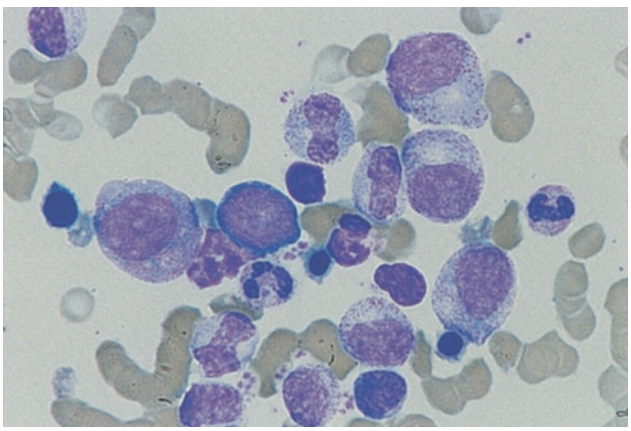


Figure 14.3 Stages of neutrophil maturation, showing a myeloblast, a promyelocyte, several myelocytes and metamyelocytes, a band cell and a segmented neutrophil.

Neutrophil clearance of infectious organisms is critical to the immune response and neutrophil function is impaired in the septic state. Peroxisome proliferator-activated receptor (PPAR)-γ, a ligand-activated nuclear transcription factor, has been shown to be constitutively expressed on isolated human neutrophils and upregulated in the presence of inflammatory cytokines. Inhibition of PPAR-γ activation has been shown to restore *in vitro* neutrophil chemotaxis.

Neutrophils exist in one of three states: quiescent, activated or primed. They circulate in the blood in the quiescent state and react weakly to stimuli, thus limiting potential damage to vascular walls. Priming of neutrophils is a process that does not immediately stimulate an effector response, but allows an exaggerated response upon later stimulation. Three main types of agonists are responsible for priming neutrophils, including chemotactic inflammatory mediators, serum immunoglobulins and complement opsonins, and inflammatory cytokines and growth factors. Upon neutrophil activation, a significant increase in

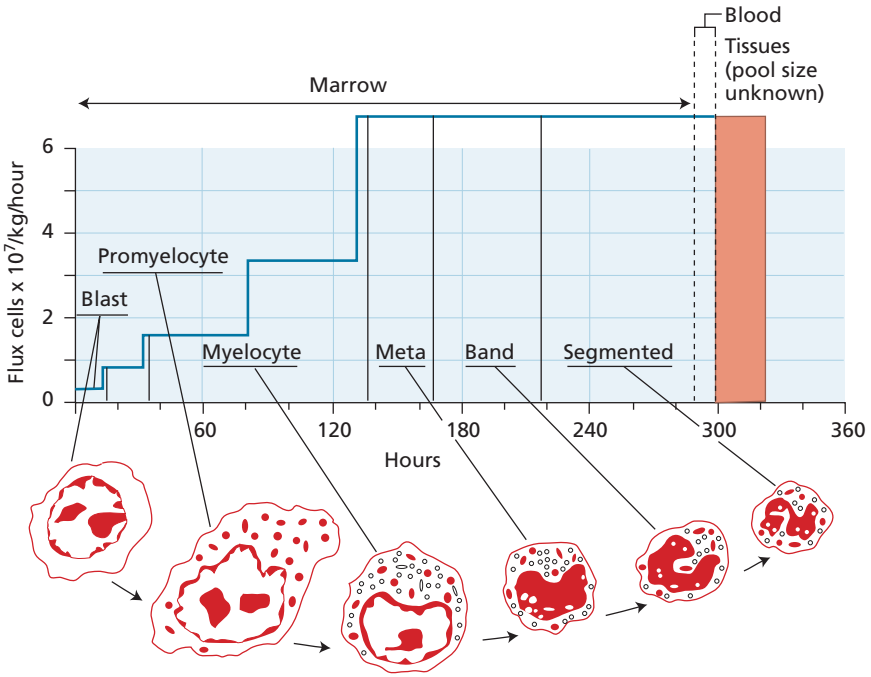


Figure 14.4 Neutrophil lifespan and stages of maturation. Of every 100 nucleated cells in the bone marrow, 2% are myeloblasts, 5% promyelocytes, 12% myelocytes, 22% metamyelocytes and bands, and 20% mature neutrophils (i.e. about 60% developing neutrophils). The times indicated for the various compartments were obtained by isotopic labelling techniques. The ordinate shows the flux, and the abscissa the time, in each compartment. The stepwise increase in cell numbers through the dividing compartments represents serial divisions. Note that no mitoses occur after the myelocyte stage.

oxygen consumption, termed the respiratory burst, occurs that leads to the production of reactive oxygen species responsible for microbial killing.

Neutrophils are the most numerous leucocytes, comprising 65% of circulating phagocytes, with a normal range in the peripheral blood of 1.5–7.7 × 10^9/L, please note that different ranges occur in different ethnic populations (see p. 256 and p. 854). The term 'granulocyte' refers to neutrophils, eosinophils and basophils. Neutrophils are made in the bone marrow at a rate of 10^{12} per day, and once released into the circulation have a half-life of 6–8 hours. The largest proportion of neutrophils are present within the marrow (reserve pool), with circulating and tissue pools comprising smaller fractions (Figure 14.4). The circulating pool itself consists of a marginated pool of cells that are loosely adherent to the vascular endothelium, and a freely circulating pool with the compartments in a constant state of dynamic equilibrium. Several factors, including corticosteroids, exercise and infection, can lead to an increase in the free circulating pool. Corticosteroids promote the release of neutrophils from the reserve pool into the circulation and prevent migration from the blood into the tissue pool. On the other hand, endotoxin and some complement components (C5a) result in increased margination and a reduction in the circulatory pool (Figure 14.4). The variations in neutrophil morphology are shown in Figure 14.5. These include (i) Barr body, a drumstick appendage to the neutrophils in females; (ii) Pelger–Huët anomaly, with bilobed nuclei; (iii) Alder–Reilly anomaly; (iv) May–Hegglin anomaly; (v) toxic granulation; (vi) hypersegmented neutrophils and (vii) Chédiak–Higashi syndrome.

Neutrophils undergo apoptosis within 24 hours of leaving the bone marrow through either intrinsic or extrinsic mechanisms. G-CSF plays an integral role in both the recruitment of neutrophils from the bone marrow and the inhibition of mature neutrophil apoptosis. Calpains are calcium-dependent cysteine proteases that activate proapoptotic factors such as Bax, and inhibit antiapoptotic factors such as X-linked inhibitor of apoptosis (XIAP). G-CSF slows the influx of extracellular calcium and thus downregulates calpain activity, resulting in inhibition of caspase-3, the executioner of apoptosis. Neutrophil apoptosis is characterized by pyknocytosis and the loss of expression of L-selectin and CD16, and the increased expression of CD11b/CD18. Certain drugs, cytokines and mediators of inflammation can delay/inhibit apoptosis in neutrophils, including glucocorticoids, G-CSF, GM-CSF, IL-3, IL-6, IL-15, endotoxin, TNF-α and IFN-γ.

Disorders of neutrophil function and number

Neutrophilia

Leucocytosis, or an increased white blood cell count, may be due to either a primary (congenital or acquired) marrow disorder or secondary to a disease process not involving marrow, toxin or drug (Table 14.5). Neutrophil counts are high in neonates and decrease to normal adult levels with ageing. Secondary leucocytosis not associated with leukaemia, but with a very high white cell count (>50 × 10^9/L), is often referred to as a 'leukaemoid reaction' and can be associated with the presence of Döhle bodies and toxic granulation within the cytoplasm (Figure 14.5), as well as with a 'left shift' (with the presence in blood of myelocytes, metamyelocytes and band forms)

Table 14.4 Neutrophil granules and their contents.

Granule	Azurophilic (primary)	Specific (secondary)	Gelatinase (tertiary)	Secretory vesicles
Marker enzyme	Myeloperoxidase	Lactoferrin	Gelatinase	Alkaline phosphatase
Membrane	CD63, granulophysin, CD68, V-type H^+-ATPase	CD15, CD66, CD67, CD11b/CD18, cytochrome *b*, fMLP-R, fibronectin-R, G-protein α-subunit, laminin-R, NB-1 antigen, 19-kDa protein, 155-kDa protein, Rap-1, Rap-2, SCAMP, thrombospondin-R, TNF-R, urokinase-type plasminogen activator-R, VAMP-2, vitronectin	CD11b/CD18, cytochrome *b*, diacylglycerol-deacylating enzyme, fMLP-R, SCAMP, urokinase-type plasminogen activator-R, VAMP-2, V-type H^+-ATPase	CD10, CD13, CD45, CD14, CD16, CD35 (CR1), CD11b/CD18, alkaline phosphatase, fMLP-R, SCAMP, urokinase-type plasminogen activator-R, V-type H^+-ATPase, VAMP-2, C1q-receptor, decay activating factor
Matrix, microbicidal	Myeloperoxidase, nitric oxide synthase, lysozyme, BPI protein, defensins, serpocidins, elastase, cathepsins, proteinase 3, azurocidin (CAP 37)	Lactoferrin, lysozyme	Lysozyme	
Matrix, hydrolases	Acid β-glycerophosphatase, α-mannosidase, β-glucuronidase, β-glycerophosphatase, *N*-acetyl-β-glucosaminidase, sialidase	Gelatinase, collagenase, histaminase, heparanase, NGAL, sialidase	Gelatinase, acetyltransferase	
Matrix, other	Acid-mucopolysaccharide, heparin-binding protein	β_2-Microglobulin, urokinase-type plasminogen activator, vitamin B_{12}-binding protein	β_2-Microglobulin,	Plasma proteins (including tetranectin)

fMLP, f-Met-Leu-Phe; R, receptor; SCAMP, secretory carrier membrane protein; VAMP, vesicle-associated membrane protein.

and an elevated leucocyte alkaline phosphatase score (compared with a low score in chronic myeloid leukaemia). In contrast with acute leukaemia, there is orderly maturation and proliferation of all normal myeloid elements in the bone marrow. Leukaemoid reactions have been described in patients with osteomyelitis, empyema, septicaemia, tuberculosis, Hodgkin disease, juvenile rheumatoid arthritis and dermatitis herpetiformis.

Leucocyte adhesion deficiency (LAD) is a congenital disorder that presents with persistent leucocytosis, delayed separation of the umbilical cord, recurrent infections, impaired wound healing and defects of neutrophil activation. The condition is caused by defects in adhesion of neutrophils to blood vessel walls. As a result, phagocytes do not migrate from the bloodstream to sites of infection. Two types of LAD have been described. In LAD type I, mutations of the gene encoding the β-subunit of the β_2 integrins (CD11b and CD18) have been detected. The molecular basis for the rare LAD type II is defective glycosylation of ligands on leucocytes recognized by the selectin family of

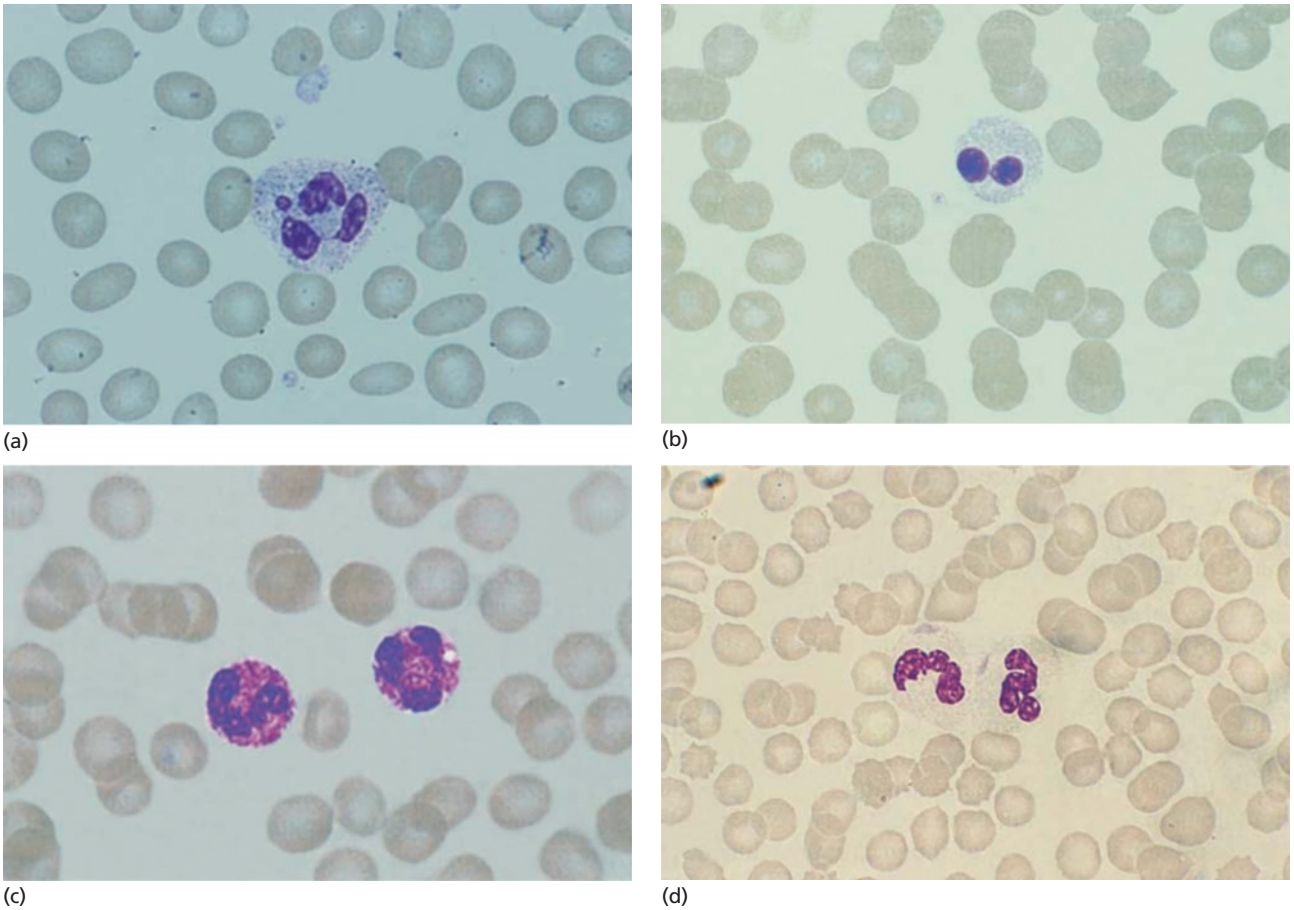


Figure 14.5 Variations in neutrophil morphology. (a) Barr body, a drumstick appendage to the nucleus. (b) Pelger–Huët anomaly with a bilobed nucleus. (c) Alder–Reilly anomaly with prominent purple granules (also in monocytes and lymphocytes). (d) May–Hegglin anomaly with Döhle bodies in the cytoplasm. (*Continued*)

adhesion molecules. The clinical features in the two types are similar, and because of the defect in neutrophil migration, abscesses and other sites of infection are devoid of pus, despite the striking neutrophilia. Treatment involves the use of prophylactic antibiotics and aggressive therapy of periodontal disease. Allogeneic stem cell transplantation from a matched related donor or an unrelated cord blood has resulted in outstanding outcomes with sufficient mixed chimerism to prevent recurrent infections.

Chronic idiopathic neutrophilia is an association of a chronically elevated neutrophil count (in the range $11–40 \times 10^9$/L) in healthy individuals without any associated clinical problems.

The neutrophilia in Down syndrome is transient, but may be exaggerated in response to stress. The transient myeloproliferative disorder, or 'transient leukaemia', can be seen in up to 10% of children born with Down syndrome and is characterized by clonal proliferation of myeloid blasts that is asymptomatic and self-limiting in most cases. Approximately 20% of neonates with Down syndrome will develop an aggressive form of acute myeloid leukaemia (mostly FAB M7-like) requiring intensive chemotherapy. Varying degrees of leucocytosis characterizes chronic myeloid leukaemia (CML), which is a myeloproliferative neoplasm expressing the Philadelphia chromosome or the *BCR–ABL1* fusion gene. Other myeloproliferative neoplasms that are characterized by isolated leucocytosis include chronic neutrophilic leukaemia and an atypical form of chronic myeloid leukaemia (BCR-ABL1 negative), which have been shown recently to be associated with autoactivating mutations in the receptor for GCSF (CSF3R) in 89% and 40% of patients, respectively.

Familial cold urticaria and leucocytosis is a syndrome of fever, urticaria, and muscle and skin tenderness on exposure to cold that appears to be dominantly inherited. The onset of the disease is in infancy, with urticaria, rash and leucocytosis generally occurring several hours after cold exposure. The skin rash is histologically characterized by intense infiltration by neutrophils.

Non-malignant causes of neutrophilia include acute infections, with elevated counts in most bacterial infections. In

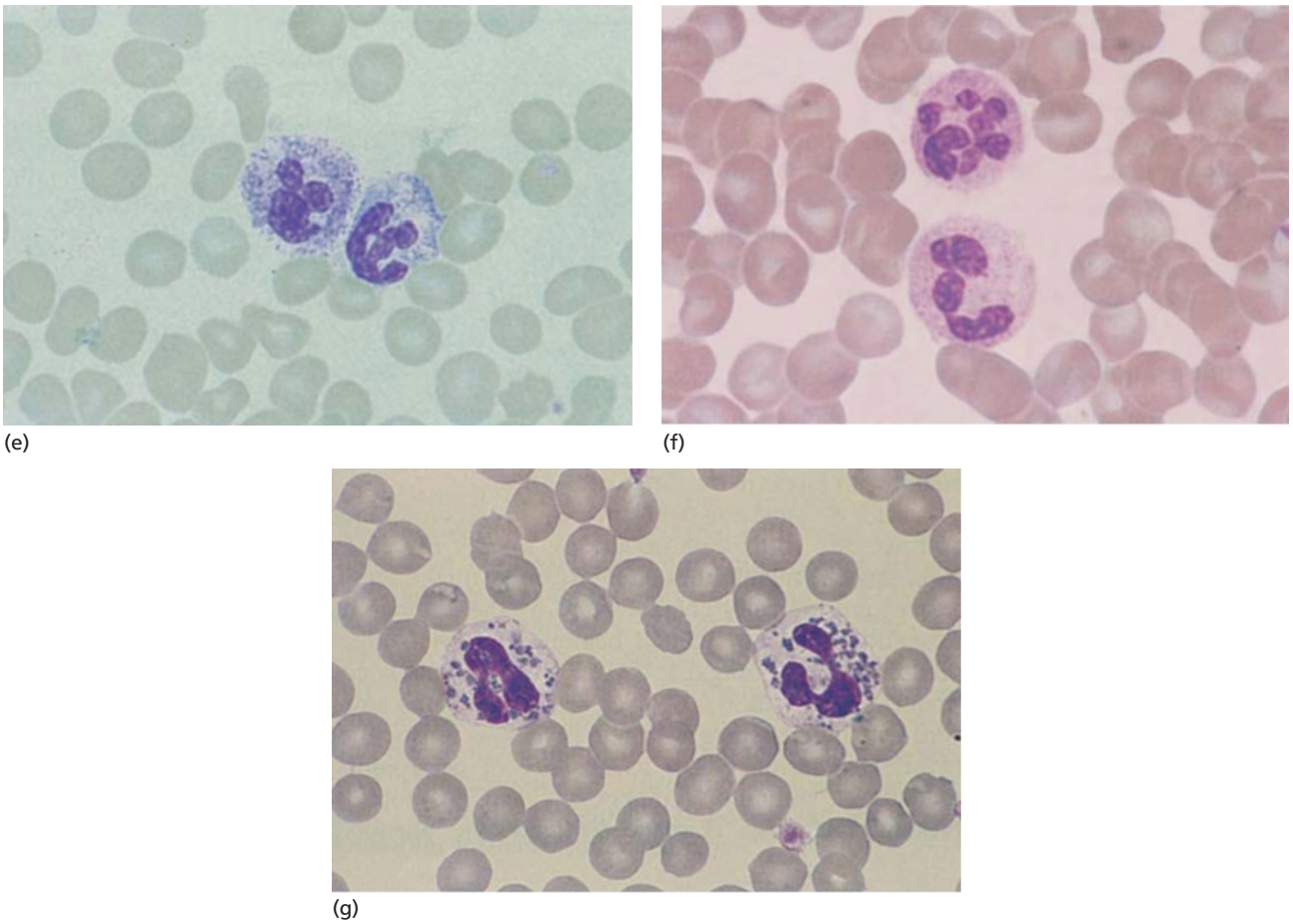


Figure 14.5 (*Continued*) (e) Toxic granulation. (f) Neutrophil with nuclear hypersegmentation and a normal neutrophil. (g) Chédiak–Higashi syndrome with giant granules.

response to overwhelming infection, marrow depletion of leucocytes can occur, resulting in neutropenia rather than neutrophilia. Neutrophilia in response to chronic inflammatory processes is usually more modest in degree and can be accompanied by monocytosis. Modest elevations of neutrophil counts are commonly seen in various forms of 'stress', such as exercise, adrenaline injection, myocardial infarction, in the postoperative period, in postictal states and with emotional distress. This is probably due to the migration of neutrophils from the marginated pool to the circulatory pool. Mild neutrophilia has also been reported with unipolar depression. A number of drugs and drug reactions are commonly associated with increased neutrophil counts. Steroids stimulate the release of neutrophils from the bone marrow and diminish their egress from the circulation, resulting in chronic neutrophilia. This can be distinguished from neutrophilia due to infection by the distinct lack of band forms in the former. The β-agonists produce an acute neutrophilia by releasing neutrophils from the marginated pool. Other drugs known to produce neutrophil leucocytosis include lithium, which increases the production of CSF and potentiates its effects on myeloid colony formation, and tetracycline, which have been associated with counts as high as 80×10^9/L.

Neutropenia

Neutropenia can be due to impaired production by the bone marrow, a shift from the circulating pool to marginated pool, increased peripheral destruction or a combination of these (Table 14.6). It has been defined as an absolute neutrophil count (ANC) of more than two standard deviations below a normal mean value. There is variation of neutrophil counts among different ethnic groups, with people of African origin generally having slightly lower counts (lower limit of normal 1.2×10^9/L) compared with Caucasians (lower limit of normal 1.5×10^9/L). The lower count in African populations has been attributed to a relative decrease in the size of the marrow storage pool. In patients whose neutropenia is related to decreased production, the propensity to develop infections is directly related to the degree and duration of neutropenia (see also p. 854).

On the other hand, in patients whose neutropenia is due to peripheral destruction or margination of neutrophils, there is

Table 14.5 Causes of neutrophilia.

Primary
Hereditary
Chronic idiopathic
Familial myeloproliferative disease
Leukaemoid reaction associated with congenital anomalies
Leucocyte adhesion deficiency (LAD) types I and II
Familial cold urticaria and leucocytosis
Secondary
Infection
Stress
Chronic inflammation
Drugs (steroids, lithium, tetracycline)
Non-haematological neoplasms
Asplenia and hyposplenism
Neoplastic
Chronic myeloid leukaemia
Other myeloproliferative neoplasms (myelofibrosis, polycythaemia vera, essential thrombocythaemia, chronic neutrophilic leukaemia, atypical CML)

Table 14.6 Causes of neutropenia.

Decreased production
Inherited
Reticular dysgenesis
Dyskeratosis congenita
Schwachman–Diamond–Oski syndrome
Cyclic neutropenia
Kostmann syndrome
Hyper-IgM syndrome
Chronic idiopathic neutropenia
Acquired
Aplastic anaemia
Bone marrow infiltration (leukaemia, lymphoma, tumours, tuberculosis, etc.)
Severe infection
Drug induced (cytotoxic chemotherapy, radiation, chloramphenicol, penicillins, cephalosporins, phenothiazine, phenylbutazone, gold, antithyroid drugs, quinidine, anticonvulsants, hydrochlorothiazide, bumetanide, alcohol)
Myelodysplastic syndrome
Vitamin B_{12} or folate deficiency
Pure white cell aplasia
T-γ lymphocytosis and neutropenia
Neutropenia associated with metabolic disorders
Acute leukaemia
Increased peripheral destruction
Hypersplenism
Immune mediated
Drug induced
Associated with collagen vascular disease (Felty syndrome, systemic lupus erythematosus)
Complement mediated (haemodialysis, cardiopulmonary bypass)
Altered distribution
Drugs
Stress

no direct correlation between the degree of neutropenia and the propensity for infections. Neutropenia due to marrow failure or associated with chemotherapy can predispose patients to severe life-threatening infections, although this is more likely in patients with neutrophil counts below 0.5×10^9/L. Common organisms encountered in this setting are *Staphylococcus aureus*, *Pseudomonas aeruginosa*, *Escherichia coli* and *Klebsiella* species. On the other hand, patients with some congenital or immune forms of neutropenia can tolerate low counts for prolonged periods without any apparent increase in the incidence of infections. Moderate asymptomatic neutropenia associated with specific ethnic groups such as African-Americans and Yemenite Jews is associated with a dominant inheritance pattern and requires no specific intervention.

Drug-induced neutropenia is probably the most common cause of isolated neutropenia. A thorough evaluation of the medication history of a patient with neutropenia is important for excluding drugs as the inciting factor. Usually the mechanisms involve suppression of bone marrow activity or are immunological in nature. The neutropenia commonly develops 1–2 weeks after initiation of the drug and resolves soon after discontinuation of the offending agent. The use of cytotoxic agents in cancer therapy is frequently associated with the development of neutropenia. Other potential causes of neutropenia are listed in Table 14.6.

Several well-defined inherited syndromes associated with neutropenia have been described and are worth mentioning.

Severe congenital neutropenia (SCN) is a multigene heterogeneous group of disorders characterized by severe neutropenia at birth (ANC $< 0.5 \times 10^9$/L) and maturation arrest at the developmental stage of the promyelocyte/myelocyte. Kostmann syndrome is an autosomal recessive form involving mutations of the *HAX1* gene encoding a mitochondrial-associated protein with structural similarity to other antiapoptotic BCL-2 family members. Autosomal dominant and sporadic cases are more closely linked to gene mutations of *ELA2*, which codes for neutrophil elastase 2. Less common mutations in genes encoding p14, G6PC3, GFI1, CSF3R, TAZ and WASP have all been documented in SCN. SCN is usually diagnosed in infancy in the setting of recurrent severe infections. Monocytosis, eosinophilia and hypergammaglobulinaemia are often present. Myelopoiesis is blocked and the increased number of promyelocytes found in the bone marrow are characterized by atypical nuclei and cytoplasmic vacuolization. SCN associated with *HAX1* and *ELA2*

has an increased risk of leukaemic transformation that is not observed in cases of SCN linked to the other gene mutations. Importantly, the risk of developing infection is independent of the underlying genetic defect and can be correlated to the ANC nadir. First-line treatment of infections in such patients is antibiotics with G-CSF. G-CSF is the cornerstone of therapy in reducing the infectious risk. Stem cell transplantation offers the potential of cure for those who do not respond to antibiotics and G-CSF. The Severe Chronic Neutropenia International Registry has identified a group of patients with SCN who were less responsive to G-CSF in achieving and maintaining an adequate ANC; these patients have a cumulative incidence of developing MDS/AML or dying of sepsis at 10 years of 40% and 14%, respectively. These patients should be considered for haemopoietic stem cell transplantation earlier in life. The leukemic transformation is associated with the acquisition of mutations in the receptor for *CSF3R*. These mutations are primarily nonsense autoactivating mutations, causing the truncation of approximately 100 amino acids of the cytoplasmic domain of the receptor. The time between the initial detection of *CSF3R* mutations and the diagnosis of MDS/AML varies between several months to up to more than 10 years.

Cyclic neutropenia is a rare autosomal dominant disorder with variable expression that is characterized by repetitive episodes of fever, pharyngitis, stomatitis and other bacterial infections attributable to recurrent severe neutropenia occurring every 15–35 days. The nadir neutrophil count, usually between zero and 200×10^9/L, lasts 3–7 days and is frequently associated with monocytosis. Cycling of red cell and platelet production is also observed in some cases. The bone marrow is characterized by transient arrest at the promyelocyte stage before each cycle. Both childhood and adult onset have been reported. Recent studies have shown that autosomal dominant and sporadic cases of this disease are due to a mutation in *ELA2*, located on chromosome 19p13.3. This enzyme is synthesized in neutrophil precursors early in the process of primary granule formation. This disorder is effectively treated with G-CSF, and transformation to AML or MDS has not been observed in these patients, with or without G-CSF therapy.

Schwachman–Diamond–Oski syndrome is an autosomal recessive disorder characterized by exocrine pancreas insufficiency, metaphyseal dysostosis and bone marrow dysfunction (see Chapter 10). Reticular dysgenesis is associated with neutropenia, lymphoid hypoplasia, and thymic hypoplasia with normal erythropoiesis and megakaryopoiesis. Patients have a hypoplastic marrow and low levels of IgM and IgG and die from overwhelming infections usually in early infancy. Dyskeratosis congenita is a rare disease characterized by abnormal skin pigmentation, nail dystrophy and mucosal leucoplakia. More than 80% of the affected individuals develop bone marrow failure, which is the major cause of death. The disorder is caused by defective telomere maintenance in stem cells. The major X-linked form of the disease is due to mutations in the *DKC1* gene located at Xq28 and coding for dyskerin, a nucleolar protein (see Chapter 10). Neutropenia has been seen with immunological abnormalities such as hyper-IgM syndrome and X-linked agammaglobulinaemia. Hyper-IgM syndrome is an X-linked disorder characterized by lymphoid hyperplasia, low concentrations of IgG and IgA but high concentration of IgM, and severe neutropenia. A genetic defect in the T-cell CD40 ligand has been implicated as the cause of the disease. Patients commonly die of overwhelming infections by the age of 5 years unless treated with intravenous immunoglobulin and long-term G-CSF.

Chronic benign neutropenia, *chronic idiopathic neutropenia* and *autoimmune neutropenia* possess very similar laboratory findings and differ only with regard to age of onset and association with other immune disorders. Chronic benign neutropenia commonly presents in older children or young adults. Patients are usually asymptomatic and have neutrophil counts in the range $0.2–0.5 \times 10^9$/L. Bone marrow examination is commonly normocellular or occasionally moderately hypocellular. They usually have a peripheral monocytosis and often a benign course, although occasional cases progressing to acute leukaemia have been reported. Antineutrophil antibodies, detected in some patients, are not commonly present, but antibodies to the progenitor cells or other precursors may be the inciting factor in these patients. As these patients generally have a benign course, treatment to increase neutrophil counts should be reserved for those who have recurrent infections. Corticosteroids, splenectomy, cytotoxic agents and G-CSF have all been used successfully in this setting. Chronic benign neutropenia of infancy and childhood is probably a related disease, with the majority of patients presenting during the first year of life.

Infants of hypertensive mothers also commonly have moderate to severe neutropenia lasting for several days. This is probably related to bone marrow suppression. Moderate to severe neutropenia can also occur in newborn infants as a result of the transfer of maternal IgG anti-neutrophil antibodies in a manner similar to rhesus haemolytic disease of the newborn. This isoimmune neutropenia develops antenatally and is due to maternal production of antibodies against antigens on fetal neutrophils.

Pure white cell aplasia is a rare condition associated with recurrent pyogenic infections and with thymoma in 70% of the affected patients. There is almost complete absence of myeloid precursors without any abnormality of erythroid or megakaryocytic precursors in the marrow. In the majority of patients, the marrow inhibitory activity is in the IgG and IgM fractions of serum, but in some the inhibition is due to the lymphocytes. The immunoglobulin is directed against progenitor cells or myeloid precursors. The disorder has been associated with therapy with ibuprofen, certain natural remedies and chlorpropamide. If associated with thymoma, surgical removal of the thymus gland can partially correct the neutropenia. Other treatment options include corticosteroids, ciclosporin, cyclophosphamide, intravenous immunoglobulin and allogeneic stem cell transplantation.

Inhibition of granulopoiesis by suppressor or cytotoxic T lymphocytes can occur in patients with collagen vascular disorders, as well as in patients with T-γ lymphocytosis. Patients with T-γ lymphocytosis commonly present with recurrent infections at a median age of 55 years. There is clonal proliferation of either $CD3^+/CD56^+$ T cells or $CD3^-/CD56^+$ natural killer (NK) cells. The characteristic findings include peripheral blood lymphocytosis, with most lymphocytes being large granular lymphocytes. Patients may also have lymphadenopathy and hepatosplenomegaly; the bone marrow is commonly normocellular with increased lymphocytes and arrested myelopoiesis at the myelocyte stage. Patients with the $CD3^+$ subset of the disease most commonly have an indolent course, although most patients need treatment for recurrent episodes of life-threatening infections. G-CSF therapy is used in managing acute infections and methotrexate, prednisolone, cyclophosphamide or ciclosporin can be effective in improving neutrophil counts for sustained periods. In contrast, the $CD3^-/CD56^+$ NK-cell disorders are clinically aggressive, occurring in younger individuals and are associated with a rapidly progressive course unresponsive to combination chemotherapy.

Neutropenia is also associated with collagen vascular diseases such as systemic lupus erythematosus (SLE) and with rheumatoid arthritis. IgG or IgM antibodies may be directed against mature neutrophils or their precursors. Target antigens for antineutrophil antibodies have been identified for both Felty syndrome and SLE. In Felty syndrome, severe neutropenia is associated with rheumatoid arthritis, splenomegaly and leg ulcers. Therapy for neutropenia with methotrexate and ciclosporin has been attempted with variable success. The efficacy of both GM-CSF and G-CSF in reversing neutropenia and decreasing the risk of infections in Felty syndrome and SLE has been well documented.

Neutrophil counts between 0.1 and 0.45 $\times$ 10^9/L are seen in WHIM syndrome, a rare disorder characterized by warts, hypogammaglobulinaemia, infections and myelokathexis. The bone marrow is usually hypercellular with normal-appearing early myeloid forms and abnormal metamyelocytes, bands and mature neutrophils with diploid and tetraploid nuclei. Autosomal dominant inheritance of a CXC-chemokine receptor (CXCR)4 mutation is associated with WHIM syndrome and leads to impaired receptor downregulation. The overexpression of CXCR4 by leucocytes in WHIM leads to impaired egress of mature neutrophils from the bone marrow and resultant neutropenia. Decreased expression of BCL-X in myeloid precursors is also a feature of WHIM and is associated with an increased rate of apoptosis. G-CSF can be used to improve the neutropenia in WHIM syndrome.

Disorders of neutrophil function

A number of congenital and acquired conditions with abnormal neutrophil morphology (see Figure 14.2a) and/or function have been recognized (Table 14.7). Some of these are associated with abnormal neutrophil numbers and have been discussed earlier.

Table 14.7 Disorders of neutrophil morphology and/or function.

Functional defect	Congenital	Acquired
Minimal	Hereditary neutrophil hypersegmentation (AD) Pelger–Huët anomaly (AD) Alder–Reilly anomaly (AR) May–Hegglin anomaly (AD) Myeloperoxidase deficiency (AR)	Megaloblastic hypersegmentation Mucopolysaccharidoses
Adherence/migration	Leucocyte adhesion deficiency (AR) Neutrophil-specific granule deficiency (AR) Schwachman syndrome (?AR) Hyper-IgE syndrome Job syndrome (AR) Familial Mediterranean fever (AR)	Renal failure Diabetes Neonates Malnutrition Leukaemia
Phagocytic killing	Chronic granulomatous disease Papillon–Lefevre syndrome (AR)	Malnutrition Vitamin E deficiency Severe iron deficiency Neonates Diabetes Viral infections Sickle cell disease

AD, autosomal dominant; AR, autosomal recessive.

Chédiak–Higashi syndrome

Chédiak–Higashi syndrome is a rare autosomal recessive disorder characterized by oculocutaneous albinism, recurrent and severe bacterial infections, giant blue-grey granules in the cytoplasm of white blood cells (see Figure 14.5g), a mild bleeding diathesis, progressive peripheral neuropathy and cranial nerve abnormalities. Morbidity results from patients succumbing to frequent bacterial infections or to an 'accelerated phase', a progressive lymphoproliferative syndrome.

Chédiak–Higashi syndrome neutrophils also have a deficiency of antimicrobial proteins as well as disordered degranulation and chemotaxis. Dysfunction of other elements of the immune system, such as cytotoxic T lymphocytes and NK cells, contribute to the propensity for infection and the development of the accelerated phase of the disease. Mutations in the lysosomal trafficking regulator gene, *LYST*, located on chromosome 1q42.1–q42.2, have been implicated as the cause of this disease. At present, treatment for the disorder is allogeneic stem cell transplantation.

Chronic granulomatous disease

Chronic granulomatous disease (CGD) is an inherited disease characterized by severe and recurrent purulent bacterial and fungal infections, including pneumonia, lymphadenitis, hepatic abscesses and osteomyelitis. The majority of patients present in the first year of life with infections with catalase-positive organisms. Phagocytic cells of CGD patients have a defect in the phagosomal and plasma membrane-associated NADPH oxidase, resulting in impaired superoxide formation necessary for efficient bacterial and fungal killing. In addition, a failure to switch off the inflammatory response leads to the formation of granulomas, the distinctive hallmark of the disorder. All the subtypes of X-linked CGD are caused by mutations in the gene for the gp91-phox subunit of cytochrome *b* (*CYBB*) located at Xp21.1. There is significant heterogeneity in the mutations in the gene, with most being family specific. This accounts for the clinical heterogeneity seen in X-linked CGD. In about 30% of patients, this disorder is transmitted in autosomal recessive fashion. Other mutations have been described in this form of the disease (p47-phox, p22-phox, p67-phox and p40-phox).

The incidence of CGD is 1 in 200,000 to 1 in 250,000 live births and the diagnosis is suggested by failure of neutrophils to reduce nitroblue tetrazolium (NBT slide test). The diagnosis can be further established by directly measuring respiratory burst activity as oxygen consumption, oxygen production or H_2O_2 production. Severe deficiency of glucose-6-phosphate dehydrogenase (G6PD) in neutrophils, is a rare X-linked disorder, which can result in a greatly attenuated respiratory burst and a condition resembling CGD. Therapy involves prevention and early treatment of infections, aggressive parenteral antibiotic therapy for established infections, and use of prophylactic trimethoprim–sulfamethoxazole. In a prospective multicentre trial of CGD patients, recombinant human IFN-γ has been shown to augment host defence. Haemopoietic stem cell transplantation is a valid therapeutic option for children with severe manifestations of CGD. In a prospective multicenter study 56 patients with median age of 12.7 years who had high-risk disease underwent reduced-intensity conditioning allogeneic stem cell transplant from related and unrelated donors. The 2 year probability of overall survival was 96% and event-free survival was 91%. Graft failure occurred in 5% of patients.

Myeloperoxidase deficiency

Myeloperoxidase (MPO) deficiency is the most common inherited disorder of phagocytes and is inherited in an autosomal recessive manner. The gene encoding MPO is located at 17q21.3–q23, near the breakpoint of the translocation in acute promyelocytic leukaemia. Despite the key role of MPO in the microbicidal function of neutrophils, persons with MPO deficiency lack any clinical symptoms except for a higher incidence of fungal infections in which case aggressive antifungal therapy is indicated. In most cases of MPO deficiency therapy is not required.

Neutrophil specific granule deficiency

Specific granule deficiency is a rare congenital disorder characterized by recurrent bacterial and fungal infections of skin and lungs. The inheritance is autosomal recessive and although the precise molecular defect has not been elucidated, recent data implicate functional loss of the myeloid transcription factor CCAAT/enhancer-binding protein, C/EBPε, as important in the development of specific granule deficiency. The neutrophils of these patients display atypical bilobed nuclei, lack expression of at least one primary and all secondary and tertiary granule proteins, and possess defects in chemotaxis, disaggregation, receptor upregulation and bactericidal activity.

Papillon–Lefevre syndrome

This is a rare autosomal recessive disorder characterized by palmoplantar keratoderma and early-onset peridontitis. Pyogenic liver abscesses and recurrent skin infections are an increasingly common complication. Consanguinity is common and patients commonly present during the first 6 months of life, with early progressive loss of both primary and secondary dentition. A phagocytic defect in microbicidal activity and degradation of ingested material is thought to be present. This is attributed to loss-of-function mutations of the *CTSC* gene located on chromosome 11q14.2, which encodes the protease cathepsin C. A total of 75 different disease-causing mutations have been published. More than 300 cases have been reported worldwide.

Eosinophils (Figure 14.2b)

Development and function

Eosinophils, which account for 5–10% of leucocytes (0.2×10^9/L), are similar to neutrophils morphologically, except

for the presence of a bilobed nucleus and numerous bright-orange cytoplasmic granules. There are three distinct granule populations: the round, uniformly electron-dense primary granules present mainly in the eosinophilic promyelocyte/myelocyte stages; secondary or specific granules; and the less well-characterized small granules (Table 14.8).

Primary granules contain eosinophil peroxidase and Charcot–Leyden crystal protein. The eosinophil peroxidase is distinct from neutrophil MPO and can mediate damage to microorganisms, and tissues and bronchoconstriction in asthma. The large specific granules contain the eosinophil's cytotoxic and proinflammatory proteins and account for more than 95% of granules in the mature eosinophils, conferring the characteristic appearance of the cell. Eosinophil granules contain a number of enzymes similar to those found in the lysosomes of the neutrophil, but eosinophils lack lysozyme. Different cationic polypeptides are the major constituents of eosinophil granules and include major basic protein (MBP), eosinophil cationic proteins (ECP) and eosinophil-derived neurotoxin (EDN). MBP is toxic to cells, including parasites and mammalian epithelial cells, and evokes release of mediators from basophils and mast cells. Eosinophils have a significant cytotoxic and proinflammatory function and play an important part in the pathogenesis of a number of allergic, parasitic and malignant disease processes.

Eosinophils are derived from bone marrow stem-cell-derived myeloid progenitors in response to a number of T-cell-derived cytokines and growth factors, including IL-3, GM-CSF and IL-5. IL-5 is the most lineage-specific factor and plays an important role in regulation of terminal differentiation and postmitotic activation of eosinophils. Therefore, IL-5 is a late-acting cytokine that is both necessary and sufficient for eosinophil development *in vivo*, a finding that has been confirmed by IL-5 transgenic and IL-5 knockout mice.

In normal individuals, eosinophils exist transiently in the circulation and localize specifically to certain tissues and organs

Table 14.8 Contents of eosinophil granules.

Granule	Protein content	Comment
Primary granule (Charcot–Leyden granule)	Charcot–Leyden crystal (CLC) protein	Weak lysophospholipase, carbohydrate-crystal protein-containing binding properties
Specific granule (secondary granule)	Eosinophil peroxidase	
	Major basic protein (MBP)	Toxic to parasites
	Eosinophil cationic protein (ECP)	Ribonuclease, bactericidal, toxic to parasites
	Eosinophil-derived neurotoxin (EDN)	Ribonuclease, toxic to parasites
	Eosinophil peroxidase (EPO)	Antibacterial
	Lysophospholipase	
	Acid phosphatase	
	Arylsulfatase B	
	Phospholipase A_2 (secretory)	Antibacterial
	BPI protein	LPS binding, antibacterial
	NAMLAA	Antibacterial
	FAD	
	Catalase	
	Urokinase	
	CD63	Tetraspanin
	Proteoglycan	
	α_1-Antitrypsin	
Small-type granule	Arylsulfatase B	Lysosomal hydrolase
Secretory vesicle	Cytochrome b_{558}	NADPH oxidase component
	CR3	β_2 integrin

BPI, bactericidal permeability-increasing protein; FAD, flavin adenine dinucleotide; LPS, lipopolysaccharide; NAMLAA, *N*-acetylmuramyl-L-alanine amidase.

exposed to the external environment. Eosinophils are recruited in response to early- and late-phase components of immediate hypersensitivity reactions, as well as other immunologically mediated reactions. Their activation and recruitment involves the interaction of several adhesion pathways and chemotactic agents, including the complement fragment C5a, PAF, IL-3, IL-5, GM-CSF, IL-2, RANTES, eotaxins and the $CD8^+$ T-cell-derived lymphocyte chemoattractant factor (LCF).

The specificity and intensity of the microbicidal function of eosinophils differ from those of other leucocytes. The major role of eosinophils in host defence is the destruction of metazoan parasites. Eosinophils readily degranulate in response to stimulation by antigens, cytokines and complexed or secretory IgA, IgE and IgG. Proteins released from eosinophils result in histamine release from basophils and mast cells, and amplify the inflammatory response. These proteins are also powerful toxins towards host cells, leading to tissue injury. A further role of eosinophils is tissue repair and remodelling through regulation of the deposition of extracellular matrix proteins.

Disorders of eosinophils

Eosinophilia

Eosinophilia is considered as an absolute eosinophil count of 0.5×10^9/L or more. Blood and tissue eosinophilia can be seen in a number of parasitic, neoplastic, collagen vascular and allergic diseases (Table 14.9). Because of the varied causes of eosinophilia, diagnostic evaluation should include complete physical examination and history, routine chemistries, serum IgE, vitamin B_{12} levels, HIV serology, stool ova and parasites, electrocardiogram, echocardiogram, pulmonary function tests, chest and abdominal computed tomography, and bone marrow biopsy and aspirate.

In these disorders, a variety of abnormal stimuli lead to the increased production and tissue localization of eosinophils. When no underlying cause can be identified, the hypereosinophilic syndrome (HES) may be present. Several reactive pulmonary and cutaneous eosinophilic syndromes (e.g. Churg–Strauss syndrome, Loeffler syndrome and eosinophilic lymphofolliculosis or Kimura disease), as well as an eosinophilia–myalgia syndrome have also been described. Eosinophilia–myalgia syndrome is related to metabolites and contaminants in the preparation of the drug L-tryptophan and presents with peripheral blood eosinophilia, fatigue and severe myalgias. HES is characterized by sustained eosinophilia of 30–70% of total leucocyte count ($> 1.5 \times 10^9$/L) for longer than 6 months, absence of other underlying causes of eosinophilia and evidence of organ dysfunction due to eosinophilic tissue infiltration. Presenting features can include anorexia, weight loss, fever, sweating, thromboembolic episodes, heart failure, splenomegaly, and skin and central nervous system (CNS) disease. Peripheral blood eosinophils have a variety of cellular abnormalities, and bone marrow eosinophils are increased (30–60%), but myeloblasts are usually not. It has been difficult to assess the clonality of HES. Cases that are clonally derived, as demonstrated by clonal karyotypic abnormalities and the performance of restriction polymorphism analysis in females, are now reclassified by the WHO (2008) as chronic eosinophilic leukaemia.

Table 14.9 Causes of eosinophilia.

Infections
Parasitic: helminthic (filariasis, strongyloidiasis, hydatid disease, onchocerciasis, etc.), visceral larva migrans
Non-parasitic: coccidiomycosis, recovery from acute infections, cat scratch disease, *Cryptococcus*
Allergic disease
Atopic diseases
Drug hypersensitivity
Bronchopulmonary aspergillosis
Pulmonary
Allergic bronchopulmonary aspergillosis
Eosinophilic pneumonia
Transient pulmonary infiltrates (Loeffler syndrome)
Prolonged pulmonary infiltrates with eosinophilia
Tropical pulmonary eosinophilia
Allergic granulomatosis (Churg–Strauss syndrome)
Cutaneous
Eosinophilic lymphofolliculosis (Kimura disease)
Bullous pemphigoid
Granulomatous dermatitis with eosinophilia (Wells disease)
Eosinophilic fasciitis (Shulman syndrome)
Atopic dermatitis
Urticaria and angioedema
Connective tissue disorder
Vasculitis
Serum sickness
Eosinophilic fasciitis
Immunological disorders
Wiskott–Aldrich syndrome
Selective IgA deficiency
Graft-versus-host disease
Neoplastic conditions
Eosinophilic leukaemia
Lymphoma (Hodgkin T cell)
Chronic myeloid leukaemia
Acute myeloid leukaemia, M4Eo
Some solid tumours
Miscellaneous
Eosinophilic myalgia syndrome
Toxic oil syndrome
Idiopathic hypereosinophilic syndrome

This clinically variable disorder has various subtypes, including the myeloproliferative variant (10–50%) (this is now classified as chronic eosinophilic leukaemia; see also Chapter 26), lymphoproliferative variant (5–50%), familial form, episodic form and benign form. Patients with the myeloproliferative variant of HES are characterized by male predominance, end-organ damage, elevated serum tryptase, splenomegaly, anaemia, thrombocytopenia and bone marrow myeloproliferation with reticulin fibrosis. This variant was historically recognized as unresponsive to steroid therapy, with a mortality greater than 50%, most commonly due to cardiac dysfunction from endomyocardial fibrosis. The myeloproliferative variant is often associated with an interstitial deletion on chromosome 4q12, resulting in a *FIP1L1–PDGFRA* fusion gene. The protein product of this fusion gene is a tyrosine kinase with enhanced activity and particular sensitivity to low-dose imatinib therapy, leading to durable complete haematological and cytogenetic remission. Even rarer cases with different haematological features are due to rearrangement of *PDGFRB* or *FGFR1* (see Chapter 26). The lymphoproliferative variant of HES is a distinct clinical syndrome characterized by hypereosinophilia occurring in response to IL-5 production by a clonal $CD3^{-}CD4^{+}CD8^{-}$ activated T-cell population. Patients with this variant are more likely to have skin, gastrointestinal and pulmonary involvement and less likely to have endomyocardial fibrosis and myelofibrosis. Unique to this group of patients is the increased incidence of progression to lymphoma. The other variants of HES are rare and include a form with cyclic episodes of eosinophilia and angio-oedema (also referred to as Gleich syndrome) and a rare familial form with predominantly autosomal dominant transmission and gene involvement at 5q31–33.

The goal of treatment in HES is to limit organ damage by controlling the eosinophil count and, except for treatment of *FIP1L1–PDGFRA* associated HES with tyrosine kinase inhibitors, steroids remain the first-line therapy. In patients who are refractory or intolerant to the side-effects of steroids, cytotoxic agents such as hydroxycarbamide (hydroxyurea), vincristine and cytarabine have been used, as well as immunomodulatory therapy with interferon α, and ciclosporin. Mepolizumab is a humanized monoclonal antibody against free IL-5. Several studies have demonstrated the benefit of mepolizumab for various forms of HES. Currently it is considered to be an investigational drug in the US and is available for compassionate use for patients with life-threatening disease who failed other therapies.

Basophils and mast cells (Figure 14.2c)

Development and function

The functions of basophils and mast cells are similar, but not identical. They express the receptor that binds with high affinity the Fc portion of IgE antibodies (FcεIR) and have large metachromatic (purple-black) granules rich in histamine, serotonin and leucotrienes. Basophils have a bilobed nucleus, in contrast to mast cells, which have a unilobed nucleus. More recent studies have demonstrated that despite their significant similarities, basophils and mast cells are terminally differentiated progeny of distinct bone marrow progenitors. Basophils develop from haemopoietic stem cells, mature in the bone marrow and circulate in the blood, whereas mast cells mature in the tissues. They both play significant roles in the development of a number of allergic and inflammatory disorders, as well as host defence mechanisms against parasites.

Mechanisms mediating the maturation of basophils and mast cells differ. Mast cells are derived from $CD34^{+}$, c-Kit-positive progenitors and leave the marrow before their terminal maturation and home to vascularized peripheral tissues where they mature. The major growth and differentiation factor for basophils is IL-3, whereas the growth and development of mast cells requires the presence of SCF.

Basophils are the least abundant leucocytes, accounting for less than 0.5% of bone marrow and peripheral blood leucocytes. Basophils arise from a common basophil–eosinophil progenitor cell, mature in the marrow over a period of 2–7 days and after their release into the circulation persist for up to 2 weeks. They are the key mediators of immediate hypersensitivity reactions such as asthma, urticaria and anaphylaxis. In addition, they have been implicated in the delayed cutaneous hypersensitivity reaction. Basophils are stimulated by a number of mediators, such as IgE, IL-3, C5a, GM-CSF, insect venoms and morphine, to release the contents of their granules such as histamine. Synergistic induction of both proinflammatory and immunoregulatory cytokines by IL-33-expressing resident tissue cells, IL-3-expressing mast cells, Th2 lymphocytes and IgE receptor activation by allergen exposure may provide the mechanism by which basophils are activated without the required involvement of the adaptive immune system.

The interaction between IgE and basophil/mast cell FcεIR and antigen bridging results in basophil degranulation and initiation of their effects of granule contents. The granules of basophils and mast cells contain sulfated glycosaminoglycans, responsible for their intense staining, as well as histamine, LTD_4, PAF, eosinophil chemotactic factor and kallikreins, responsible for type I immediate hypersensitivity reactions and chronic inflammation (Table 14.10). Histamine is derived from histidine by decarboxylation and is stored as a complex with heparin or chondroitin sulfate proteoglycans. The primary protease present in mast cells, tryptase, is mainly released during the early phase of allergic response and is a marker of mast cell activation in chronic inflammatory diseases. Other neutral proteases such as carboxypeptidase B, chymase and sulfatases are also released and degrade extracellular matrix proteins. Basophil and mast cell activation also leads to the elaboration of cytokines such as GM-CSF, TNF-α, IFN-γ, IL-3, IL-4 and IL-5, which serve to amplify the inflammatory response.

Table 14.10 Basophil and mast cell granules, and their contents.

Component	Function	Main physiological role	Other properties	Cell specificity
Protein				
Histamine	Binds to H_1, H_2, H_3 receptors	Hypersensitivity reactions and inflammation		Basophils, mast cells
Proteoglycan				
Heparin	Package of basic proteins into granules		Binds and stabilizes proteases	Connective tissue mast cells
Chondroitin sulfates	Package of basic proteins into granules		Binds and stabilizes proteases	Basophils
Enzymes				
Chymase	Inactivates bradykinin	Affects microcirculation		Connective tissue mast cells
	Activates angiotensin 1	Modulates microcirculation		
	Activates precursor IL-1β	Modulates skin inflammation		
Tryptase	Cleaves C3 to C3a and C3b	Proinflammatory, stimulates neutrophil chemotaxis and adherence	Tetrameric when bound to heparin, monomer inactive, restricted substrate specificity, raised levels in mast cell disorders	Mast cells
	Activates metalloproteinase 3, inactivates fibrinogen, degrades calcitonin gene-related peptide	Regulates collagenase, attenuates fibrin deposition		
Cathepsin G-like protease				Connective tissue mast cell
Carboxypeptidase	Inactivates bradykinin	Affects microcirculation		Connective tissue mast cell
Other				
Charcot–Leyden crystal protein	Lysophospholipase	Phospholipid metabolism	Neutralizes pulmonary surfactant	Basophil
Major basic protein		Disrupts membranes		Basophil
Sulfatase				
Exoglycosidase				

Disorders of basophils and mast cells

High basophil numbers are commonly seen in patients with myeloproliferative neoplasms, in particular CML. Basophil number can be strikingly elevated in patients with CML, accounting for over 20% of circulating leucocytes in the more advanced stages of the disease (Chapter 24). In other myeloproliferative neoplasms, elevation of basophil numbers is usually more modest. Cases of AML with high levels of immature

basophils have also been reported. Rarely, basophils may constitute over 80% of circulating leucocytes, a condition sometimes referred to as basophilic leukaemia. Other causes of basophilia include ulcerative colitis, myxoedema, recovery from acute illness and drug allergies, although these conditions are usually associated only with modest elevations of circulatory basophils.

As discussed earlier, SCF or c-Kit ligand is an important factor in mast cell development. An increased number of tissue mast cells can be seen in a number of disorders, including atopy, parasitic diseases, Hodgkin lymphoma and other lymphoproliferative disorders, certain neoplasms and rheumatoid arthritis.

Several conditions, ranging from isolated cutaneous mastocytomas to mast cell leukaemia, are associated with mast cell proliferation. Solitary mastocytomas generally regress spontaneously. The more common cutaneous mastocytosis or urticaria pigmentosa typically presents with multiple, small, round, reddish-brown maculopapular lesions that, when subjected to minimal trauma, lead to intense pruritus. In some patients, this disease progresses to the systemic variety, with involvement of bone marrow, spleen, liver and the gastrointestinal tract. Systemic mastocytosis can also occur without prior or concurrent cutaneous disease, and in association with haematological disorders, including leukaemias and lymphomas. Organ dysfunction may be secondary to the release of biochemical mediators by mast cells, such as peptic ulcer disease secondary to histamine release. Mast cell leukaemia, a rare condition, presents with circulating mast cells of abnormal morphology (accounting for up to 95% of circulating nucleated cells), peptic ulcer disease, constitutional symptoms, anaemia and hepatosplenomegaly. It should be distinguished from AML, which can develop in association with systemic mastocytosis.

Management of patients within all categories of mastocytosis includes careful counselling of patients and care providers, avoidance of factors triggering acute mediator release, treatment of acute and chronic mast cell mediator release, an attempt to treat organ infiltration by mast cells, and treatment of any associated non-mast-cell haematological disorder. The agents and modalities commonly used in treating patients with mastocytosis include histamine H_2-receptor blockers, adrenaline, steroids, cromolyn sodium, proton pump inhibitors, psoralen with ultraviolet light (PUVA), chemotherapy, radiation, interferon α, ciclosporin, 2-chlorodeoxyadenosine and splenectomy. With increased availability of small-molecular-weight inhibitors of intracellular signalling pathways, targeting of the constitutively active mutated c-*kit* has attracted more attention. Two classes of constitutive activating c-*kit* mutations have been reported. The more frequent mutation occurs in the catalytic pocket coding region, with substitutions at codon 816, and the other in the intracellular juxtamembrane coding region. Unlike in other disorders with *c-kit* mutation, imatinib therapy is not effective. Therefore, kinase inhibitors that block mutated c-*kit* activity are being evaluated as therapeutic agents in systemic mastocytosis.

Monocytes and macrophages (Figure 14.2d,e)

The mononuclear phagocyte system has been defined as a family of cells comprising bone marrow progenitors, blood monocytes and tissue macrophages. Monocytes originate from a pluripotent haemopoietic stem cell that commits to hematopoietic progenitor cells termed CFU-GEMM (colony-forming unit – granulocyte, erythrocyte, monocyte, megakaryocyte) and the more committed CFU-GM (granulocyte, monocyte). The progenitor cells can commit to both the neutrophil and monocytic pathways. Cytokines and growth factors, such as GM-CSF and IL-3, allow commitment along monocytic pathways. Monocyte colony-stimulating factor, also known as colony-stimulating factor (CSF)-1, is the most important factor in the development of monocytes and macrophages, and is necessary, but not sufficient, for their activation. Macrophages stimulated by macrophage colony-stimulating factor can also clear apoptotic cells by macropinocytosis.

Following their release into the circulation, monocytes rapidly partition between the marginating and circulating pools. The circulating monocytes have a highly convoluted surface and a lobulated nucleus. They can be further characterized by non-specific esterase staining and contain a folded, multilobulated nucleus. After migration into tissues, they become larger and acquire the characteristics of tissue macrophages. Monocytes contain lysosomal hydrolases and the intracellular enzymes elastase and cathepsin. After transformation into tissue macrophages they produce predominantly metalloproteases and metalloprotease inhibitors, lose expression of hydrolases, and express macrophage-specific genes and products such as inducible NOS and IFN-γ. Tissue macrophages are long-lived and self-sustaining cells.

Kupffer cells, phagocytic cells residing within the lumen of hepatic sinusoids, represent up to 90% of fixed tissue macrophages and are the first phagocytes to encounter bacteria originating from the colon. Kupffer cells are also implicated in the removal of neutrophils after the clearance of an organism, downmodulating the inflammatory response and abrogating the tissue destruction sometimes seen in overwhelming sepsis.

Macrophage activating factor, identified as IFN-γ, as well as IL-2, IL-4, macrophage colony-stimulating factor and GM-CSF are either directly or indirectly, through IFN-γ, responsible for macrophage activation. They stimulate monocyte/macrophage proliferation, increase adhesion receptor expression and stimulate the production of proteolytic agents responsible for pathogen clearance. The hypothalamic–pituitary–adrenal axis and autonomic nervous system communicate with the inflammatory system via catecholamine-induced activation of macrophage NF-κB and subsequent release of macrophage-derived proinflammatory cytokines.

Although production of IFN-γ by Th1 cells results in a cytocidal macrophage state, IL-4 and IL-13 produced by the Th2 population of T lymphocytes stimulate the antigen-presenting

Table 14.11 Causes of monocytosis.

Inflammatory conditions
Infections
Tuberculosis
Bacterial endocarditis
Fever of unknown origin
Syphilis
Other
Systemic lupus erythematosus
Rheumatoid arthritis
Temporal arteritis
Polyarteritis
Ulcerative colitis
Sarcoidosis
Myositis
Malignant disorders
Acute myeloid leukaemia
Hodgkin lymphoma
Non-Hodgkin lymphomas
Histiocytoses
Carcinomas
Myelodysplastic syndrome
Miscellaneous
Cyclic neutropenia
Chronic idiopathic neutropenia
Kostmann syndrome
Post splenectomy

cell state. These cytokines then enhance macrophage stimulation of T cells by inducing class II MHC antigen and costimulatory molecule expression. Activated macrophages, in turn, produce cytokines that stimulate both types of helper T cells.

Several cytokines, including IL-4, IL-10, IL-13 and TGF-β, can inhibit different aspects of macrophage function. Furthermore, prostaglandin (PG)E_2, elaborated by macrophages, as well as corticosteroids can suppress various actions of macrophages. These provide a feedback mechanism for control of the immune response.

Disorders of monocyte/macrophages

Monocytosis and monocytopenia

Chronic inflammatory conditions, both infectious and immune in nature, are associated with monocytosis (Table 14.11). These include tuberculosis, bacterial endocarditis, syphilis, collagen vascular disease, sarcoidosis and ulcerative colitis. Monocytosis is also commonly seen in a number of haematological malignancies such as AML, Hodgkin lymphoma, non-Hodgkin lymphoma and histiocytosis. A decreased number of circulating monocytes has been reported with endotoxaemia, corticosteroid administration and hairy cell leukaemia.

Histiocytic disorders

Dendritic cells originate in the bone marrow and share a common progenitor with macrophages. Precursors of dendritic cells are released from the bone marrow and enter tissues in which they differentiate into functional antigen-presenting dendritic cells. Tissue-based dendritic cells comprising the dendritic cell system lack phagocytic capacity or Fc receptors and are predominantly antigen-presenting cells. The dendritic Langerhans cells are found in virtually all tissues except the brain and are the major immunological cellular components of the skin and mucosa. Their racquet-shaped ultrastructural inclusions (Birbeck bodies) distinguish them from other tissue cells. They interact with and process antigen, then migrate to lymphoid organs where, through interaction with T cells, they generate cellular and humoral immune responses. This ability of dendritic cells to interact with T cells and other inflammatory cells contributes to the often varied clinical manifestations of the histiocytic disorders.

The histiocytic disorders comprise various haematological disorders, with cells of the mononuclear phagocytic system or the dendritic cell system involved in their pathogenesis. In general, diseases associated with proliferation of histiocytes can be grouped into two different categories: inflammatory disorders and neoplastic (clonal) disorders (Table 14.12). In the more recent classifications by the World Health Organization Committee on Histiocytic/Reticulum Cell Proliferations, other disorders in which histiocytes are implicated, such as storage diseases (Gaucher and Niemann–Pick), have been excluded.

Langerhans cell histiocytosis

The offending cells in the disorders previously referred to as histiocytosis X (including eosinophilic granuloma, Letterer–Siwe disease and Hand–Schüller–Christian disease) have the characteristics of epidermal Langerhans cells. These disorders, now collectively referred to as Langerhans cell histiocytosis (LCH), have variable clinical features depending on the organs infiltrated by the responsible Langerhans and accompanying cells. The true prevalence of these disorders is seven cases per million and typically occur in children under 15 years of age.

It is now clear that LCH is characterized by clonal proliferation of CD1a-positive cells. The Langerhans cells from the lesions of patients have several phenotypic properties that distinguish them from their normal counterparts. Differences in staining by the lectin peanut agglutinin, expression of placental alkaline phosphatase, expression of the IFN-γ receptor, and expression of costimulatory receptors such as CD86 and CD80 allow one to distinguish between LCH lesional cells and their normal counterparts. There is extensive expression of GM-CSF, IL-1, IL-3, IL-4, IL-8, TNF and leukemia inhibitory factor (LIF) in the lesions

Table 14.12 Histiocytic disorders.

Disorders of varied biological behaviour
Related to dendritic cells
Langerhans cell histiocytosis
Juvenile xanthogranuloma and related disorders
Solitary dendritic cell histiocytoma
Related to macrophages
Haemophagocytic syndromes
Primary: familial haemophagocytic histiocytosis
Secondary: infectious, tumour associated, drug associated (e.g. phenytoin)
Rosai–Dorfman disease (sinus histiocytosis with massive lymphadenopathy)
Solitary macrophage histiocytoma
Clonal disorders
Related to monocytes
Leukaemia, acute myelomonocytic and monocytic leukaemia, chronic myelomonocytic leukaemia, extramedullary monocytic tumours
Related to dendritic cells
Histiocytic sarcoma (malignant histocytosis): localized or disseminated
Related to macrophages
Histiocytic sarcoma (malignant histiocytosis): localized or disseminated

of LCH, suggestive of activation of T lymphocytes as well as recruitment of macrophages, eosinophils and granulocytes.

The diagnosis of LCH is based on a biopsy of the involved organs, with the key diagnostic feature being the presence of pathological Langerhans cells, which can be identified by demonstration of either CD1a surface antigen or the presence of Birbeck granules on electron microscopy. Early lesions are generally cellular and locally destructive, with an abundance of essentially normal Langerhans cells. As the lesions mature, there are fewer Langerhans cells with occasional necrosis.

The lesions of LCH occur in skin, bone, lymph nodes, liver, spleen, bone marrow, lungs, the CNS and the gastrointestinal tract. Clinical features are varied and depend on the organs involved. It may involve single organs or involve multiple systems, and assessment of organ function is important as it can have prognostic significance.

Langerhans cells play an integral role in antigen recognition and participate in immune responses within many tissues. Pulmonary Langerhans cell histiocytosis (PLCH) is a non-neoplastic collection of reactive Langerhans cells that causes interstitial lung disease in adult smokers. PLCH is uncommon and comprises 5% of all interstitial lung disease, predominantly afflicting middle-aged men and women smokers. PLCH typically presents with dyspnoea, non-productive cough and constitutional symptoms, and radiographic features of reticulonodular changes with cyst formation. Langerhans cells in PLCH can be identified by positive immunostaining for CD1a, langerin, E-cadherin and S-100, as well as the presence of Birbeck granules. The treatment of PLCH is smoking cessation. Reports can be found in the literature supporting the use of corticosteroids and chemotherapy for progressive disease. Prognosis is variable and cases of spontaneous regression, even without intervention, have been documented.

Solitary or multifocal eosinophilic granuloma occurs mainly in older children and young adults and accounts for the majority of cases of LCH. Hand–Schüller–Christian disease occurs in younger children (2–5 years) and often presents with exophthalmos due to a tumour mass in the orbital cavity. Letterer–Siwe disease is the rarest and often most severe form of LCH, typically presenting with a scaly, seborrhoeic, eczematoid and, occasionally, purpuric or ulcerative rash in infants younger than 2 years. Bone involvement in LCH can lead to a tender swelling (as a result of infiltration of adjacent tissues) and inability to bear weight. Radiographically, lesions appear as 'punched-out' holes, sometimes with sclerotic edges. Other clinical manifestations include rashes, which may be maculopapular, nodular or vesicular; ear discharge; lymphadenopathy; diabetes insipidus, due to involvement of the hypothalamus and the pituitary; respiratory symptoms, with radiographic changes such as micronodular infiltrates due to lung involvement; hepatomegaly with laboratory evidence of liver dysfunction; splenomegaly; CNS disease, with ataxia, dysarthria and cranial nerve palsies and, rarely, gut involvement with diarrhoea, malabsorption and protein-losing enteropathy.

Spontaneous resolution can occur in a significant proportion of patients with LCH, and patients with limited disease usually do not require systemic therapy. In contrast, patients with multifocal disease generally benefit from systemic therapy. Treatment options have included low-dose radiation for symptomatic single lesions, local injection of steroids, topical steroids, PUVA, non-steroidal anti-inflammatory drugs, high-dose systemic steroids and combination chemotherapy, including agents such as prednisolone, vinblastine, vincristine, etoposide and 6-mercaptopurine. Other treatments that have been tested in patients with disease progression include ciclosporin, antithymocyte globulin, 2-chlorodeoxyadenosine, thalidomide, TNFα inhibitors, anti-CD1a antibodies and haemopoietic stem cell transplantation. BRAF V600E mutations have been described in 38–69% of patients with LCH. In patients that have this mutation and are resistant to treatments, the BRAF inhibitor vemurafenib has produced promising results in small studies.

Haemophagocytic lymphohistiocytosis

These disorders include genetic (familial and immunodeficiency-related syndromes) and acquired forms.

Familial forms affecting neonates and infants occur in 1 in 50,000 live births. Males and females are equally affected and over two-thirds of cases occur in siblings. The familial form is an autosomal recessive disease without a well-defined genetic defect. Recently, several defects in genes with important immune functions have been reported in patients with familial haemophagocytic lymphohistiocytosis (HLH) including mutations in the genes for perforin (*PRF1*), Munc13-4 (*UNC13D*) and syntaxin 11 (*STX11*). The acquired form is associated with Chédiak–Higashi syndrome, Griscelli syndrome and X-linked lymphoproliferative syndrome.

Mutations of the perforin gene have been reported in several patients with familial HLH and result in defective lymphocyte cytotoxic activity. The manifestations of the disease are thought to be mediated by hypersecretion of proinflammatory cytokines such as IFN-γ, TNF-α, IL-6, IL-10 and macrophage colony-stimulating factor (M-CSF). Such excess of proinflammatory cytokines results in tissue infiltration by lymphocytes and macrophages, leading to haemophagocytosis and the characteristic laboratory abnormalities including cytopenias, coagulopathies, hyperferritinaemia and hypertriglyceridaemia.

Both familial and acquired forms of HLH are commonly precipitated by viral (particularly Epstein–Barr virus and other herpes viruses), bacterial, fungal and protozoan infections, occurring frequently in an immunocompromised host. Other factors that have been associated with acquired HLH include malignancies (particularly lymphoproliferative disorders), drugs (such as phenytoin) and rarely with inborn errors of metabolism (lysinuric protein intolerance and multiple sulfatase deficiency). Symptoms of HLH commonly include high fever, anorexia, malaise, irritability and vomiting. Hepatosplenomegaly, lymphadenopathy and neurological signs such as cranial nerve palsies and seizures are also common The marrow is often hyperplastic with increased numbers of haemophagocytic histiocytes. Histopathological features of the lymph nodes or other involved tissue are often diagnostic, showing infiltration by lymphocytes and histiocytes and characteristic prominent erythrophagocytosis and haemophagocytosis, features required for diagnosis. During the acute phase of the illness, the plasma concentrations of inflammatory cytokines and the α-chain of the soluble IL-2 receptor as well as impaired NK cell activity are abnormal. Revised criteria from the Histiocyte Society for the diagnosis of HLH are shown in Table 14.13.

Treatment of familial HLH has included the use of corticosteroids (dexamethasone is preferred over prednisolone due to its ability to cross the blood–brain barrier), immunoglobulin infusions, ciclosporin and etoposide. Although early responses are observed, disease recurrence within months is common. Adequate control of CNS disease is important and intrathecal therapy using methotrexate with or without corticosteroids is merited in patients with recurrent CNS disease. Patients are at risk of developing opportunistic infections due to their underlying immune dysfunction and the effects of therapy. Haemopoietic stem cell transplantation from a matched sibling or unrelated donor remains the definitive treatment modality in patients with genetic forms of HLH.

Table 14.13 Diagnostic criteria for haemophagocytic lymphohistiocytosis.

Familial disease/known genetic defect
Clinical and laboratory criteria (5/8 criteria)
Fever
Splenomegaly
Cytopenia in two or more cell lines
Haemoglobin < 90 g/L (below 4 weeks < 120 g/L)
Platelets < 100×10^9/L
Neutrophils < 1×10^9/L
Hypertriglyceridaemia and/or hypofibrinogenaemia
Fasting triglycerides ≥ 3 mmol/L
Fibrinogen < 1.5 g/L
Ferritin ≥ 500 μg/L
sCD 25 ≥ 2400 U/mL
Decreased or absent NK cell activity
Haemophagocytosis in bone marrow, CSF or lymph nodes

Supportive evidence includes cerebral symptoms with moderate pleocytosis and/or elevated protein, elevated transaminases and bilirubin, lactate dehydrogenase >1000 U/L. sCD 25, soluble interleukin-2 receptor.

Sinus histiocytosis with massive lymphadenopathy or Rosai–Dorfman syndrome is characterized as a benign, frequently chronic, painless lymphadenopathy involving the cervical nodes and less commonly other nodal areas and has no known aetiology. Other features may include fever, weight loss and extranodal disease in skin, soft tissues, orbits, upper respiratory mucosa, bone and other organs. Although the disease commonly occurs in the first two decades of life, all ages can be affected. Diagnostic evaluation of involved lymph nodes reveals infiltration by histiocytes and multinucleated giant cells associated with erythrophagocytosis. The proliferating histiocytes are morphologically distinguished from the Langerhans cells of LCH by the absence of Birbeck granules on electron microscopy, as well as their surface phenotype. Nodal disease is often self-limited and regresses without intervention. However, bulky extranodal disease can cause symptoms and even organ impairment. Therapy depends on the extent of involvement and ranges from observation to surgical debulking and/or radiation. Chemotherapy plays no role in treatment.

Erdheim–Chester disease is a rare non-Langerhans cell histiocytosis that occurs in adults and may be of clonal origin. Bony pain with symmetric osteosclerosis of the metaphyses and diaphyses of long bones is characteristic of the disease. Extraosseous lesions can be found in approximately 50% of cases. Because of the very rare nature of this disease, only anecdotal reports of treatment with steroids, interferon, chemotherapy,

radiation and stem cell transplantation with variable rates of success are found in the literature. Prognosis is dependent on the extent of extraosseous disease and many patients die within a few years due to progressive pulmonary and retroperitoneal involvement.

Selected bibliography

Borregaard N, Cowland JB (1997) Granules of the human neutrophilic polymorphonuclear leukocyte. *Blood* **89**(10): 3503–21.

Chan CY, St John AL, Abraham SN (2012) Plasticity in mast cell responses during bacterial infections. *Current Opinion in Microbiology* **15**(1): 78–84.

Gordon S (2007) The macrophage: past, present and future. *European Journal of Immunology* **37**(1): S9–17.

Gotlib J, Maxson JE, George TI, Tyner JW (2013) The new genetics of chronic neutrophilic leukemia and atypical CML: implications for diagnosis and therapy. *Blood* **122**(10): 1707–11.

Hogan SP, Rosenberg HF, Moqbel R, *et al*. (2008) Eosinophils: biological properties and role in health and disease. *Clinical and Experimental Allergy* **38**(5): 709–50.

Ochs HD, Hagin D (2014) Primary immunodeficiency disorders: general classification, new molecular insights, and practical approach to diagnosis and treatment. *Annals of Allergy, Asthma and Immunology* **112**(6): 489–95.

Ogawa M (1993) Differentiation and proliferation of hematopoietic stem cells. *Blood* **81**(11): 2844–53.

Stuart LM, Ezekowitz RA (2005) Phagocytosis: elegant complexity. *Immunity* **22**(5): 539–50.

Schwartz C, Voehringer D (2011) Basophils: important emerging players in allergic and anti-parasite responses. *BioEssays* **33**(6): 423–6.

Vaiselbuh SR, Bryceson YT, Allen CE, Whitlock JA, Abla O (2014) Updates on histiocytic disorders. *Pediatric Blood and Cancer* **61**(7): 1329–35.

Ward DM, Shiflett SL, Kaplan J (2002) Chediak-Higashi syndrome: a clinical and molecular view of a rare lysosomal storage disorder. *Current Molecular Medicine* **2**(5): 469–77.

CHAPTER 15

Lysosomal storage disorders

15

Atul B Mehta[1] and Derralynn A Hughes[2]

[1]Royal Free Hospital, London, UK
[2]Royal Free and University College Medical School, London, UK

Lysosomes

Lysosomes are membrane-bound intracellular organelles that represent the major degradative compartment of mammalian cells. They are morphologically heterogeneous and are distinguished from other organelles by the presence of a range of catalytic enzymes and lysosome-associated membrane proteins (LAMPs). More than 50 lysosomal hydrolases, active at acidic pH, catalyse the degradation of macromolecules (lipids, proteins, nucleic acids, glycosaminoglycans and oligosaccharides). These include breakdown products of blood cell membranes, gangliosides from neurones, and glycosaminoglycans derived from connective tissue and extracellular matrix. Substantial flows of substrate of intracellular origin are also generated as a product of intermediary metabolism and turnover of macromolecules. Substrate is delivered to the lysosome by the processes of endocytosis, pinocytosis, phagocytosis or autophagocytosis, or is imported from the cytoplasm using a system of biological chaperones.

The lysosomal membrane and its LAMPs protect cytoplasmic components from the acid hydrolases. LAMPs are transmembrane proteins with highly glycosylated intralysosomal domains, among the most densely *N*-glycosylated proteins so far reported. The lysosomal membrane is also involved in fusion with other organelles, in maintenance of an acidic intralysosomal pH and in the efflux of amino acids and monosaccharides and oligosaccharides produced by the lysosomal hydrolases.

Soluble lysosomal enzymes are synthesized on the rough endoplasmic reticulum and acquire N-linked oligosaccharide side-chains in the Golgi apparatus. Here, asparagine-linked high-mannose residues are phosphorylated at position 6. This modification is specific for lysosomal enzymes and is utilized to route enzymes to the lysosomal compartment via mannose 6-phosphate receptors. Lysosome sorting of β-glucocerebrosidase, the enzyme deficient in Gaucher disease, is mediated by a complex of the lysosomal integral membrane protein LIMP-2 and the mannose 6-phosphate receptor.

Pathophysiology of lysosomal storage disorders

In most lysosomal storage disorders, an inherited deficiency of a specific lysosomal enzyme results in the accumulation of undegraded substrates within the lysosome. In others, accumulation of storage product results from deficiency or malfunction of activator proteins or transport proteins. Individual mutations of the relevant genes give rise to variable levels of residual enzyme activity. The resulting diseases are grouped according to the major stored substance, for example the mucopolysaccharidoses, sphingolipidoses and glycoproteinoses. Storage product within the lysosomes causes disruption of cellular organization and disturbance of normal membrane functions including signal transduction and ion transport. Different lysosomal storage diseases have the characteristic organ distribution patterns of the stored metabolites. However, even within a discrete storage disorder there are often wide-ranging clinical manifestations and considerable interindividual heterogeneity. Furthermore, genotype–phenotype relationships are generally unclear and even siblings with the same mutation can have widely disparate clinical manifestations.

Although knowledge of the genetics and biochemistry of the disorders has recently improved, little is known of the pathological processes that actually result in end-organ damage. These

Postgraduate Haematology, Seventh Edition. Edited by A Victor Hoffbrand, Douglas R Higgs, David M Keeling and Atul B Mehta.

processes do not simply relate to the burden of storage product, but extend to involve a complex host reaction to abnormal cells. This may result in cytokine secretion, cellular proliferation, disturbed calcium homeostasis, exaggerated inflammation and perturbed control of apoptosis. In Gaucher disease, where storage cells are macrophages that play an essential role in host physiology and pathogenesis of inflammatory and immunological responses, a wide variety of enzymes, cytokines and coagulation factors are perturbed (Table 15.1). Interleukin (IL)-1, IL-6, IL-10 and tumour necrosis factor (TNF) have been implicated in the pathogenesis of Gaucher disease. Elevated serum levels of IL-6, which is produced by macrophages, endothelial cells, fibroblasts and T cells, have been found to correlate with an index of severity in patients with Gaucher disease, and may be instrumental in the pathogenesis of B-cell dysregulation and bone disease. IL-10, which in general inhibits the synthesis of other inflammatory cytokines, is also elevated and may represent an abnormal state of immune activation.

Table 15.1 Enzyme and cytokine disturbance in type I Gaucher disease.

	Elevated	Reduced
Plasma	Glucosylceramide	Total plasma cholesterol
	Apolipoprotein E	Low-density lipoprotein
	Transcobalamin II	High-density lipoprotein
	Ferritin	Apolipoprotein A-1
	β-Hexosaminidase	Apolipoprotein B
	α-Mannosidase	Factor XI
	β-Glucuronidase	Factors V and VIII (normalize post splenectomy)
	Angiotensin-converting enzyme	
	Lysozyme	Factors II, VII, X, XII
	Tartrate-resistant acid phosphatase	
	Chitotriosidase	
	Thrombin–antithrombin	
	Plasmin–antiplasmin	
	D-Dimer	
	Immunoglobulins	
	Paraprotein	
	IL-1	
	IL-6	
	IL-8	
	IL-10	
	TNF-α	
	M-CSF	
	Soluble CD14	
	CCL18/PARC	
Tissue	Glucosylceramide	
	β-Hexosaminidase	
	β-Glucuronidase	
	Galactocerebrosidase	
	Tartrate-resistant acid phosphatase	
	Non-specific esterase	
Hepatic	Alkaline phosphatase	
	Transaminases	

EET, enzyme enhancement therapy; ERT, enzyme replacement therapy; SCT, stem cell transplantation; SRT, substrate reduction therapy.

Prevalence

Recent studies suggest that the true prevalence of lysosomal storage disorders may be higher than previously thought. This is because patients with minimal symptoms, such as homozygotes for the Gaucher N370S mutation, may not come to medical attention and also because of misdiagnosis of multisystem disorders such as Fabry disease. Studies in Europe, the USA and Australia suggest that although the diseases are individually rare, their combined overall birth prevalence is 1 in 5000–8000.

Diagnosis

In the absence of an informative family history, diagnosis of lysosomal storage disorders requires a high degree of clinical suspicion. Storage product may be identified within biopsy specimens, plasma or urine. These may have been requested incidentally, for example the finding of Gaucher disease in the bone marrow biopsy of a thrombocytopenic patient or Fabry disease in a renal biopsy, or may be a specific directed examination such as a skin biopsy in a patient with the characteristic rash of Fabry disease. Levels of lysosomal enzymes may be measured in plasma or leucocytes using commercially available synthetic or naturally occurring labelled substrates. Once a candidate enzyme is identified, analysis of the corresponding gene may identify a specific mutation and facilitate rapid screening of other family members. As effective treatment becomes available for a larger number of disorders, it is increasingly important that patients should be diagnosed as early as possible as presymptomatic individuals may be candidates for early intervention.

General aspects of therapy

Effective treatment of lysosomal storage disorders self-evidently involves reduction of the stored compound and prevention of its re-accumulation. This has been achieved by elevation of enzyme activity by stem cell transplantation, infusion of the missing enzyme (enzyme replacement therapy or ERT) or stabilization of protein folding variants by pharmacological chaperones. Conversely, reduction of the substrate can also be achieved by

limiting its synthesis through inhibition of anabolic enzymes, so-called substrate reduction therapy (SRT).

Many patients with Gaucher disease, metachromatic leucodystrophy, Krabbe disease and MPS type I have undergone stem cell transplantation. Donation of stem cells has largely been from HLA-identical siblings, but as the highest level of enzyme activity will be supplied by donor cells from non-carriers, cord blood from unrelated donors has recently been used with good results.

ERT is now available or under investigation for several lysosomal storage disorders. A range of production platforms has been developed recently. Recombinant human enzyme can be produced in Chinese hamster ovary (CHO) cells or plant (e.g. carrot) cells. Gene activation technologies have been applied to cultured human fibroblasts. The resulting translated proteins are chemically modified, purified and prepared for infusion. The effectiveness of ERT depends on the accessibility of the site of the pathology to exogenous enzyme and the ability of affected cells to internalize the enzyme. ERT is effective in reducing the non-neuronopathic manifestations of a number of disorders, including Gaucher disease type I, Fabry disease, Pompe disease and MPS types I, II, III IV and VI. However, infused enzymes do not cross the blood–brain barrier and ERT does not prevent the onset and progression of neurological symptoms. Intrathecal administration is under investigation. Furthermore, ERT for systemic therapy has to be given by intravenous infusion. Antibodies readily develop to infused enzyme and this certainly attenuates effectiveness in some diseases (e.g. Pompe disease).

Recently, other treatment strategies have been developed. One approach, SRT, is to decrease the rate of synthesis of the stored component so that it is approximately equal to the rate of degradation. Analogues of small molecules are used to inhibit the activity of synthetic enzymes. Thus, glucose and ceramide analogues can be used separately to reduce the activity of the enzyme glucosylceramide synthetase. This approach is only feasible in individuals with later-onset forms of the diseases who harbour mutations associated with detectable levels of residual enzyme activity. The small molecules used in SRT can be given orally, are easily absorbed and have a wide tissue distribution. Furthermore, SRT may have therapeutic utility in crossing the blood–brain barrier.

Some mutations lead to reduced levels of residual enzyme activity by causing misfolding of peptide chains or abnormal intracellular transport. The oral administration of small molecules (pharmacological chaperones) that can rescue misfolded or unstable enzymes is currently under active investigation for several lysosomal storage disorders. Such chaperones are designed to bind specifically to mutant peptides, aid their passage through the endoplasmic reticulum, rescue them from proteosomal degradation and guide them to the lysosome. They would dissociate from the enzyme within the acidic lysosomal environment. Combinations of ERT and chaperone therapy are also being evaluated in clinical trials.

Agents that bind to ribosomes (e.g. derivatives of gentamicin) and which modify the processing of RNA transcripts to allow 'read-through' of premature stop codon mutations are another class of mutation-specific therapy currently being evaluated in lysosomal storage disorders. Finally, gene therapy is under investigation in animal models of lysosomal storage disorders and has been reported as an adjunct to ERT in Gaucher disease.

Prognosis

The clinical course of infants diagnosed with a lysosomal storage disorder usually follows a predictable path of loss of learned skills and neurological deterioration until death results from infection and progressive organ damage. When onset is later, in adolescents and adults, the clinical course is more varied and the prognosis depends on the major organ systems affected. However, patient heterogeneity is such that it is not possible to predict which patients are most likely to experience significant morbidity or early mortality.

Clinical manifestations

In view of recent advances in therapy, Gaucher disease, Fabry disease and Pompe disease are presented in greater detail below.

Gaucher disease

Gaucher disease is due to deficiency of the enzyme β-glucocerebrosidase, a lysosomal enzyme that hydrolyses glucosylceramide to glucose and ceramide (Figure 15.1). It is an autosomal recessive condition arising as a result of mutation within the *GBA* gene. Affected individuals have a mutant enzyme with reduced activity, resulting in accumulation of the substrate (glucosylceramide) in lysosomes of reticuloendothelial cells. It is one of the commonest lysosomal storage disorders, with an estimated incidence of 1 in 60,000 to 1 in 80,000 individuals.

Clinical features

Clinical manifestations are due to cellular and tissue damage consequent upon accumulation of sphingolipid-laden macrophages in reticuloendothelial organs. Three main clinical phenotypes are observed (Table 15.2), determined in large part by the residual activity of the mutant enzyme. All three types are progressive disorders. Residual enzyme activity in type II is so low that abnormal cells accumulate in the central nervous system (CNS). This is the acute neuronopathic form of the disease, which presents with neurological complications in early infancy and usually leads to death before the age of 2 years. Type III disease is the subacute neuronopathic form that leads to a slowly progressive neurodegenerative disorder.

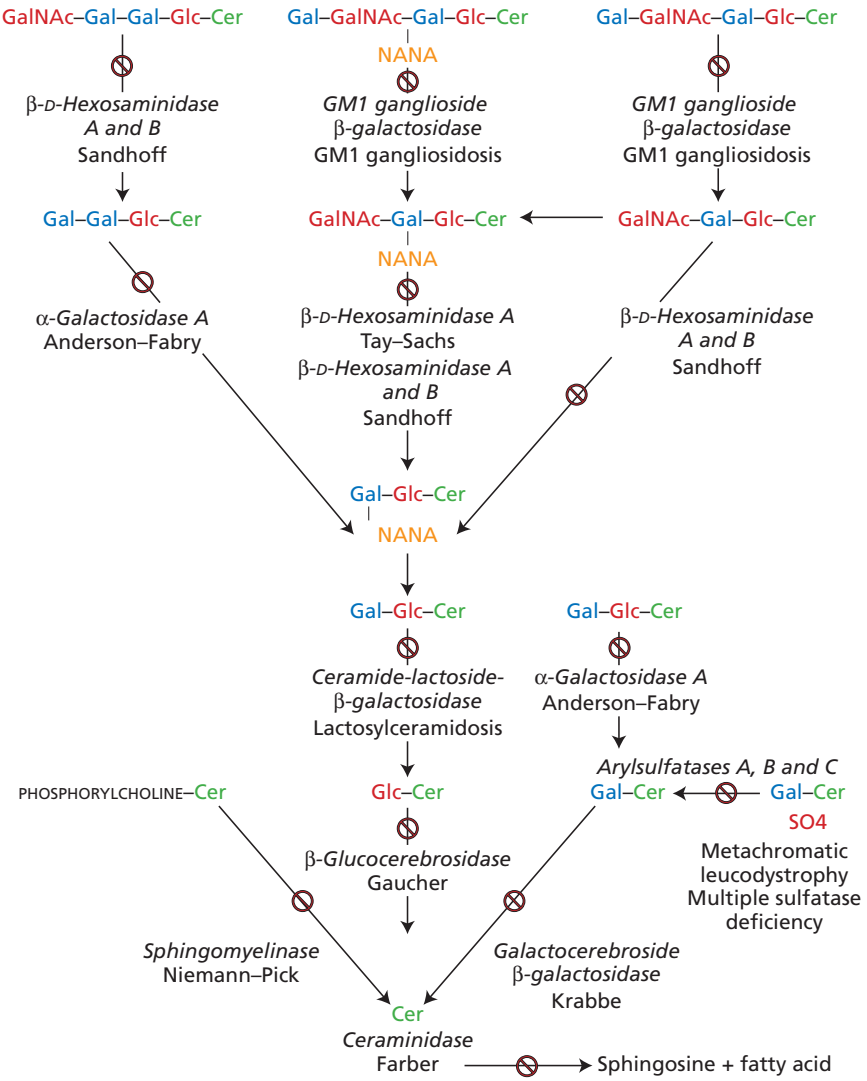


Figure 15.1 Biochemistry of lysosomal storage disorders.

Type I is the commonest form of Gaucher disease and typically does not cause neurological disease. It is particularly common among subjects of Ashkenazi Jewish origin; within this community, as many as 1 in 15–20 subjects are carriers, and approximately 1 in 800–1000 subjects are homozygous. Type I Gaucher disease is a heterogeneous disorder that may present in childhood or in late adult life. It is likely that many undiagnosed subjects are asymptomatic. Symptomatic individuals have hepatosplenomegaly, skeletal disease and bone marrow infiltration, leading to pancytopenia. Rarer manifestations of type I Gaucher disease include pulmonary disease, skin involvement and peripheral neuropathy. Patients with type I Gaucher disease have an increased incidence of malignancy generally and an increased incidence of haematological malignancies, especially B-lymphocyte disorders (myeloma, monoclonal gammopathy of undetermined significance) and myelodysplasia. There is an increased incidence of Parkinson's disease among affected individuals and carriers. The underlying mechanism is unknown, but experimental and animal models suggest that diminished activity of glucocerebrosidase may lead to excess deposition of α-synuclein in the basal ganglia, which is a pathological hallmark of Parkinson's disease.

Table 15.2 Clinical manifestations of Gaucher disease.

Manifestation	Type I	Type II	Type III
Onset	1 year	<1 year	2–20 years
Hepatosplenomegaly	++	+/−	+
Bone disease	++	−	+/−
Cardiac valve disease	−	−	+
CNS disease	−	+++	+/−
Oculomotor apraxia	−	+	+/−
Corneal opacities	−	+/−	+/−
Age at death	60–90 years	<5 years	<30 years

Laboratory features

Affected individuals have mutations within the *GBA* gene; more than 300 different mutations have been described. The commonest mutation causing type I disease is a single base-pair substitution in codon 370 (N370S), which accounts for approximately 70% of mutant alleles in affected Ashkenazi Jewish subjects. A base-pair substitution in codon 444 (L444P) is the commonest mutation underlying neuronopathic Gaucher disease. Diagnosis is confirmed by enzymatic assay of β-glucocerebrosidase activity in leucocytes and fibroblasts. However, enzymatic assay does not always identify heterozygote subjects and measured enzyme activity correlates poorly with clinical severity.

Splenic enlargement and marrow infiltration frequently lead to anaemia, thrombocytopenia and leucopenia. Changes in serum immunoglobulins are frequent. Polyclonal hypergammaglobulinaemia is found in more than one-third of patients, and monoclonal gammopathy of undetermined significance is seen in up to 20%. Liver function tests are often abnormal, reflecting infiltration of the liver by Gaucher cells leading to necrosis, fibrosis and, occasionally, even frank cirrhosis. There is an increased incidence of gallstones. The serum cholesterol level is typically lowered. Gaucher disease is associated with a bleeding diathesis, attributable to abnormal platelet function and thrombocytopenia. Factor XI deficiency is commonly observed and is largely due to co-inheritance of other genetic abnormalities that are also common among Ashkenazi Jews.

The abnormal lipid-laden macrophages are readily detected on tissue biopsy (e.g. bone marrow aspirate; Figure 15.2), although biopsy is no longer considered necessary to make the diagnosis. The serum levels of ferritin, angiotensin-converting enzyme (ACE), acid phosphatase and transcobalamin 2 are typically elevated. The enzyme chitotriosidase is derived from macrophages and is typically grossly elevated in untreated Gaucher disease, and declines progressively with treatment.

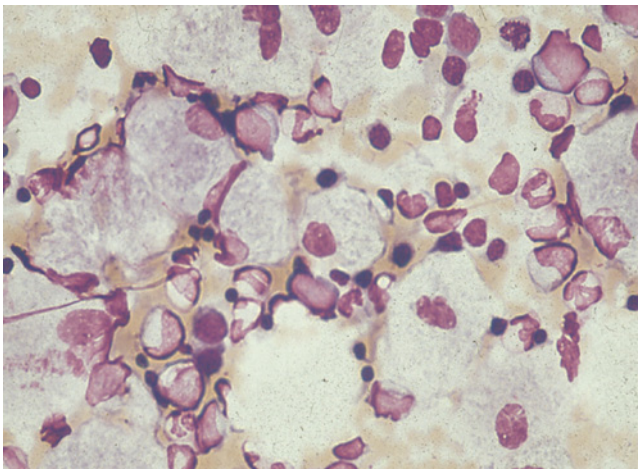


Figure 15.2 Gaucher cells in the bone marrow.

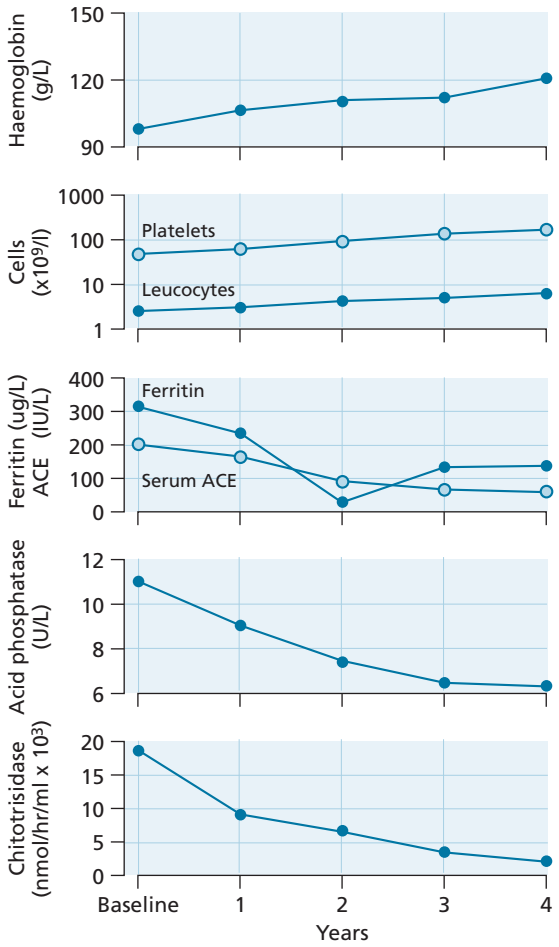


Figure 15.3 Clinical course in Gaucher disease treated with enzyme replacement therapy.

Levels may be as high as 30 000 U/L (normal range < 150 U/L); values below 1000 U/L generally indicate stable disease, and with prolonged ERT, values may even come down into the normal range (Figure 15.3). However, up to 6% of the population are deficient in this enzyme because of a 24-bp duplication in the chitotriosidase gene. These individuals cannot be monitored by measurement of plasma chitotriosidase activity. A new surrogate marker, pulmonary activation-regulated cytokine (PARC, CCL18), is also elevated in plasma of patients with Gaucher disease and is useful for monitoring those patients with chitotriosidase deficiency. PARC is a member of the CC chemokine family and its over-expression may be relevant to some of the pathophysiological features of Gaucher disease, such as abnormalities in neutrophil chemotaxis.

Treatment

All patients with Gaucher disease should be evaluated by experienced physicians. Gaucher disease is the first lysosomal storage disorder to be treated by ERT, and a number of different

ERT preparations are now available for the treatment of type I Gaucher disease. The most widely used preparations are Cerezyme® (Sanofi Genzyme), which is a recombinant preparation purified from CHO cells, and VPRIV® (Shire), which is derived by gene activation of cultured human fibroblasts. Trials indicate that these two preparations are equally efficacious. A third ERT preparation (Taliglucerase®, Protalix Pfizer) is derived from recombinant technology applied to carrot cells, and is also efficacious, but less well tolerated and more likely to provoke antibody production. ERT is expensive and 'bio-similar' generic ERT preparations are anticipated in the near future. Indications for ERT include significant pancytopenia (e.g. Hb < 100 g/L, platelets < 100×10^9/L), skeletal disease and significant hepatosplenomegaly. ERT has no discernible impact on CNS disease; type II Gaucher disease is therefore unresponsive to ERT and ERT is indicated for the systemic, but not the CNS, manifestations of type III.

ERT is administered by intravenous infusion (typically in the patient's home) at a dose of 45–60 units per kg every 2 weeks. However, some patients with advanced type I and III disease may benefit from more frequent infusions (e.g. weekly) for the first year or more. The dose of therapy is titrated against the severity of clinical and laboratory changes and is gradually lowered as the disease burden declines. Patients with less advanced disease may require lower doses, and some patients may only require monthly infusions. ERT is well tolerated and easily administered through a butterfly needle, although children and occasionally adults may require an indwelling catheter. A small proportion of patients (<10%) develop antibodies, but these are not usually neutralizing and do not affect treatment efficacy. Infusion reactions are rare and easily managed. Patients are monitored regularly in specialist units with blood tests (the chitotriosidase assay is particularly helpful) and serial skeletal magnetic resonance scans.

An oral form of therapy (miglustat, Zavesca) is licensed for mild to moderate type I disease. It is a form of SRT (see above), a glucose analogue that reduces the amount of substrate (glucosylceramide) being produced within lysosomes, such that patients with reduced residual enzyme activity will benefit. It is being evaluated at present in other lysosomal storage disorders (e.g. Niemann–Pick type C) as its administration leads to reduction of a range of substrates in addition to glucosylceramide. It frequently causes diarrhoea and is often poorly tolerated. Eliglustat is a ceramide analogue that has recently been developed as SRT. It is more active and has a better side-effect profile than miglustat. Once licensed, it may prove to be an effective oral alternative to ERT for patients with type I Gaucher disease.

Supportive therapy is frequently required. The skeletal disease in Gaucher disease (Figure 15.4) is painful and patients may require analgesia. Regular orthopaedic assessments are advised. Prior to the use of ERT, patients frequently developed acute 'bone crisis' – episodic severe pain, typically in the limbs and often precipitated by dehydration. Blood component therapy may be required for pancytopenic patients, particularly at the time of orthopaedic or other surgical intervention. Splenectomy should be avoided if possible as splenectomized subjects are more likely to develop tissue infiltration in other organs (e.g. liver, lungs, skeleton). Allogeneic stem cell transplantation is curative, although not without risk and has a definite role in carefully selected children with neuronopathic Gaucher disease. The costs of ERT and SRT mean that stem cell transplantation approaches warrant further investigation, especially in the developing world, where there are large numbers of children who are currently unable to access treatment.

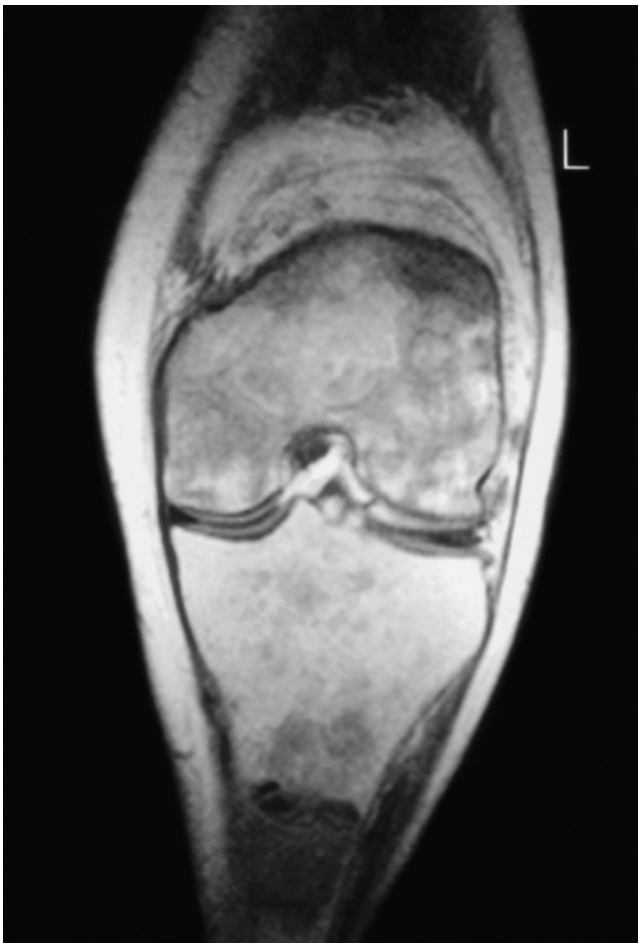


Figure 15.4 MRI scan showing skeletal changes in Gaucher disease.

Fabry disease

Fabry disease is an X-linked lysosomal storage disorder due to mutation within the gene for α-galactosidase A (*GLA*) (see Figure 15.1). The resulting inability to catabolize glycosphingolipids leads to progressive accumulation of the substrate globotriasylceramide in a range of tissues. In contrast to Gaucher disease,

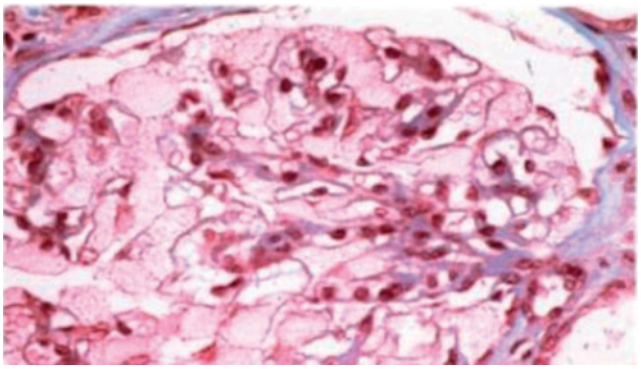


Figure 15.5 Renal biopsy in Fabry disease showing accumulation of globotriaosylceramide in the lysosomes of podocytes. (Source: Mehta *et al*.,2010. Reproduced with permission of Oxford University Press.)

the lipid accumulation in Fabry disease affects a range of cells (e.g. endothelial cells, epithelial cells, myocytes) within a broad range of tissues and organs, particularly the kidneys (leading to renal failure), heart (causing ventricular hypertrophy and conduction disturbances) and vasculature of the CNS. It is one of the commonest lysosomal storage disorders, with an incidence of approximately 1 in 100,000. It is panethnic. Milder variants of Fabry disease may be much commoner and may present as left ventricular hypertrophy, stroke/transient ischaemic attack and renal failure. Although it is X-linked, Fabry disease typically causes manifestations in females, though these occur later and are usually less severe than in males. Diagnosis is by assay of α-galactosidase A activity in leucocytes and detection of molecular abnormalities within the *GLA* gene. Tissue diagnosis is by skin, cardiac or renal (Figure 15.5) biopsy.

Clinical features and treatment

Clinical features are legion. Although females are heterozygous, they are usually symptomatic and may be as severely affected as males. A skin rash (angiokeratoma) and pain in limbs (acroparaesthesia) are early symptoms (under 10 years old). In late childhood, reduced sweating, abdominal symptoms and lymphoedema are characteristic. Renal failure, cardiac failure, stroke, epilepsy and CNS/sensory organ involvement are later features. Life expectancy is 40–50 years for men and 50–65 years for most women.

ERT for Fabry disease has been available since 2001 and two formulations are available: a recombinant galactosidase that is translated in Chinese hamster ovary cells and mannose-terminated (agalsidase β, Sanofi Genzyme Corporation, MA, USA); and an enzyme of identical amino acid sequence that is translated in a human fibroblast cell line wherein post-translational modification is performed within the human cell itself (Shire, MA, USA). The infused enzyme in Fabry disease must be taken up by lysosomes within cells in diverse organs and tissues, so targeting is of crucial importance. Beneficial clinical effects have been observed in renal and cardiac function, pain, hearing loss and gastrointestinal symptoms.

Pompe disease

Pompe disease (glycogen storage disease type II) is caused by deficiency of acid α-glucosidase. It is characterized clinically by an early infantile subtype associated with cardiomyopathy, hypotonia and reduced survival, and a later-onset form with features consistent with a limb girdle muscle disease phenotype and survival into adulthood. On blood film examination lymphocytes show glycogen vacuolation, which is positive on PAS staining. The prevalence of Pompe disease has been estimated to be 1 in 146,000. In untreated patients with the late-onset form of Pompe disease, muscle strength and pulmonary function deteriorate over time.

Alglucosidase-α (Myozyme) is a recombinant formulation of human acid α-glucosidase generated from the transduction of Chinese hamster ovary cells. In infantile-onset patients, enzyme therapy has been shown to extend overall and ventilator-free survival and improve lung function and mobility in adults with later onset Pompe.

Niemann–Pick disease

Niemann–Pick disease is divided into subtypes A and B resulting from sphingomyelinase deficiency and subtype C resulting from defects in cellular cholesterol trafficking. Patients with type A exhibit neurodegenerative disease resulting in death in infancy, those with type B have lung, but not neurological involvement, and those with type C show slowly progressive neurological disturbance. Hepatosplenomegaly may be found in all three types. Over 100 sphingomyelinase mutations causing Niemann–Pick disease type A or B have been described, and DNA-based carrier screening has been implemented in the Ashkenazi Jewish community. Murine knockout models have also been constructed and used to investigate disease pathogenesis and treatment. Based on these studies in the mouse model, a clinical trial of ERT has been initiated in adult patients with non-neurological sphingomyelinase-deficient Niemann–Pick disease. The use of inhibitors of glycolipid synthesis, including miglustat, appears promising in the therapy of Niemann–Pick disease type C.

Selected bibliography

Beck M (2007) New therapeutic options for lysosomal storage disorders: enzyme replacement, small molecules and gene therapy. *Human Genetics* **121**: 1–22.

Grabowski GA (2008) Phenotype, diagnosis, and treatment of Gauchers disease. *Lancet* **372**: 1264–71.

Hughes D, Cappellini MD, Berger M *et al*. (2007) Recommendations for the management of the haematological and onco-haematological aspects of Gaucher disease. *British Journal of Haematology* **138**: 676–86.

Lenders M, Karabul N, Dunting T *et al*. (2015) Thromboembolic events in Fabry disease and the impact of Factor V Leiden. *Neurology* **84**: 1009–16.

Mehta A, Beck M, Eyskens F *et al*. (2010) Fabry disease: a review of current management strategies. *Quarterly Journal of Medicine* **103**: 641–59.

Mistry PK, Lukina E, Ben Turkia H *et al*. (2015) Effect of oral eliglustat on splenomegaly in patients with Gaucher disease type 1: the ENGAGE randomized clinical trial. *Journal of the American Medical Association* **17**: 695–706.

Thomas AS, Mehta A, Hughes DA (2014) Gaucher disease: haematological presentation and complications. *British Journal of Haematology* **165**: 427–40.

Van der Ploeg AT, Reuser AJJ (2008) Pompe's disease. *Lancet* **372**: 1342–53.

Zarate AY, Hopkin RJ (2008) Fabry disease. *Lancet* **372**: 1427–35.

CHAPTER 16

Normal lymphocytes and non-neoplastic lymphocyte disorders

16

Paul Moss and Mark Drayson

University of Birmingham and Queen Elizabeth Hospital, Birmingham, UK.

Introduction

The immune system has evolved to provide protection against infection and is also believed to play an important role in the control of malignant disease. It is often considered to comprise two functional components, known as the *innate* and *adaptive* arms of the immune response, but in reality these have a broad overlap. The innate immune system is responsible for the initial response to infection and inflammation and is mediated by cells such as macrophages, neutrophils and dendritic cells, as well as lymphoid subsets such as natural killer (NK) cells, γδ T cells and B1 B cells. The adaptive or specific immune system is mediated by antigen-specific lymphocytes that are selected and expanded following recognition of antigen on antigen-presenting cells.

The anatomy of the immune system

The cells of the immune system are of haemopoietic origin and derive ultimately from the haemopoietic stem cell in the bone marrow. The myeloid and lymphoid lineages diverge during differentiation, with separation of a common myeloid progenitor and a common lymphoid progenitor. There are three broad classes of lymphocyte – B cells, T cells and NK cells – and these have different developmental pathways.

T cells are generated in the thymus following the migration of prothymocytes from bone marrow to thymus followed by selection of thymocyte precursors. Over 95% of thymocytes die in the thymus, but the minority population emerges from the thymus as single-positive $CD4^+$ or $CD8^+$ T cells and enters the lymphoid system as naive precursors that can survive for many years. The antigen receptor on T cells, the T-cell receptor (TCR), exists only as a cell-surface molecule and is not secreted. T cells have extremely diverse functions including: (i) providing signals that help induce T cells and B cells to proliferate and differentiate, (ii) specifically deleting virally infected cells or foreign cells, (iii) activating macrophages to enhance cellular cytotoxicity and (iv) regulation of established immune responses.

Most *B cells* are generated from within the bone marrow, the 'B' in their name referring to an obscure avian structure called the bursa of Fabricius in which the B cells of birds develop. Naive mature B cells enter the lymphoid circulation, but, if triggered by antigen in the periphery, a proportion of cells will return to the bone marrow as long-lived plasma cells that secrete immunoglobulin. B lymphocytes are the precursors of antibody-producing cells. Each B cell produces, and expresses on its surface, immunoglobulin with a distinct specificity for antigen. The specificity of the immunoglobulin is determined by the way the immunoglobulin variable-region genes are rearranged during B lymphopoiesis. B cells that bind antigen through their surface immunoglobulin have to obtain accessory signals if they are to proliferate and differentiate into antibody-secreting cells. These can be provided by helper T cells, which recognize antigen that has been taken up and presented by the B cell.

NK cells similarly appear to develop from within the environment of the bone marrow. They are able to kill cells that fail to express major histocompatibility complex (MHC) class I

Postgraduate Haematology, Seventh Edition. Edited by A Victor Hoffbrand, Douglas R Higgs, David M Keeling and Atul B Mehta.

molecules on their surface or that express molecules associated with cellular stress.

The bone marrow and thymus are therefore the sites of lymphocyte development and are known as the *primary lymphoid organs*. However, immune responses are initiated when lymphocytes encounter antigen and this occurs primarily in *secondary lymphoid tissues* such as lymph nodes and the spleen.

Lymphocytes circulate around the body tissues via the blood and lymphatic vasculature. Lymphatic vessels drain extravascular spaces and lymph nodes are collections of lymphoid tissue in lymphatic vessels, which are organized to optimize encounters between lymphocytes and antigen. Afferent lymph drains into the lymph nodes, bringing circulating lymphocytes and populations of antigen-loaded dendritic cells from regional tissue. Efferent lymph returns lymphocytes to the bloodstream, where naive cells continue this circulatory pattern in a continuing quest for antigenic encounter. Antigen-experienced lymphocytes migrate to a variety of tissues in order to mediate their effector functions. The pattern of homing is largely determined by the chemokine receptor profile on the lymphocyte (see below).

Nature of the antigen-specific receptors on T and B cells

Antigen recognition by lymphocytes

Each B or T lymphocyte expresses approximately 20,000–100,000 antigen-specific receptors on their surface and all the receptors on a single T cell or B cell are identical. However, although the antigen receptors on a single cell are identical, they are different from the receptors on other B or T lymphocytes. During lymphoid development, a repertoire of billions of B and T cells is generated and all have slightly different antigen receptors on their surface. At the initiation of an immune response, only a few of these cells are able to recognize the antigen and these are then expanded during their differentiation into effector cells. This is the basis of the *clonal selection* theory of immunology. The mechanisms that generate this great diversity of antigen receptors on lymphocytes are described in the next section.

Antibodies are immunoglobulins produced by B lymphocytes in response to an antigen for which they exhibit specific binding. There are five classes of immunoglobulins: IgM, IgG, IgA, IgD and IgE. Figure 16.1 shows a diagram of the basic

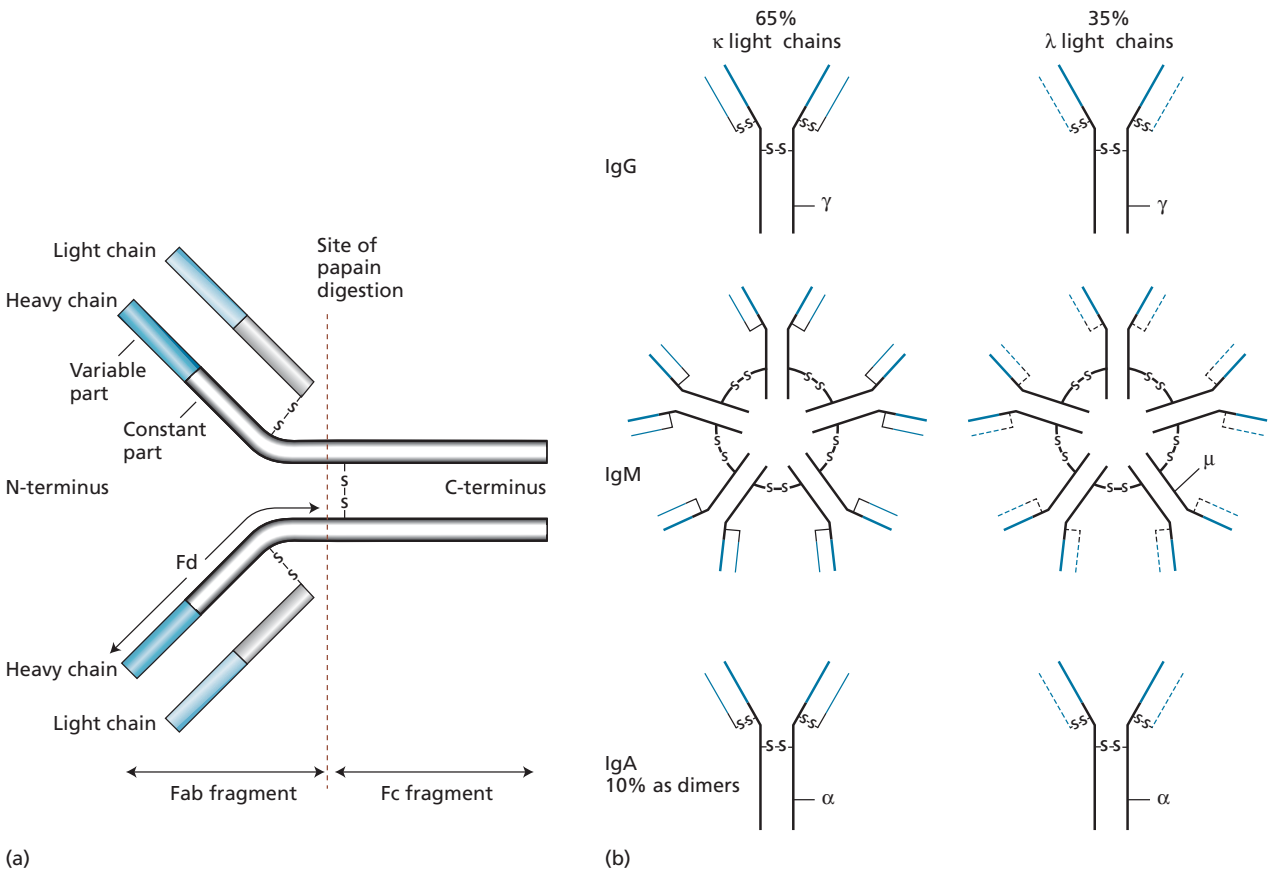


Figure 16.1 (a) Structure of the basic immunoglobulin molecule. (b) Structure of IgG, IgM and IgA molecules.

immunoglobulin molecule, which consists of four polypeptide chains arranged as two light (L) chains and two identical heavy (H) chains. The light and heavy chains are held together by disulfide (S–S) bonds. IgG and serum IgA molecules are monomers of this basic immunoglobulin structure, whilst secretory IgA is dimeric. IgM molecules are pentamers, with the basic immunoglobulin molecules held together by S–S bonds and a joining (J) chain (Figure 16.1).

There are two distinct types of light chain, kappa (κ) and lambda (λ), which are common to all immunoglobulins. Either of these chains may combine with any heavy chain, but in any one immunoglobulin molecule both light chains are identical. κ chains occur in about 65% and λ chains in about 35% of the normal immunoglobulins in each class. Each class has an immunologically distinct heavy chain: γ for IgG, μ for IgM, α for IgA, δ for IgD and ε for IgE. There are four subclasses of human IgG (IgG1, IgG2, IgG3 and IgG4, with γ1, γ2, γ3 and γ4 heavy chains), and two IgA subclasses (IgA1 and IgA2, with αl and α2 heavy chains). Analyses of various light chains from different sources show that the amino acid sequence differs in half of the chain (variable region), whereas in the other half the sequence remains remarkably constant (constant region) between light chains of the appropriate κ or λ groups. Similarly, in the corresponding heavy chain, there is a variable region and a constant region when different chains are analysed. Papain can split the basic immunoglobulin molecule into three fragments at a site near the S–S bonds that hold the heavy chains together. One fragment contains the C-termini of the heavy chains and is called the Fc fragment. The other two are called Fab fragments, each of which consists of the N-terminus of the heavy chain and the whole of the light chain, and contains the antigen-binding site of the molecule.

Repetition of amino acid sequences within the heavy chain constant regions indicates that there are either three (for IgG and IgA) or four (for IgM, IgD and IgE) constant region domains for H chains. These are designated CH1, CH2, CH3 and CH4. In contrast, there is only one constant region domain for light chains and only one variable region domain for heavy and light chains. The segment between CH1 and CH2 is called the hinge region. This area imparts flexibility to the immunoglobulin molecule so that antigen-binding sites can span varying distances.

The vast majority of the differences between antibodies of various specificities occurs in three or four short amino acid sequences in the L- and H-chain variable regions. These hypervariable sequences contact the antigen on binding and provide the basis of antibody specificity. The remaining sequences within the variable region are known as the framework determinants. These are believed to provide the general skeleton of the antigen-recognizing region, within which variations in and between hypervariable sequences generate specificity for the different epitopes bound by different antibodies. Figure 16.1 shows the structure of surface immunoglobulin on B cells.

Free light chains are kappa and lambda proteins that are produced by B cells but do not become attached to a heavy chain. They are secreted into the serum and can be measured in the blood of healthy individuals, where the normal serum free light chain (SFLC) values are 3.3–19.4 mg/L of kappa, 5.71–26.3 mg/L of lambda and kappa/lambda ratio: 0.26–1.65. SFLC production is increased from malignant B cell clones, particularly plasma cells, and an increase in either kappa or lambda SFLC is a sensitive test for paraproteinaemia. Levels are also increased in many cases of chronic leukaemia or non-Hodgkin lymphoma. SFLC are filtered from the blood by the kidney but are largely reabsorbed by renal tubules, and so urinary free light chains are generally not detectable in healthy individuals. However, they can become detectable in urine when the serum level of FLC gets so high that the kidney cannot reabsorb all the protein that is filtered, and this is often the case in patients with myeloma. Indeed, urinary FLCs are the basis of Bence–Jones protein, which was a test for myeloma discovered in 1848.

Biological and physical properties of immunoglobulins

Immunoglobulins are essentially multifunctional, as not only do they bind antigen, but they also perform a variety of other functions depending on their class. Most of these additional functions reside in the Fc fragment and are listed in Table 16.1. The most important include the following:

1 Complement fixation (IgM > IgG3 > IgG1 > IgG2). IgA does not bind complement in the classical pathway.

2 Opsonization. After binding to antigen, Fc receptors can bind to mononuclear phagocytic cells, particularly monocytes and macrophages, leading to uptake and phagocytosis of the bound cell or antigen. This property is particularly marked in the IgG3 class. After IgG-coated cells bind to Fc receptors on mononuclear phagocytic cells, they are destroyed by phagocytosis or cytotoxicity (Chapters 9 and 14).

3 Transplacental passage: there is preferential active transport of IgG1 relative to the other IgG subclasses. As only IgG passes the placental barrier, only IgG blood group antibodies can cause haemolytic disease of the fetus and newborn (HDFN) and only IgG1 and IgG3 will mediate significant immune red cell destruction. The IgG level in cord blood will be much the same as the level in the mother. Passively transferred maternal IgG gradually disappears from the infant after birth and is almost gone by 3 months of age. The serum of newborn infants contains a small amount of IgM of fetal origin and almost no IgA; the production of IgA and IgG starts at about 1–2 months of age.

The immunoglobulin proteins on the surface of the B cell form a signalling complex with a range of different molecules (Figure 16.2). The signalling pathways that are activated during recognition of antigen are now being targeted by drugs that are proving very valuable in the treatment of leukaemia and lymphoma. In particular, intracellular Bruton's tyrosine kinase

Table 16.1 Effector functions of the immunoglobulin isotypes.*

	IgG1	IgG2	IgG3	IgG4	IgA1	IgA2	Secretory IgA	IgM	IgE
Complement fixation: classical pathway†	+	(+)	++	–	–	–	–	++++	–
Complement fixation: alternative pathway	–	–	–	–	+‡	++	–	–	–‡
Placental transfer/lymphocyte binding	++	+	+	+	–	–	–	–	–
Monocyte/macrophage Fc receptor binding	++	–	+++	–	–	–	–	–	?§
Mast cell binding	–	–	–	(?)	–	–	–	–	+++
Binding to *Staphylococcus aureus* protein A	+	+	–	+	–	–	–	–	–

*No biological function has been ascribed to IgD, but it might be intimately involved in maturation of B cells into competent effector cells and/or memory cells.
†IgG molecules fix complement only up to C3.
‡Aggregated molecules can activate the alternative pathway.
§Human IgE has been reported to bind to macrophages.

(BTK) is activated following engagement of immunoglobulin with antigen and BTK inhibitors are proving of great value in the treatment of CLL and mantle cell lymphoma (Figure 16.3a). The delta isoform of PI3kinase is also relatively selectively expressed in lymphocytes and this pathway is also activated following antigen-receptor signalling. Again, selective inhibitors of PI3Kd are widely used in a several lymphoid disorders (Figure 16.3b).

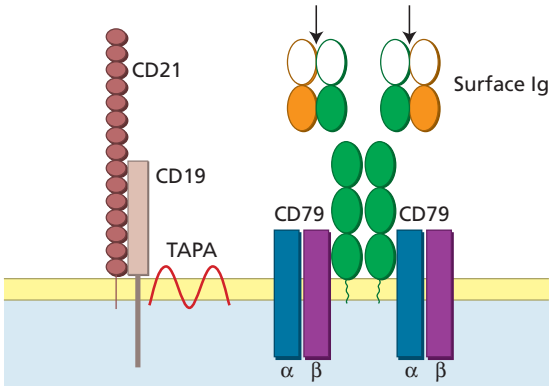


Figure 16.2 The structure of the antigen-specific receptor of B cells (BCR). Antibody (immunoglobulin) molecules on the surface of B cells provide their antigen-specific receptors. The green structures are the heavy (H) chains and the orange the light (L) chains. The antigen-combining (variable) regions are shown as open circles; the locations of the antigen-combining sites are indicated by arrows. The binding of antigen by B cells through their surface immunoglobulin can lead to antigen being internalized and can result, indirectly, in proliferation by the B cell and its differentiation to become an antibody-secreting cell or a memory cell. Signals delivered to the B cell when the surface immunoglobulin binds antigen are delivered through the α and β CD79 transmembrane signalling molecules and other surface immunoglobulin-associated molecules, including the complex of CD21 and CD19 with Tyr-d-Arg-Phe-β-Ala (TAPA).

T cells recognize peptides presented in association with MHC molecules. In humans, the MHC molecules are also known as human leucocyte antigens (HLAs) and there are two classes of MHC molecule. MHC class I molecules (Figure 16.4) are expressed by all nucleated cells except germ cells. They are not expressed by erythrocytes, but are found on the surface of erythroblasts. The peptides recognized by T cells in association with class I MHC molecules are, in most circumstances, derived from proteins produced from within the cell. This places MHC class I molecules in an excellent position to present peptides derived from viral proteins following intracellular viral infection (Figure 16.5). It is now clear that antigen-presenting cells are also able to take up proteins from outside the cell and process them such that peptides gain access to the MHC class I presentation. This process is called *cross-presentation* and may be particularly important in generating $CD8^+$ T cell responses to tumour-associated proteins.

MHC class I antigen presentation starts with the intracellular breakdown of proteins by a multimolecular proteolytic complex known as a *proteasome*. These peptides are actively transported by TAP (transporter associated with antigen processing) proteins into the endoplasmic reticulum, where empty MHC class I molecules are being assembled. The nascent MHC molecules are able to 'fold' around the peptides, which make non-covalent interactions with the peptide-binding groove at the top of the molecule. This complex is then stabilized by the association of β_2-microglobulin before being transported to the cell surface. In this way, the cells are continuously advertising the peptide composition of the proteins that they are producing.

The selection of T cells during their development in the thymus involves the processes of negative selection and positive selection. T cells that have high affinity for self peptides held in the groove of a self MHC molecule are deleted by apoptosis in a process known as negative selection. T cells with lower affinity for self-peptide–MHC complexes are positively selected and survive to become peripheral T cells. T cells that recognize

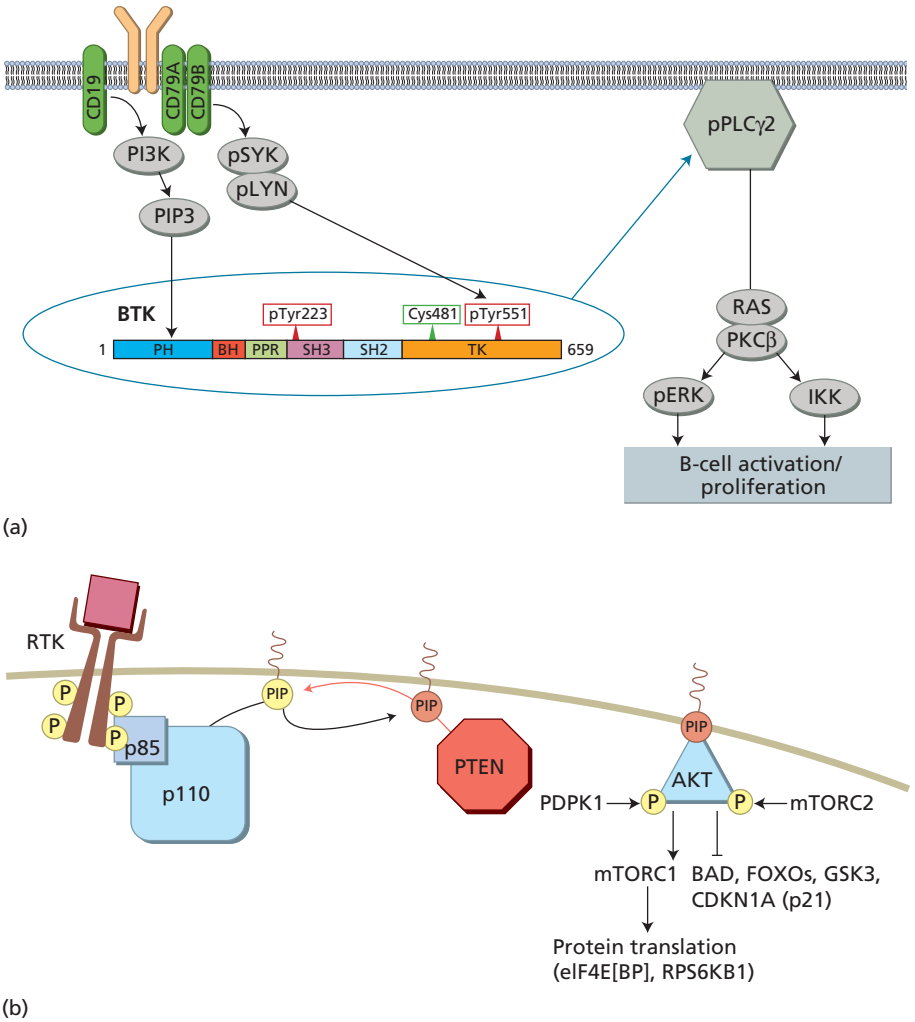


Figure 16.3 (a) BTK and its involvement in BCR signalling. Bruton's tyrosine kinase (BTK) comprises several domains including the: (i) pleckstrin homology (PH) domain, which targets BTK to the plasma membrane, (ii) BTK homology (BH) domain, (iii) polyproline (PPR) domain, (iv) Src homology (SH3 and SH2) domains (SH2 domain of BTK binds to the B cell adapter protein BLNK, which is required for full BTK activation) and (v) tyrosine kinase (TK) domain. Activation of BTK requires phosphorylation and binding of PIP3. Following BCR engagement and activation, and following PIP3 binding, BTK translocates to the plasma membrane, where it is phosphorylated at Tyr551 by LYN and SYK, leading to autophosphorylation of Tyr223. Once bound to BLNK, BTK phosphorylates PLCG2 at several sites, leading to calcium mobilization and activation of the protein kinase C (PKC) family members and other effectors, including RAS, and to NFκB activation and phosphorylation of ERK. Most BTK inhibitors bind irreversibly to Cys481 within the ATP binding site of the BTK kinase domain, thus preventing BTK activation. Apart from its role in BCR signalling, BTK is also implicated in cytokine and Toll receptor signalling pathways. (Source: Hutchinson and Dyer, 2014 [*Br J Haematol*. 2014;**166**:12–22]. Reproduced with permission of John Wiley & Sons.) (b) A schematic representation of the phosphoinositol-3 kinase pathways. Extracellular ligands bind receptor tyrosine kinases which are subsequently (auto)phosphorylated. The p85 regulatory subunit of PI3K then binds to one of these phosphates and activates the p110 catalytic subunit. PI3K takes a phosphate from a donor ATP and adds to an acceptor phosphatidylinositol-4,5-bisphosphate (PIP2) to generate phosphatidylinositol-3,4,5-triphosphate (PIP3). This action is opposed by PTEN, which dephosphorylates PIP3. PIP3 can recruit activated (phosphorylated) AKT to the membrane; once PIP3 is bound to the PH domain of activated AKT a variety of downstream substrates are phosphorylated. In the B cell these targets include BTK (see Figure 16.2a) (Source: Blachly and Baiocchi, 2014. Reproduced with permission of John Wiley & Sons.)

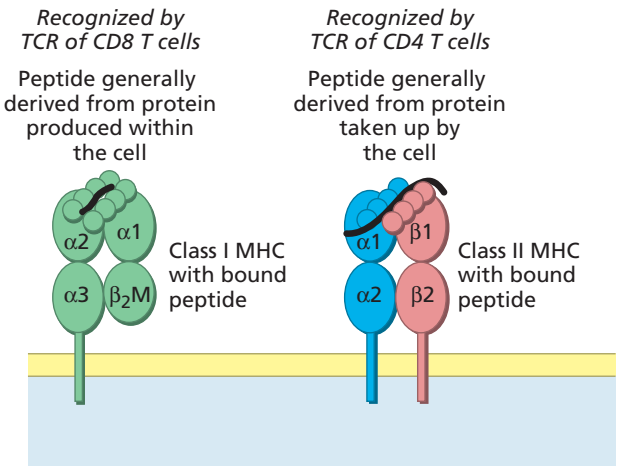


Figure 16.4 T cells recognize peptides held in MHC molecules. A class I MHC molecule is depicted on the left and a class II MHC molecule on the right. Some of the protein produced within each cell is broken down to peptides, which are presented on the cell surface in the peptide-binding groove of MHC molecules, usually class I. Extracellular molecules are taken up by antigen-presenting cells, broken down within the cell and presented on the cell surface in the peptide-binding groove of MHC molecules, usually class II. There are three isotypes of class II molecules, known as DP, DQ and DR, and three main class I isotypes, A, B and C. All of these are encoded within the MHC gene complex at 6p21.3. The genes encoding the peptide-binding grooves of each of these isotypes show extraordinary variability within the human population, so-called 'allelic polymorphism'. The range of peptides that can be held by different MHC molecules varies. Consequently, this polymorphism is reflected in differences between individuals in the ability to recognize specific peptides. Any one T-cell receptor (TCR) will only recognize a particular peptide within a particular groove structure. β_2-Microglobulin is a non-polymorphic immunoglobulin-like domain that is non-covalently associated with HLA class I MHC molecules; it stabilizes peptide binding and is essential for the expression of class I on the cell surface.

peptide presented by MHC class I molecules express a molecule, CD8, that binds to the α3 domain of the MHC class I molecule (Figure 16.4). When any cell in the body presents immunogenic peptides, these may be recognized by a cytotoxic CD8 T lymphocyte that can then kill the target cell. This is most likely to occur as the result of virus infection when virus-encoded proteins are produced; it may also occur following the acquisition of a genetic mutation within a cell.

The other class of MHC molecule, MHC class II (Figure 16.4), is less widely expressed. The only cells that constitutively express large amounts of this class of MHC molecule are specialized antigen-presenting cells collectively known as dendritic cells, B lymphocytes and thymic epithelial cells. Dendritic cells share a pattern of haemopoietic differentiation with monocytes and in the laboratory they can be derived from blood mononuclear preparations by culture with granulocyte/macrophage colony-stimulating factor (GM-CSF) and interleukin (IL)-4. They migrate to many tissues, particularly epithelia, where they remain until activated by local tissue injury. On activation, they take up fluid and particles from their surrounding environment (Figure 16.5) and digest proteins, which they present as MHC class II–peptide complexes at the cell surface.

The scrutiny of antigen-presenting cells by T cells starts when dendritic cells move from peripheral tissues to the T-cell-rich areas of adjacent secondary lymphoid organs. In this site they are known as interdigitating dendritic cells (IDCs), which express molecules associated with T-cell activation such as CD40, CD80 and CD86. In the T zones of secondary lymphoid organs recirculating T cells are constantly migrating to the surface of IDCs, which they scrutinize for the presence of an MHC–peptide complex to which they can bind with their TCR. In this way, the T cells continually screen for the presence of peptides derived from both extracellular and intracellular antigens (see section Immune responses).

The TCR (Figure 16.6) has certain similarities to immunoglobulin in that it is composed of two non-identical polypeptide chains that have constant and variable regions. In addition, as described in the next section, the genes that encode for TCR and immunoglobulin are remarkably similar. There are two types of TCR, with αβT cells expressing a heterodimer of an α-chain and a β-chain, and the minority population of γδT cells expressing the γδ TCR heterodimer, which recognize different forms of antigen. A number of different molecules are associated with the two TCR polypeptide chains in order to generate the TCR complex. This transmembrane signalling complex of molecules is collectively known as CD3 (detailed in Figure 16.6) and is linked to second messenger signalling molecules, whose expression varies between different T-cell subsets and at different stages of T-cell differentiation.

Complement

Complement consists of a series of proteins that are present in fresh plasma as inactive precursors, and which react sequentially with each other to form products that are important in the immune destruction of cells, including bacteria. In complement activation there are two stages of relevance. The first is the generation of the active form of C3, C3b, which leads to the coating, or opsonization, of the cell with a large amount of protein. There are two pathways by which C3b may be generated – the classical and the alternative pathways (Figure 16.7). The second stage is the lytic stage in which activation of the proteins of the membrane attack complex, comprising components C5 through to C9, leads to the destruction of red cells in the circulation. In general terms, the complement cascade is analogous to the clotting sequence. Activation of one

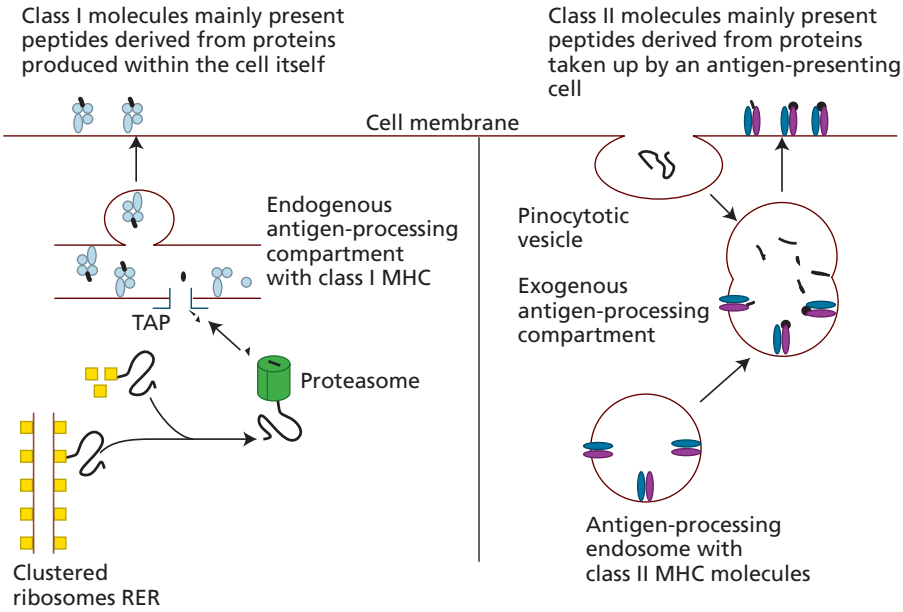


Figure 16.5 Antigen processing and presentation in association with MHC molecules. (Left): Antigen presentation of proteins produced within cells is mainly the property of class I MHC molecules. A proportion of proteins (black ribbons) produced within a cell on ribosomes (yellow) are broken down to peptides (black fragments) within a cytoplasmic molecular complex known as a proteasome. The resulting peptides are actively transported by TAP proteins (blue) through the wall of specialized endoplasmic reticulum (ER) that has MHC class I molecules in its wall. The peptides and β_2-microglobulin associate with class I molecules and are then expressed on the cell surface. (Right): Antigen presentation in association with class II MHC molecules involves pinocytosis of antigen (black ribbon) and fusion with an antigen-processing endosome, which has class II MHC molecules (magenta and blue) inserted in its wall. The antigen-presenting grooves of the class II MHC molecules are kept empty by the association with invariant chain (CD74). Fusion of the pinosome with the endosome heralds the activation of proteolytic enzymes; the invariant chain and the ingested proteins are broken down to peptides. The resulting peptides (black fragments) are assembled into the antigen-presenting groove of class II MHC molecules that are held in the endosomal wall. The HLA class II molecules with bound peptides are then carried to the cell surface.

component or group of components leads to the generation of enzyme activity for activation of the next components.

Complement components in the classical pathway are designated C1 to C9 in their native form. Apart from C4, which is activated before C2 and C3, the components are activated sequentially according to their numerical order. Complement components in the alternative pathway include factor B, factor D and properdin. During activation of the 'classical' components, small-molecular-weight fragments (C4a, C3a, C5a) are released, which have important chemotactic and anaphylactic activity.

The opsonization phase of the complement sequence (Figure 16.7)

The classical pathway

The classical pathway can be activated by many factors, including antigen–antibody complexes and endotoxins. The first component of complement, C1, is a complex of three protein molecules, C1q, C1r and C1s, which activates C4 and C2, generating C4b2b, known as C3 convertase. The cell-bound C4b2b can optimally activate several thousand C3 molecules by splitting them into C3a and C3b. As C3 is present in large amounts in the serum (1–1.5g/L), the fixation of C3b can considerably increase the globulin coating of the target cell. The active phase of C3 is transient because C3b is rapidly degraded by an enzyme (C3b inactivator or factor I) and its accelerator (βIH or factor H). C4b2b3b will 'trigger' the fixation of C5, C6, C7, C8 and C9, leading to the formation of the membrane attack complex on normal red cells and bringing about their lysis.

The alternative pathway

The alternative pathway does not necessarily involve antibody and represents non-specific 'innate' immunity. The proteins of the alternative pathway form a feedback loop for the conversion of C3 to C3b; the latter is both a product and reactant of this loop.

Initiation of the alternative pathway is a two-step process involving binding of C3b to an activator and its interaction

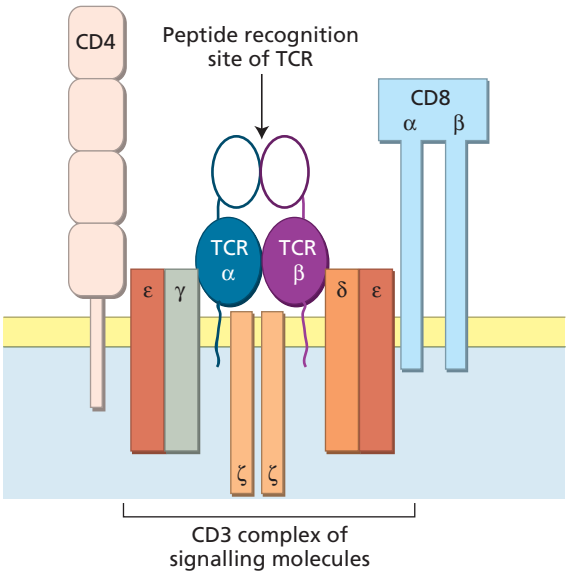


Figure 16.6 The αβ T-cell receptor (TCR) complex. The αβ TCR is composed of two polypeptide chains, each with a variable (open ovoid) and constant (closed ovoid) domain. Peptide plus MHC is recognized by the combined variable regions. The TCR is surrounded by the CD3 complex of transmembrane signalling molecules. This is composed of four types of polypeptide chain, γ, δ, ε and ζ. Most peripheral T cells express CD4 *or* CD8 with their TCR.

with neighbouring surface structures. Fluid-phase C3b is generated spontaneously and interacts with factor B to form a complex. Factor B is then activated by factor D, releasing Ba into the plasma and yielding a transient alternative C3 convertase, C3bBb. The latter can be stabilized by properdin, which is essential for preventing the dissociation of C3bBb by factor H. The alternative C3 convertase splits serum C3 into C3a and C3b. C3b attaches to the cell surface and can then combine with more factor B and D, thereby restarting the feedback loop.

The lytic phase of the complement sequence (Figure 16.7)

The lytic phase starts with the activation of C5 by C3b, yielding membrane-bound C5b and fluid-phase C5a. This step is followed by non-enzymatic interaction of C5b with C6, C7, C8 and C9. These molecules adhere to each other to form the membrane attack complex and insert themselves into the lipid bilayer of the target cell. C8 catalysed by C9 produces protein-lined cylinders to form pores through which ions and water can enter. The importance of the lytic phase of the complement cascade in the pathogenesis of paroxysmal nocturnal haemoglobinuria is apparent when considering the efficacy of eculizumab IgGk anti-C5 monoclonal antibody therapy in the management of this disease.

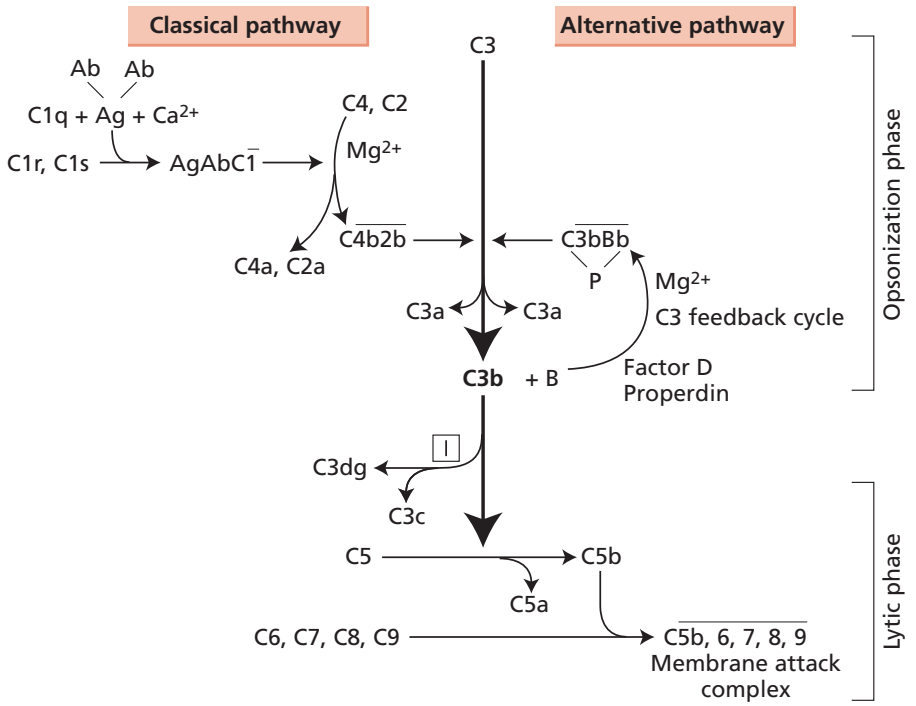


Figure 16.7 The complement cascade. Component C3b plays a central role in the classical and alternative pathways.

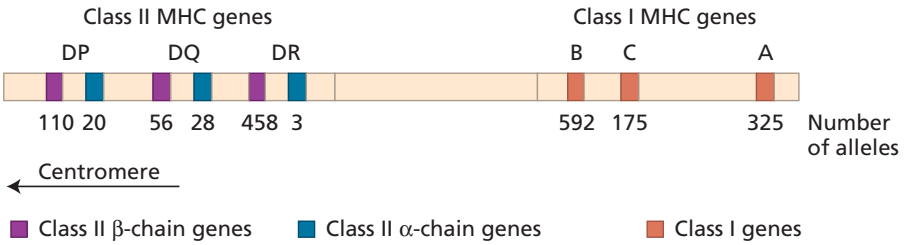


Figure 16.8 Schematic representation of the organization of the HLA gene complex to demonstrate the extent of MHC polymorphism. This is a simplified diagram of the main genes that encode MHC class I and MHC class II molecules and their exceptional polymorphism. For the latest information on HLA alleles go to http://hla.alleles.org (accessed May 2015).

Polymorphism of MHC molecules

There is extensive polymorphism of MHC class I and class II molecules which is presumed to have been driven by evolutionary selection against infectious disease (Figure 16.8). These were first recognized as targets for allograft rejection and the allelic forms of the MHC molecules differ in the fine structure of their MHC-binding grooves. This is reflected in differences in the range of peptides that different MHC alleles can present to T cells. The alleles on each chromosome 6 comprise the MHC haplotype of that chromosome. It follows that approximately one in four siblings share the same MHC haplotype on both chromosomes. For the purposes of stem cell transplantation, the patient and donor must usually be fully matched for HLA alleles and thus only 25% of siblings are appropriate for donation.

Certain MHC alleles are associated with relative protection against specific infections. Conversely, some alleles, or combinations of alleles, are associated with a greater chance of developing autoimmunity. Many diseases, including diabetes mellitus, Graves disease and ankylosing spondylitis, are distinctly more common in individuals with a particular MHC allele or MHC haplotype.

Generation of antigen-specific receptors on T and B lymphocytes

There are six pairs of genes that encode antigen-specific receptors: three for immunoglobulin (κ and λ light chains, and heavy chains) and three for TCRs (β, γ and a combined α and δ locus; see Table 16.2). These genes show marked similarities, indicating that the gene complexes evolved from a common precursor gene. The genetic organization of the variable-region genes and the way in which they are rearranged can generate a huge diversity of antigen recognition structures for subsequent display on the surface of mature B and T lymphocytes. For T cells, this is the only way in which diversity of the variable-region structure is achieved. In B cells, there is an additional mechanism that increases the variable-region gene repertoire. This is called *somatic hypermutation*, which is activated during B-cell maturation in germinal centres and which introduces mutations into the rearranged immunoglobulin variable-region genes.

Both the immunoglobulin and TCR variable-region genes have to undergo a process of *gene rearrangement* from their germline configuration before they can encode an antigen recognition structure. As an example, the immunoglobulin heavy chain gene is located at 14q32.3 and the germline organization of the part of the gene that encodes for the variable region of IgH is shown at the top of Figure 16.9. The variable-region component of the immunoglobulin heavy chain gene is divided into three types of *gene segment*: V segments, D segments and J segments. A large number of individual V gene segments are encoded within the genome. The V gene segments are longer than J or D segments and encode much of the framework of the variable-region domain, together with the first and second hypervariable regions (known as the complementarity-determining regions, CDR1 and CDR2). The CDR1 and CDR2 regions encode two of the three parts of the variable region that determine the antigenic specificity of the heavy-chain V region.

There are fewer D and J gene segments. The third hypervariable region (CDR3) is encoded at the site of joining of one of the functional D segments with any one of the functional J segments and includes the downstream end of one of the V segments. Heavy-chain rearrangement involves two looping-out manoeuvres (Figure 16.9). In the first of these, one of the J segments becomes spliced to one of the D segments and the intervening sequences are deleted. Next, one of the two rearranged D–J pairs becomes linked to one of the V segments and again the intervening sequences are deleted. The association of segments appears to occur at random and the theoretical number of different variable region genes that might be generated in this way is the product of the number of functional V, D and J segments, i.e. about 8262. In practice, D to J and V to D–J joining is not exact and additional random nucleotides may be added at the point where the gene segments join. This results in very much greater diversity, which is seen only in CDR3 and which includes both the D to J and V to D–J junctions.

The diversity of CDR1 and CDR2 is therefore much less than that of CDR3. Junctional diversity in CDR3 is sufficiently great to allow the conclusion that B cells with the same CDR3

Table 16.2 The variable region genes of human T-cell and B-cell antigen receptors.

Gene complex	Chromosomal location	Gene segments	
		Type	Approx. number
Immunoglobulin heavy chain	14q32.3	V_H	51
		D_H	~27
		J_H	6
		C_H	10
Immunoglobulin κ light chain	2p12	$V_κ$	40
		$J_κ$	5
		$C_κ$	1
Immunoglobulin λ light chain	22q11	$V_λ$	~29
		$J_λ$	4
		$C_λ$	4
TCR α-chain	14q11.2 (contains TCR δ locus)	$V_α$	~70
		$J_α$	61
		$C_α$	1
TCR δ-chain	14q11.2 (between Vα and Jα of TCR α)	$V_δ$	~4
		$D_δ$	3
		$J_δ$	3
		$C_δ$	1
TCR β-chain	7q32.5	$V_β$	52
		$D_β$	2
		$J_β$	13
		$C_β$	2
TCR γ-chain	7p15	$V_γ$	12
		$J_γ$	5
		$C_γ$	2

C, constant regions; D, diversity segments; J, joining segments; V, variable sements.
Where the number of functional gene segments is uncertain, this is denoted by '~'. There are many non-functional gene segments (pseudogenes); these are disregarded in this table. Because TCR α and δ genes are encoded in the same gene complex on chromosome 14, successful rearrangement of the TCR α genes inevitably results in looping out of the δ genes so that α and δ genes cannot be coexpressed.

sequence are almost certainly derived from the same clone and this fact is used widely to identify the origin and relationship of malignant B cells.

B lymphopoiesis

B-cell production starts in the fetal liver at the end of the first trimester of pregnancy and normally ceases at this site later in pregnancy. Subsequently, B cells are also produced in the bone marrow and production in this tissue continues throughout life such that approximately 2% of adult marrow mononuclear cells are B-cell progenitors.

The events that occur as cells differentiate towards B cells are summarized in Figure 16.10. The associated phenotypic changes are set out in Table 16.3. The earliest signs associated with differentiation of haemopoietic cells towards the B lineage are the expression of CD19, CD24 and MHC class II molecules on the cell surface and CD22 inside the cell. These changes are followed by the expression of molecules such as the recombinase-activating genes *RAG1* and *RAG2*, which are involved in immunoglobulin gene rearrangement.

Almost always, heavy-chain gene rearrangement precedes light-chain gene arrangement. The first rearrangements to occur are in the joining of one of the J segments to one D segment, with the looping out of the intervening sequences. If this is successful, the rearranged V–D–J sequence will be transcribed together with the genes encoding the μ constant region. After translation of this transcript, the cell has cytoplasmic μ heavy chain and is known as a *pre-B cell*. During all the preceding stages of B lymphopoiesis, cells showing differentiation towards B cells are known as *pro-B cells*.

Only one functional VDJ rearrangement is made by each B cell and this *allelic exclusion* of antigen receptor genes means that only one functional antigen-binding protein is expressed on the surface of the cell. Cells that fail to make a single functional rearrangement die by apoptosis.

The enzyme terminal deoxynucleotidyltransferase (TdT) is expressed during variable-region gene rearrangement and introduces non-encoded (N) sequences at the junctions of V to D–J and D to J. The enzyme is expressed both in pro-B cells in the marrow and in cortical thymocytes during TCR gene rearrangement. B cells express only a single light chain and the ratio of κ-expressing B cells to λ-expressing B cells is close to 60:40. Disturbance of the κ/λ B-cell ratio is a reliable indicator of the presence of a neoplastic clone.

Intact heavy chains can be exported to the surface of a B cell only if they are complexed with light chain or surrogate light chain. This does not apply to free light chains, which are able to pass out of B cells and plasma cells. Physiological light chain production is always greater than that required to complex all the available heavy chain, and measurement of serum free light chain is now a useful tool in the management of paraproteinaemia. The heavy chains that are secreted in heavy chain disease are able to leave the cell only because they are truncated heavy chains.

T-cell production and selection in the thymus

Although the earliest progenitors of T cells are produced in the bone marrow, the development and selection of most mature

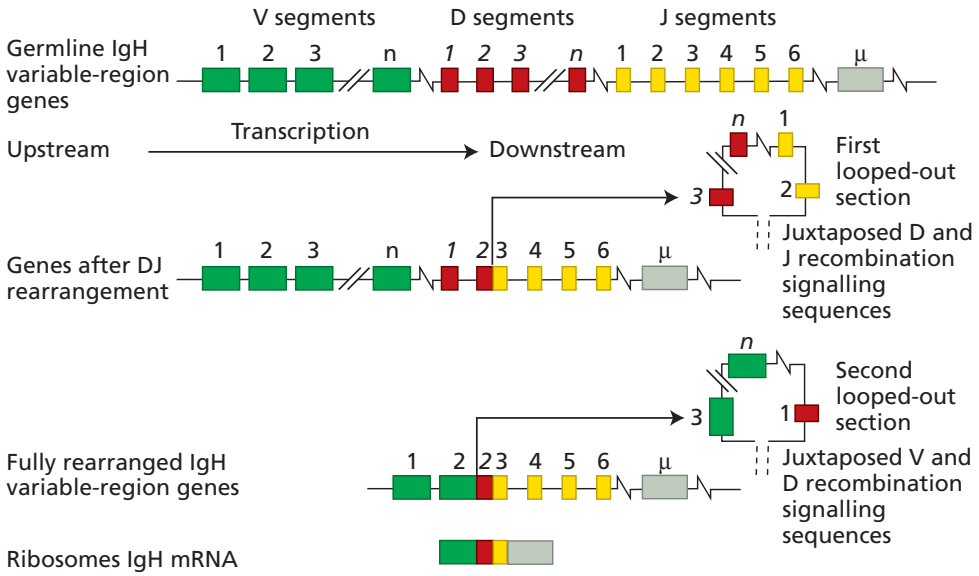


Figure 16.9 Immunoglobulin heavy-chain variable-region genes and their rearrangement. The germline structure of the variable-region gene complex is shown in the top line. The genes are present in this form in haemopoietic stem cells. The approximate number (*n*) of V_H, D_H and J_H segments are given in Table 16.1. The constant-region genes are downstream of the V-region genes. The first of these, μ, encodes the IgM heavy chain constant-region domain. This is followed in sequence by δ, γ3, γ1, a non-functional pseudo-ε gene, α1, γ2, γ4, ε and α2. The boxes represent exons and the lines introns. During rearrangement, first one of the J segments becomes aligned with one of the D segments and the intervening sequences are deleted. The aligned DJ pair then becomes linked to one of the V segments and again intervening sequences are deleted. If this V to DJ rearrangement is able to encode a variable region, then there is no V to DJ rearrangement on the other chromosome; if it has been unsuccessful (e.g. the rearrangement is out of frame) the cell goes on to attempt to rearrange a V segment to the DJ on the other chromosome. The D to J and V to DJ alignment is made possible by the presence of recombinase signalling sequences that flank: (i) the upstream end of each J segment and the downstream end of each D segment and (ii) the upstream end of each D segment and the downstream end of each V segment. Additional diversity at the junctions between the rearranged V, D and J segments results in part from imprecise splicing and partially through the insertion of additional non-encoded (N) base pairs at the D to J and V to DJ junctions through the action of terminal deoxynucleotidyltransferase.

immunologically competent T cells occurs in the thymus. The thymus is an encapsulated gland that is organized into lobules by capsular septa. Within each lobule, there is a complex meshwork of epithelial and other cells that are responsible for regulating the development of prothymocytes into mature T cells. The subcapsular region of the thymus is divided into the more peripheral cortex and the deeper medulla.

CD8 and CD4 thymocytes are selected on the basis of their potential to recognize peptides held in association with either MHC class I or MHC class II molecules, respectively. During the development from prothymocytes to mature lymphocytes, T cells pass from the outer cortex, through the inner cortex and on through the medulla before emerging as immunocompetent T cells in the peripheral circulation. T cells whose TCR molecules fail to engage with MHC–peptide complexes within the thymus die by neglect, whereas those whose TCR can interact with these complexes are subject to positive and negative selection (see above). It has been estimated that 95% of all thymocytes die within the thymus either as a result of failure to rearrange their TCR genes in an expressible form or as a consequence of elimination during the T-cell selection process.

Thymic epithelial cells in the cortex are critical in the positive selection of thymocytes that have moderate affinity for MHC–peptide complexes. The boundary of the cortex and medulla is populated by macrophages (often called sentinel macrophages) that phagocytose cells undergoing apoptosis. The medulla contains medullary thymic epithelial cells (mTEC) which mediate the process of negative selection. Interestingly, these cells express the autoimmune regulator (Aire) protein which leads to transcriptional expression of a wide range of cellular genes, therefore ensuring that tolerance is achieved to proteins expressed in a range of peripheral organs. Structures known as Hassall corpuscles, which are whorled aggregates of epithelial cells, can also be seen in the medulla though their function is uncertain.

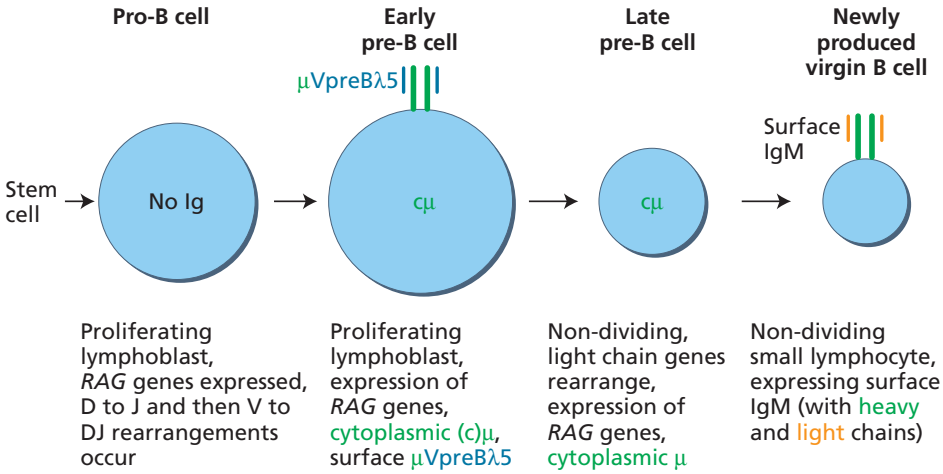


Figure 16.10 Outline of the main stages of B lymphopoiesis. The central process in B-cell formation is the rearrangement of immunoglobulin variable-region genes. For rearrangement of V-region genes to occur in either T or B cells, the recombinase-activating genes *RAG1* and *RAG2* have to be expressed. Absence of these genes totally blocks further differentiation towards B or T cells. After a successful heavy-chain VDJ has been made, a B cell must express the heavy chain with the surrogate light chain composed of V_{preB} and λ5 if further differentiation is to occur. Cells that fail to make either a productive heavy- or light-chain rearrangement destroy themselves by apoptosis.

TCR gene rearrangements and phenotypic changes

Bone marrow-derived T-cell progenitors (prothymocytes) seed the subcapsular region of the thymic lobule (Figure 16.11). At this stage of ontogeny the cells have not started to rearrange their TCR genes, do not express the mature T-cell markers CD3, CD4 or CD8, and may be identified by expression of CD7 and CD34. Interaction with the thymic stroma is accompanied by expression of CD2 and cytoplasmic expression of CD3 genes. By this stage, the cell is committed to the T-cell lineage and TCR gene rearrangement is underway.

The configuration of the TCR and its genes is discussed in the section on antigen-specific receptors on T and B cells and is summarized in Figure 16.5 and Table 16.2. During early fetal development, the first cells to leave the thymus as mature T cells have successfully rearranged their γ and δ genes and express the γδ form of the TCR. During later stages of human fetal development and throughout the rest of life, 85–98% of T cells that leave the thymus have undergone successful α and β gene rearrangement and express the αβ TCR.

Rearrangement of the gene encoding the TCR β-chain occurs first with D to J rearrangements followed by V to D–J rearrangements (Figure 16.11). TCR α-chain gene rearrangement is then initiated.

The B-cell repertoire

B cells constitute around 10–20% of the peripheral lymphoid pool and two broad lineages have been recognized, *B1 cells* and

Table 16.3 Phenotypic changes during B-cell lymphopoiesis.

	Surface CD34	Cytoplasmic CD22 with surface CD19 and CD24	Nuclear TdT	Surface CD10	Surface CD20	Cytoplasmic μ	Surface IgM
Pro-B cell	+	+	+	−	−	−	−
Early pre-B cell	+	+	+	+	−	−	−
Late pre-B cell	−	+	−	+	+	+	−
Virgin B cell	−	+	−	−	+	+	+

TdT, terminal deoxynucleotidyltransferase.

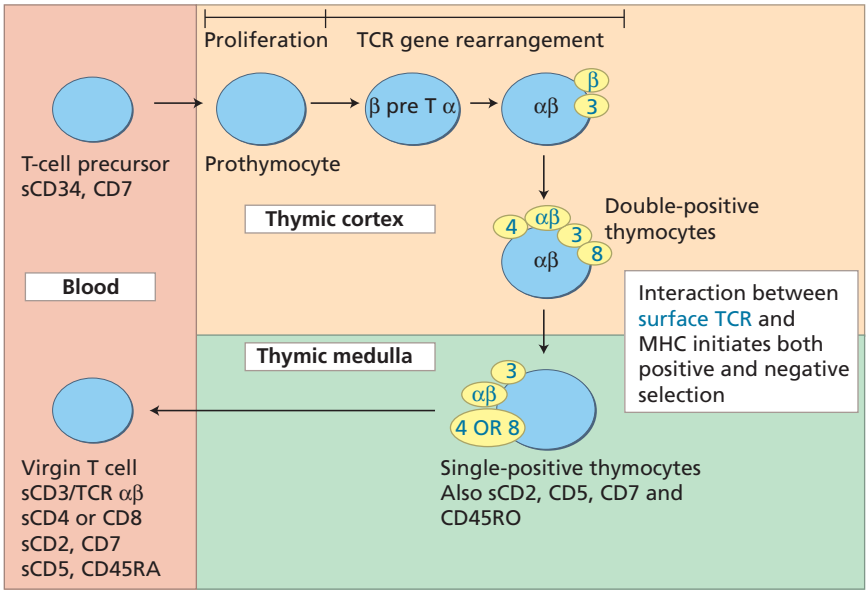


Figure 16.11 T-cell maturation in the thymus. Molecules in yellow are depicted without the prefix CD. T-cell progenitors enter the thymus from the marrow or other primary lymphoid site to become a prothymocyte. At this stage their T-cell receptor (TCR) genes are in germline configuration. Prothymocytes proliferate in the outer cortex and then TCR rearrangement starts: first, β-chain genes are rearranged, then these are expressed with a surrogate α-chain (pre-T α–); after this, α-chain genes rearrange. As with B cells, these rearrangements require *RAG1* and *RAG2* genes to be expressed, and junctional diversity is increased by the addition of N sequences using terminal deoxynucleotidyltransferase.

conventional (or B2) cells. B1 cells are the predominant B-cell subtype in fetal life and during the first year after birth, and are capable of antigen-independent self-renewal in the periphery. The antigen-binding specificity of the antibodies produced by these cells tends to have broad reactivity including autoreactivity. The function of B1 cells and the background IgM antibody they produce is still far from clear. It has been suggested that they may play a regulatory role in selection of the B-cell repertoire and they are highly conserved during evolution. B1 B cells often express CD5 and share many features in common with the malignant cells of B-cell chronic lymphocytic leukaemia.

The majority of B cells are conventional (B2) B cells and are produced in the marrow after the first year of life. Three stages of conventional B-cell differentiation can be identified: newly produced virgin B cells, recirculating B cells and marginal zone B cells (Table 16.4). Most recirculating B cells are virgin cells, whereas the B cells of the marginal zones do not recirculate and are a mixture of virgin and memory cells.

Naive B cells are small non-dividing virgin B lymphocytes that classically are positive for surface IgM and IgD. They are in a constant state of migration between the follicles of secondary lymphoid organs. On their way to the follicles, they migrate through the T-cell-rich zones that contain antigen-presenting cells and have an average lifespan of 4 weeks or more.

Marginal zone cells, like recirculating B cells, respond to T-cell-dependent antigens and to bacterial cell wall lipopolysaccharides (thymus-independent type 1, or TI-1, antigens). Unlike recirculating B cells, they will also respond to bacterial capsular polysaccharides. Capsular polysaccharides are thymus-independent type 2 (TI-2) antigens. These antigens do not evoke antibody responses until several months after birth, and levels of antibody produced in response to these antigens do not reach adult levels until 5 years of age.

The T-cell repertoire

The majority (around 95%) of peripheral T cells express αβ TCR. When these αβ T cells leave the thymus, most express either CD4 or CD8 on their surface and are restricted to the recognition of peptides in the context of MHC class II or class I molecules, respectively. A minority of T cells express a γδ TCR and exhibit antigenic specificity for non-peptide molecules.

CD4⁺ and CD8⁺ T cells and their functions

CD4⁺ T cells can have one or more of a variety of functions. Their major role is in the provision of signals that induce proliferation or differentiation of T or B cells, generally known as *T-cell help*. They can also activate macrophages and may mediate HLA class-II-restricted cytotoxic activity. Increasing attention is being given to their regulatory role in T-cell immune responses and this is mediated through a FoxP3-positive subset of T regulatory cells.

CD8⁺ T cells have a predominant role in the recognition of virally infected cells and recognize antigen in association with HLA class I. Antigen recognition is followed by lysis of the target cell and secretion of cytokines such as interferon (IFN)-γ. There are two main killing mechanisms, which both induce apoptosis in the target cell. Firstly, enzymes known as *granzymes* and pore-forming molecules termed *perforins* are transferred from the T cell to the target cell. Granzymes and perforins are present in granules that stain crimson and are easily seen against the featureless pale cytoplasm of cytotoxic T cells and NK cells in Jenner–Giemsa preparations. These features have given rise to the term 'large granular lymphocyte'. The second cytotoxic mechanism involves the expression of the Fas ligand (FasL) by

Table 16.4 Differences between the three main types of human B cell found in adults.

	Marrow IgM^+, IgD^-	Recirculating	Marginal zone
Diameter		~8 μm	~10 μm
Chromatin	Condensed	Condensed	Open
Cytoplasm volume/basophilia	Scanty/little	Scanty/little	Moderate/moderate
Proliferating	No	No	No
Lifespan	About 3 days	4 weeks or more	3 weeks or more
Antigen-independent movement	Migrate from marrow to secondary lymphoid tissue	Migrate between the secondary lymphoid tissues	Remain in the marginal zones
Surface immunoglobulin	IgM	IgM and IgD	IgM or IgG or IgA
Memory or virgin	All virgin	Almost all virgin	Variable mixture of virgin and memory
Immunoglobulin V-region mutations	None	Not present	Present in memory cells
Molecules expressed on the cell surface			
CD19, CD20, CD37, CD40 and class II MHC	Positive	Positive	Positive
CD21, CD39	Negative	Positive	Positive
CD5	Negative	Some positive	Negative
CD23	Negative	Positive	Negative
CD25	Negative	Negative	Positive
CD38	Negative	Negative	Negative
Capacity to respond to different classes of antigen			
Bacterial cell wall lipopolysaccharide	+	+	+
Bacterial capsular polysaccharide	−	−	+
Protein-based antigens	+	+	+

IgM^+, IgD^- cells of the marrow are representative of newly produced virgin B cells.

the effector cell. The resulting engagement of Fas on the target cell induces the target to undergo apoptosis. Fas is a member of the TNF receptor family and FasL is an analogue of TNF-α and TNF-β.

Once the peripheral T-cell pool is established, it can maintain itself without further input from the thymus. Even neonatal thymectomy in humans during cardiac surgery does not cause clinically noticeable immunodeficiency, indicating that T-cell clones can be very long-lived. Unlike the peripheral pool of recirculating B cells, which cannot replenish itself when depleted, small numbers of transferred recirculating T cells will proliferate to fill a depleted peripheral T-cell pool in a process known as *homeostatic proliferation*.

Natural killer cells

NK cells are cytotoxic lymphocytes that lack expression of a TCR and have a range of functional activities in the immune system. They express a range of inhibitory and activating receptors and it is the balance of signals received through these molecules that determines whether the target cell will be recognized. Inhibitory signals include: (i) killer inhibitory receptors (KIRs) that bind to HLA-C or HLA-B alleles on the surface of the target cell and (ii) CD94–NKG2 heterodimers that bind to HLA-E. Activating receptors include: (i) activating forms of KIR molecules, whose ligands are uncertain at present, (ii) NKG2D, which binds to proteins such as MICA and ULBP expressed on cells under physiological stress, and (iii) the immunoglobulin receptor (FcR) CD16.

HLA class I expression is often downregulated on the surface of cells following viral infection and is also a common feature of malignant cells. This provides some protection from recognition by $CD8^+$ T cells, but renders the cell susceptible to lysis by NK cells (Figure 16.12). Thus, $CD8^+$ T cells and NK cells can be viewed as having complementary recognition systems based on the level of HLA class I on the target cell. The ability of NK cells to recognize ligands such as MIC-A, which are expressed on the

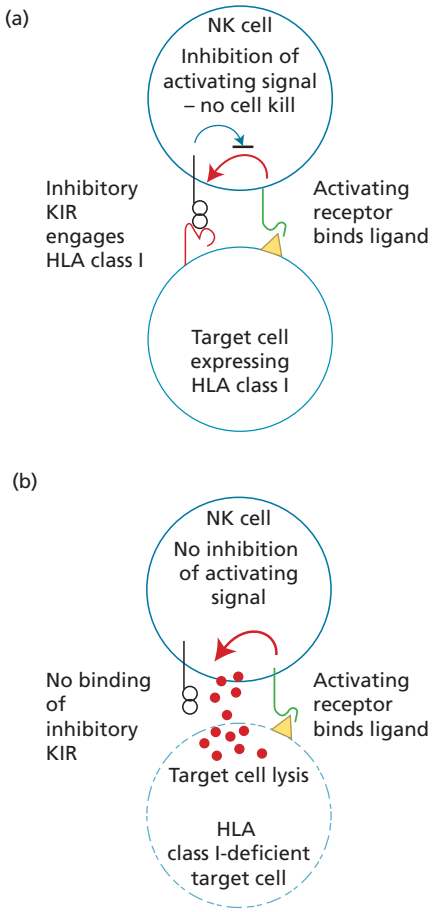


Figure 16.12 Mechanism by which NK cells kill target cells that fail to express HLA class I. NK cells express two classes of receptors which either inhibit (a) or activate (b) NK cell killing. Activating receptors can bind to a range of ligands on the target cell whose expression is often constitutive. In contrast, the major forms of inhibitory receptor bind to HLA class I molecules. If HLA class I expression is downregulated on the target cell, no inhibitory signal is delivered to the NK cells and the target cell is killed.

surface of damaged or stressed cells, demonstrates that NK cells may also have the ability to target cells at the site of inflammation, irrespective of HLA class I expression level.

The lytic mechanisms of NK cells seem to be the same as those used by cytotoxic T cells. NK cells proliferate in the presence of IL-2 and their activity can be augmented by exposure to IFN-γ. They characteristically express a range of receptors such as CD56 (NCAM) and CD57, but these are also expressed on a subset of T cells.

Natural killer T cells

Natural killer T (NKT) cells represent a small (< 1%) population of peripheral blood T cells and appear to be positioned

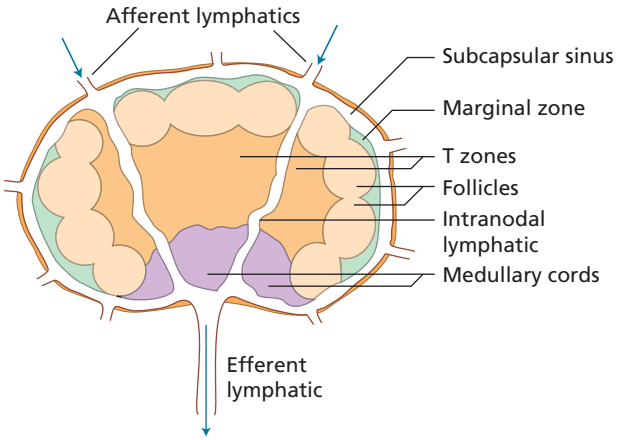


Figure 16.13 The main compartments of a lymph node. Note that the size of the marginal zone is variable; although it is often obvious in mesenteric lymph nodes, it may not be obvious in small nodes such as the popliteal nodes, particularly if these have not been sites of recent immune responses.

somewhere between conventional T cells and NK cells in that they express a TCR, but also a range of receptors typically associated with NK cells. Moreover, the TCR is invariant, with the same Vβ11Vα24 heterodimer being expressed on all cells. NKT cells recognize lipid antigens presented on the surface of CD1d molecules, CD1d being a protein with homology to MHC class I. It appears that NKT cells are activated very early in an immune response, although their functional significance is uncertain.

The immune response

In this section we consider how an immune response is developed within the secondary lymphoid tissue. The general structure of a lymph node is shown in Figure 16.13 to provide an anatomical context for immune physiology.

Lymph nodes have an afferent lymphatic supply that is fed by lymph draining the extravascular tissue spaces and this provides the main source of antigen for the node. Dendritic cells activated by local disturbance take up and process antigen, and then pass through afferent lymph into lymph nodes (Figure 16.14). The afferent lymph passes into the subcapsular sinus, which forms a lake of lymph that covers the cortical surface of the node. From the subcapsular sinus, lymph passes through intranodal lymph sinuses that surround and separate the cone-like segments that make up the solid tissue of the node. Attached to these and the walls of the tissue cones are macrophages and other poorly defined cells. Increased numbers of these cells in the intranodal lymph sinuses and similar cells in the subcapsular sinus are described by histopathologists as *sinus histiocytosis*. In the medulla, the intranodal lymph sinuses feed into the efferent lymphatic vessel that returns the lymph to the venous

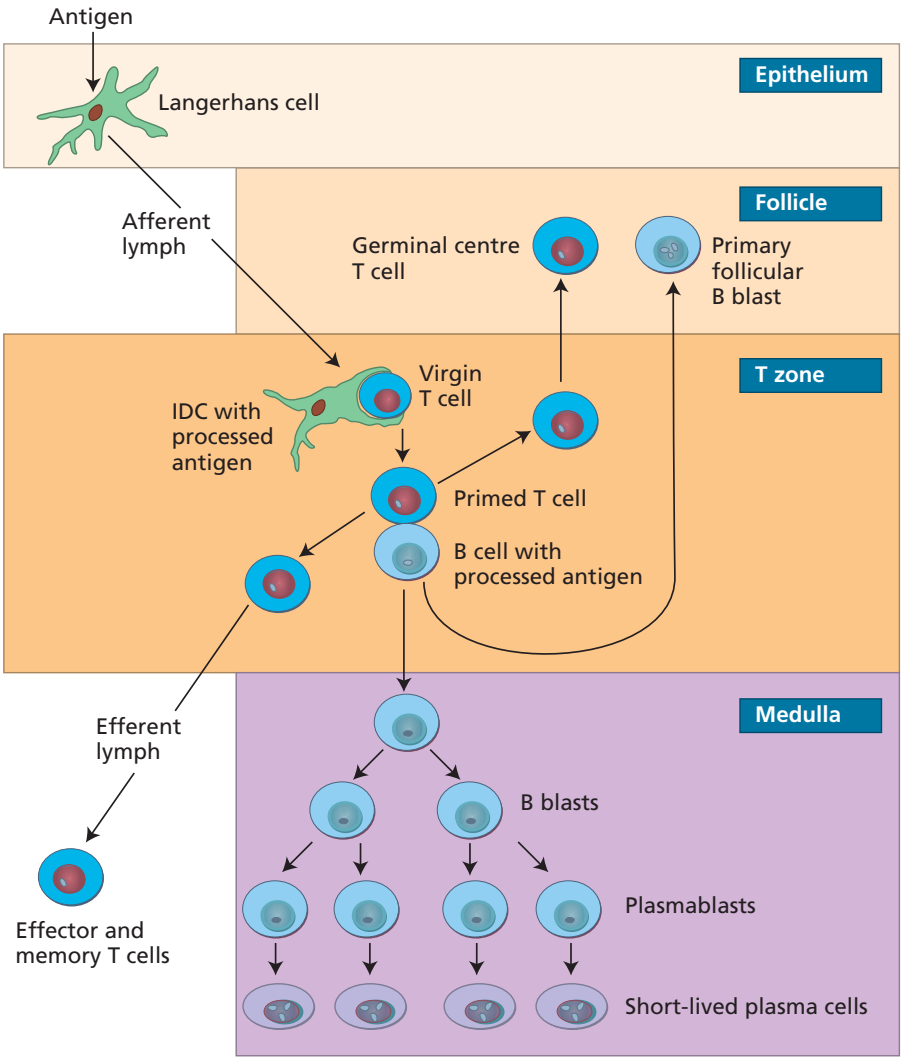


Figure 16.14 T-cell priming and T-cell-dependent B-cell activation in a lymph node. Local disturbance induces dendritic cells in the tissues to take up material from their surroundings. Ingested proteins are processed to peptides inside the cell. Dendritic cells that have taken up antigen migrate through afferent lymphatics to draining lymph nodes or through the blood to the spleen. By the time they reach a T zone, they have differentiated into interdigitating cells (IDCs), which are specialized at presenting antigenic peptides, held in MHC molecules, to recirculating T cells. Virgin T cells migrating through the T zones move over the surface of the interdigitating cells and are activated if they meet antigen they recognize. As the result of this priming process, they move to the outer T zone and become targets for B cells that have taken up and processed antigen. B cells activated by these T cells migrate to extrafollicular foci of B-cell proliferation – the medullary cords in lymph nodes – where they generate short-lived plasma cells. Other B cells migrate to follicles where they may form germinal centres. T cells, after a brief period of proliferation in the T zone, either leave the node as effector cells/recirculating memory T cells or migrate to follicles, where they proliferate further and participate in the selection of B cells that have mutated the immunoglobulin variable-region genes in germinal centres.

blood supply; in the case of the gut and lower half of the body, via the thoracic duct to the left subclavian vein. The solid tissue of the node is composed of variable numbers of roughly cone-like segments (Figure 16.13). The base of each cone abuts onto the subcapsular lymph sinus in the cortex of the node and the apex is in the medulla. These cones fit together, but are separated by the intranodal lymphatic sinuses, to form the roughly kidney-shaped structure of lymph nodes. The cones have three main zones: the follicles in the cortex, the T zones and the medullary cords. The medullary cords form a convoluted apex to the cone. The contents and functions of each of these zones are described in detail in subsequent sections. The

blood supply to the node enters and leaves the node through the medulla, and the specialized small blood vessels termed *high endothelial venules* are located in the T zones. Virgin B and T cells, and some memory cells, enter the T zone by passing across the high endothelial venules. Here they encounter IDCs (Figure 16.14), which present peptides on MHC molecules and initiate immune responses. The series of stages leading to T-cell-dependent B-cell activation in the T zones is shown in Figure 16.14. If a T cell is able to recognize a peptide–MHC complex, it is activated through TCR signalling. Co-stimulation is provided by interaction between molecules such as CD80 and CD86 expressed on IDCs and CD28, which is constitutively expressed by T cells. The effect of this interaction is to bring about changes in the T cell that are collectively called *T-cell priming* (Figure 16.15).

Cytokines are produced by the T cell following this interaction and the nature of the cytokines produced in different situations is considered later. Short-term proliferation of the T and B cells is also induced, and most B cells migrate to local sites of antibody production. In the spleen this is the red pulp, and in lymph nodes the medullary cords. The lifespan of most of these plasma cells is 3 days. The extent of immunoglobulin class switching will depend on the conditions of dendritic cell activation and T-cell priming. In primary antibody responses, the plasma cells generated by B-cell activation in T zones do not have somatic mutations in their immunoglobulin V-region genes. On the other hand, in secondary responses, marginal-zone memory B cells that have somatic mutations in their V-region genes can be induced to migrate to T zones on contact with antigen and give rise to short-lived plasma cells.

The other pathway of migration of T and B cells activated in T zones is to the follicles. Both antigen-specific B blasts and T blasts migrate to the follicles at an early stage in T-zone responses and give rise to *germinal centres*. Germinal centres are present in the first 3 weeks following an immune response and build around B blasts that migrate to follicles and undergo massive clonal expansion such that the spaces in the follicular dendritic cell (FDC) network become filled with blasts. At this stage, changes occur whereby the classical germinal centre structures of dark and light zones develop. The dark zone is formed by the blasts moving to the edge of the FDC network next to the T zone. These blasts, now termed *centroblasts*, activate the somatic hypermutation mechanism that acts on their rearranged immunoglobulin V-region genes. Centroblasts continually give rise to *centrocytes*, non-dividing cells that migrate into the FDC network that forms the light zone of the germinal centre. Centrocytes either leave the light zone within 12 hours or die *in situ*. They can leave the light zone only if they receive antigen-specific selection signals.

Centrocytes pick up antigen from FDCs in order to process peptides for T-cell recognition. This requirement for the B cells to receive T-cell help in the germinal centre protects against potential autoimmune responses, as those B cells that have developed reactivity against autoantigens are unlikely to receive selection signals from germinal centre T cells. Centrocytes that survive selection within germinal centres leave the light zone as either plasmablasts or memory B cells. The plasmablasts migrate to bone marrow or the lamina propria of the gut and differentiate into plasma cells.

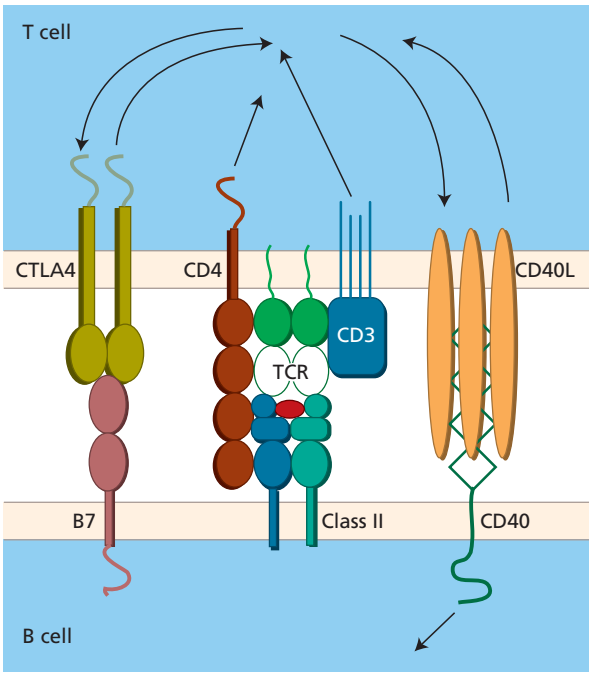


Figure 16.15 Surface molecules involved in T-dependent B-cell activation in T zones. B cells take up antigen that they bind specifically through their surface immunoglobulin. This is internalized, broken down to peptides and the peptides are presented on the B-cell surface, held in the peptide-binding grooves of MHC class II molecules. Cross-linking of surface immunoglobulin by antigen induces the endocytosis of the antigen–antibody complex and signals upregulation of CD40 expression and *de novo* B7.1 and B7.2 expression. If this B cell interacts with a primed T cell that recognizes the peptide complex with class II MHC molecules, there will be co-stimulation through the molecule CD28, which is constitutively expressed by CD4 T cells, and this can result in further signalling through co-stimulatory molecules that are transiently expressed on the T-cell surface; CD40 ligand exemplifies these transiently expressed signalling molecules. These interactions can lead to B- and T-cell proliferation and differentiation and may also induce cytokine secretion by the cells. Cytokine receptor expression by the B cell and the T cell is initiated or upregulated. The arrows indicate that TCR engagement induces CD40 ligand expression and that engagement of these molecules by their counterstructures on the B cell delivers further signals to the T cell. CD40 ligation induces immunoglobulin class switching in the B cell and migration as indicated in Figure 16.12.

Immunoglobulin class switching

Most plasma cells and memory B cells undergo *switch recombination* in which rearranged V-region genes become linked to heavy chain constant region genes downstream from IgD. This process is similar to immunoglobulin variable gene rearrangement, with the DNA forming loops between complementary switch region genes that lie upstream of each set of heavy chain gene exons. The order of the heavy chain constant-region genes that encode the different heavy chain isotypes is located downstream from the variable-region genes: μ, δ, γ3, γ1, a non-functional pseudo-ε gene, α1, γ2, γ4, ε and α2. Thus, switching to γ2 involves looping out μ, δ, γ3, γ1, pseudo-ε and α1. Switch recombination within B cells is driven by cytokines produced following recognition by antigen-specific T cells, and variation in the distribution of immunoglobulin isotypes underlies disorders such as allergy in which there is excess IgE production.

Differentiation of primed T cells into effector cells

Following recognition of antigen, most $CD8^+$ T cells differentiate into cytotoxic effector cells. Differentiation of primed $CD4^+$ cells is directed into one of four major pathways (Figure 16.16) and it appears that factors such as the nature of the antigen, associated inflammatory stimuli and the affinity of T-cell engagement can influence the subtype of CD4 T cell that is produced. $CD4^+$ T cells stimulated in the presence of IL-12, which is produced by macrophages and dendritic cells, are more likely to differentiate into Th1 cells. In contrast, IL-4 produced by either NKT cells or Th2 cells promotes production of Th2 cells and inhibits Th1 cell formation.

Th1 cells produce cytokines that are associated with macrophage activation, granuloma formation and delayed hypersensitivity. These cytokines are principally IL-2, IFN-γ and TNF-β. Macrophages work in concert with Th1 cells to provide a major mechanism by which mycobacteria and other pathogens are destroyed. Inflammatory $CD4^+$ T cells enhance this activity through the action of cytokines following recognition of peptides presented by MHC class II molecules on macrophages. Some Th1 cells have cytolytic potential and can kill infected macrophages through granzyme/perforin or Fas ligand.

Th2 cells produce cytokines associated with antibody production, particularly IL-4, IL-10 and, in some instances, IL-6, although the last of these is mainly produced by macrophages and osteoclasts.

Th17 cells produce large amounts of the cytokine IL-17 and are associated with inflammatory disorders. They may play an important role in the control of fungal disease and could represent an important target for future immunotherapeutic interventions.

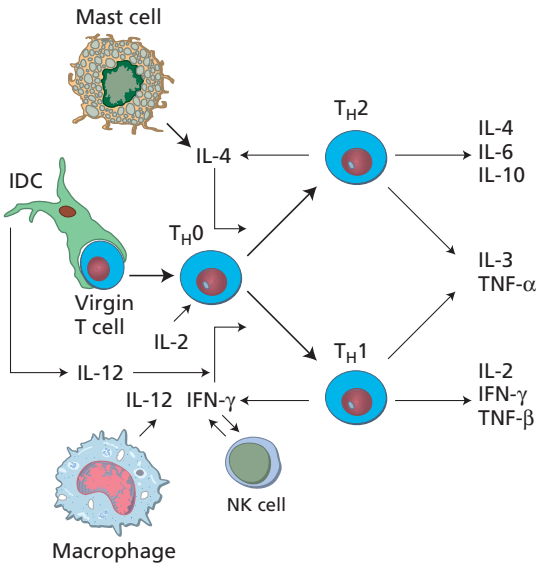


Figure 16.16 Maturation pathways of $CD4^+$ T helper cells. $CD4^+$ T cells that have been activated by antigen acquire the capacity to produce cytokines. The cytokines that they produce depends on the environment in which they are activated. Two major subsets of effector Th cell are recognized, Th1 and Th2 cells. The cytokines produced by Th1 cells tend to promote further Th1 cell formation and inhibit Th2 cell formation, and IL-4 produced by Th2 cells promotes further differentiation towards Th2 cells. Th1-promoting cytokines are also produced by activated macrophages, dendritic cells and NK cells, whereas mast cells produce IL-4. Th17 cells are an additional effector subset that is induced through the activity of IL-23 and produce large amounts of IL-17 and IL-22. T regulatory cells can be induced through the activity of transforming growth factor (TGF)-β and play an important role in the control of autoimmune responses. IDC, interdigitating dendritic cell.

The balance between Th1 and Th2 cells can be an important determinant of immunopathology. Tuberculoid leprosy exemplifies apparent dominance of a Th1 response, while atopic disease seems to reflect an imbalance towards Th2. Nevertheless, manipulation of these cytokine networks provides a potential means of modifying established immune responses to avoid the complications associated with over activity of either Th1 or Th2 cells.

Regulatory $CD4^+$ T cells

Immunological tolerance is mediated by a number of mechanisms, including deletion of self-reactive B and T cells in the bone marrow and thymus respectively. In addition, it is now clear that a specialized population of T lymphocytes can actively suppress immune responses. These cells have been termed regulatory T (Treg) cells and their dominant phenotype is that

of expression of the transcription factor FoxP3 in association with CD4. In addition, these cells usually express high levels of CD25, the IL-2 receptor α-chain, as well as low levels of CD127, a component of the IL-7 receptor. Treg cells are involved in maintenance of peripheral T-cell tolerance to self antigens and protection against autoimmunity. They constitute approximately 1–5% of peripheral blood CD4 lymphocytes and may be derived either from a discrete thymic lineage (natural Treg cells) or following the differentiation of $CD4^+CD25^-$ T cells in the presence of cytokines such as transforming growth factor (TGF)-β (induced Treg cells). The role of Treg cells in a range of human diseases is being explored at present and typically they are deficient, in either number or function, in autoimmune disorders, whilst relative overactivity may be apparent in malignant disease. Immunotherapeutic strategies that involve cellular therapy with Treg cells, or which utilize reagents such as anti-CTLA-4 antibody to suppress their function, are now in clinical trial.

The role of co-stimulatory blockade in the treatment of malignant disease

Whenever a T cell or NK cell recognises a potential antigen on a target cell, the decision as to whether or not the cytotoxic cell is activated will depend on the balance between the activatory and inhibitory signals that the effector cell receives. The immunoglobulin superfamily and the tumour necrosis factor–TNF receptor superfamily are the two major gene families that encode such molecules, and many different proteins are expressed at various stages of cell differentiation and activation. The major co-inhibitory molecules include PD1, CTLA-4, LAG-3 and BTLA, whereas co-stimulation is provided through proteins such as CD28, ICOS, 4-1BB and OX40. A common finding in patients with cancer is that lymphocytes that have the potential to recognize and attack the tumour are 'exhausted' or functionally inactivated due to an alteration in the balance of inhibitory or activatory molecules that they express on their surface. This is believed to result from exposure to immune evasion mechanisms that are mediated by the tumour cell. As such, therapeutic monoclonal antibodies such as anti-PD1 and anti-CTLA-4, which are able to block inhibitory signalling, are now widely used in the treatment of cancer where they are showing great promise. Indeed, antibody therapy against PD1 or its ligand PDL1 is achieving apparent long-term cure in a subset of patients with solid tumours (Figure 16.17) and is also showing huge promise in the treatment of Hodgkin lymphoma.

Chimeric antigen receptors in the treatment of haemopoietic malignancy

As discussed earlier, in recent years there have been considerable advances in the understanding of the structure and biochemistry of signalling within immune receptor complexes. This has led to the development of artificial 'chimeric antigen receptors' (CARs) which can be used in the treatment of malignant disease. In most cases these involve grafting a monoclonal antibody onto an intracellular signalling domain, and then transfecting this into effector lymphocytes using a viral or non-viral vector. This process redirects the specificity of the lymphocyte such that it is now able to kill target cells that express the target for the monoclonal antibody. The 'first generation' CARs typically attached the antibody to the TCR zeta domain, whereas the 'second generation' variants attached an additional intracellular signalling domain such as one from CD28 or ICOS. Third generation CAR molecules are proving to be the most powerful in clinical practice and they include multiple intracellular signalling domains to maximise the proliferative and cytotoxic capacity of the transfected cell. The greatest success so far has been in the use of CD19-targeted CARs which are used in the treatment of B cell malignancy (Figure 16.18).

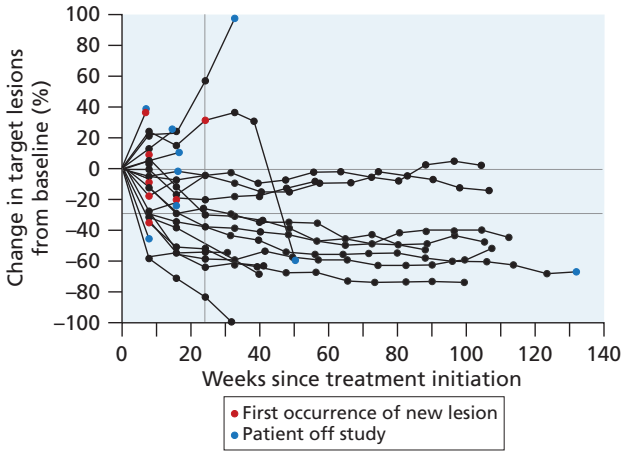


Figure 16.17 Activity of anti-PD-1 antibody in patients with treatment-refractory melanoma. The plot shows change in tumour burden in 27 patients with melanoma who received anti-PD-1 every 2 weeks. Patients who did respond usually showed an effect within 16 weeks, which was durable in many cases. (Source: Topalian *et al*., 2012. Reproduced with permission of the Massachusetts Medical Society.)

Cytokines and their classification

Cytokines are soluble proteins produced by leucocytes and other cells that influence the behaviour of cells that carry cytokine receptors. Many are secreted, but others, such as TNF-α, are cell membrane proteins that are active when bound to the cell that produced them, but which also have activity as a soluble protein. The potency of cytokines *in vivo* has been exemplified clearly by the dramatic effects of recombinant granulocyte colony-stimulating factor (G-CSF) in inducing accelerated recovery of neutrophil counts after the administration of myelotoxic therapy

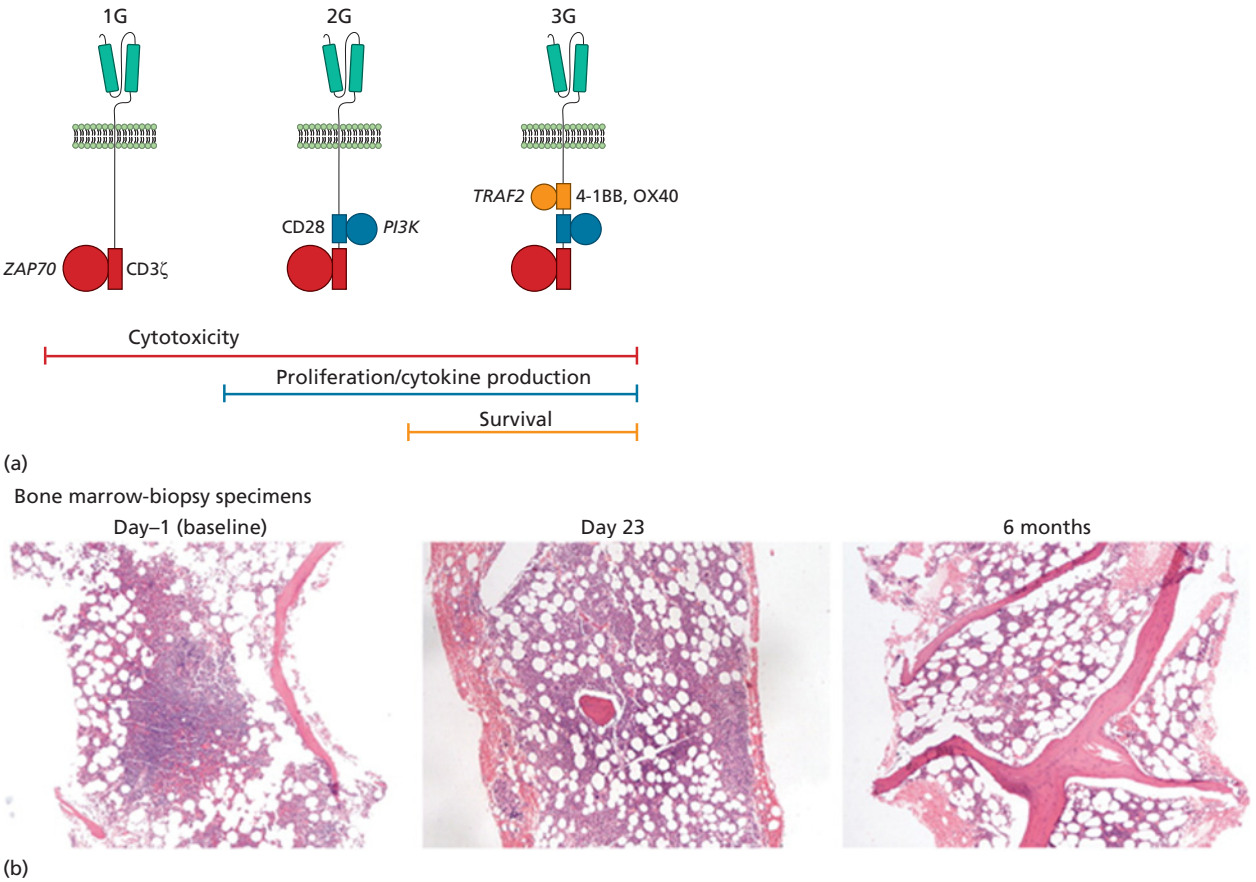


Figure 16.18 Chimeric antigen receptors are useful in the treatment of B cell malignancy. (a) Schematic representation of the first, second and third generation CAR vectors. (b) Example of bone marrow response in the treatment of chronic lymphocytic leukaemia following the infusion of CD19-targeted CAR-transfected autologous T cells. (Source: Porter *et al.*, 2011. Reproduced with permission of Massachusetts Medical Society.)

and the effect of interferon-α in the management of haematological malignancies. It seems likely that many cytokines act only at very short range and some appear to be more potent as membrane-bound forms than as released proteins. The families of cytokines are summarized in Table 16.5.

Chemokines and their classification

Chemokines are a class of cytokines with chemoattractant properties and are all related in sequence. There are two groups, the CC chemokines, which have two adjacent cysteine residues in their sequence, and the CXC chemokines, in which these two cysteine residues are separated by another amino acid. Chemokine receptors are integral membrane proteins linked to G-proteins and have seven membrane-spanning domains. They are classified according to the type of chemokine that they bind, i.e. CCR1–9 and CXCR1–5. Cells that express chemokine receptors are attracted towards an increasing concentration of chemokine molecules, and these interactions are critical to many functions of the innate and adaptive immune response (Table 16.6). Some chemokines are expressed in a constitutive fashion, whereas others are released in response to inflammation. Antibodies and drugs that block chemokines or their receptors are being used in a range of different disorders, including prevention of graft versus host disease after stem cell transplantation.

Interpretation of blood lymphocyte counts

Blood provides the most accessible view of the lymphoid system, but it must be remembered that peripheral blood contains only around 2% of total body lymphocytes. The cells present in the blood are in transit and many are recirculating T and B lymphocytes that will pass rapidly into secondary lymphoid organs in less than 30 minutes.

Table 16.5 The cytokine and cytokine receptor families.

Cytokine family	Members of family	Type of receptor
β-Trefoil	IL-1α, IL-1β	Immunoglobulin family
Haemopoietins	IL-2, IL-3, IL-4, IL-5, IL-6, IL-7, IL-9, IL-13, IL-15, GM-CSF	Class I cytokine receptor
	IL-10, interferons	Class II cytokine receptor
Tumour necrosis factor	TNF-α, TNF-β	Nerve growth factor receptor family
Cysteine knot	NGF	Nerve growth factor receptor family
	TGF-β	Serine threonine kinase
Chemokines	IL-8, MIP-1α, MIP-1β, I-309, MCP-1, MCP-2, MCP-3, γIP-10	Rhodopsin family

Table 16.6 The major chemokines within the CXC and CC subgroups.

Chemokine	Production	Receptors	Cells that are attracted	Effects
CXC subgroup				
IL-8	Monocytes Macrophages Fibroblasts	CXCR1 CXCR2	Neutrophils T cells	Inflammation Angiogenesis
β-TG	Platelets	CXCR2	Neutrophils	Inflammation
GROα,β,γ	Monocytes Endothelium	CXCR2	Neutrophils T cells Fibroblasts	Inflammation Angiogenesis
IP-10	Endothelium Monocytes T cells Fibroblasts	CXCR3	T cells NK cells Monocytes	Promotes Th1 immunity Immunostimulation
SDF-1	Stromal cells	CXCR4	Stem cells Lymphocytes	Stem cell homing Haemopoiesis
CC subgroup				
MIP-1α	Monocytes T cells Fibroblasts	CCR1, 3, 5	Monocytes NK and T cells Dendritic cells	Th1 immunity
MIP-1β	Monocytes Macrophages Neutrophils Endothelium	CCR1, 3, 5	Monocytes NK and T cells Dendritic cells	Th immunity
MCP-1	Monocytes Macrophages Fibroblasts Keratinocytes	CCR2B	Monocytes NK and T cells Dendritic cells	Th2 immunity
RANTES	T cells Endothelium Platelets	CCR1, 3, 5	Monocytes NK and T cells Dendritic cells	Inflammation T-cell activation
Eotaxin	Endothelium Monocytes Epithelium	CCR3	Eosinophils Monocytes T cells	Allergy

Table 16.7 Normal ranges for lymphocyte subsets in the blood.*

	0–2 months		2–3 months		4–8 months		1–2 years		2–5 years		5–12 years		Adults	
Percentile	*5*	*95*	*5*	*95*	*5*	*95*	*5*	*95*	*5*	*95*	*5*	*95*	*5*	*95*
Total lymphocytes	3.2	8.5	2.9	8.8	3.6	8.8	2.2	8.3	2.4	5.8	1.8	5.8	1.0	3.4
$CD3^+$ (all αβ and γδ T cells)			2.1	6.5	2.3	6.5	1.5	5.4	1.6	4.2	0.9	2.6	0.6	2.5
$CD4^+$ (class II MHC-restricted αβ T cells)	1.2	5.3	1.4	5.6	1.4	5.7	1.0	3.6	0.9	2.9	0.5	1.4	0.35	1.5
$CD8^+$ (class I MHC-restricted αβ T cells, some γδ T cells and NK cells)			0.7	2.5	0.7	2.5	0.6	2.2	0.6	1.9	0.4	1.2	0.23	1.1
CD4/CD8 ratio	1.1	4.5	1.1	4.4	1.1	4.2	1.0	3.0	0.9	2.7			0.66	3.5
$CD3^-CD57^+$ or $CD56^+$ NK cells					0.3	0.7			0.2	0.6			0.2	0.7
$CD19^+$ or $s\kappa^+$ or $s\lambda^+$ total B cells					0.5	1.5			0.5	1.3			0.04	0.7

There is considerable variation in the normal ranges reported from different studies and this table is only intended to be illustrative.
*All values (except CD4/CD8 ratio) are $\times 10^9$/L.

The numbers of different lymphocyte subsets normally found in the blood in different age groups is given in Table 16.7. These numbers are derived from studies in which whole blood was labelled by fluorescent dye/monoclonal antibody conjugates, followed by red cell lysis and flow cytometry. When this method is used to measure the proportion of lymphocytes that belong to different subsets, much of the interlaboratory variation that is observed in calculating absolute numbers of subsets can be attributed to the method of measuring the total white cell count and the percentage of lymphocytes. A wide range of results is to be expected between different individuals. The most consistent variation is seen in childhood, but from adolescence onwards age-related changes are small and there is also little difference in relation to ethnicity or gender.

Alteration of the lymphocyte count can result from an absolute change in the number of cells or alteration in the distribution of lymphocytes within tissues. Redistribution of lymphocytes accounts for much of the variation in lymphocyte subset numbers found in serial measurements within a healthy individual. Some of these changes in lymphocyte number follow a diurnal pattern, with peak levels at night and nadir in the morning; accordingly, time of sampling should be taken into account.

Increased number of effector cells in the blood usually reflects an active immune response. Analysis of the phenotype of these effector cells provides some information on the type of immune response, especially in differentiating between cytotoxic and inflammatory $CD4^+$ T cells. Particularly in viral infections, this response may be of sufficient magnitude to cause a lymphocytosis. Some non-malignant causes of lymphocytosis are given in Table 16.8.

Redistribution of lymphocytes is probably the cause of the lymphocytosis seen in *Bordetella pertussis* infection. Although lymphocytosis is uncommon in bacterial infections, in children over the age of 6 months the second and third weeks of infection with pertussis are usually associated with a lymphocytosis in excess of 10×10^9/L (in some cases $> 50 \times 10^9$/L). The lymphocytosis consists of small lymphocytes and is believed to be caused by a protein toxin from *B. pertussis* that prevents migration of lymphocytes across endothelium into tissue.

Table 16.8 Non-malignant causes of lymphocytosis.

Viral infections
Infectious lymphocytosis, infectious mononucleosis, cytomegalovirus infection; occasionally rubella, hepatitis, adenoviruses, varicella, HIV, human herpesvirus 6, mumps, chickenpox, dengue

Bacterial infections
Pertussis; occasionally healing tuberculosis, brucellosis, secondary and congenital syphilis, cat scratch fever, typhoid fever, diphtheria

Protozoal infections
Toxoplasmosis; occasionally malaria

Other conditions
Serum sickness, allergic drug reactions, splenectomy, dermatitis herpetiformis, metastatic melanoma, hyperthyroidism, congenital adrenal hyperplasia

Acute infectious lymphocytosis is a benign disease, usually of children. In most cases there are no symptoms, but in some there is fever and in a small proportion gastrointestinal symptoms. Increase in the size of secondary lymphoid organs, anaemia and thrombocytopenia are rare. There is an increased number of small lymphocytes, persisting for 3–7 weeks, with an average peak level of 30–40 $\times$ 10^9/L. This is usually associated with an eosinophilia (average 2 $\times$ 10^9/L), but the aetiology of the condition is unknown.

Some *lymphopenias* predominantly reflect redistribution rather than a depletion of total body lymphocyte numbers. A dramatic short-term lymphopenia is induced by corticosteroids. This causes the retention of lymphocytes in secondary lymphoid organs but these are released again after about 2 days and the blood lymphocyte count returns to near-normal levels. Endogenous secretion of corticosteroids during acute illnesses may be partly responsible for the lymphopenias often seen in conditions such as heart failure or pneumonia. In many other conditions, lymphopenia reflects an increased rate of death of lymphocytes and/or a reduction in their rate of formation. Some of the causes of secondary lymphopenia are listed in Table 16.8. A normal absolute lymphocyte count can belie an underlying deficit of one or more lymphocyte subsets. This is often seen in HIV infection, when a severe deficit of $CD4^+$ T cells may be disguised by expansion of $CD8^+$ T cells.

Infectious mononucleosis

Infectious mononucleosis (IM) is caused by a primary infection with Epstein–Barr virus (EBV), a herpesvirus which infects B lymphocytes. EBV enters B cells via CD21, a surface receptor for the C3d component of complement. After the acute infection has been resolved, lifelong subclinical infection is maintained, with a low frequency of infected B cells and detectable virus in the saliva – a main vehicle for contagion.

EBV infection of children usually results in immunity without development of the typical clinical manifestations of IM. This immunity can be detected serologically and is associated with lifelong protection. Usually only after the age of 10 years is infection by EBV associated with the clinical manifestations of IM. In developing countries, the rate of seroconversion before the age of 10 years can be so high that clinically evident IM is rare. IM has its highest prevalence in young adults and a study of medical students showed that the syndrome was seen in 25% of individuals who underwent seroconversion. It is uncommon after the age of 30 years and rare after the age of 40 years. The determinants of clinical symptoms are unknown, but may relate to host genotype, viral load or viral polymorphisms.

Clinical features

The symptoms of IM usually develop abruptly, with fatigue, malaise and fever after an incubation period of up to 7 weeks. These symptoms last for about 3 weeks. Sore throat occurs in over 80% of cases and is usually accompanied by anorexia and nausea. The sore throat develops in the first week and subsides in the second week, rarely generating severe symptoms or massive tonsillar/pharyngeal oedema. Sharply defined red spots at the junction of the soft and hard palates are of diagnostic value. Positive throat swabs for β-haemolytic streptococci are frequently found. Bilateral non-inflammatory cervical lymphadenopathy is almost invariable, and inguinal and axillary lymphadenopathy is usual. The spleen is palpable in more than half of cases, although only occasionally does it extend to the iliac crest. These secondary lymphoid organs increase in size in the first week and subside slowly after the second week. Slight hepatomegaly and jaundice occurs in about 10% of cases. Fever is present in most cases, but of no characteristic type and may be transient. A few patients develop a fine macular rash, but rashes are more usually found as temporary reactions to penicillin and especially ampicillin.

Blood picture

In most patients, IM is associated with a leucocytosis; this peaks in the second and third weeks and usually persists for 1–2 months (the first week is occasionally associated with a leucopenia). In two-thirds of patients, the leucocytosis ranges from 10 to 20 $\times$ 10^9/L, but in some cases may substantially exceed these levels. The leucocytosis is attributable to an absolute increase in numbers of both normal small lymphocytes and of activated T cells (atypical lymphocytes). Most of the activated cells are $CD8^+$ T cells, but they also include $CD4^+$ T cells and $CD3^-$ NK cells. Most of these activated lymphocytes are cytotoxic for virus-infected cells and target viral peptides presented on MHC class I molecules. Although infection of B cells by EBV stimulates their proliferation, this appears to be controlled by the T-cell response such that the proportion of blood mononuclear cells that is EBV infected rarely exceeds 0.1%.

A peripheral neutrophilia may be seen early in the disease, but a neutropenia is equally common and eosinophilia is not unusual. Thrombocytopenia may occur and is occasionally severe. Anaemia is rare and then usually associated with anti-i antibodies. EBV infection may trigger a haemophagocytic syndrome in rare cases.

Serological changes

Three categories of antibody are produced as a result of EBV infection: virus-specific, heterophile and autoimmune. The first virus-specific antibodies to appear are directed against the EBV capsid antigen (VCA). IgM anti-VCA antibodies probably develop during the incubation period and peak in the second week of the illness followed by a rapid decline. IgG anti-VCA antibodies peak in the second to third weeks and persist for life. Most patients also have a transient response to EBV early

antigen (EA), which peaks in weeks 2–3. Antibodies to EBV nuclear antigen (EBNA) do not develop until some weeks into the illness, but are present lifelong in all patients by 6 months. Serological diagnosis of acute IM is most accurately made by the presence of IgM anti-VCA and anti-EA antibodies and the absence of anti-EBNA antibodies.

Paul and Bunnell demonstrated that patients with IM have serum agglutinins directed against sheep erythrocytes (heterophile antibodies) and that a serum titre in excess of 1:112 is highly suggestive of IM. Similar agglutinins are found in low titre in healthy individuals (directed against Forssman antigen) and in some leukaemias and lymphomas, as well as serum sickness. However, in these conditions the heterophile antibody can be absorbed onto guinea pig red cells. Formalin-treated horse erythrocytes appear to be agglutinated exclusively by heterophile antibodies of IM, and this forms the basis of the *Monospot test*. Heterophile antibodies provide the routine serological test for IM, but are commonly negative, particularly in children and in patients over the age of 25 years. EBV-specific serodiagnostic tests should be applied in cases with strong clinical suspicion, but negative heterophile antibodies. Total serum immunoglobulin levels increase around 4 weeks following onset of symptoms, and raised levels may persist for many months. The greatest proportional increase is in IgM, but IgG may also be raised. The specificity of most of these immunoglobulins is unknown, but a variety of autoantibodies may be found, including cold-reactive anti-i antibodies, Donath–Landsteiner cold haemolysins and, occasionally, antibodies against smooth muscle, thyroid, stomach, rheumatoid factors and antinuclear antibodies.

Differential diagnosis and treatment

The diagnosis and course of IM are usually uncomplicated. Signs of significant respiratory, cardiovascular, intestinal, urinary or joint disease make consideration of other diseases mandatory; some of these other diseases are listed in Table 16.9. Perhaps the commonest problem is when the patient is heterophile antibody negative. In this situation other viral infections, particularly cytomegalovirus (CMV), should be considered, with assay for CMV-specific IgM. Primary EBV infection is rare in older patients and there may not be conspicuous lymphadenopathy. Occasionally, the blood picture may raise the suspicion of a leukaemia, in which case immunophenotyping of the blood mononuclear cells may be appropriate. Persistent lymphadenopathy beyond a few weeks suggests the need for diagnostic biopsy, particularly if heterophile antibodies are negative, but the possibility of a false-positive Monospot test should also be considered. Virus-specific serology may be helpful in both these situations.

There is no specific therapy for IM. In patients with severe fever or lymphadenopathy, corticosteroids produce prompt lysis of fever and reduction of lymph node hyperplasia. Steroids may be indicated in management of associated haemolytic anaemia, thrombocytopenia, progressive neurological complications and incipient airway obstruction. Patients should be advised of the small risk of splenic rupture from minor abdominal trauma and contact sports should be avoided for several months.

Table 16.9 Causes of secondary lymphopenia.

Infections
Influenza; occasionally other viral infections, Colorado tick fever, miliary tuberculosis, pneumonia, septicaemia, malaria, HIV
Loss of lymphocytes
Intestinal lymphangiectasia, Whipple disease, severe right-sided heart failure, rarely inflammatory bowel disease, lymphatic fistula
Therapeutic procedures
Radiotherapy, anti-lymphocyte globulin, corticosteroids, cytotoxic drugs, purine analogues
Neoplastic conditions
Metastatic carcinoma, advanced Hodgkin disease
Nutritional/metabolic
B_{12} or folate deficiency, zinc deficiency, uraemia
Other conditions
Systemic lupus erythematosus and other collagen vascular diseases, myasthenia gravis, aplastic anaemia, graft-versus-host disease, pancreatic necrosis, sarcoidosis, idiopathic

X-linked lymphoproliferative syndrome (Duncan syndrome) is a rare inherited X-linked condition in which there is specific immunodeficiency against EBV. These individuals may die as a result of primary infection or the subsequent development of an EBV-driven B-cell lymphoma. Patients receiving immunosuppressive therapy for allografts and who are carriers of EBV can develop proliferations of B lymphocytes that carry the EBV genome. These cases of *post-transplant lymphoproliferative disease* are heterogeneous, and B-cell proliferations vary from a polyclonal diffuse B-cell hyperplasia to monoclonal B-cell lymphomas. There is evidence for a causal role of EBV in the development of Burkitt lymphoma, in association with both HIV infection and in the form endemic in African children. In less than 20% of sporadic cases of Burkitt lymphoma, EBV-associated DNA can be demonstrated in the lymphoma cells.

Secondary associations of infectious mononucleosis

In the 1950s it was reported that there was an increased incidence of Hodgkin lymphoma in individuals with a history of IM. Although the epidemiological challenges of confirming this association are profound, several very large studies have revealed

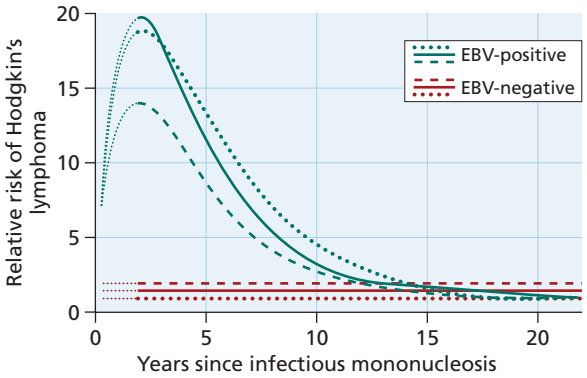


Figure 16.19 Relative risk of Epstein–Barr virus (EBV)-positive and EBV-negative Hodgkin lymphoma after infectious mononucleosis. Solid lines represent the relative risks of EBV-positive (blue) and EBV-negative (red) Hodgkin lymphoma, given that EBV status was determined in an unbiased way and that the missing data on viral status in 11 tumours were uninformative with respect to their true EBV status. Short dashed lines represent the relative risks of EBV-positive and EBV-negative Hodgkin lymphoma given that all tumours whose EBV status was unknown were EBV-positive. Long dashed lines represent the relative risks given that all tumours whose EBV status was unknown were EBV-negative. The analyses were restricted to the period 2 years or more after infectious mononucleosis. (Source: Hjalgrim *et al*., 2003. Reproduced with permission of Massachusetts Medical Society.)

a definite association. Hjalgrim and colleagues studied 38,555 people with a confirmed diagnosis of IM and identified 29 cases of Hodgkin lymphoma, of which 16 had evidence of EBV. The study revealed a 4.1-fold increased risk of EBV-associated Hodgkin lymphoma in patients with a history of IM, with a median incubation time of just over 4 years to development of lymphoma (Figure 16.19).

IM is also associated with an increased risk of multiple sclerosis later in life. Thacker and colleagues have estimated this relative risk at 2.3 and this is likely to reflect the growing appreciation that EBV infection is associated with the pathogenesis of multiple sclerosis, although the mechanisms involved are unclear.

It is not yet clear whether individuals who suffer from IM carry a genetic predisposition to a variety of immunopathological disorders or if these secondary events are a direct consequence of the IM syndrome. However, these observations do offer important clues as to the pathogenesis of these important disorders.

Selected bibliography

Blachly JS, Baiocchi RA (2014) Targeting PI3-kinase (PI3K), AKT and mTOR axis in lymphoma. *British Journal of Haematology* **167**(1): 19–32.

Chapel H, Haeney M, Misbah S, Snowden N (2014) *Essentials of Clinical Immunology*, 6th edition. Wiley-Blackwell, Hoboken, NJ.

Hjalgrim H, Askling J, Rostgaard K *et al*. (2003) Characteristics of Hodgkin's lymphoma after infectious mononucleosis. *New England Journal of Medicine* **349**: 1324–32.

Hutchinson CV, Dyer MJ (2014) Breaking good: the inexorable rise of BTK inhibitors in the treatment of chronic lymphocytic leukaemia. *British Journal of Haematololgy* **166**(1): 12–22.

Maude SL, Teachey DT, Porter DL, Grupp SA (2015) CD19-targeted chimeric antigen receptor T-cell therapy for acute lymphoblastic leukemia. *Blood* **125**(26): 4017–23.

Porter DL, Levine BL, Kalos M, Bagg A, June CH (2011) Chimeric antigen receptor-modified T cells in chronic lymphoid leukemia. *New England Journal of Medicine* **365**(8): 725–33.

Reshef R, Luger SM, Hexner EO *et al.* (2012) Blockade of lymphocyte chemotaxis in visceral graft-versus-host disease. *New England Journal of Medicine* **367**(2): 135–45.

Delves PJ, Martin SJ, Burton DR, Roitt IM (2011) *Roitt's Essential Immunology*, 12th edition. Wiley-Blackwell, Hoboken, NJ.

Thacker EL, Mirzaei F, Ascherio A (2006) Infectious mononucleosis and risk for multiple sclerosis: a meta-analysis. *Annals of Neurology* **59**: 499–503.

Topalian SL, Hodi FS, Brahmer JR *et al*. (2012) Safety, activity, and immune correlates of anti-PD-1 antibody in cancer. *New England Journal of Medicine* **366**: 2443–54.

Topalian SL, Drake CG, Pardoll DM (2015) Immune checkpoint blockade: a common denominator approach to cancer therapy. *Cancer Cell* **27**: 450–61.

CHAPTER 17

The spleen

17

Paul Moss

Queen Elizabeth Hospital and University of Birmingham, Birmingham, UK

The spleen performs a number of important roles in homeostasis. Its function is affected in a range of primary blood diseases and systemic disorders. These haematological effects are usually a minor phenomenon, but in some cases may come to dominate the clinical presentation.

Evolution of the spleen

The spleen evolved around 500 million years ago with the appearance of the adaptive immune system between the origin of vertebrates and the development of the jawed vertebrates. It has undergone more functional and structural diversification than the thymus and this may reflect its varied physiological roles. The spleen shares a capsule of fibromuscular tissue in all animals and this extends inwards as a reticular network. The red pulp is present in all species, but there have been marked changes in the anatomy of the white pulp and this is likely to reflect increasing sophistication of cellular immunity.

The spleen is derived from a condensation of mesenchymal cells that arise in the mesentery close to the pancreatic rudiment. The mesenchymal cells differentiate into reticulum cells, pluripotent stem cells and colony-forming units. Together with the liver, the spleen has a transient role in haemopoiesis from the third month, continuing until birth (Chapter 11). From about 20 weeks the bone marrow becomes a site of haemopoiesis and this increases rapidly during the last trimester of pregnancy, whereas haemopoietic activity in the spleen disappears. There is no evidence for a specific inhibitor of haemopoiesis and the spleen remains a potential site for blood production in the adult, particularly for the maintenance of erythropoiesis. This occurs during pathological states and the development of such extramedullary haemopoiesis after birth is described below.

Structure and function

The major functions of the spleen are: (i) filtration and 'quality control' of red cells within the circulation, (ii) capture and destruction of blood-borne pathogens and (iii) generation of adaptive immune responses. In order to achieve these aims the spleen has evolved a unique anatomical structure that is based on the filtering of blood through two main systems. These consist of a *white pulp*, which is concerned mainly with immunological function, and a *red pulp*, which regulates the selection of red cells for re-entry into the circulation. A major feature is the presence of both '*open*' and '*closed*' *circulatory systems* and these are described in detail in the subsequent sections.

The spleen lies in the left hypochondrium, with its long axis beneath the proximal half of the tenth rib. Its convex surface rests under the diaphragm whereas the visceral surface is in contact with the stomach and left kidney, with the tail of the pancreas reaching the hilum at the medial side. The normal spleen weighs about 150–250 g, but there is considerable variation between normal individuals and at various ages in the same individual. At puberty it weighs about 200–300 g but after the age of 65 years this decreases to 100–150 g or less. In the adult its length is 8–13 cm, width 4.5–7.0 cm, surface area 45–80 cm^2 and volume less than 275 cm^3. Interestingly, up to 10% of people have accessory spleens, normally as a single piece of tissue that can be found in a variety of sites, either locally or more distantly in the abdomen. A spleen greater than 14 cm long is usually palpable (see later).

Postgraduate Haematology, Seventh Edition. Edited by A Victor Hoffbrand, Douglas R Higgs, David M Keeling and Atul B Mehta.

It enlarges in a wide range of diseases and has been measured at up to a massive 2 kg or more in some blood disorders.

The spleen has a complicated structure and several different functions. It is enclosed by a connective tissue framework that extends inwards to form a fibrous network. Blood enters at the pelvis and the majority of vessels open into these open networks (the *red pulp*) before re-entering the closed venous system. There is no afferent lymphatic to the spleen and the efferent lymphatic system leaves along the route of the splenic vein. The spleen contains a large amount of lymphatic tissue that is mostly concentrated in concentric rings around the arterioles (*white pulp*). Between the red pulp and white pulp is an *intermediate marginal zone*, which lies at the periphery of the white pulp, blending into the red pulp.

Splenic blood flow and the red pulp

The circulation within the spleen is illustrated in Figure 17.1. Blood is brought to the spleen via the splenic artery, which branches into the trabecular arteries and then arborizes in a pattern that lacks interarterial connections and thus effectively generates end arteries. The central arteries acquire a coaxial sheath of lymphocytes containing lymphoid follicles and this together constitutes the white pulp (see below). These central arteries then split into many arterioles and capillaries, some of which terminate in the white pulp, while others go on to enter the red pulp.

In the red pulp there are two major forms of blood circulation: a *closed system*, typical of the rest of the vascular system, in which arteries and veins communicate through endothelial-lined vessels, and an *open circulation*, in which arterioles terminate in free endings on *splenic cords* (also known as the cords of Billroth) and from which cells must subsequently cross an endothelial layer to re-enter the circulation. The cords consist of a fibroblast-like reticular meshwork containing numerous macrophages and erythrocytes and they are critical to the filtration function of the spleen. Red cells in the cords need to gain entry to a venous sinus if they are to be allowed to re-enter the systemic circulation. The sinuses, 20–40 mm in diameter, are lined by endothelial and adventitial cells with a basement membrane and possess narrow interendothelial spaces in the sinus wall through which flexible red cells may pass. Red cells with inflexible membranes, mostly elderly, are not able to pass through these gaps and are ingested by macrophages in the cords. The great majority of blood flows in this unique open

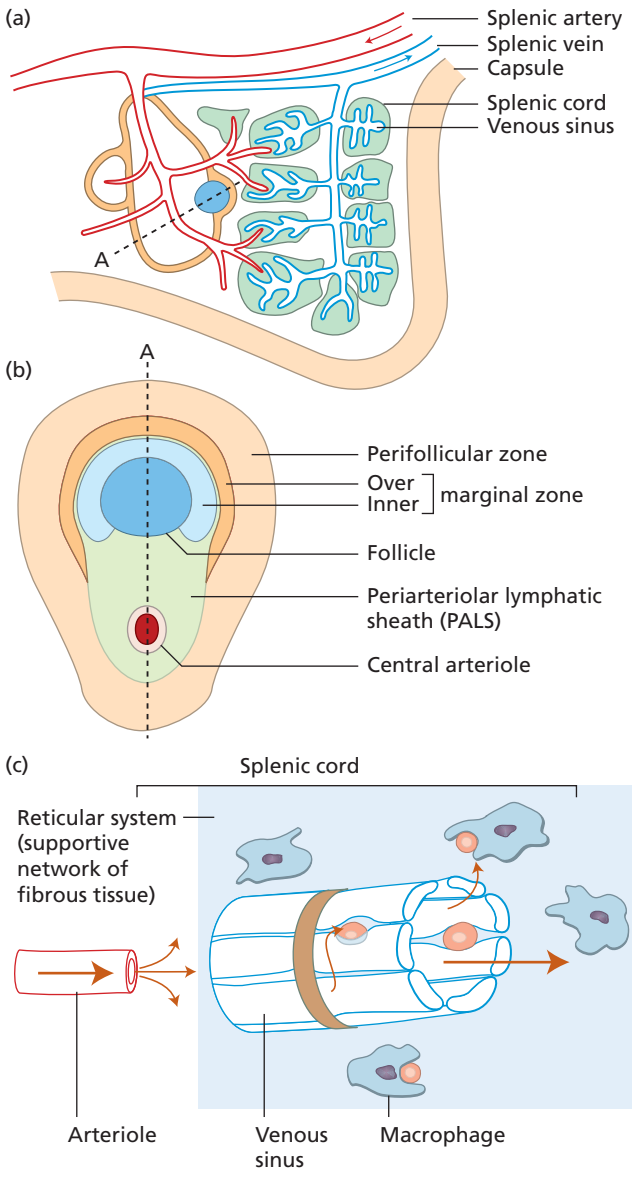


Figure 17.1 The vascular and lymphatic organization of the spleen. (a) Schematic representation of blood flow in the spleen. Blood enters through the splenic artery and then breaks up into splenic arterioles. The white pulp is an aggregation of lymphoid tissue around these vessels. The arterioles then open into splenic cords lined with macrophages. Blood must then re-enter the circulation by passing into a venous sinus. (b) Cross-section of white pulp at section A–B in (a). The central arteriole is surrounded by a periarteriolar lymphoid sheath of T lymphocytes and occasional follicles that are rich in B cells. The marginal zone is divided into an inner and outer zone around which there is a large perifollicular zone. (c) Schematic representation of the organization of the red pulp. Blood cells are released by a splenic arteriole into a reticulum-lined splenic cord. Red cells that have flexible membranes are able to gain entry into the venous sinus by passing through gaps in the endothelial lining. Cells that are unable to achieve this are 'filtered' from the blood through ingestion by macrophages. (Source: Mebius and Kraal, 2005. Reproduced with permission of Nature Publishing.)

system and only a minority of arteries connect directly to the trabecular vein in a typical vascular closed system.

There is thus both a rapid and a slow transit component in the splenic circulation. The rapid transit is of the order of 1–2 min, whereas the open circulation has a circulation time of 30–60 min or even longer. In normal subjects the blood flows through the spleen at a rate of about 5% of blood volume per minute, so that each day the blood has repeated passages through the spleen. During the flow, by a process called *plasma skimming*, the plasma and leucocytes pass preferentially to the white pulp, while the red cells remain in the axial stream of the central artery. The passage of cells into the sinuses is controlled by their ability to squeeze through the interendothelial spaces, assisted by contraction of the reticular cells.

Blood pooling

The normal red cell content of the spleen is 30–70 mL, which represents less than 5% of the total red cell mass. When the spleen is enlarged, expansion of the vascular bed occurs. This results in a considerable pool with a high haematocrit and only a slow exchange of red cells with the general circulation. In states of massive hypersplenism, such as myelofibrosis, as much as 40% of the red cell mass may be pooled in the spleen. This pooling will functionally exclude a relatively large volume of red cells from the main arteriovenous circulation, and thus be an important cause of anaemia. In such cases, it should be noted that the red cell mass, as measured by a radionuclide labelling technique, may give a misleadingly normal result, whereas the peripheral blood packed cell volume will give a more reliable measurement of the effectively circulating red cell mass. In splenomegaly due to cellular infiltration, the pool is less prominent. Conversely, in the congestive splenomegaly of portal hypertension, an increased red cell pool is a dominant feature.

The normal spleen contains a reservoir of granulocytes that is in dynamic equilibrium with the circulating granulocytes. It represents 30–50% of the total marginating pool, with a mean transit time through the spleen of about 10 min. Splenic sequestration of granulocytes is thought to be responsible for the neutropenia that often occurs in patients with splenomegaly. Platelets have also been shown to have a significant reservoir in the spleen and are rapidly interchangeable with the circulation. In normal subjects, 20–40% of the total platelet mass is pooled in the spleen and the platelets spend up to one-third of their lifespan there. The pool increases when the spleen is enlarged. This pooling and temporary sequestration must be distinguished from the destruction of platelets in the spleen that occurs in many cases of thrombocytopenia.

There is no evidence that the normal spleen is involved in the regulation of plasma volume, but splenomegaly is frequently associated with an increased plasma volume that may lead to a dilutional pseudo-anaemia. The clinical importance is that in splenomegaly the blood count may give an exaggerated impression of anaemia and measurement of red cell and plasma volumes can be valuable.

Role of the spleen in ensuring quality control of red cells

The spleen is the body's largest filter of blood and a major function of the spleen is the quality control of red cells. Red cells are normally flexible, whereas cells with abnormal membranes, or those with inclusions that render them relatively inflexible, remain in the cords where they are repaired or destroyed. *Sequestration* is a reversible process whereby cells are temporarily trapped by adhesion to the reticular meshwork of the cords on their passage through the spleen. Phagocytosis is the irreversible uptake by macrophages of particulate matter, non-viable cells and viable cells that have been damaged by prolonged sequestration or by antibody coating.

In the presence of metabolically active macrophages, the densely packed red cells are deprived of oxygen and glucose. This stress increases membrane rigidity and reduces the natural deformability of the biconcave cell. This effect is particularly marked if there is an underlying abnormality of the red cell metabolic system, if cells are coated with antibody or if they are fragmented or misshapen in other ways. In these situations they remain trapped in the cord space and undergo phagocytosis. Siderotic granules, Howell–Jolly (DNA) bodies, nuclear remnants and Heinz bodies are removed by culling or pitting during temporary sequestration. After removal of the inclusions the red cells return to the circulation. Reticulocytes may be retained in the splenic cords for a considerable proportion of the 2–3 days of their maturation period and during this time they lose their intracellular inclusions, alter the lipid composition of their surface and become smaller in size. It is not clear whether the spleen has any special role in the removal of normal aged red cells and it seems more likely that such cells are removed by the general reticuloendothelial system, which includes both the spleen and the bone marrow.

Immunological function

The spleen is the largest single accumulation of lymphoid tissue in the body and is estimated to contain 25% of the T-lymphocyte pool and 10–15% of the B-lymphocyte pool. T cells, mainly of the $CD4^+$ subtype, are found predominantly in the periarteriolar lymphatic sheath, whereas B cells are located in the follicles and marginal zones of the white pulp. There is also an abundance of macrophages and dendritic cells and the architecture is maintained through a complex anatomical organization that includes a tubular conduit system that transports a range of chemokines and other molecules. The spleen has a unique role in acting as an immunological filter for the bloodstream. Lymphocytes and dendritic cells enter the splenic tissue by initial entry into the marginal zone, and from there can pass into the periarteriolar

lymphatic sheath by crossing a lining of sinus cells. There is a constant flow of both T and B cells through the spleen, with T cells typically staying in the spleen for a few hours, whereas B cells stay in the follicle and marginal zone for prolonged periods. At present it is not certain how lymphocytes exit the white pulp.

One interesting feature of the immunological function of the spleen is that it is able to act as an important site for both innate and adaptive immune responses. The marginal zone acts as the site for both of these processes, whereas the white pulp is restricted to adaptive immunity. Blood is filtered directly into the marginal zone and here the predominant cell populations are macrophages and marginal-zone B cells. The macrophages express a range of pattern-recognition receptors that recognize bacterial molecules such as lipopolysaccharide. These cells thus act as an important filter for the clearance of blood-borne pathogens. The marginal-zone B cells are a unique B-cell subset that bridges the innate and adaptive immune responses. They are able to recognize bacterial pathogens without the help of T cells and this T-independent antibody response consists largely of low-affinity IgM and can play an important role in limiting bacterial replication. In addition, these B cells can act as antigen-presenting cells and are able to migrate to the periarteriolar lymphatic sheath, where they present antigen to T cells and facilitate the production of high-affinity IgG antibody. B-cell maturation and clonal expansion occurs in the follicles and plasmablasts, and plasma cells subsequently migrate to the red pulp where they are retained through their expression of CXCR4, which binds to CXCL12 in red pulp tissue.

The T-cell-independent antibody response produces a spectrum of low-affinity IgM antibody clones that provide a first line of immune defence against bacterial sepsis, especially from *Streptococcus pneumoniae*, *Haemophilus influenzae* and *Neisseria meningitidis*. The spleen also appears to act as a defence against viral infections and intraerythrocyte parasitic infections such as *Plasmodium* and *Babesia*. The important role of the spleen in this regard is revealed by immunization of splenectomized individuals with T-independent antigens where antibody titres are typically only 10% of those seen in control subjects. Given these unique features it is no surprise that splenectomy is associated with a degree of relative immunosuppression to encapsulated bacterial species.

Extramedullary haemopoiesis

As indicated above, the spleen is an important site of haemopoiesis *in utero* and retains the ability to re-activate this process after birth. This can occur as a compensatory erythroblastic hyperplasia in severe anaemia, such as chronic haemolysis, megaloblastic anaemia and thalassaemia major, or as a more generalized haemopoiesis often seen in primary myelofibrosis or other malignant disorders in the bone marrow. The mechanisms involved in this *extramedullary haemopoiesis* are poorly understood. It is not clear if pluripotent stem cells are present in the spleen or migrate from the bone marrow. It is even conceivable that changes in the stroma or primary haemopoietic cells that arise due to the underlying pathological process are involved in favouring initiation of the haemopoietic process.

Splenomegaly and hypersplenism

Spleen size

An enlarged spleen is a frequent and important clinical sign. It is thus essential to have a reliable picture of the presence and extent of splenomegaly. In the adult, an enlarged spleen is usually palpable when its length exceeds 14 cm. However, the measurement of spleen size by means of a physical examination of the abdomen is unreliable, as minor enlargement is often undetected by palpation and even a grossly enlarged spleen may be missed in an obese person. Conversely, a lax phrenic–colic ligament or loss of tone of the abdominal wall may give rise to a 'wandering spleen' which will be palpable, as will one that is pushed downwards by a flattened diaphragm in obstructive airways disease.

Reliable information is obtained by radiology through ultrasound, magnetic resonance imaging (MRI) and computed tomography (CT), all of which give an accurate representation of the anatomy of the spleen and its position in relation to adjacent organs (Figure 17.2). PET-CT may be used to detect tumours such as lymphoma in the spleen. Another method of scanning the spleen is to use gamma cameras for scintigraphy of ^{99m}technetium-labelled heat-damaged autologous red cells (Figure 17.3). Although laborious, this technique does provide information on the functional activity of the spleen and is also useful in identifying abnormally positioned and accessory splenic tissue. It is, however, not widely available.

Causes of splenomegaly

Enlargement of the spleen can occur in a wide range of conditions (Table 17.1). The relative incidence of each cause of splenomegaly is subject to geographical variation. In Western countries, leukaemia and lymphomas, myeloproliferative disorders, haemolytic anaemias, infectious mononucleosis and portal hypertension account for most cases. In tropical countries the major causes are parasitic infections such as malaria, leishmaniasis and schistosomiasis.

The underlying pathological basis for the enlargement can be of value in diagnosis and includes such features as reactive increase of white pulp in inflammation and infection, congestive expansion of the red pulp compartment, increased blood pool, cellular infiltration or extramedullary haemopoiesis.

Specific causes of splenomegaly

Malaria

Several pathogenic mechanisms are involved in splenomegaly of malaria, including reactive lymphoid change, dilation of red

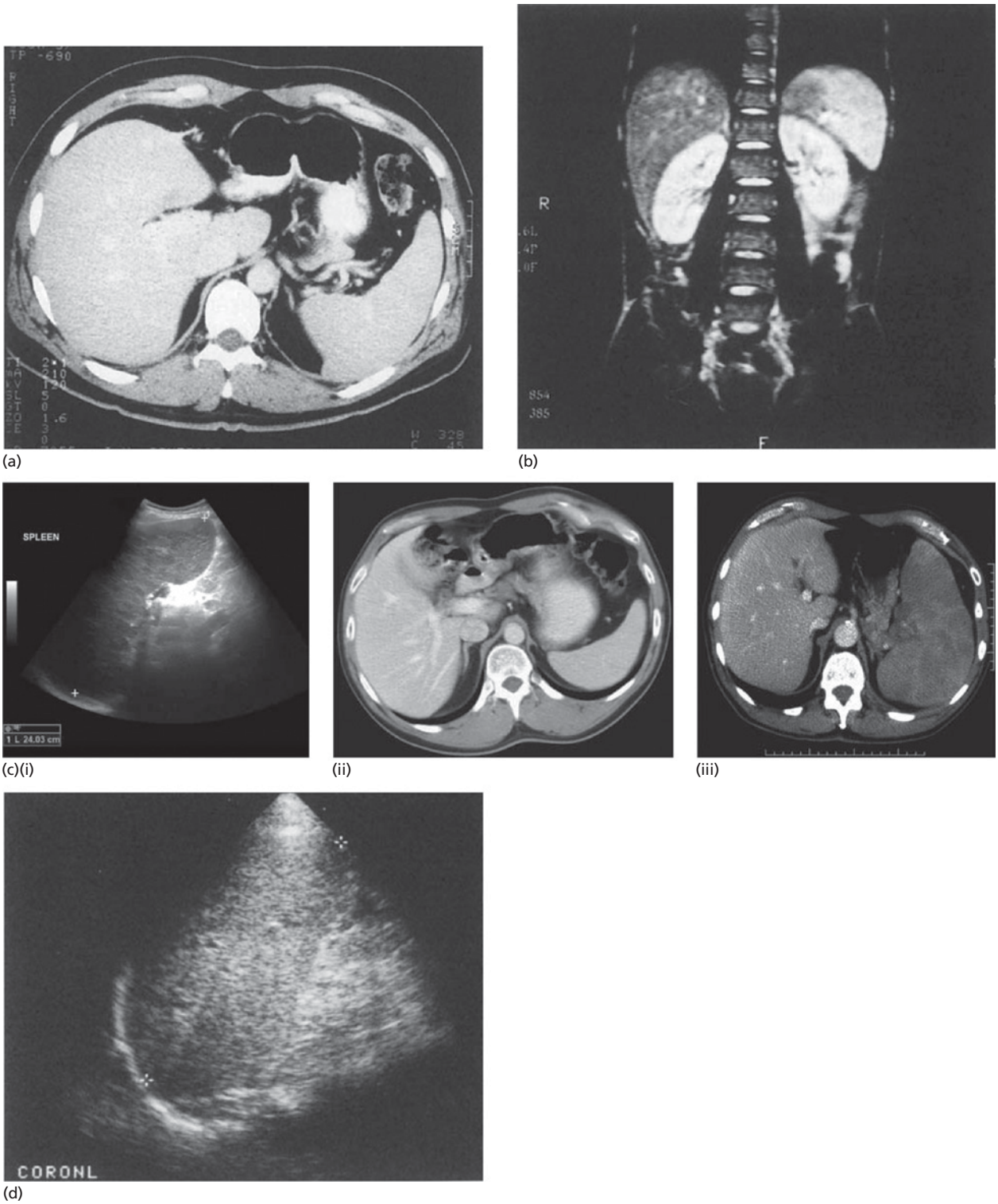
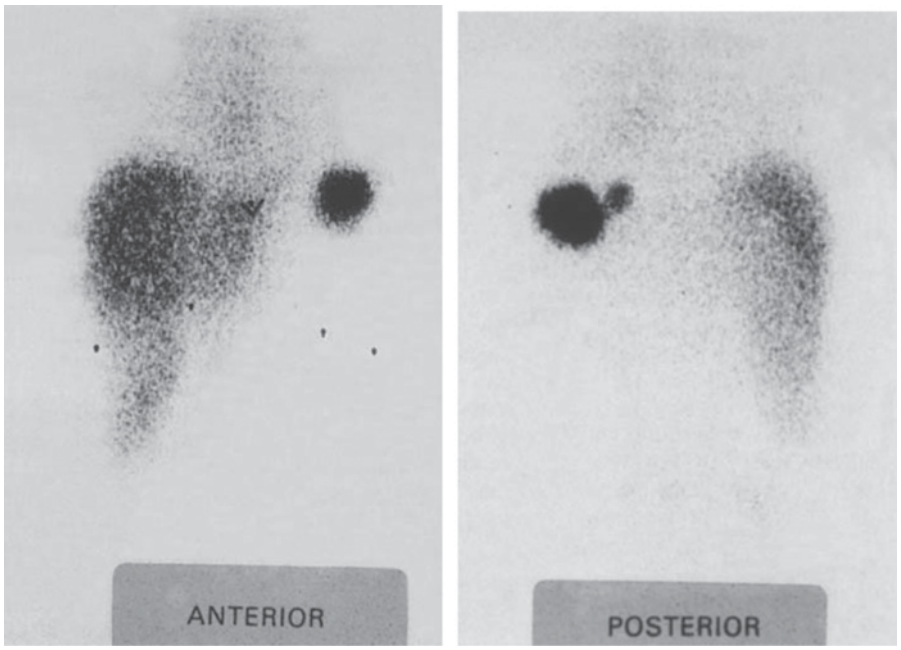


Figure 17.2 Imaging of the spleen by various methods. (a) CT: transverse section showing liver (left) and spleen (right). (b) MRI: coronal (longitudinal) section showing liver, spleen and kidneys. (c) (i) Ultrasound of spleen showing splenomegaly (15.3 cm); (ii) normal spleen (10 cm) on computed tomography (CT) scan; (iii) CT scan: the spleen is enlarged and shows multiple low density areas. A diagnosis of diffuse large cell B lymphoma was made histologically after splenectomy. (d) Ultrasound scan of enlarged spleen. (Source: Dr T. Ogunremi [parts (c)(i) and (ii)]. Reproduced with permission.)



(a) (b)

Figure 17.3 Demonstration of residual splenunculus by a scan of heat-damaged isotope-labelled red cells: (a) anterior view; (b) posterior view. Uptake is predominantly in the liver.

pulp sinuses and expanded phagocytic activity. The tropical splenomegaly syndrome, associated with recurrent malaria, is particularly common in New Guinea and Central Africa.

Haemoglobin disorders

Splenomegaly is associated with HbC disease in West Africa, with HbE disease in the Far East and with thalassaemia syndromes, which have a wide distribution throughout the Mediterranean area and tropics. HbSS sickle cell disease is usually associated with splenic atrophy, but the spleen usually remains enlarged in adults with HbS/C and HbS/β-thalassaemia syndromes.

Malignant haematological disorders

In polycythaemia vera the increase in spleen size is mainly due to vascularity, with expansion of the red pulp and an increased red cell pool. In primary myelofibrosis, the red cell pool is remarkably increased and the spleen size is further augmented by myeloid metaplasia and expansion of the reticular elements. In contrast, in chronic myeloid leukaemia and lymphoproliferative disorders the increase in size is attributed mainly to cellular infiltration.

Primary splenic tumours

Primary splenic tumours are rare. Metastatic carcinoma, such as from the breast or lung, is also a rare event.

Portal hypertension

This may be both a cause and a consequence of splenomegaly. In a spleen that is massively enlarged from any cause, the huge increase in blood flow can lead to an increased portal pressure that results in backpressure on the spleen and a spiral of compression leading to widening and fibrosis of the red pulp cords. On the other hand, portal vein obstruction due to liver disease can also enlarge and damage the spleen.

Hypersplenism

Hypersplenism is a clinical syndrome and does not imply a specific causal mechanism. It has the characteristic features of:

- Enlargement of the spleen
- Reduction in one or more of the cell lines in the peripheral blood and
- Normal or hyperplastic cellularity of the bone marrow.

Many disorders are associated with hypersplenism (Table 17.1) and where the association is uncertain the diagnosis of hypersplenism may ultimately be confirmed by the response to splenectomy.

Splenectomy

The role of splenectomy in individual diseases is discussed in the relevant chapters, but when splenectomy is contemplated for any reason, the preoperative evaluation of the patient requires close co-operation between the surgeon and the haematologist. Vaccinations should be performed and it is valuable to check liver function, obtain appropriate imaging and potentially evaluate the hepatic and portal blood flow by Doppler ultrasound examination.

Table 17.1 Causes of splenomegaly.

Haematological
Acute leukaemia
Chronic myeloid leukaemia*
Chronic lymphocytic leukaemia*
Malignant lymphomas*
Myelofibrosis*
Polycythaemia vera*
Essential thrombocythaemia (some cases)
Hairy cell leukaemia*
Gaucher disease*, Niemann–Pick disease, Langerhans cell histiocytosis X
Primary splenic hyperplasia (i.e. unknown cause)
Thalassaemia
Sickle-cell disease, HbSC disease and other haemoglobinopathies
Haemolytic anaemias
Megaloblastic anaemia (rare)
Systemic
Acute infections: septicaemia, typhoid, infectious mononucleosis, cytomegalovirus
Subacute and chronic infections: tuberculosis, syphilis, brucellosis, subacute bacterial endocarditis, HIV
Tropical parasitic infections (tropical splenomegaly*): malaria*, leishmaniasis*, schistosomiasis*, trypanosomiasis
Collagen diseases: systemic lupus erythematosus, rheumatoid arthritis (Felty)
Sarcoidosis
Amyloidosis
Cysts
Haemangiomas
Carcinoma (rare)
Congestive splenomegaly
Portal hypertension*
Splenic/portal/hepatic vein obstruction
Congestive cardiac failure

*Common causes of splenomegaly.

Splenic injury is common after blunt abdominal trauma and must be assessed in order to consider the need for emergency splenectomy. Focused abdominal sonography for trauma (FAST) is a form of ultrasound that is rapid to perform and valuable in patients who are haemodynamically unstable. Abdominal CT remains the investigation of choice when the pulse and blood pressure remain stable. Low-grade injury may be managed conservatively, and angioembolization can be very useful for controlling haemorrhage, whilst preserving splenic function (Figure 17.4). In addition to trauma, splenic rupture may also occur rarely in a range of conditions such as infectious mononucleosis or pregnancy.

Surgical excision of the spleen has been a standard treatment for the diagnosis and management of disorders associated with an enlarged or hyperactive spleen and also when an otherwise normal spleen mediates the clinical problems associated with an extrasplenic defect such as hereditary spherocytosis or autoimmune acquired haemolytic anaemia. Splenectomy can bring about significant disease control in some malignant conditions, such as splenic marginal zone lymphoma, where the bulk of the tumour population is located in the spleen. Finally, the spleen is often removed incidentally as part of another surgical procedure. At a surgical level, laparoscopic splenectomy has become the treatment of choice in the absence of portal hypertension or significant medical co-morbidity.

Complications of splenectomy

Bleeding can be a complication after surgery and usually comes from the peritoneal and diaphragmatic surfaces rather than from identifiable blood vessels. Frequently, no specific bleeding source is found at re-operation.

Thrombocytosis may be seen in the immediate postoperative period and the platelet count often rises up to $600–1000 \times 10^9$/L, with a peak at around 7–12 days. In a number of patients, moderate thrombocytosis persists indefinitely after splenectomy. Although a reactive thrombocytosis is not usually associated with thromboembolic problems, postoperative prophylaxis with heparin is often given. It is advisable to give antiplatelet therapy (e.g. aspirin 75 mg daily) as long as thrombocytosis is present.

Post-splenectomy sepsis

Splenectomy increases the risk of developing an overwhelming bacterial infection, (sometimes termed *overwhelming post-splenectomy sepsis*; OPSS), which in historical cohorts can be fatal in up to 50% of patients. The onset is often acute and fulminant with a non-specific presentation that may include fever, muscle ache, chills, sore throat, diarrhoea or vomiting. Septicaemia or meningitis can be seen, and features such as disseminated intravascular coagulation and hypoglycaemia may develop rapidly. *Streptococcus pneumoniae* is the most common cause and bacteraemia can be so severe that cocci are seen directly in stained blood films. *Haemophilus influenzae*, *Neisseria meningitidis* and *Escherichia coli* may also be implicated. The underlying cause is a defective reticuloendothelial clearance of encapsulated organisms, combined with delayed and impaired development of the IgM antibody response.

The risk of infection is related to the underlying disease for which splenectomy was performed. A Danish population-based study confirmed that splenectomized patients had an increased rate of hospital contact for infection compared to the general population (Figure 17.5a). However, when the data were

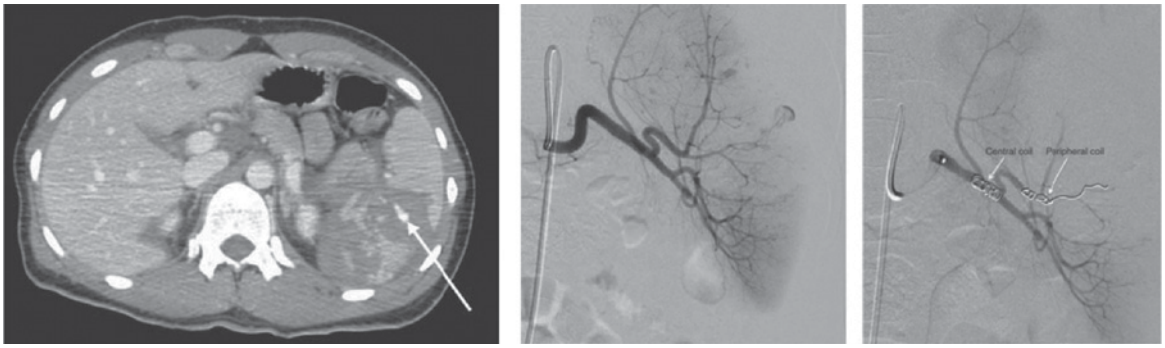


Figure 17.4 (a) Computed tomogram showing splenic injury (Organ Injury Scale grade 4) with contrast extravasation (white arrow) (b) Angiography (i) with signs of ongoing bleeding before embolization and (ii) after distal and proximal embolization. (Source: Skattum *et al*., 2012. Reproduced with permisson of John Wiley & Sons.)

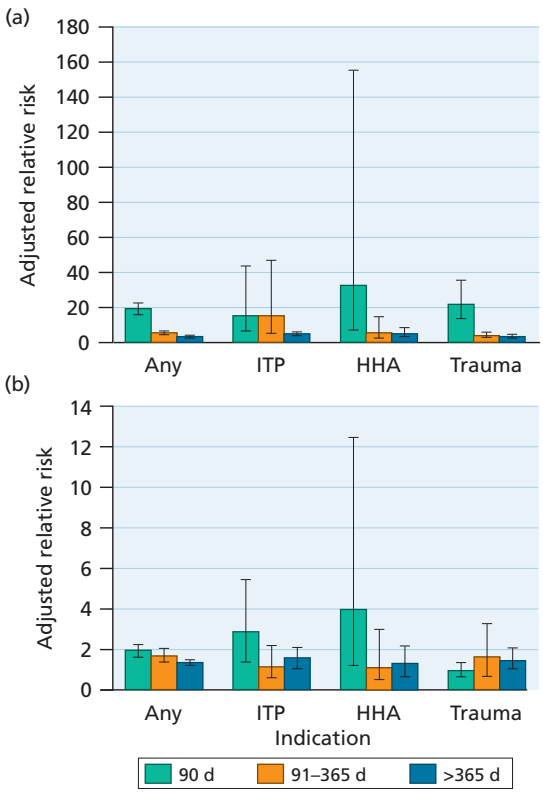


Figure 17.5 Relative risk of hospital contact for any infection in splenectomized patients according to the time since splenectomy. Data based on a Danish population-based study (Thomsen *et al*., 2009). Results are shown for: any indication for splenectomy, immune thrombocytopenia, hereditary haemolytic anaemias and trauma, in comparison with age- and sex-matched subjects from the general population (a) and up to five matched-indication non-splenectomized patients (b). (Source: Rodeghiero and Ruggeri, 2012. Reproduced with permisson of John Wiley & Sons.)

compared to patients with the same disease, but who had not had splenectomy, the relative risk of infection fell significantly (Figure 17.5b). In particular, patients with ITP or hereditary haemolytic anaemias showed a relative risk of 2.6 and 3.6 during the first 90 days after splenectomy, but not after this timepoint (Figure 17.5).

Infection may occur at any age, but children are at increased risk, particularly within the first few years of life. Splenectomy should therefore be postponed until after the age of 5 years of age wherever possible. The prevention of fatal postsplenectomy sepsis has four components:

- Patient education
- Vaccination
- Antibiotic prophylaxis
- Early empirical use of antibiotics for febrile episodes.

Patient education

Education of patients or parents is perhaps the most important aspect of management, to ensure that they are aware of the possibility of infection and know how to react appropriately. Patients should be advised to carry a card or bracelet to alert health professionals to the potential risk of overwhelming infection. When travelling to tropical areas, asplenic patients are at increased risk of severe *Plasmodium* infection and must adhere scrupulously to antimalarial prophylaxis. Patients should also be advised that severe infection can occur after animal bites, typically due to *Capnocytophaga Canimorsus* which is a commensal bacterium of cats and dogs.

Vaccination

Normal inoculations, including live vaccines, can be given safely after splenectomy. When splenectomy is being planned, the patient should be considered for immunization against pneumococcus, *H. influenzae* type B (Hib) and meningococcal infection. To obtain the maximum immune response, patients should,

Table 17.2 Suggested schedule for vaccine immunization in individuals with asplenia or splenic hypo function.

Age at which asplenia or splenic dysfunction acquired	Vaccination schedule*		
	Month 0	**Month 1**	**Later**
Under 2 years	Complete according to national routine childhood schedule, including booster doses of Hib/MenC and PCV13	A dose of MenACWY conjugate vaccine should be given at least 1 month after the Hib/MenC and PCV13 booster doses	After the second birthday, one additional dose of Hib/MenC and a dose of PPV should be given
Over 2 years and under 5 years (previously completed routine childhood vaccinations with PCV7)	HibMenC booster PCV13	MenACWY conjugate vaccine	PPV (at least 2 months after PCV13)
Over 2 years and under 5 years (previously completed routine childhood vaccinations with PCV13)	HibMenC booster PPV	MenACWY conjugate vaccine	
Over 2 years and under 5 years (unvaccinated or previously partially vaccinated with PCV7)	HibMenC vaccine First dose of PCV13	MenACWY conjugate vaccine	Second dose of PCV13 and then PPV (at least 2 months after PCV13)
Over 5 years (regardless of vaccination history)	HibMenC vaccine PVV	MenACWY conjugate vaccine	

*Where possible, vaccination course should ideally be started at least 2 weeks before surgery or commencement of immunosuppressive treatment
PCV, pneumococcal conjugate vaccine; PPV, pneumococcal polysaccharide vaccine
Source: Davies *et al*., 2011 [*Br J Haem* **155**, 2011, 308–17]. Reproduced with permission of John Wiley & Sons.

if possible, be immunized at least 2 weeks before splenectomy. One approach to vaccine policy for hyposplenic patients is shown in Table 17.2.

Streptococcus pneumoniae is the major concern and there are two major forms of vaccine available. Polyvalent polysaccharide pneumococcal vaccines (PPV) include a wide range of different serotypes, typically at least 23, but their efficacy in hyposplenic individuals is poor. Conjugated vaccines consist of a bacterial polysaccharide linked (conjugated) to a carrier protein and thereby enlist a T cell immune response. The 13-valent pneumococcal conjugate vaccine (PCV13) is more effective than PPV in hyposplenic patients, although the breadth of serotype coverage is not as great. The optimal schedule of vaccination with both subtypes is not yet known and in some countries PCV vaccines are not recommended, whereas others suggest a schedule such as PCV13 followed 2 months later by PPV23. The recommendations for vaccination in children under 5 years will depend on the prior vaccination history (Table 17.2). Where possible, the levels of protective antibodies should be measured to guide vaccination requirement. The World Health Organization (WHO) recommends a serotype-specific IgG level of ≥0.35 μg/mL as a potential protective threshold following conjugate immunization in young children, but it is not clear if this extends to adults where a higher threshold (e.g. ≥1.0 μg/mL) may be more appropriate.

There are six different serotypes of *Haemophilus influenzae* (a–f) and serotype b is the most virulent, although vaccination in infants has been highly effective in recent years. The Hib vaccine should be given to asplenic adults who have not received it.

The need for meningococcal vaccination in asplenic patients is uncertain, but it is recommended. Vaccines are effective against types A and C. The quadrivalent meningococcal ACWY conjugate vaccine is coming to replace single MenC vaccination or quadrivalent polysaccharide vaccine, and a two-dose vaccine course is a reasonable option, with boosters every 5 years.

Annual influenza vaccine is also recommended for asplenic or hyposplenic patients, primarily as influenza may predispose patients to secondary bacterial infection.

Antibiotics

Long term antibiotic prophylaxis should also be taken into consideration. There is a relatively poor evidence base for this approach in the era of pneumococcal vaccination and different regimens are sometimes advised. The British Committee for Standards in Haematology advises prophylaxis for patients at high risk of infection. These are defined as:

- Age less than 16 years *or* greater than 50 years
- Inadequate serological response to pneumococcal vaccination

- History of previous invasive pneumococcal disease
- Splenectomy for underlying haematological malignancy particularly in the context of ongoing immunosuppression.

Those at lower risk should be advised about the risks and benefits of lifelong antibodies and may wish to discontinue them. Oral penicillin 250 mg b.d. is usually recommended, and patients who are allergic to penicillin should be offered erythromycin 250 mg b.d.

Pre-emptive antibiotics should be given to patients so that they may start them if they develop a fever or infective symptoms such as pyrexia, malaise or shivering and cannot get to hospital within 2 hours. Suitable options are 500 mg amoxicillin or 750 mg of levofloxacin. In all such cases, the patient should also seek immediate medical help and in hospital, ceftriaxone (IV or IM), with or without vancomycin, is a reasonable empirical choice after cultures have been taken.

Recurrence of symptoms

Accessory splenic tissue may be overlooked at operation; after splenectomy, it may enlarge and cause a recurrence of the symptoms for which the original operation was carried out. The haematological features of hyposplenism (see further on) such as Howell–Jolly bodies and increased pitting may be absent; CT or radionuclide scanning (Figure 17.3) will demonstrate the presence of a 'splenunculus' and identify its location for subsequent surgical removal should this be required.

Hyposplenism

Hyposplenism occurs in a range of medical conditions, as well as following surgical splenectomy or therapeutic splenic embolization. In some disorders such as sickle cell disease, severe gluten-induced enteropathy and untreated HIV infection, it can be seen quite frequently, whereas it occurs only occasionally in a range of other conditions (Table 17.3). In sickle cell anaemia there is functional asplenia by 1 year of age and autoinfarction leads to a state of anatomical asplenia after 6–8 years of age. It has been estimated that there are around 1 million asplenic people in the US, including some 100,000 with sickle cell disease. Congenital absence of the spleen is rare and may be associated with organ transposition and severe malformations of the heart and lungs. At the other extreme of life, atrophy of the spleen is seen after around 65 years of age, with a rapid decrease in weight in older age.

Patients with functional hyposplenism have impaired immunity to blood-borne bacterial and protozoal infections, and persistent thrombocytosis. Management is similar to that required after splenectomy. It includes prophylactic antibiotics and vaccines (see above) and advice to the patient to seek medical attention immediately in the event of illness or fever. Antiplatelet therapy is advisable when the platelet count is high. Characteristic changes in the blood count occur following splenectomy or when the spleen atrophies to less than 20% of normal size. There is no simple assay to measure the severity of hyposplenism, and red cell scintigraphy is probably the current gold standard.

Table 17.3 Causes of hyposplenism.

Congenital aplasia syndrome
Ageing
Haematological disorders
Sickle cell disease
Thrombocythaemia
Myelofibrosis
Malaria
Lymphomas
Circulatory
Splenic arterial/venous thrombosis
Autoimmune disease
Systemic lupus erythematosus
Rheumatoid arthritis
Hyperthyroidism
Sarcoidosis
Chronic graft-versus-host disease
Combined immunodeficiency
Gastrointestinal (? immune basis)
Gluten-induced enteropathy
Dermatitis herpetiformis
Crohn's disease
Ulcerative colitis
Tropical sprue
Infiltrations
Lymphomas
Sézary syndrome
Myeloma
Amyloidosis
Secondary carcinomas, especially breast
Cysts, e.g. hydatid
Nephrotic syndrome
Drugs
Methyldopa
Intravenous immunoglobulin
Corticosteroids
Irradiation
Splenectomy and splenic embolization

Red cell changes

The changes in red cell morphology include the presence of Howell–Jolly bodies, siderotic granules and target cells. In a proportion of subjects, irregularly contracted or crenated acanthocytic forms are also a feature (Figure 17.6). There is usually an increase in the number of reticulocytes in the circulation and occasionally isolated erythroblasts are seen. However, there is no alteration in red cell survival.

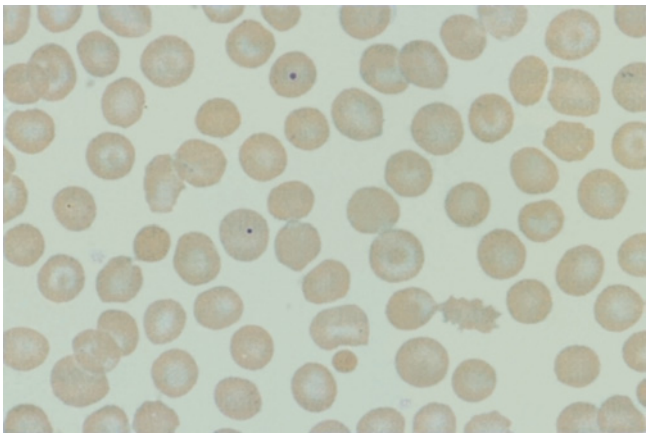


Figure 17.6 Blood film showing features of hyposplenism: Howell–Jolly bodies, target cells and contracted cells.

Because the number of cells with siderotic granules that enter the circulation is related to the sideroblastic percentage in the bone marrow, the siderocyte count in the peripheral blood is increased in haemolytic anaemias, thalassaemia and sideroblastic anaemia. The number of Howell–Jolly bodies is also variable and is most marked in conditions characterized by dyserythropoiesis. Other red cell inclusions may be prominent in the hyposplenic state: Heinz bodies are found following oxidative injury by drugs and in patients who have glucose-6-phosphate dehydrogenase deficiency or an unstable haemoglobin, precipitated β-chains are found in α-thalassaemia and crystalline deposits of haemoglobin C in HbC disease.

Leucocyte changes

After splenectomy there is a rise in the total leucocyte count. A neutrophil leucocytosis in the immediate postoperative period is later replaced by a significant and permanent increase in both lymphocytes and monocytes. Usually the total white cell count stabilizes at between 10 and 15 × 10^9/L, but occasionally it may rise to twice this level. Minor increases in blood eosinophils and basophils have been noted after splenectomy, but this is not a regular feature.

In response to infection, splenectomized subjects produce a much greater leucocytosis than persons with intact spleens. Often there is a marked left shift in the differential leucocyte count, with myelocytes and occasionally more primitive cells.

Platelet changes

As indicated above, the thrombocytosis after splenectomy is usually transitory and falls to normal or near-normal values over the following 1–2 months. However, even if the platelet count has returned to normal values, occasional large and bizarre platelets can be seen in the blood films of many splenectomized subjects. Their presence suggests that the spleen normally removes these particular platelets.

Immunological effects

The spleen plays an important role in immunoglobin synthesis; a fall in the IgM fraction of the serum immunoglobulins is commonly found after splenectomy. IgG levels do not change, while IgA and IgE increase.

Selected bibliography

de Back DZ, Kostova EB, van Kraaij M, van den Berg TK, van Bruggen R (2014) Of macrophages and red blood cells; a complex love story. *Frontiers in Physiology* **30**(5): 9.

Davies, JM, Lewis, MPN, Wimperis, J *et al.* (2011) Review of guidelines for the prevention and treatment of infection in patients with an absent or dysfunctional spleen: prepared on behalf of the British Committee for Standards in Haematology by a working party of the Haemato-Oncology task force. *British Journal of Haemetology* **155**(3): 308–17.

De Porto APNA, Lammers AJJ, Bennink RJ *et al.* (2010) Assessment of splenic function. *European Journal of Clinical Microbiology and Infectious Disease* **29**: 1465–73.

Gamme G, Birch DW, Karmali S (2013) Minimally invasive splenectomy: an update and review. *Canadian Journal of Surgery* **56**(4): 280–5.

Mebius RE, Kraal G (2005) Structure and function of the spleen. *Nature Reviews Immunology* **5**: 606–16.

Orthopoulos GV, Theodoridou MC, Ladis VA, Tsousis DK, Spoulou VI (2009) The effect of 23-valent pneumococcal polysaccharide vaccine on immunological priming induced by 7-valent conjugate vaccine in asplenic subjects with beta-thalassemia. *Vaccine* **27**: 350–4.

Rodeghiero F, Ruggeri M (2012) Short- and long-term risks of splenectomy for benign haematological disorders: should we revisit the indications? *British Journal of Haematology* **158**(1): 16–29.

Rubin LG, Schaffner W (2014) Care of the asplenic patient. *New England Journal of Medicine* **371**(4): 349–56.

Skattum J, Naess PA, Gaarder C (2012) Non-operative management and immune function after splenic injury. *British Journal of Surgery* **99**(1): 59–65.

Saksobhavivat N, Shanmuganathan K, Chen HH *et al.* (2015) Blunt splenic injury: use of a multidetector CT-based splenic injury grading system and clinical parameters for triage of patients at admission. *Radiology* **274**(3): 702–11.

Vancauwenberghe T, Snoeckx A, Vanbeckevoort D *et al.* (2015) Imaging of the spleen: what the clinician needs to know. *Singapore Medical Journal* **56**(3): 133–44.

William BM, Corazza GR (2007) Hyposplenism: a comprehensive review. Part I: basic concepts and causes. *Hematology* **12**: 1–13.

CHAPTER 18

The molecular basis of haematological malignancies

18

Niccolo Bolli[1] and George Vassiliou[2]

[1]Fondazione IRCCS Istituto Nazionale dei Tumori, University of Milan, Milan, Italy
[2]Haematological Cancer Genetics, Wellcome Trust Sanger Institute, Cambridge, UK

Introduction

Cellular growth and differentiation are carefully controlled processes that are regulated by several interconnected pathways in order to facilitate diverse biological phenomena ranging from development, responses to normal and abnormal stimuli and the replacement of dying cells. In this context, the growth of individual cells is restrained in the interest of their cognate tissue and by extension of the organism as a whole. Oncogenesis represents the progressive corruption of this order and the stepwise escape of an individual cell and its progeny from checks on their growth. This corruption occurs in the form of the serial acquisition of genetic mutations, which disrupt the genome of the fateful cell and morph it into a *cancer genome* (Table 18.1). In turn, the ontogeny of the cell of origin and the nature of the mutations are the primary determinants of the cancer phenotype, including its histological type, biological behaviour, clinical characteristics and responsiveness to therapies.

Over the last few years, we have witnessed unprecedented progress in our understanding of cancer genomes, driven primarily by extraordinary advances in DNA sequencing technologies. The impact of cancer genomics on the study of haematological malignancies has been particularly dramatic, in large part because of our deep prior acquaintance with normal blood cell development and hierarchy, as well as our detailed characterization and classification of haematological cancers. Importantly, these developments are not only enhancing our understanding of blood cancer biology, but are already having an impact on the way we diagnose and treat patients.

This chapter will describe important advances in genomics and review key themes and paradigms relating to the pathogenesis and evolution of haematological cancers.

The cancer genome

The identification of the Philadelphia chromosome in patients with chronic myeloid leukaemia (CML) by Nowell and Hungerford in 1960 was the first report of a genetic lesion associated with a human malignancy and forms a landmark in cancer studies. It has since become clear that cancer is a genetic disease, which develops through a process of Darwinian-like clonal evolution involving the stepwise acquisition of somatic mutations in individual cells subject to natural selection on the basis of their ability to outgrow their peers through increased proliferation and survival. Within a healthy adult human there are many small clones, most of which have limited abnormal growth potential and are invisible or manifest as common benign growths such as skin moles. Occasionally, however, a single cell acquires a set of mutations that allows it to proliferate to the extent that it generates a malignant phenotype. The nature of such mutations and the way in which they collaborate to produce the malignant phenotype can only be fully understood through the integrated study of the large set of variants within a tumour.

Next generation sequencing (NGS) technologies allow such analysis through parallel sequencing of thousands to millions of DNA molecules in a relatively short period of time. Compared to classical Sanger methodology this represents an increase in

Postgraduate Haematology, Seventh Edition. Edited by A Victor Hoffbrand, Douglas R Higgs, David M Keeling and Atul B Mehta.

Table 18.1 Glossary

Acquired mutation: see 'somatic mutation'

Allele burden: The fraction of alleles with a specific sequence in relation to the total number of alleles for the same region of the genome. For example a heterozygous mutation in a pure population of leukaemia cells has an allele burden of 0.5. If 70% of cells are leukaemic and 30% of cells are normal, the mutant allele burden of the heterozygous mutation would be $0.7 \times 0.5 = 0.35$.

Amplification: A genetic modification producing an increased number of copies of a genomic region

Branching evolution (of cancer): A form of clonal evolution of cancer which leads to the generation of more than one clone of cells characterized by distinct somatic mutations, but which share at least one mutation traceable back to a single ancestral cell.

Cancer genome: The genome of a cancer cell, which differs from the germline genome as a result of somatic mutations.

Chromosomal translocation: see 'genomic rearrangements'.

Chromothripsis: a single catastrophic event by which hundreds to thousands of chromosomal rearrangements occur in confined genomic regions of one or a few chromosomes (from Greek θρύΨις = shattering into small pieces).

Clonal evolution (of cancer): The stepwise acquisition of mutations in a founder cell and its progeny leading towards the development of a cancer.

Clonal mutation: A mutation present in a population of related cells derived from a single cell.

Constraint hypothesis: A hypothesis about clonal evolution proposing that the observed order of acquisition of somatic mutations during cancer evolution reflects a requirement for a specific mutation to occur before another for a growth advantage to be gained by the host cell(s).

Co-occurrence (of cancer mutations): The occurrence of two or more mutations in the same type cancer more often than would be expected by chance.

Convergent evolution (of cancer): A pattern of cancer evolution during which independent clones expand after acquiring the same or very similar mutation. This is likely to reflect the fact that such a mutation is particularly advantageous to the specific cancer cell, giving a marked growth advantage when acquired by chance.

Deletion: A genetic modification leading to the loss of a genomic region.

Dominant negative mutation: A heterozygous mutation that leads to marked or complete loss of function of the coded protein and of the normal protein coded by the other (wild-type) copy of the gene.

Driver gene or driver mutation: A mutated gene that confers a selective growth advantage to a cancer cell

Epigenetics: The study of changes to DNA and chromatin, other than those that alter the DNA nucleotide sequence, that alter the transcriptional potential of a cell and are usually heritable

Exome: The collection of all exons in a genome.

Exome sequencing: Sequencing of all exons in a genome. This has referred to exons of protein-coding genes, but increasingly non-protein-coding genes are included (e.g. long non-coding RNAs).

Gain-of-function mutation: A mutation that gives the coded protein a novel or markedly enhanced function

Genome-wide association studies (GWAS): studies of many common and uncommon genetic variants in different individuals to determine if any variant is associated with a disease or trait. The primary outcome of these studies is the identification of variants such as SNPs which are *associated* with, but do not necessarily cause the disease in question.

Genomic rearrangement: A mutation that juxtaposes nucleotides that are normally distant from each other, such as a chromosomal translocation, inversion or deletion.

Germline genome: An individual's genome as formed at the time of conception (fertilized oocyte). This genome is shared by all cells in the body.

Germline mutation: Mutations present in the germline genome. Sporadic mutations acquired in the germ cells of parents are also included in this category.

Germline variants: Variations in sequences or copy number of DNA segments observed between different individuals that are responsible of the phenotypic variation between people. Two unrelated individuals differ by approximately 3 million such variants.

Haplotype: A haplotype is a contiguous region of the genome containing a set of tightly linked genes that are likely to be inherited together.

Indel: A mutation that results in insertion or deletion of one or a few nucleotides to DNA.

Kataegis: Localized hypermutation of a region of the genome, thought to be mediated by APOBEC enzymes (from Greek καταιγίς = storm).

Linear evolution (of cancer): A form of clonal evolution of cancer that generates a single final clone of cancer cells which harbours all mutations that ever arose during its evolution.

Loss-of-function mutation: A mutation that leads to marked or complete loss of function of the coded protein.

Loss-of-heterozygosity (LOH): A genetic modification leading to the loss of the maternally- or paternally-derived copy of a genomic region. This can happen as a result of deletion or uniparental disomy (uPD).

Methylation: Covalent addition of a methyl group to a DNA, RNA, protein or other molecule.

Missense mutation: A nucleotide substitution (e.g. G to T) that results in an amino acid change (e.g. valine to phenylalanine).

Mutational signature: A recurrent pattern of DNA mutations attributable to a particular type of mutagen or mutational process, characterized by certain nucleotide mutations in a specific 5′ and 3′ nucleotidic context.

Mutual exclusivity: The occurrence of two or more mutations in the same cancer type less often than would be expected by chance.

Next-generation sequencing (NGS): DNA sequencing using one of the methodologies developed since 2005 and which allow massively parallel sequencing of thousands or millions of fragments of DNA simultaneously.

Nonsense mutation: A nucleotide substitution that results in the generation of a stop codon (i.e. TAA, TGA or TAG).

Non-synonymous mutation: A mutation that alters the encoded amino acid sequence of a protein. These include missense, nonsense, splice site, gain of translation start, loss of translation stop and indel mutations.

Proto-oncogene: A gene that when activated by mutations, becomes an **oncogene** and imparts a growth advantage of its host cell.

Opportunity hypothesis: A hypothesis proposing that the observed order of acquisition of somatic mutations during cancer evolution reflects the statistical likelihood that mutations are acquired in this order. This likelihood is determined by the earlier mutation influencing the opportunity for acquiring the next. Implicit in this hypothesis is that the reverse order can also be observed, albeit less often.

Table 18.1 (*Continued*)

Passenger mutation: A mutation that does not give a selective growth advantage to its host cell.
Single nucleotide polymorphism (SNP): A DNA sequence variation occurring commonly within a population (e.g. 1%) in which a single nucleotide in the genome differs between individuals.
Somatic mutation: A mutation that occurs in any cell of the body after conception. Sometimes called *acquired mutation*.
Substitution (or nucleotide substitution): A DNA mutation leading to the replacement of a native nucleotide with another.
Splice sites: DNA sequences flanking exons which are important for mRNA splicing.
Subclonal mutation: A mutation that exists in only a subset of the neoplastic cells within a tumour.
Transition (mutation): Change of a nucleotide to another of the same group such as C>T (both pyrimidines) or G>A (both purines).
Transversion (mutation): Change of a nucleotide to another of the opposite group such as A>C (purine to pyrimidine).
Uniparental disomy (uPD): A genetic modification leading to the loss of the maternally or paternally derived copy of a genomic region as a result of replacement of this sequence with the equivalent sequence derived from the other parent.
Untranslated region (UTR): Exonic region located before the start (5' UTR) or after the stop (3'UTR) codon of a gene and which do not encode amino acids.
Whole genome sequencing: Sequencing of the entire sequence of an individual genome using germline DNA, tumour-derived DNA or DNA from another cellular source such as cell lines, single cells etc.
Whole exome sequencing: Sequencing of all the exons of all the genes in an individual genome (there are approximately 30,000 coding genes in a mammalian genome). As with whole genome sequencing this could be done using germline DNA, tumour-derived DNA or DNA from another cellular source.

speed and capacity in excess of 1 billion-fold and a dramatic reduction in costs. NGS output is in a digital form, where each sequencing read, corresponding to an individual genomic DNA fragment, is captured separately. Reads are then mapped back to the reference genome using computational aligning algorithms, allowing the sequence of the *whole genome* to be analysed in one sequencing run. Alternatively, genomic regions of interest can be enriched for prior to sequencing, allowing the investigator to limit analysis to the coding portion of the genome (*whole exome sequencing*), or to specific parts of the genome such as the exons of a selected set of genes. Importantly, NGS is quantitative and can be used to discriminate between homozygous, heterozygous and subclonal events based on the number of DNA molecules supporting each variant (*allele burden*).

The accumulation of somatic mutations is an inevitable consequence of the passage of time and this is true for any cell, including the germ cells through which acquired mutations are passed to the offspring, generating diversity in the human and any other living species. However, mammalian genomes are afforded an extraordinary level of protection from mutations such that each nucleotide has only a 1 in 10^{10} chance of mutation per generation. Nevertheless, mutations do accumulate with time at a rate that is influenced by the abundance and nature of mutagens acting on cellular DNA and its proof-reading machinery.

A cancer cell, just like all the other cells in the body, is a direct descendant of the fertilized egg through a lineage of mitotic cell divisions. However, the genome of a cancer cell and indeed that of a normal cell harbour a set of somatic mutations acquired during the cell's own life and the life of its ancestors back to the fertilized egg (Figure 18.1). These are collectively termed **somatic mutations** to distinguish them from **germline mutations.** In the following paragraphs we will describe examples of germline as well as somatic variants associated with cancer and discuss the processes leading to acquisition of somatic variants.

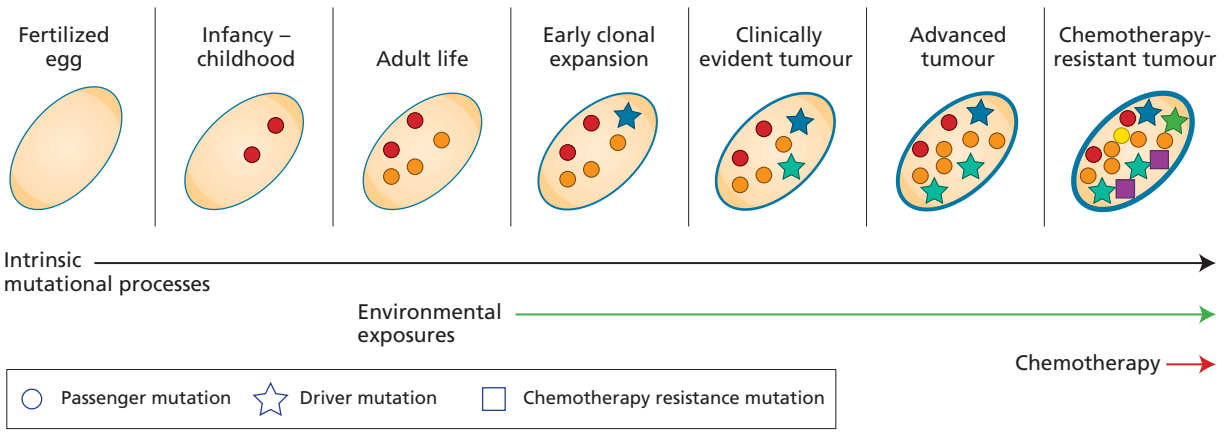


Figure 18.1 The sequence of somatic mutations acquired during cell divisions from the fertilized egg to a single chemoresistant cancer cell. The mutational processes causing these mutations are also highlighted. Modified from Stratton *et al*. 2009.

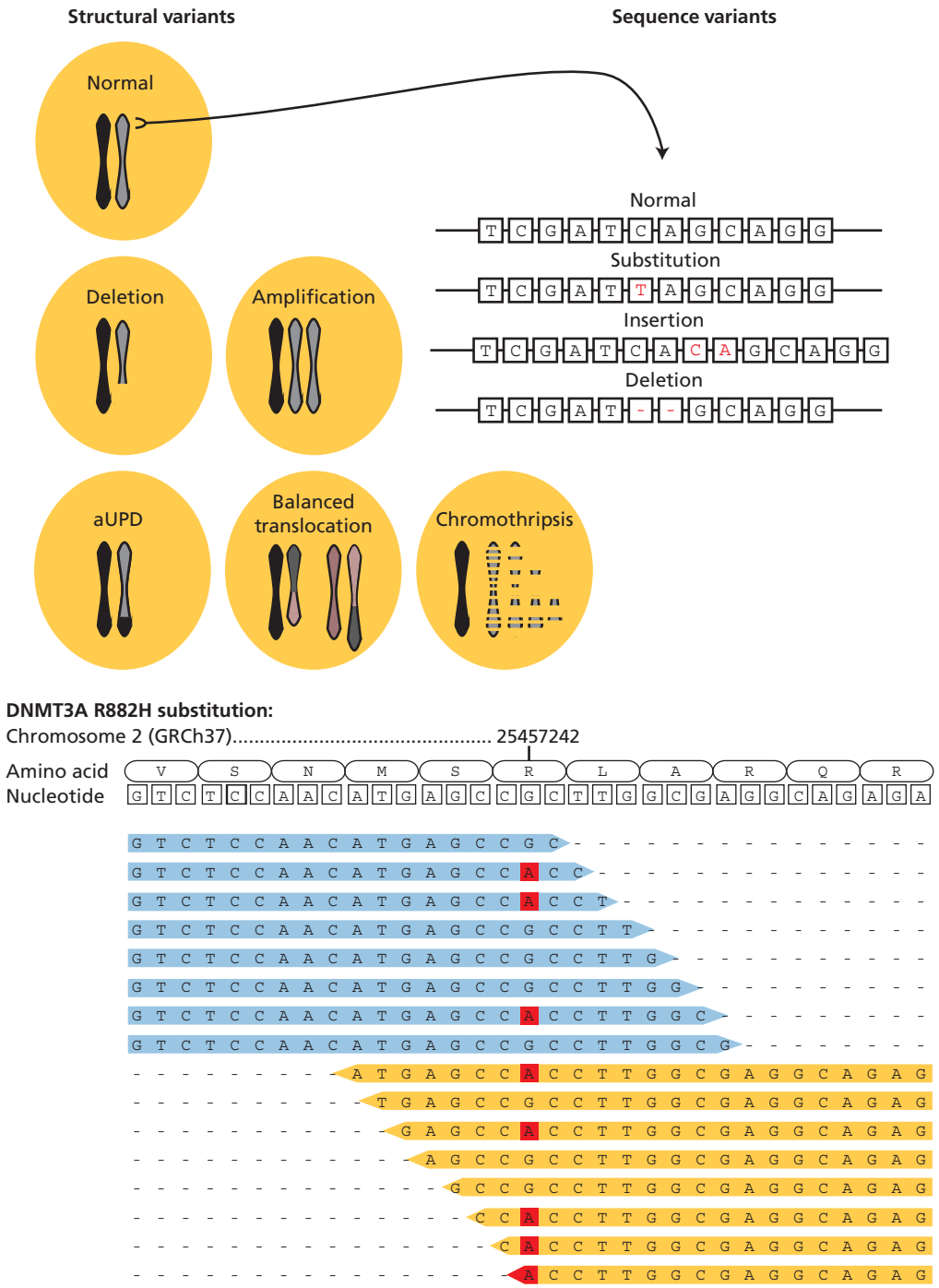


Figure 18.2 Different types of mutations and chromosomal rearrangements

Classes of DNA mutations

The structure and sequence of genomic DNA can be altered in several ways by mutagens. The resulting changes belong broadly to one of two categories: those that change the nucleotide sequence at a local scale and those that affect large genomic regions such as parts of chromosomes or entire chromosomes (Figure 18.2). Mutations that change the nucleotide sequence of DNA are broadly classified as *substitutions* and *indels* (insertions or deletions). When substitutions affect coding exons, they can be *synonymous* (no change in the coded amino acid), *missense* (change of the coded amino acid to another amino

acid) or *nonsense* (change of the coded amino acid to a stop codon). When indels affect coding exons they can have similar effects. However, if the size of an indel is not a multiple of three nucleotides the resulting change is a *frameshift* mutation. Frameshift mutations lead to the translation of a novel aberrant sequence of amino acids terminating at an endogenous stop codon sequence.

Both substitutions and indels do occur outside coding exons where their effects are more difficult to decipher. When they affect splice junctions of exons, they can lead to reduced usage or complete exclusion of an exon from the final mRNA. When they occur elsewhere in the genome they may not have a significant functional effect, but in some cases they can affect gene regulatory sequences. Examples of such mutations having an oncogenic effect were recently described.

Mutations altering the copy number of regions of the genome are referred to as *amplifications* and *deletions*. They can affect part of a gene, several genes, large chromosome segments or whole chromosomes. With these mutations, the net gain or loss of genes has an oncogenic effect, although for large amplifications or deletions it is often difficult to pinpoint which of the several genes involved are most important. *Chromosomal translocations* and *inversions* are a recurrent type of structural genomic rearrangement in haematological and other cancers. With these mutations, two regions of the genome situated far away from each other are juxtaposed as a result of breakages in their host chromosomes, followed by illegitimate fusions between them. These events can result in the formation of a fusion oncogene, such as *BCR-ABL1* in the t(9;22) in CML or lead to marked over-expression of a gene such as *CCND1* or *MYC* in the t(11;14) and t(8;14) respectively (Figure 18.3). In each of the latter two cases, the target gene is over-expressed as a result of coming under the control of the strong IGH promoter in mantle and Burkitt lymphomas, respectively. *Chromothripsis* is a very dramatic type of structural change affecting the genomes of some cancers. In chromothripsis, tens to hundreds of chromosomal rearrangements involving localized genomic regions can be acquired in an apparently single catastrophic event.

Finally, an important concept in cancer genomics is that of *loss-of-heterozygosity (LOH)*. As each of us inherits a maternal and a paternal allele for every gene or genomic region, we are heterozygous throughout our genome. Our heterozygosity can be verified using polymorphisms that distinguish the maternal and paternal alleles. In cancer it is common to find that areas of the genome have 'lost their heterozygosity', as they only display one of the two alleles and this happens if one of the two alleles is lost. This can be due either to a deletion of that allele or its replacement by the other allele (which is now duplicated). The duplication of one allele and associated loss of the other is known as acquired *uniparental disomy (aUPD)* and is the result of mitotic recombination (or mitotic cross-over) between two homologous chromosomes. As there is no net loss of genetic material, aUPD is also known as *copy neutral LOH*. Mitotic recombination occurs at a very low rate and the reason it appears a lot more frequently in cancer is because just like any other mutation, aUPD can give an additional growth advantage to the host cell, often by duplicating a previously acquired oncogenic mutation. For example, *FLT3* internal tandem duplications are commonly duplicated by aUPD.

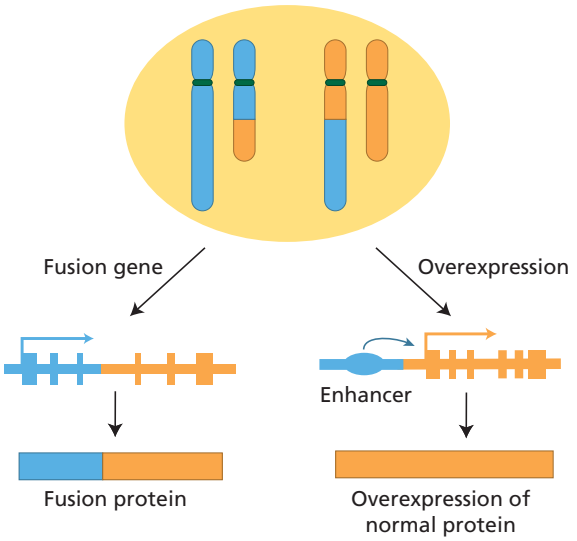


Figure 18.3 Common mechanisms of gene activation by chromosomal translocations: (1) formation of a chimaeric gene by fusion of the two genes spanning the chromosomal breakpoints and (2) over-expression of an intact gene through the action of a potent enhancer translocated upstream.

Inherited predisposition to haematological cancers

Most haematological malignancies are sporadic and not attributable to identifiable heritable or environmental risk factors, with some important exceptions (Table 18.2). For example, the risk of developing blood neoplasms is higher in individuals with certain uncommon inherited syndromes. In most instances, what is inherited is a genetic predisposition to develop cancer(s) through transmission of genetic variants that are either oncogenic themselves or accelerate the acquisition of somatic mutations that cause overt cancer. As cancer is a multistep process, acquiring the first oncogenic variant through germline transmission significantly increases the chances of acquiring the full set of variants that generate a cancer during the person's lifetime, hence the increased risk of cancer in these individuals. Also, as germline mutations are present in all cells of the body, they can have developmental effects in several organs and tissues. For this reason, many of the common mutations acquired somatically in cancer are never seen as heritable mutations. For example the acute myeloid leukaemia

Table 18.2 Examples of risk factors for the development of leukaemia and lymphoma

Germline syndromes
- Oncogenic variants
  - *RUNX1* (familial platelet disorder with propensity to myeloid malignancy)
  - *CEBPA* (rare cases of familial AML)
  - *NF1* (Type I neurofibromatosis)
  - *PTPN11* (Noonan's syndrome)
  - *CBL* (rare cases of JMML)
- Variants that increase mutagenesis
  - *BLM* (Bloom's syndrome)
  - *ATM* (Ataxia Telangiectasia)
- Predisposition alleles/haplotypes
  - 46/1 haplotype in PV/ET
  - Several alleles in multiple myeloma, e.g. rs603965 in t(11;14)
  - Several alleles in chronic lymphocytic leukaemia, e.g. rs13397985

Endogenous mutagenic processes
- Spontaneous deamination of methylated cytosines at CpG islands
- Cytidine deaminases (AID, APOBEC)
- RAG-driven deletions, e.g. in B-ALL carrying t(12;21)

Exogenous mutagenic processes
- Biological
  - EBV infection (EBV-driven lymphoproliferative disorder)
  - HTLV-1 infection (adult T-cell leukaemia/lymphoma)
  - HHV-6 (Castelman's disease and lymphomas of serous cavities)
- Chemical
  - Benzene
  - Chemotherapeutic agents
- Physical
  - Ionizing radiation

(AML)–associated *NPM1* mutations lead to early embryonic lethality in mice and unsurprisingly are never seen as germline mutations in humans.

In familial platelet disorder with propensity to myeloid malignancy (FPD/AML), heterozygous inheritance of a *RUNX1* oncogenic variant causes an autosomal dominant syndrome characterized by mild to moderate thrombocytopenia, abnormalities of platelet function and a propensity to develop myelodysplasia (MDS) and/or AML. *RUNX1* is a transcription factor critical for the establishment of definitive haemopoiesis and a target of somatic mutations including *deletions*, *translocations*, *substitutions* and *indels* in leukaemia. Most familial *RUNX1* mutations in FPD/AML lead to premature stop codons or affect the Runt domain of the gene involved in DNA binding. The median incidence of myeloid malignancy in FPD/AML ranges between 20 and 65% between families, but even within the same family affected individuals may present with variable clinical severity and at varying ages. This highlights how *RUNX1* mutations are insufficient for carcinogenesis in isolation, but require co-operating somatic mutations for cancer development. In fact mutations affecting the other copy of *RUNX1* as well as unrelated chromosomal deletions and translocations are commonly seen at the time of leukaemia progression.

An increased risk of malignancy can also arise from inheritance of variants in genes not directly implicated in cancer. Bloom's syndrome is an autosomal recessive disorder characterized by predisposition to the development of haematological and solid cancers, along with a constellation of physical findings. In this disorder, biallelic mutations in the gene for the DNA helicase *BLM* result in a widespread genomic instability and a highly increased rate of mitotic recombination leading to aUPD. This can lead to rapid homozygosity of oncogenic mutations and an increased growth advantage to the host cell.

As well as genetically well-defined syndromes such as the ones described above, a predisposition to cancer is also seen in some families with haematological malignancies, including CLL, myeloma and lymphoma. The increase in risk is generally small and the genetic basis unclear, however *genome-wide association studies* (GWAS) have identified areas of the genome within which the genetic risk 'resides'.

An intriguing example of inherited risk comes from the observation that JAK2 mutations, found in most cases of polycythaemia vera (PV) and also in other myeloproliferative neoplasms, are acquired preferentially within a particular JAK2 *haplotype* called 46/1. This haplotype, whose frequency in Europeans is approximately 25%, appears to 'invite' the acquisition of somatic JAK2 V617F mutations and gives a 3–4 times increased risk of developing JAK2-V617F-positive myeloproliferative neoplasm. It is not clear what underlies this phenomenon, but one possibility is that acquisition of JAK2 V617F on this haplotype gives a greater growth advantage to the host cell than its acquisition on any other JAK2 haplotype ('fertile ground hypothesis'). An alternative hypothesis is that the mutation occurs more frequently on this haplotype (hypermutability hypothesis).

Finally, another phenomenon of predisposition to a specific acquired mutation occurs in children with Down syndrome. In the first three years of life, up to 10% of these children develop transient myeloproliferative disorder (TMD), a disease that mimics AML, but typically resolves spontaneously within a few months. Almost all cases of TMD carry somatic mutations affecting the X-linked transcription factor GATA1, which drive the proliferation of leukaemic cells in TMD. The mutations disappear with the resolution of the TMD.

Acquired DNA mutations in haematological cancers

Although a lot can be learned from studying genetic susceptibility to cancer, neoplasms arise as a result of somatically acquired DNA mutations. The accumulation of somatic mutations is an inevitable consequence of the passage of time and this is true for all cells of all living organisms. However, mammalian genomes are afforded an extraordinary level of protection from

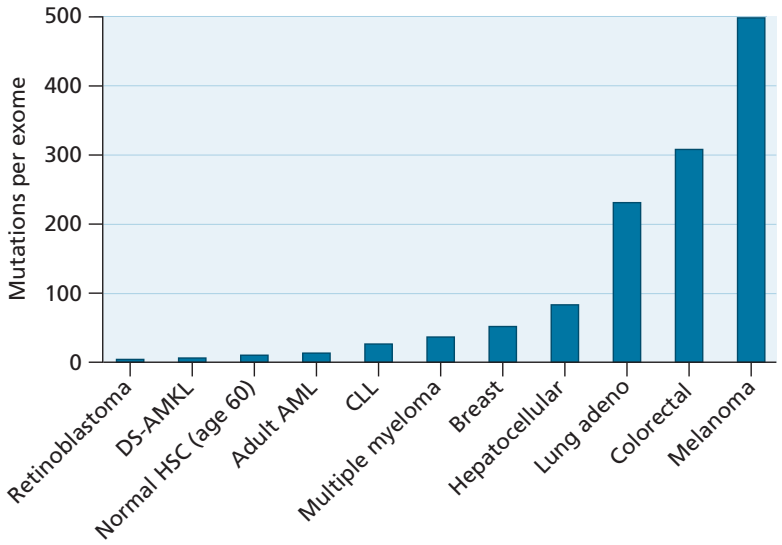


Figure 18.4 Median number of mutations per exome in various haematological and solid cancers.

mutations. Nevertheless, mutations do accumulate with time at a rate which is determined by the very small infidelity of the DNA replication and repair machinery and by the degree of exposure to endogenous or exogenous mutagens. Therefore, each somatic cell and by extension each cancer cell, harbours a variable number of somatic mutations, which accrued during its lifetime and that of its ancestors from the fertilized egg onwards (Figure 18.1). This cargo of mutations in the cancer cell genome can be generated through a number of different mutational processes. The number and the nature of the mutations can provide important pathogenic information and in some cases identify the responsible mutagen and the extent of its impact on the genome. For example, it is evident that the total number of mutations identified in the genomes of different forms of cancer varies widely and reflects the type and length of exposure to different mutagens (Figure 18.4). Also, studies of the particular type of change in DNA sequence can give important information about the mutagens and processes involved. More than 30 such *mutational signatures* have now been described and for many of these the responsible biological processes or mutagens have been identified. Below we discuss some of these processes.

Endogenous processes

A specific type of DNA mutation that affects all cells is the spontaneous deamination of methylated cytosines primarily at CpG islands which, if not corrected by editing enzymes, leads to a C > T *transition*. This process causes mutations at a very slow, but constant, rate and is thought to account for most of the spontaneous background mutation rate and for the underrepresentation of CpG dinucleotides in mammalian genomes. It is thought to play a role in the evolution of species and can lead to cancer development if it causes oncogenic mutations. In fact it is the only recognizable mutational process in most AML genomes.

A different set of mutagenic signatures relevant to haematological cancers are caused by the cytidine deaminases APOBEC and AID (activation-induced deaminase). APOBEC proteins are a group of RNA editing enzymes that under certain circumstances cause specific types of DNA mutations, including a signature common to B-lymphoid neoplasms including ALL, CLL, MM and DLBCL. These mutations are in the form of N > C at TpCpN trinucleotides (where N is any nucleotide) and are often associated with local hypermutation of regions of the genome, a phenomenon known as '*kataegis*'. AID is the master regulator of secondary antibody diversification and initiates the class-switch recombination and hypermutation of immunoglobulin genes. Ectopic action of AID has been detected at a number of secondary sites, some of which have been implicated in the pathogenesis of CLL and DLBCL (e.g. *BCL-6* mutations and *IGH* translocations).

Another group of DNA editing enzymes, the RAG endonucleases *RAG1* and *RAG2,* mediate the physiological V(D)J recombination process during B-cell development. In cases of B-ALL carrying t(12;21) and expressing the chimeric ETV6-RUNX1 fusion protein, the RAG enzymes can be aberrantly activated to drive a mutational process characterized by deletions scattered across the genome at regions flanked by recombination signal sequence motifs. These deletions can be oncogenic and collaborate with ETV6-RUNX1 to drive B-ALL.

Exogenous processes

The best known mutagen, tobacco smoke, causes mutations with a particular signature which abound in the genomes of lung and some other epithelial cancers from smokers. So far, no evidence of this signature has been described in haematological cancers and the same can be said for UV-induced mutations. However, exposure to a number of exogenous biological, chemical and physical agents can lead to haematological cancers. One of the

best studied examples is the correlation between Epstein–Barr virus (EBV) infection and development of cancers, including lymphomas. Asymptomatic EBV infection can be demonstrated in >90% of the world population, in whom the virus can establish a latent infection in B-lymphocytes. The latent virus expresses a set of genes/antigens to maintain the viral genome and to successfully evade host immune surveillance. The same antigens have the ability to manipulate several cellular pathways and promote B-cell proliferation. This proliferation is normally kept in check by EBV-specific T lymphocytes that continuously survey and suppress infected B cells. However, in some highly immunosuppressed individuals, such as bone marrow or solid organ transplant recipients, this suppression fails and an EBV-driven B-cell lymphoma develops (post-transplant lymphoproliferative disorder). EBV is also implicated in the pathogenesis of Burkitt lymphoma and Hodgkin lymphoma, although the mechanism leading to neoplastic transformation is less understood. Similarly, infection with the human T-lymphotropic virus type 1 can lead to development of adult T-cell leukaemia/lymphoma, whilst HIV infection significantly increases the risk of B-cell lymphomas through reduction of anti-tumour immuno-surveillance, a risk that can be mitigated by combination antiretroviral therapy.

Among oncogenic chemical and physical agents in haematology, exposure to benzene, ionizing radiation, alkylating agents and other chemotherapeutic agents have been clearly linked to the development of leukaemia and lymphoma with variable latency. A particular example is that of the association between exposure to topoisomerase inhibitors, such as etoposide and anthracyclines, and the development of acute leukaemia associated with chromosomal translocations involving the mixed lineage leukaemia (*MLL*) gene on chromosome 11q23. Topoisomerase II (Topo II) creates double-stranded breaks during execution of its function of relaxing overwound DNA, and it appears that Topo II inhibitors may stabilize complexes that are formed between the enzyme and free DNA ends, thus increasing the likelihood that those ends might participate in a translocation. The possibility has been raised that naturally occurring Topo II inhibitors have a causative role in some cases of leukaemia, but there is no good evidence to support this.

From genotype to phenotype

Regardless of how mutations are generated, most of them will not impart a growth advantage and are deemed to be '*passengers*'. In most cancers only a minority of mutations are 'drivers', a term used to signify mutations that impart oncogenic properties to the host cell(s) and drive cancer growth; either by activating oncogenes or inactivating tumour suppressor genes. While a definitive categorization of any mutation as a driver would require functional evidence that the variant can actually cause cancer, this is impractical for the ever-increasing number of variants that are identified by NGS studies. Instead, the 'driver' label is commonly given to mutations that are seen in any cancer type more often than would be expected by chance. Implicit in this working definition is the fact that it relies on some statistical calculation of 'chance' and is therefore prone to some error. In other words, some of the less frequently observed true drivers may not make the cut, whilst some passengers may fortuitously do so.

Another important concept to highlight here relates to the fact that each cancer has several driver mutations that work together to endow the cancer cells with their neoplastic phenotype. When studying these mutations, what becomes clear is that certain mutations co-occur regularly in the same cancer (e.g. NPM1 and FLT3 mutations in AML), whilst others almost never co-occur (e.g. JAK2 and CALR mutations in MPD). When these observations are combined with what we know about the function of individual mutations, it becomes apparent that mutations that co-occur more often than expected by chance (*co-occurrence*), complement each other in transforming the host cell. By contrast, mutations with similar function do not co-occur or do so less often than expected by chance (*mutual exclusivity*), as only one or the other is sufficient to cause a specific effect or subvert a particular pathway.

Below we discuss some important examples of driver mutations from different haematological malignancies and highlight some of the better understood genetic pathways subverted in these cancers.

NOTCH-signalling and lymphoid malignancies

NOTCH1 is a cell surface receptor with a central role in T-lymphopoiesis. It regulates interactions between adjacent cells and is activated by ligand binding, which triggers heterodimerization of the receptor and proteolytic cleavage of an intracellular peptide (ICN1=intracellular domain of NOTCH1) that migrates to the nucleus, where it affects DNA transcription. In the bone marrow, NOTCH1 signalling is required to promote the generation of T-lymphoid precursors from early haemopoietic progenitors, at the expense of B-lymphopoiesis. T-cell precursors subsequently migrate to the thymus, where NOTCH1 signalling is essential for survival and differentiation of T-cells with a successfully rearranged ab T-cell receptor. Furthermore, NOTCH1 is required for T-cell proliferation in the presence of antigen presenting cells in peripheral tissues.

Given its physiological roles above, it is not surprising that in the multistep process leading to T-cell lymphoblastic leukaemia (T-ALL), constitutive activation of NOTCH1 signalling is one of the most prominent genetic events. It is present in 60% of T-ALL cases, and it invariably results in constitutive activation of ICN1 and of its transcriptional activity. Interestingly, deregulated ICN1 activity can be achieved through different molecular mechanisms, all of which represent subversions of the physiological pathway (Figure 18.5). Point mutations affecting

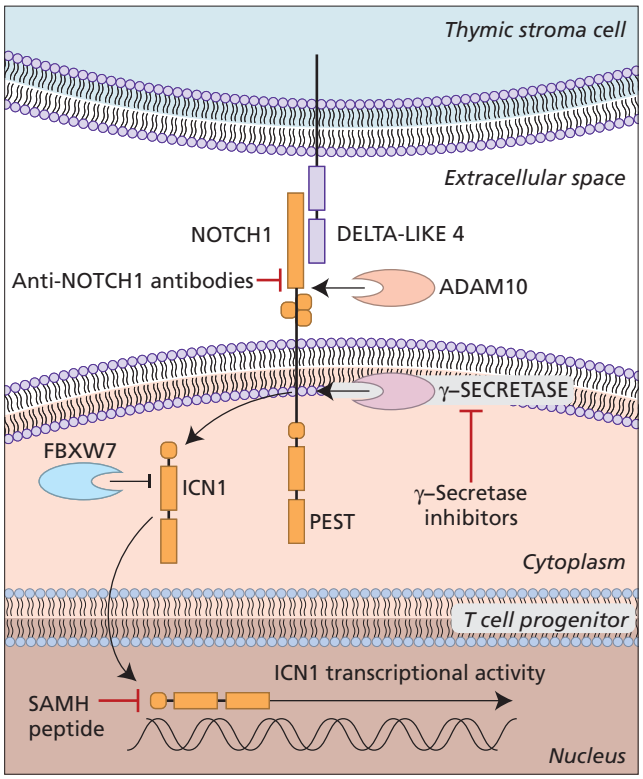


Figure 18.5 Schematic representation of NOTCH1 signaling and possible inhibitory strategies. NOTCH1 undergoes a double proteolytic cleavage upon interaction with the delta-like 4 ligand. The first is by the ADAM10 metalloprotease and can be inhibited by anti-NOTCH antibodies. The second cleavage is by the gamma secretase complex, that can be targeted by small molecules. This double cleavage results in the release of the intracellular domain of NOTCH1 (ICN1) which activates transcription of target genes in the nucleus. FBXW7 antagonizes this activity by targeting activated NOTCH1 to the proteasome. Lastly, small SAMH1 peptides disrupt the NOTCH1 nuclear transcriptional complex.

the heterodimerization domain of NOTCH1 result in ligand-independent activation, while truncating mutations of the C-terminal regulatory PEST domain result in increased half-life of ICN1. Rarely, *NOTCH1* is involved in the chromosomal translocation t(7;9), which results in the expression of a truncated and constitutively active form of the protein. Finally loss-of-function mutations of FBXW7, an E3 ubiquitin protein ligase that targets ICN1 for degradation, similarly result in its constitutive activation. Studies have shown that ICN1 over-expression causes leukaemia by altering the transcriptional profile of the cell and in particular by activating the oncogene MYC. Among co-operating recurrent events are deletions of the *CDKN2A* tumour suppressor locus, encoding a negative regulator of the cell cycle and a positive regulator of p53, present in up to 70% of T-ALL. In addition, T-ALLs characteristically show translocations and aberrant expression of transcription factor oncogenes, and deletions of other tumour suppressor genes. Clearly, different T-ALL cases will display different arrays of such genetic lesions; however, it appears that T-ALL cells rely on NOTCH1 signalling independent of its mutational status, making the pathway an attractive therapeutic target.

The well-defined role of NOTCH1 in T-cell development explains the relative specificity of these mutations to T-ALL. However, recently NOTCH1 PEST domain mutations were identified in approximately 10% of CLL, where they constitute a poor prognostic indicator. In most cases *NOTCH1* mutations occur in the subgroup of CLL with unmutated IGHV genes and cases with trisomy of chromosome 12 as the sole cytogenetic abnormality. Unlike T-ALL, *NOTCH1* mutations in CLL are almost exclusively frameshift or nonsense events, clustering within a hotspot in exon 34 of the gene and generating a truncated protein. The identification of *NOTCH1* mutations in CLL may be somewhat surprising, given the role of this gene in T-cell development; however, NOTCH1 signalling has a well-defined role in promoting terminal differentiation of mature B-lymphocytes to antibody-secreting cells. The theme of cancer mutations targeting genes involved in physiological pathways important to different types of cells in the body is very common in both haematological and solid tumours.

Epigenetics and leukaemia

The regulation of transcription is the fundamental regulatory node for controlling gene expression. The process requires physical access of specialized proteins to genomic DNA in order to read and transcribe the code into RNA. However, every human cell contains about 2 meters of linear genetic information, which is packaged into a nucleus with a 6 μm diameter. This packaging is achieved by 'wrapping' genomic DNA around histones, a family of ultra-highly conserved nuclear proteins. This moulds genomic DNA into a beads-on-string configuration, which is then further folded and tightly compacted into a macromolecular complex known as 'chromatin'. This makes access to DNA very challenging, but also turns histones into the custodians of such access; a role which they conduct primarily through their N-terminal protein 'tails'. In fact, the process is controlled by different post-translational modifications of these histones, including methylation, acetylation and phosphorylation, which can lead to increased or decreased access to DNA and by extension to a higher or lower level of gene transcription, respectively. Additionally, changes to the DNA itself have a further role in controlling gene expression. The area of biology concerned with these phenomena is known as 'epigenetics' and corruption of epigenetic processes is a frequent event in leukaemogenesis.

Amongst the better understood examples of leukaemogenic mutations operating through their aberrant effects on chromatin are those affecting the *MLL* gene. MLL is a lysine-specific

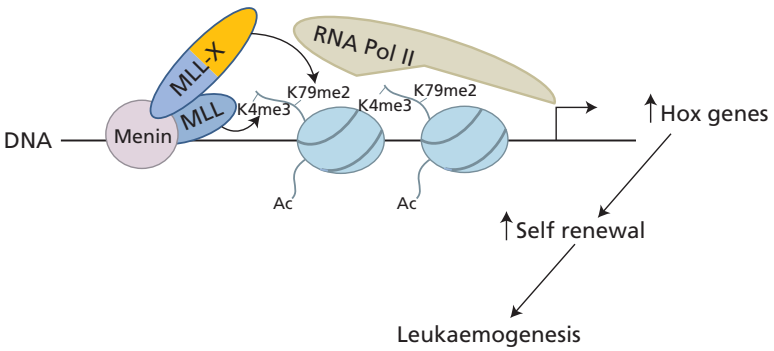


Figure 18.6 Functional consequences of disrupted MLL activity by an MLL fusion protein (MLL-X). The fusion protein acts in a dominant manner recruiting the normal MLL protein to lysine residues (K4 and K79) of histones that are not normally targeted by MLL and whose methylation promotes active and aberrant transcription of the downstream genes (such as HOXA genes).

methyltransferase which normally methylates specific lysine residues in histone tails to alter chromatin accessibility and increase gene expression. In the haemopoietic system, the MLL protein activates transcription of posterior *HOXA* genes (*HOXA7-A10*), which promote the self-renewal of HSCs and immature progenitors. In normal haemopoiesis, their expression is maintained by MLL in the HSC/immature progenitor compartments and diminishes as cells differentiate. However, *MLL* is a recurrent target of chromosomal translocations, which lead to an aberrant fusion of this gene to one of more than 80 gene partners. MLL-fusion proteins hijack the function of MLL and misdirect it to histone lysine residues, which are not normally targeted by the wild-type MLL. For example, the MLL-AF9 fusion protein was shown to be co-recruited with normal MLL (expressed from the intact MLL allele) to the *HOXA9* gene locus, where the two proteins methylate different lysines on local histones to 'open' the chromatin and increase *HOXA9* mRNA expression (Figure 18.6). This results in aberrant self-renewal of haemopoietic progenitors, leading to leukaemia.

Recent developments in this area have identified that BET (bromodomain and extraterminal) proteins are important for recruiting MLL fusion proteins to chromatin. Inhibitors of the interaction between MLL fusion proteins and BRD4, a BET family member, have shown potent antileukaemic activity against MLL-fusion leukaemias and have recently entered clinical trials.

As well as effects on chromatin, leukaemogenic mutations can alter the normal processes involved in the other important set of epigenetic events, namely modifications of DNA itself. The most important DNA modifications are those affecting cytosine residues, which can be methylated by DNA methyltransferases such as DNMT3A and DNMT3B to form methylcytosine, a residue that is generally inhibitory to gene transcription. Additionally, cytosine methylation is an important process for suppressing the expression of the large numbers of genomic parasites found in mammalian genomes, such as transposons and endogenous retroviruses. However, cytosines at particular parts of the genome rich in cytosine-guanine dinucleotides and known as CpG islands, are not always methylated. CpG islands are frequently located at the promoter regions of genes and methylation of cytosines in these islands is associated with reduced gene expression. In fact, unmethylated CpG cytosines are frequently found in regions of 'open' chromatin where histones allow the transcription machinery access to the genomic DNA.

So how do cytosines become unmethylated? In fact, with the exception of germ cells, the process was poorly understood until recently. It appears that instead of demethylation, methylcytosines can undergo hydroxylation to hydroxymethylcytosines by enzymes known as TET dioxygenases; namely TET1, TET2 and TET3. These enzymes can further modify hydroxy-methylcytosine to formyl- and then carboxylcytosine, which can then be converted back to a cytosine. Importantly, the hydroxyl group for the initial hydroxylation by TETs comes from α-ketoglutarate produced by the action of isocitrate dehydrogenase 1 or 2 (IDH1 or IDH2) as an intermediate in the tricarboxylic acid (Krebs) cycle (Figure 18.7).

A number of key genes involved in the above processes are recurrently mutated in myeloid and less often in some lymphoid malignancies. In particular, *DNMT3A* mutations are found in 30%, *TET2* in 10%, whilst *IDH1* and *IDH2* are each mutated in 7–8% of cases of AML. Additionally, all of these genes are mutated in other myeloid neoplasms including MDS, MPN and CMML. Both *DNMT3A* and *TET2* are targeted by nonsense or other loss-of-function mutations that can be homozygous or heterozygous. However, the most frequent *DNMT3A* mutation is a missense mutation leading to substitution of arginine 882 with histidine (R882H) or cysteine (R882C) and less commonly serine (R882S) or phenylalanine (R882P). These substitutions are almost always heterozygous, but they are thought to have a dominant negative effect, i.e. as well as lacking methylase activity themselves, they act to inhibit the methylase function of the remaining normal DNMT3A allele, although they may also have other effects. Mutations affecting *IDH1* and *IDH2* have the effect of disrupting the enzymatic function of TET proteins by generating the 'oncometabolite' 2-hydroxyglutarate (2-HG) instead of α-ketoglutarate. There is now good evidence that 2-HG mediates the oncogenic properties of these mutations, by inhibiting the function of TETs and also that of a number of chromatin-modifying enzymes such as histone demethylases. In keeping with this, *IDH1, IDH2* and *TET2* mutations do not occur together in the same AML, supporting the premise that they influence similar genetic pathways (see mutual exclusivity

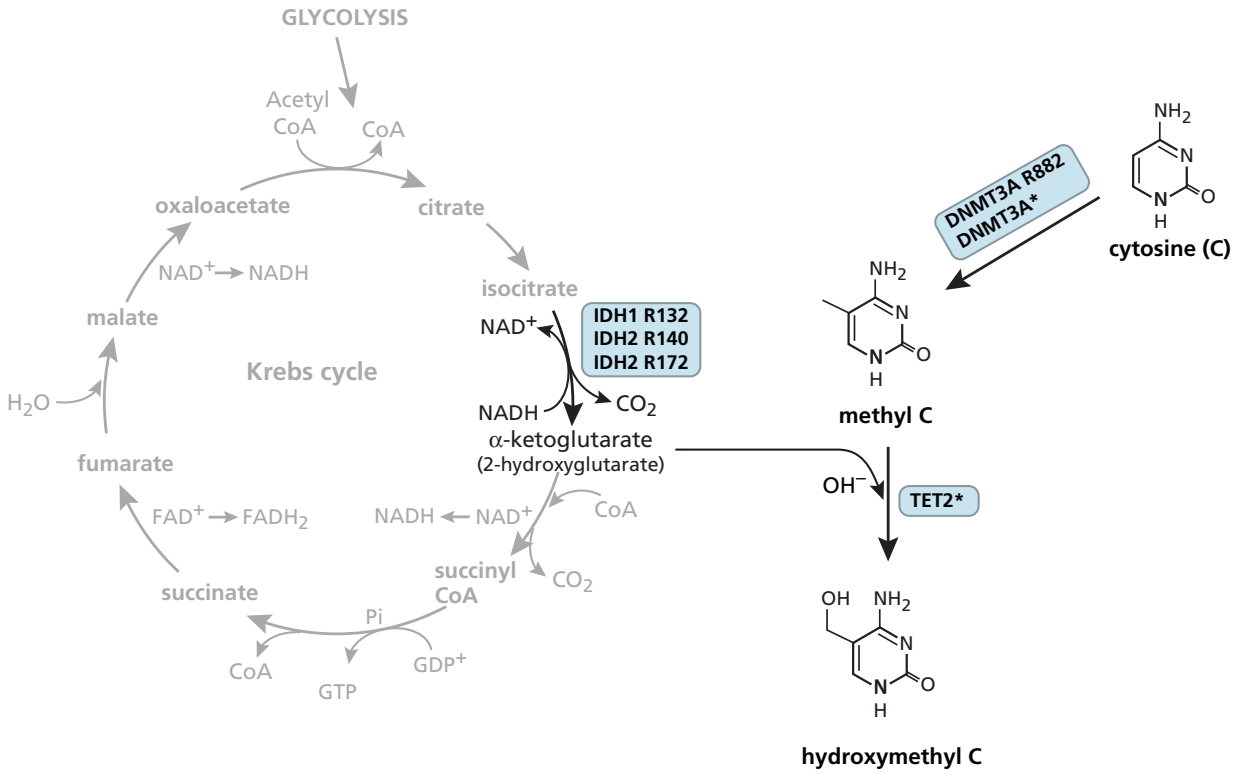


Figure 18.7 Enzymes involved in DNA methylation whose mutations are implicated in myeloid disorders. The metabolic reaction catalysed by each enzyme is indicated along with the most frequent leukaemogenic mutations (blue boxes + see text). The oncometabolite 2-hydroxyglutarate is produced, instead of α-ketoglutarate, by the action of mutant IDH1 or IDH2. 2-hydroxyglutarate inhibits TET dioxygenases (including TET2) and other epigenetic enzymes.

and co-occurrence earlier in this chapter). Studies to understand the molecular consequences of *DNMT3A, TET2* and *IDH* mutations have demonstrated that their effects on DNA methylation are genome-wide rather than being specific to leukaemia genes. In other words, it is probable that they change gene expression of some genes that promote leukaemia and others that do not. However, their net effects are pro-leukaemic.

Mutations affecting IDH1 and IDH2 are *gain-of-function* changes, which in contrast to loss-of-function mutations are usually more amenable to targeting by new therapies. The specific mutations affecting codons IDH1 R132, IDH2 R140 and IDH2 R172 (Figure 18.7) are seen commonly in myeloid malignancies, but also occur in non-haemopoietic cancers such as glioblastoma and cholangiocarcinoma. IDH1 and IDH2 inhibitors are now in early phase clinical trials and showing significant promise against AMLs carrying these mutations.

Tyrosine kinases and myeloproliferation (see also Chapter 26)

Tyrosine kinases are critical for the responses of haemopoietic progenitor cells to external growth stimuli. The binding of a ligand to the extracellular surface of a receptor tyrosine kinase (RTK) promotes receptor dimerization, which in turn stimulates autophosphorylation of specific tyrosine residues within the intracellular part of the protein. The dimerization and associated increase in kinase activity result in recruitment of effector molecules and activation of downstream signalling pathways. Many cytoplasmic tyrosine kinases are also thought to be activated by phosphorylation and dimerization. Individual myeloproliferative disorders are associated with activation of different tyrosine kinases, with specific correlation between gene mutations and the resulting neoplastic phenotype. For example, the kinase activity of JAK2 is increased by point mutations (e.g. Val617Phe) in almost all cases of polycythaemia vera and 50% of essential thrombocytemia, while point mutations in *KIT* (Asp816Val) are present in 90% of cases of systemic mastocytosis. As a consequence of these changes, the corresponding signalling pathway is activated and this provides the transformed cell with a proliferative or survival advantage.

Tyrosine kinases can also be activated by translocations, like in CML where the *BCR* gene becomes fused in-frame with the *ABL1* tyrosine kinase gene as a consequence of the Ph translocation (Figure 18.8). This usually results in the formation of a fusion protein of 210 kDa (hence p210 BCR-ABL). The fusion protein displays constitutive ABL1 kinase activity promoted by dimerization through the coiled-coil domain of

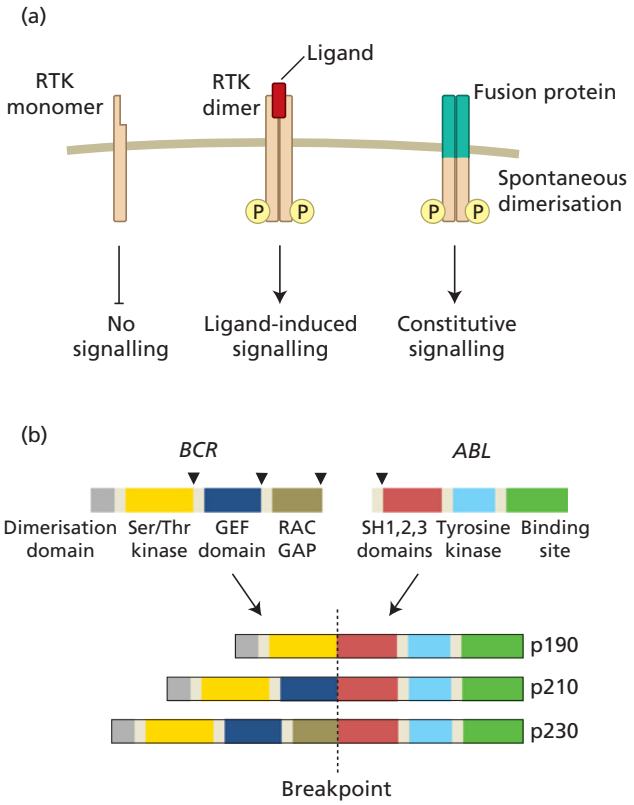


Figure 18.8 Fusion tyrosine kinases. (a) In normal cells, receptor tyrosine kinases (RTKs) exist as inactive monomers which, upon ligand binding, become activated by dimerization and autophosphorylation. By contrast fusion tyrosine kinases dimerize spontaneously and are constitutively active. (b) Different forms of *BCR-ABL1* fusion genes contain varying lengths of the *BCR* gene and result in different molecular effects (see text).

the BCR protein. Numerous downstream pathways are activated by the BCR-ABL1 protein. For example, activation of the RAS pathway is thought to contribute to the increased cell division and proliferation seen in CML cells, altered interactions with the actin cytoskeleton may underlie the diminished adhesion to inhibitory bone marrow stromal cells and STAT5-dependent upregulation of BCL-XL is implicated in the reduced apoptosis of CML progenitors. Transgenic mouse models have demonstrated that *BCR-ABL* is sufficient to induce dramatic expansion of myeloid precursors *in vivo* leading to a phenotype that resembles the human disease.

As shown in Figure 18.8, two other versions of the BCR-ABL1 protein exist in which the breakpoint occurs in different introns of the BCR gene than those usually seen in CML. The smaller protein, p190 BCR-ABL, is found in Ph-positive ALL and has a higher level of constitutive tyrosine kinase activity than the p210 BCR-ABL protein. This is thought to contribute to the more aggressive behaviour of the ALL. A p230 BCR-ABL isoform is found rarely in patients with CML. In some cases, it is associated with a morphological picture resembling chronic neutrophilic leukaemia and possibly a lower rate of acute transformation. These variants illustrate how small structural differences in a fusion protein can cause significant differences in disease phenotype.

A group of related tyrosine kinases (*PDGFRA*, *PDGFRB* and *FGFR1*) can participate in fusion genes found in eosinophilic myeloproliferative disorders. A remarkable feature of these syndromes is the large number of fusion partners described for each of these kinases, with more than 10 partners currently known for *FGFR1*, more than 5 for *PDGFRA* and more than 20 for *PDGFRB*. As with *BCR* in *BCR–ABL1*, most of these partners harbour dimerization domains that facilitate constitutive activation of the tyrosine kinase. However, reflecting their molecular heterogeneity, these disorders exhibit significant clinical diversity. For example, cases associated with the *FIP1L1-PDGFRA* fusion exhibit marked eosinophilia, while those associated with *ETV6-PDGFRB* often show dysplastic features and a lesser degree of eosinophilia. Additionally, cases involving *FGFR1*, collectively known as the *8p11 myeloproliferative syndrome*, are usually associated with a co-existing non-Hodgkin lymphoma. Notably, cases associated with rearrangements of the *PDGFR* genes respond well to imatinib, while those associated with *FGFR1* fusions do not. Newly developed FGFR inhibitors are currently being tested against the latter.

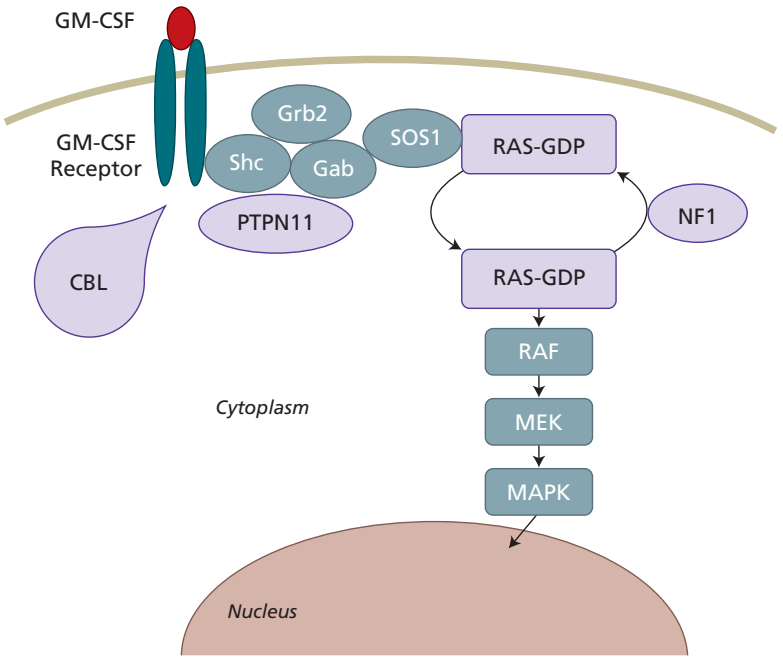


Figure 18.9 The RAS pathway and juvenile myelomonocytic leukaemia (JMML). The Ras pathway mediates many of the downstream effects of cytokine receptor activation. In most cases of JMML, the pathway is activated by mutations in one of four genes depicted in black text, which are involved in signalling downstream of the GM-CSF receptor: RAS (25% of cases), PTPN11 (also known as SHP2, 35% of cases), NF1 (15% of cases) and CBL (10–15% of cases). All of these mutations lead to RAS activation, which signals to a number of downstream effectors, including the RAF-MEK-MAPK pathway, to alter gene transcription and promote cell proliferation.

GM-CSF and RAS signalling in juvenile myelomonocytic leukaemia

Juvenile myelomonocytic leukaemia (JMML) is a rare disorder of childhood characterized by a marked expansion of both myeloid and monocytic lineages, with varying degrees of dysplasia and exquisite sensitivity to granulocyte–macrophage colony-stimulating factor (GM-CSF) in colony assays. These features reflect the fact that components of the GM-CSF receptor signalling and its key downstream effector, the RAS pathway, are frequently mutated in JMML (Figure 18.9).

The Ras genes (*KRAS, NRAS* and *HRAS*) code for small GTPases that operate as molecular switches, which are 'on' when bound to GTP and 'off' when the GTP is hydrolysed to GDP. Ras has only weak GTPase activity and in order to hydrolyse GTP and turn itself 'off' it requires activation by a GTPase-activating protein (Ras-GAP). Once GTP is hydrolysed to GDP, Ras re-activation requires that GDP is exchanged for GTP. The latter process is catalysed by guanine nucleotide exchange factors (GEFs), which release GDP and allow Ras to bind the more abundant GTP. Once active, Ras stimulates cell proliferation by activating a number of cellular pathways, including mitogen-activated protein kinase (MAPK), PI3 kinase (PI3K) and phospholipase C epsilon (PLCε).

Mutations in the *NRAS* and *KRAS* genes are seen in many solid cancers, but also in several haematological malignancies, including AML, MDS, chronic myelomonocytic leukaemia (CMML), myeloma and others. Analyses of samples from patients with JMML show point mutations in *NRAS* and *KRAS* in approximately 25% of cases. The point mutations cause single amino acid substitutions, which result in constitutively active forms of the protein and thus mimic markedly amplified signalling from the GM-CSF receptor. The neurofibromin (*NF1*) gene has Ras-GAP activity and is therefore a Ras inhibitor. Interestingly, germline inactivating mutations in the *NF1* gene cause neurofibromatosis type 1, an autosomal dominant disorder characterized by protean clinical features, including increased risk of developing JMML and other tumours. Patients who have both neurofibromatosis type 1 and JMML have a high frequency (60%) of somatic inactivating mutations of the second *NF1* allele in the leukaemic cells. Furthermore, inactivating *NF1* mutations are found in 15% of sporadic cases of JMML without the clinical syndrome of neurofibromatosis. It appears, therefore, that *NF1* acts as a classical tumour suppressor gene, where two-step inactivation of both alleles allows unchecked signalling through the RAS pathway. A similar story holds true for the gene *PTPN11*, which encodes the tyrosine phosphatase SHP2. Germline mutations of this gene can cause the developmental disorder Noonan's syndrome, which is associated with an increased incidence of JMML. Up to 35% of non-syndromic patients with JMML have somatic mutations of *PTPN11* in their leukaemic cells. The mutations all cause amino acid substitutions clustering in the N-terminal SH2 domain of the protein, which are thought to result in enhanced phosphatase activity and activation of the downstream RAS pathway, giving the host cells a growth advantage over their normal counterparts.

Most recently, mutations in the gene CBL were identified in 10–15% of JMML cases. Their identification followed the observation that some JMML cases exhibited *aUPD* of the long arm of chromosome 11. Further studies showed that these patients were born with germline heterozygous loss-of-function

mutations in CBL, which became homozygous by duplicating the part of chromosome 11 carrying the mutation. CBL is an E3 ubiquitin ligase that controls proliferative signalling networks by downregulating the growth factor receptor signalling cascades in various cell types. JMML-associated mutations thus lead to complete loss of E3 ligase activity and increased GM-CSF receptor kinase activity and downstream signalling.

Thus, the common theme underlying most JMML is enhanced GM-CSF receptor pathway signalling, and this can be achieved by activating mutations of a positive regulator (such as *RAS* or *PTPN11*) or inactivating mutations of a negative regulator (*NF1* or *CBL*). Not surprisingly, mutations in *RAS*, *NF1*, *PTPN11* or *CBL* are mutually exclusive, suggesting that mutation of just one component is sufficient to activate the signalling pathway and that additional mutations in the pathway offer little additional advantage. Other mutations that have been reported to occur less frequently in JMML target the genes include *ASXL1, FLT3, SETBP1* and *JAK3*, with the last two conferring a poorer prognosis.

Other Ras pathway mutations in haematological malignancies

The Ras protein superfamily comprises more than 100 related small GTPases with diverse functions, including the Ras, Rho, Rab, Rap and Arf subfamilies. Mutations in *HRAS*, discovered in 1982, were the first example of a human transforming oncogene. The closely related *KRAS* and *NRAS* were also found to be mutated in a plethora of different solid and haematological cancers. However, mutations in members of some of the other Ras subfamilies and also in genes coding for proteins in their downstream signalling cascades have been identified more recently. Most famous amongst them are mutations in *BRAF*, found frequently in melanoma and colorectal cancer amongst others. Amongst haematological cancers, mutations in *BRAF* were recently found in virtually all cases of hairy cell leukaemia and also in about 10% of cases of myeloma and less frequently in other lymphoid tumours. Also, mutations in the phosphatase PTPN1 are present in 20–25% of Hodgkin and primary mediastinal B-cell lymphomas; whilst mutations in *RHOA* were recently identified in 25–30% of peripheral T-cell lymphomas including 75% of angioimmunoblastic lymphoma.

Multiple myeloma

Unlike any other somatic cell type, B and T lymphocytes undergo physiological DNA rearrangements during their development. These events are confined to the loci encoding their antigen receptor components, and are required to generate the necessary antigen receptor diversity. In the case of B lymphocytes, the antigen receptor specificity is first generated through rearrangement of the V(D)J region mediated by the RAG1 and RAG2 enzymes when the developing B-cell is still in the bone marrow. In the mature circulating B-cell, the activation-induced (DNA-cytosine) deaminase (AID) protein then mediates two further processes, class-switch recombination and somatic hypermutation in the germinal centre of secondary lymphoid organs. While the former is required to change the class of the antibody produced by the B-cell (e.g. IgM to IgG), the latter introduces random point mutations in the antigen-binding site of the immunoglobulin in an attempt to increase its specificity. V(D)J rearrangements and class-switch recombination are achieved through generation of DNA double-stranded breaks (DSB) at specific consensus sites, deletions of part of the locus, and subsequent non-homologous end joining of the two DNA ends.

Multiple myeloma is a neoplasm of antibody-producing post-germinal centre B-lymphocytes. As an unwanted side-effect of their development, B-cells can occasionally acquire translocations caused by illegitimate joining of a DSB in the Ig region with a second DSB elsewhere in the genome, both caused by the activity of RAG or AID. In fact, about 40% of multiple myeloma cases show deregulated expression of oncogenes caused by their translocation next to highly active immunoglobulin (Ig) regulatory elements. This is a common theme in lymphoid malignancies in general, and underscores the tight link between normal developmental processes and cancer. Notably, Ig translocations in multiple myeloma do not usually result in fusion proteins, but rather in deregulated over-expression of a full-length proto-oncogene under the control of highly active Ig enhancer elements. Deregulation of the G1/S transition is a key early molecular abnormality in myeloma, and frequently a D-group cyclin is deregulated in Ig translocations. This can be either through a direct mechanism, where *CCND1* and *CCND3* expression is driven by Ig enhancers in the t(11;14) and t(6;14), respectively, or it can occur indirectly, such as in t(14;16) where over-expression of *MAF* leads to activation of *CCND2*. The t(4;14), present in 15–20% of multiple myeloma cases, is by contrast a rare example of a translocation resulting in the simultaneous deregulation of two oncogenes, *FGFR3* on der(14) and *MMSET* on der(4). Over-expression of FGFR3 protein occurs in only approximately 70% of patients with the translocation, while MMSET is over-expressed in all cases. The MMSET gene product is a histone methyltransferase with affinity for H3K36 and H4K240, and its deregulated expression can influence DNA damage repair processes and the overall transcriptional profile of the cell. FGFR3 is a transmembrane tyrosine kinase receptor that is involved in key embryonic and adult processes including regulation of proto-oncogenic pathways such as the mitogen-activated protein kinase (RAS-RAF-MAPK), phosphatidylinositol 3-kinase (PI3K-AKT-mTOR), phospholipase Cg (PLCg), protein kinase C (PKC) and signal transducer and activator of transcription (STAT). Interestingly, activating mutations of FGFR3 and MMSET are also found in multiple myeloma and other malignancies, indicating that they may be more broadly implicated in oncogenesis.

Multiple myeloma shows multistep evolution and is almost invariably preceded by the premalignant condition, monoclonal gammopathy of unknown significance (MGUS). Chromosomal aberrations are regularly present in MGUS, indicating that Ig translocations are early events in the pathogenesis of multiple myeloma, whilst transition towards full-blown myeloma is achieved through the step-wise acquisition of additional mutations. Among those are deletions of chromosomes 1p, 13 and 17p, and amplification of chromosome. Furthermore, mutations of genes involved in the MAPK signalling pathway, including *KRAS*, *NRAS*, *NF1* and the downstream protein BRAF are present in about 50% of cases, and their frequency increases in more advanced cases, again underscoring the broad relevance of this pathway to carcinogenesis.

Clonal evolution and subclonal architecture of haematological cancers

Clonal evolution of cancer

Cancer develops through the serial acquisition of somatic mutations in an individual cell and its progeny. The mutations are usually acquired in a stepwise manner, such that each time a new driver mutation is acquired, this can be seen as a cloning event. In other words, the new mutation en route to neoplasia can be seen as an event that extricates that cell from its peers, such that the onward path towards cancer continues in this individual cell's progeny. The next step in cancer evolution occurs when a new driver mutation arises in this progeny and a new clone is formed and so on. The process continues until a cell acquires a complement of driver mutations for frank malignancy to become manifest (*linear evolution*, Figure 18.10). In reality, the length of time required for this process is long enough for different individual cells within each of the serially appearing clones to acquire different mutations and continue along their own independent paths towards cancer (*branching evolution*), although one clone will usually get there first and represent the bulk of the disease. This process has been very well worked out in AML, ALL and multiple myeloma, where it is clear that although many cases follow a predominantly linear evolution, others follow a branching pattern. Of note, sometimes a very similar or even the same mutation can be acquired independently by two different subclones. This reflects a 'craving' for that mutation by the evolving cancer and the process is often referred to as *convergent evolution*. The subclonal structure of the tumour and the relationship between subclones is also relevant with regards to disease relapse and is discussed below.

Another important conundrum in cancer evolution is the order of acquisition of specific mutations. A number of studies have demonstrated such an order and AML, ALL and myeloma are again amongst the best studied paradigms for this phenomenon. For example, in myeloma there is good evidence that chromosomal translocations and aneuploidy are usually acquired first, whilst mutations in proliferative pathways are acquired later. Also in AML, mutations affecting genes involved in DNA methylation such as DNMT3A and TET2 are acquired first, whilst mutations in NPM1 are acquired second and proliferative mutations are again acquired late.

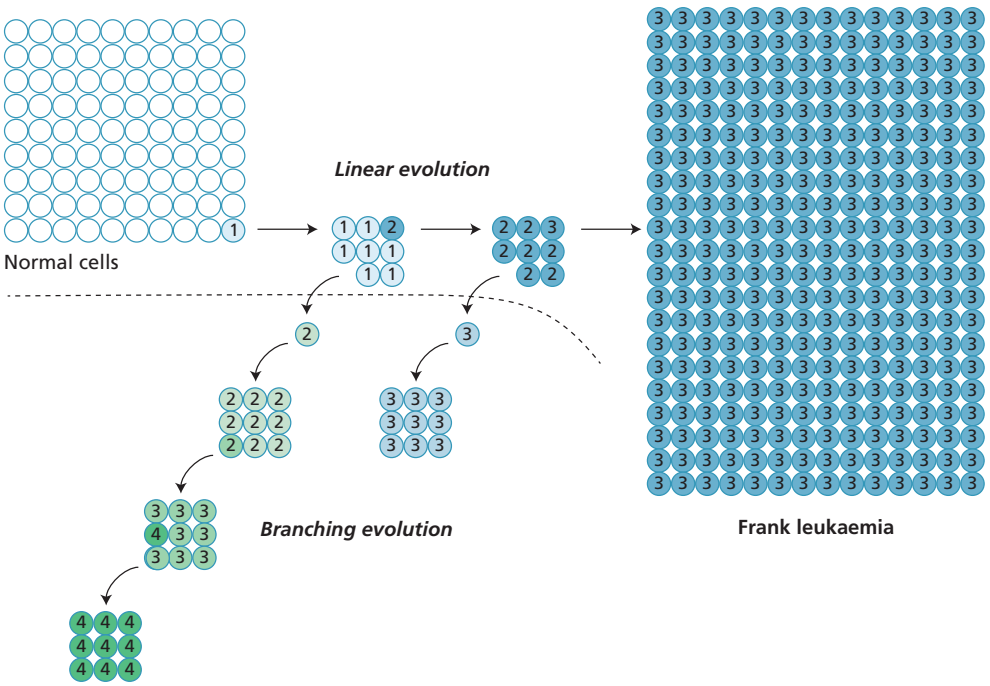


Figure 18.10 Linear and branching evolution.

There are two theories proposed for this recurrent order of acquisition: the *constraint theory* argues that a certain mutation has to happen first in order to 're-wire' the cell so that the second can give additional growth advantage, which would not be present in the absence of the first mutation. By contrast, the *opportunity theory* argues that the first mutation imbues features such as increased proliferation or accelerated level of mutagenesis, that confer greater opportunity for the second mutation to be acquired. Additionally, in the latter theory the order can be reversed, but this will happen a lot less often as there will be less opportunity. It is not yet clear which of the two theories holds true, or if indeed they both do, depending on the type of mutation; however, there is some evidence directly supporting the second. Namely, DNMT3A is usually acquired before NPM1c in AML; however, occasional cases of AML may acquire the NPM1 first and DNMT3A mutation second. Another interesting observation about mutations that are usually acquired first in clonal evolution is the fact that they are found frequently in related, but distinct, neoplasms. For example, *TET2* mutations are seen in AML, MDS, PV, ET, CMML, systemic mastocytosis (SM), chronic neutrophilic leukaemia (CNL) and even in B- and T-cell lymphomas. This suggests that these mutations provide a fertile ground for 'second' mutations to be acquired that will eventually dictate the resulting phenotype of the cell. In fact recent studies have revealed that mutations in particular genes, most commonly *DNMT3A*, *TET2*, *ASXL1* and *JAK2*, can establish a detectable clone of haemopoietic cells in heatlhy individuals lacking any evidence of haematological cancer. This phenomenon of "clonal haemopoiesis", becomes commoner with age and can be detected in the majority of people aged 90 years or older. Cells with such clone-founder mutations can then acquire additional mutations, such as those affecting *NPM1* and *FLT3* in the case of mutant *DNMT3A*-driven clones, that can generate an acute leukaemic expansion. Treatment of the acute leukaemia may then result in remission, but original clone harboring *DNMT3A* mutations only can persist untouched by chemotherapy and may give rise to late disease relapses.

Impact of clonal structure on treatment and relapse

The parallels between the clonal development of cancer and Darwinian evolution by natural selection are particularly instructive when studying anticancer treatment. Having escaped physiological controls, cancer cells effectively take charge of their own destiny, not unlike a species of individuals doing so in nature. The administration of chemotherapy can then be seen as a serious menace that threatens the 'species' with extinction. Species evade such extinction by leveraging their heterogeneity; in other words they rely on some individuals resisting the onslaught and surviving to regenerate the species. The simultaneous presentation of multiple different hazards to a species increases the likelihood of extinction, as witnessed by the effectiveness of combination antiretroviral therapy against HIV and combination chemotherapy regimens against paediatric ALL.

Nevertheless, only a minority of adult haematological cancers are curable. Clinically, treatment failure can present itself as resistant disease, partial response or relapse, but effectively in all instances this amounts to variable numbers of cancer cells surviving chemotherapy. Advances in genomics have demonstrated that tumours represent an admixture of competing subclones with distinct genotypes, which have different abilities to resist chemotherapies. The different types of treatment failure reflect the numbers of cells resisting therapy, with resistant disease and partial response indicating that the bulk of cells were resistant or partially resistant to treatment. Relapse after complete remission indicates that only a small number of cells survived and this talks directly to the subclonal architecture of cancers.

As first shown in AML and subsequently in multiple myeloma and CLL, sequencing of serial samples at diagnosis and relapse reveals tumour subclones that are present only in the relapsed sampled, only in the diagnostic sample or in both (Figure 18.11). Despite the fact that treatment can eradicate subclones in some cases, mutations representing the founding clone are invariably presented at relapse, often associated with subclones containing new mutations not apparent at diagnosis. This clearly shows how the same treatment can have a differential effect on tumour subclones. From this perspective, chemotherapy can be seen as an "evolutionary bottleneck" that either allows low-frequency chemoresistant tumour cells to survive and reconstitute the bulk of disease at relapse, or promotes evolution of clones through acquisition of new mutations associated with chemoresistance.

Usually, novel mutations identified first in chemoresistant clones can be retrospectively identified in the pretreatment sample when sensitive techniques are used, although in some instances such mutations appear to be unique to the relapsed clone. Arguably this is only of academic significance, as in both instances the disease relapses with mutations associated with chemoresistance. However, there is some evidence that relapsed clones may harbour signatures of mutational processes attributable to chemotherapy, suggesting therefore that if non-mutagenic treatment had been used, relapse may not have ensued. It is not as yet clear if this applies to some or many cases of relapsed haematological cancers.

When trying to optimize treatment against relapsed disease, one can consider that tumours relapsing after a certain latency and without any change in their subclonal structure may be sensitive to the same treatment they responded to previously, whereas samples showing early relapse in association with acquisition of new driver/resistance mutations may require different therapy. Furthermore, the identification of clinically actionable mutations may lead clinicians to favour a particular targeted treatment over another. Finally, large studies correlating the genotype of subclones with treatment response could predict which combination of drugs will be most effective, and thus

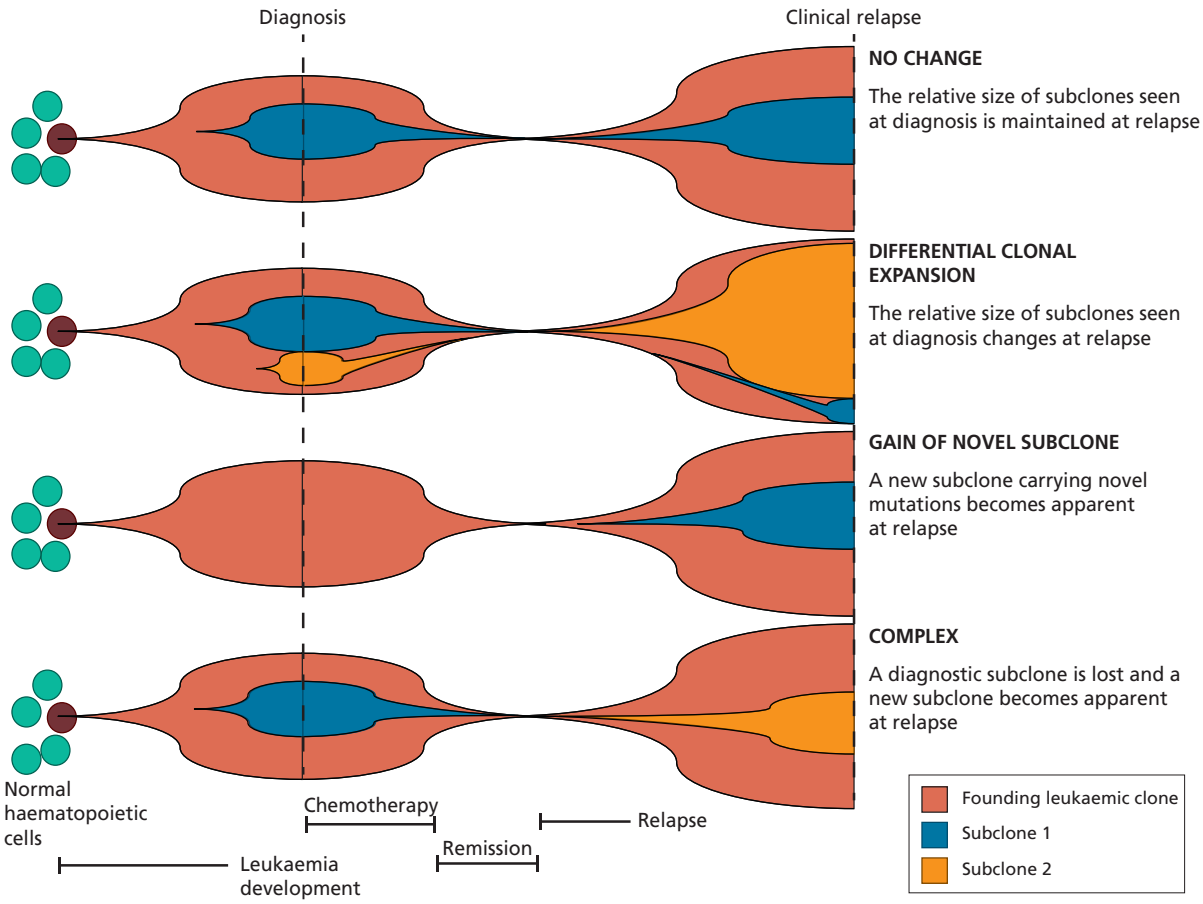


Figure 18.11 Clonal evolution of cancer subclones in relation to treatment.

allow the clinician to deliver a 'tailored' treatment for maximal efficacy to individual patients. Based on its substantial potential for guiding treatment, it is likely that the usage of NGS techniques for diagnosis and prognosis will increase significantly to assist the future management of haematological malignancies.

Selected bibliography

Alexandrov LB, Nik-Zainal S, Wedge DC *et al*. (2013) Signatures of mutational processes in human cancer. *Nature* **500**: 415–21.

Bolli N, Avet-Loiseau H, Wedge DC *et al*. (2014) Heterogeneity of genomic evolution and mutational profiles in multiple myeloma. *Nature Communications* **5**: 2997.

Cancer Genome Atlas Research Network (2013) Genomic and epigenomic landscapes of adult de novo acute myeloid leukemia. *New England Journal of Medicine* **368**: 2059–74.

Conway O'Brien E, Prideaux S, Chevassut T (2014) The epigenetic landscape of acute myeloid leukemia. *Advances in Hematology* **2014**, 103–75.

Ding L, Let TJ, Larson DE *et al*. (2012) Clonal evolution in relapsed acute myeloid leukaemia revealed by whole-genome sequencing. *Nature* **481**: 506–10.

Garraway LA (2013) Genomics-driven oncology: framework for an emerging paradigm. *Journal of Clinical Oncology* **31**: 1806–14.

Genovese G, Kahler A, Handsaker RE *et al*. (2014) Clonal hematopoiesis and blood-cancer risk inferred from blood DNA sequence. *New England journal of medicine* **371**: 2477–2487.

Grove CS and Vassiliou GS (2014) Acute myeloid leukaemia: a paradigm for the clonal evolution of cancer? *Disease Models and Mechanisms* **7**: 941–51.

Jaiswal S, Fontanillas P, Flannick J *et al*. (2014) Age-related clonal hematopoiesis associated with adverse outcomes. *New England journal of medicine* **371**: 2488–2498.

Jones AV, Chase A, Silver RT *et al*. (2009) JAK2 haplotype is a major risk factor for the development of myeloproliferative neoplasms. *Nature Genetics* **41**: 446–9.

Krivtsov AV, Armstrong SA (2007) MLL translocations, histone modifications and leukaemia stem-cell development. *Nature Reviews Cancer* **7**: 823–33.

Kurahashi H, Hara J, Yumura-Yagi K *et al*. (1991) Monoclonal nature of transient abnormal myelopoiesis in Down's syndrome. *Blood* **77**: 1161–3.

Ley TJ, Ding L, Walter MJ *et al*. (2010) DNMT3A mutations in acute myeloid leukemia. *New England Journal of Medicine* **363**: 2424–33.

Loh ML (2011) Recent advances in the pathogenesis and treatment of juvenile myelomonocytic leukaemia. *British Journal of Haematology* **152**: 677–87.

McKerrel T, Park M, Moreno T *et al*. (2015) Leukemia-associated somatic mutations drive distinct patterns of age-related clonal hemopoiesis. *Cell Reports* **10**: 1239–1245.

Morgan GJ, Johnson DC, Weinhold N *et al*. (2014) Inherited genetic susceptibility to multiple myeloma. *Leukemia* **28**: 518–24.

Morgan GJ, Walker BA, Davies FE (2012) The genetic architecture of multiple myeloma. *Nature Reviews Cancer* **12**: 335–48.

Nik-Zainal S, Alexandrov LB, Wedge DC *et al*. (2012) Mutational processes molding the genomes of 21 breast cancers. *Cell* **149**: 979–93.

Ortmann CA, Kent DG, Nangalia J *et al*. (2015) Effect of mutation order on myeloproliferative neoplasms. *New England journal of medicine* **372**: 601–612.

Papaemmanuil E, Rapado I, Li Y *et al*. (2014) RAG-mediated recombination is the predominant driver of oncogenic rearrangement in ETV6-RUNX1 acute lymphoblastic leukemia. *Nature Genetics* **46**: 116–25.

Pasqualucci L, Bhagat G, Jankovic M *et al*. (2008) AID is required for germinal center-derived lymphomagenesis. *Nature Genetics* **40**: 108–12.

Shlush LI, Zandi S, Mitchell A *et al*. (2014) Identification of pre-leukaemic haematopoietic stem cells in acute leukaemia. *Nature* **506**: 328–33.

Steensma DP, Bejar R, Jaiswal S, Lindsley RC, Sekeres MA, Hasserjian RP *et al*. Clonal hematopoiesis of indeterminate potential and its distinction from myelodysplastic syndromes. *Blood* Apr **30**: 1–9.

Stratton MR, Campbell PJ, Futreal PA (2009) The cancer genome. *Nature* **458**: 719–24.

Super HJ, McCabe NR, Thirman MJ *et al*. (1993) Rearrangements of the MLL gene in therapy-related acute myeloid leukemia in patients previously treated with agents targeting DNA-topoisomerase II. *Blood* **82**: 3705–11.

Tomasetti C and Vogelstein B (2015) Cancer etiology. Variation in cancer risk among tissues can be explained by the number of stem cell divisions. *Science* **347**: 78–81.

Van Vlierberghe P, Ferrando A (2012) The molecular basis of T cell acute lymphoblastic leukemia. *Journal of Clinical Investigation* **122**: 3398–406.

Vassiliou GS, Cooper JL, Rad R *et al*. (2011) Mutant nucleophosmin and cooperating pathways drive leukemia initiation and progression in mice. *Nature Genetics* **43**: 470–5.

Vogelstein B, Papadopoulos N, Velculescu VE *et al*. (2013) Cancer genome landscapes. *Science* **339**: 1546–58.

Yang L, Rau R, Goodell MA (2015) DNMT3A in haematological malignancies. *Nat Rev Cancer*.

CHAPTER 19

Laboratory diagnosis of haematological neoplasms

19

Torsten Haferlach[1] and Barbara J Bain[2]

[1]MLL Munich Leukemia Laboratory, Munich, Germany
[2]St Mary's Hospital, London, UK

Introduction

The diagnosis of a haematological neoplasm usually starts from a clinical suspicion, although for chronic leukaemias the diagnosis is sometimes an incidental one. A blood count and blood film is an essential first step whenever a haematological neoplasm is suspected. The next step in the diagnostic process depends on the clinical features and the specific condition that is suspected. In this chapter we discuss the laboratory techniques that are available and how they are integrated into an efficient diagnostic pathway (Figure 19.1). It is of crucial importance that all laboratory investigations are done with an awareness of the medical history and physical findings. It is also essential that a conclusion as to diagnosis is based on integrating the results of all laboratory investigations and imaging in the context of the clinical features. Achieving this can be a challenge if investigations are done in different laboratories on different sites. One possible solution, if all procedures are not performed in a single laboratory, is that a haematologist/haematopathologist should take the lead in making a final diagnosis when the diagnosis depends primarily on the blood and bone marrow and that a haematopathologist/histopathologist should take the lead when the diagnosis depends primarily on a biopsy of another tissue. Three important factors contribute to accurate diagnosis:

1 The provision of very detailed clinical information to laboratories

2 The correct sample, e.g. blood or bone marrow aspirate specimen (appropriately anticoagulated), blood and bone marrow films, trephine or fine-needle biopsy specimen, should all be sent to the laboratory in the shortest possible time

3 The integration of information technology across different sites dealing with diagnostic samples from a single patient.

Consideration should also be given to the need for centralized diagnosis by a panel of experts when experience has shown that misdiagnosis is common, as for example in lymphoma. It is also necessary to ensure that a diagnosis is achieved in a timely manner, particularly when treatment may be urgent, as in acute leukaemias and aggressive lymphomas.

Blood count and blood film

The most important information that must be extracted from the blood count is the white cell count, the haemoglobin concentration and the platelet count. The red cell indices should also be noted since macrocytosis and, less often, microcytosis can occur in haematological neoplasms. The automated instrument will generally provide a differential count and 'flags' indicating the likely presence of blast cells, abnormal lymphocytes or granulocyte precursors. An abnormal total or differential count should always be verified on a blood film and a blood film should also be examined when there are relevant 'flags'. Even if the blood count is normal, a blood film should be examined for the presence of abnormal cells in any patient with lymphadenopathy, splenomegaly, skin infiltration or any other reason to suspect a haematological neoplasm.

The blood film has two major roles. In some circumstances it provides strong evidence of a specific diagnosis that can be confirmed on further testing. In others it suggests a differential diagnosis and indicates the appropriate direction of further

Postgraduate Haematology, Seventh Edition. Edited by A Victor Hoffbrand, Douglas R Higgs, David M Keeling and Atul B Mehta.

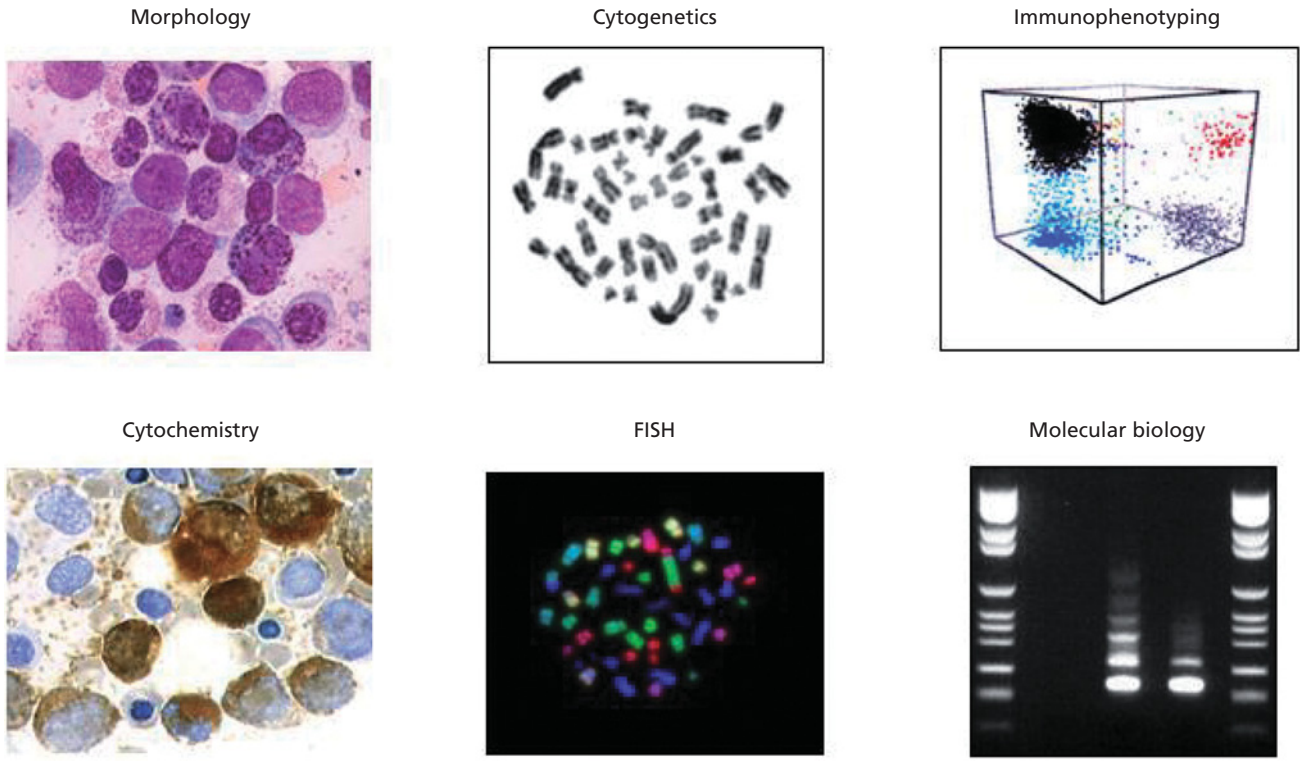


Figure 19.1 The principal methods employed in haematological diagnosis.

tests. A blood film also has the advantage that a rapid provisional diagnosis can be made, something that is important when treatment is urgent. For example, when there are circulating neoplastic cells, a rapid provisional diagnosis of Burkitt lymphoma (Figure 19.2) and acute hypergranular promyelocytic leukaemia (Figure 19.3a) can be made. Acute microgranular/hypogranular promyelocytic leukaemia (Figure 19.3b and 19.4) should also be suspected from the blood film. Even very infrequent neoplastic cells can suggest the correct diagnosis, for example in hairy cell leukaemia. Other diagnoses that may be strongly suggested by the blood film include most subtypes of acute myeloid leukaemia (AML), chronic lymphocytic leukaemia (CLL) (Figure 19.5), chronic myeloid leukaemia (CML) (Figure 19.6), follicular lymphoma, splenic marginal zone lymphoma (when the cytological features of splenic lymphoma with villous lymphocytes are present), large granular lymphocyte leukaemia and plasma cell leukaemia.

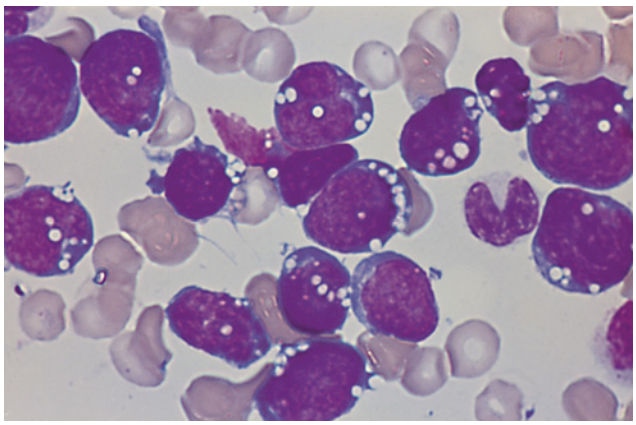
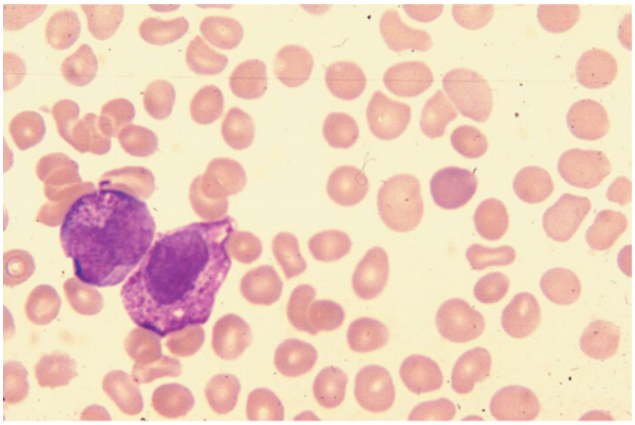


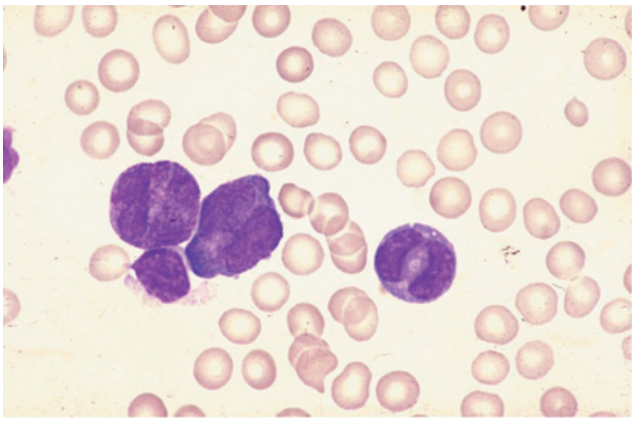
Figure 19.2 Burkitt lymphoma showing blast-like cells with strongly basophilic, heavily vacuolated cytoplasm. The cytological features were once designated 'FAB L3 ALL' but the immunophenotype is that of a mature B cell. Bone marrow film, May–Grünwald–Giemsa (MGG).

Sometimes the blood film suggests only a differential diagnosis. Although the cytological features of the previously used FAB (French–American–British) L1 category of acute lymphoblastic leukaemia (ALL) (Figure 19.7) are very likely to indicate that the diagnosis is ALL, this is not so of FAB L2 type (Figure 19.8), which can be cytologically very similar to AML of FAB M0 or M1 types (Figures 19.9 and 19.10). Similarly, acute monoblastic leukaemia (Figure 19.11) can be cytologically indistinguishable from large-cell lymphoma or even plasma cell leukaemia.

In suspected myeloid malignancy the blood film is important for the detection of myelodysplastic features, particularly hypogranular and pseudo-Pelger–Huët neutrophils (Figure 19.12). Sometimes the blood film shows only non-specific, but nevertheless useful, features, for example eosinophilia, rouleaux formation, red cell agglutinates or cryoglobulin deposition.



(a)



(b)

Figure 19.3 Two patients with acute promyelocytic leukaemia: (a) acute hypergranular promyelocytic leukaemia; (b) hypogranular/microgranular variant of promyelocytic leukaemia showing typical bilobed nuclei. These two subtypes of AML fall into a single category of 'AML with recurrent genetic abnormalities' in the 2008 WHO classification. Peripheral blood films, MGG.

Whatever the findings on the blood count and film, confirmatory tests are needed. The nature of these will vary according to circumstances, but several useful algorithms are guided by information that can be gained from the peripheral blood.

Bone marrow aspirate

A bone marrow aspirate is indicated in virtually all patients with suspected ALL, AML (Figure 19.13), CML, myelodysplastic syndrome (MDS) (Figure 19.14) or multiple myeloma (Figure 19.15). However, a bone marrow aspirate is not necessary for the diagnosis of CLL. The diagnosis of promyelocytic leukaemia may be more readily made on an aspirate than on the peripheral blood since, particularly in the hypergranular form, there may be few circulating leukaemic cells and in the microgranular/

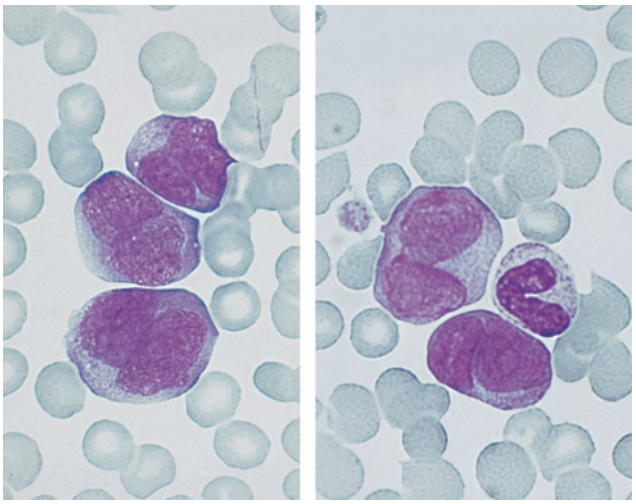


Figure 19.4 Acute hypogranular/microgranular variant of promyelocytic leukaemia showing a range of nuclear forms, but note the typical bilobed nucleus in which the two lobes are joined by an isthmus. Bone marrow film, MGG.

hypogranular form there may be more hypergranular cells in the bone marrow, thus facilitating the diagnosis. The characteristic cytological features of AML with inv(16) or t(16;16) are apparent in the bone marrow (Figures 19.16 and 19.17), but not usually in the peripheral blood. In the acute leukaemias and in CML and other myeloproliferative neoplasms, a bone marrow aspirate provides material for cytogenetic analysis as well as for morphological assessment (Figure 19.18). Successful cytogenetic investigations from peripheral blood are only worth considering if immature leukaemic cells are present or, particularly when investigating suspected myeloproliferative neoplasms, if the white cell count is at least 10×10^9/L, as mature cells will not enter metaphase.

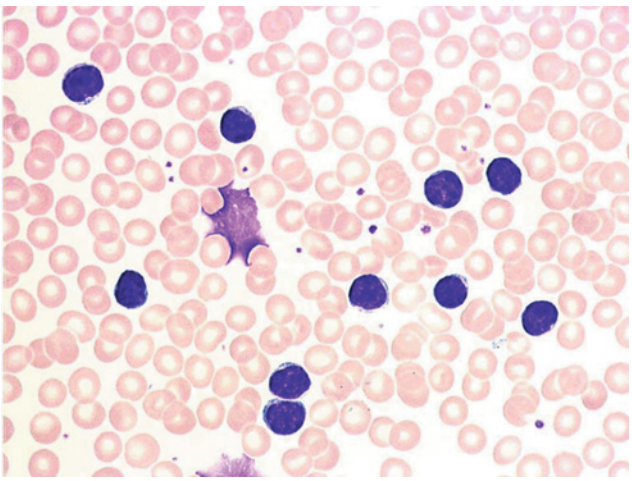


Figure 19.5 Chronic lymphocytic leukaemia showing mature small lymphocytes and a smear cell. Peripheral blood film, MGG.

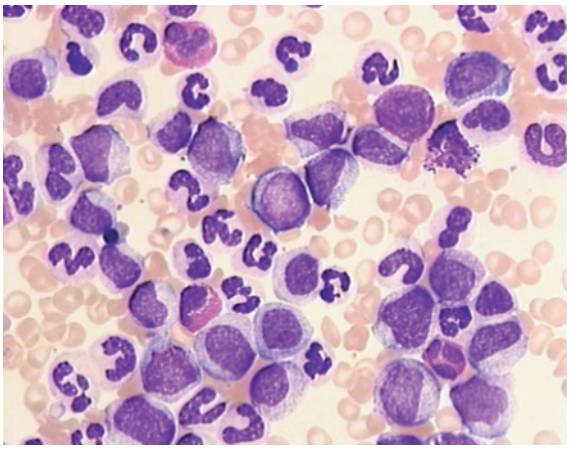


Figure 19.6 Chronic myelogenous leukaemia showing the typical spectrum of granulocytic cells from immature to mature, and including a basophil and some eosinophils.

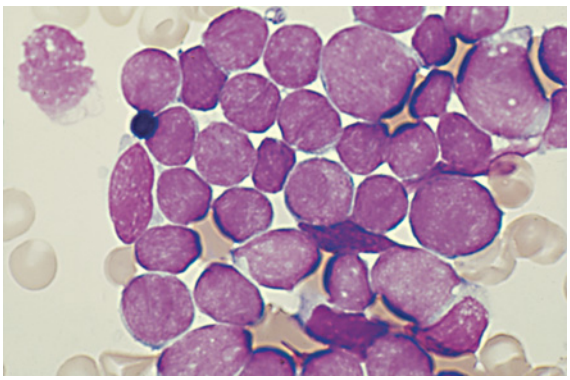


Figure 19.7 Acute lymphoblastic leukaemia FAB L1 subtype showing blast cells that vary in size, but which otherwise have fairly uniform cellular characteristics; the nucleocytoplasmic ratio is high and there are small inconspicuous nucleoli. Bone marrow film, MGG.

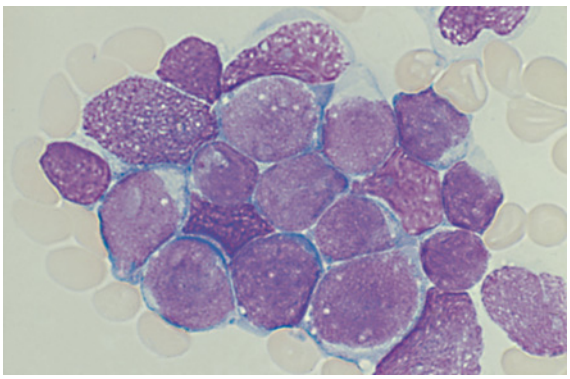


Figure 19.8 Acute lymphoblastic leukaemia FAB L2 subtype showing medium-sized and large pleomorphic blast cells with one or two prominent nucleoli; there are no characteristics that identify the lineage. Bone marrow film, MGG.

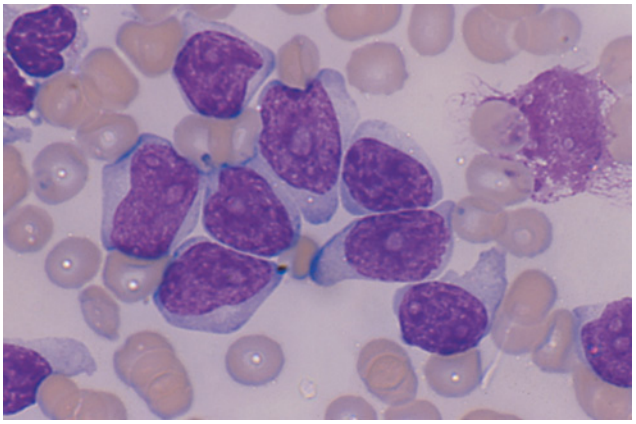


Figure 19.9 Acute myeloid leukaemia FAB M0 subtype showing large blast cells with large prominent nucleoli; there are no cytological features that identify the lineage and Sudan black B and myeloperoxidase cytochemical stains were negative. On immunophenotyping, all B and T lineage markers were negative, but there was expression of CD13 and CD33. Without immunophenotyping such a case could not be distinguished from acute lymphoblastic leukaemia L2 subtype. Bone marrow film, MGG.

Bone marrow aspiration may also be indicated in patients with unexplained cytopenia or a leucoerythroblastic blood film. An aspirate can be useful in lymphoma diagnosis, particularly if accompanied by a trephine biopsy, and can provide suitable material for immunophenotyping and for fluorescence *in situ* hybridization (FISH).

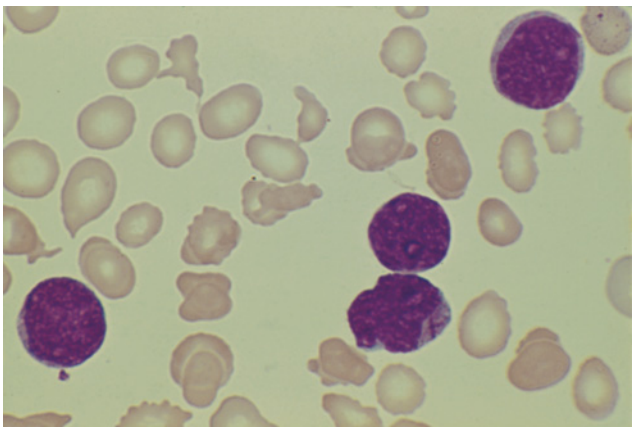


Figure 19.10 Acute myeloid leukaemia FAB M1 subtype showing small- to medium-sized blast cells with a high nucleocytoplasmic ratio; one contains an Auer rod. Sudan black B and myeloperoxidase were positive. In the absence of Auer rods and granules, M1 acute myeloid leukaemia may be indistinguishable from acute lymphoblastic leukaemia without the aid of cytochemistry or immunophenotyping. Peripheral blood film, MGG.

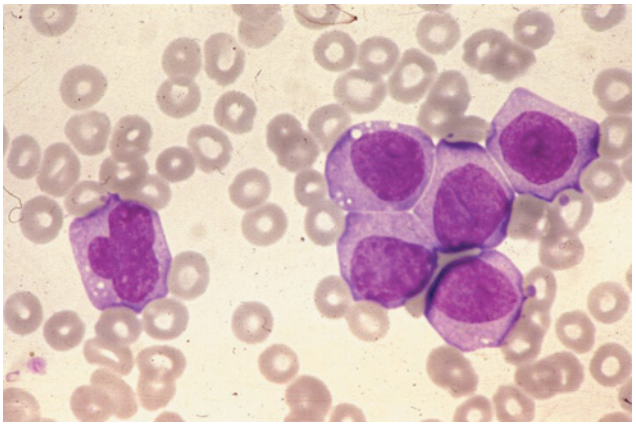


Figure 19.11 Acute monoblastic leukaemia FAB M5a subtype associated with t(9;11)(p22;q23) showing large cells with plentiful cytoplasm; one cell shows nuclear lobulation (defining it as a promonocyte in the WHO classification). In the WHO classification this type of acute myeloid leukaemia is a specific entity within the group 'AML with recurrent genetic abnormalities'. Bone marrow film, MGG.

Bone marrow trephine biopsy

A bone marrow trephine biopsy is indicated whenever there is a 'dry tap' or the aspirate obtained is aparticulate or dilute. A core biopsy is often necessary in hairy cell leukaemia, acute panmyelosis with myelofibrosis, acute megakaryoblastic leukaemia (Figure 19.19) and in patients being investigated for pancytopenia or a leucoerythroblastic blood film. If the bone marrow is being examined for suspected non-Hodgkin lymphoma

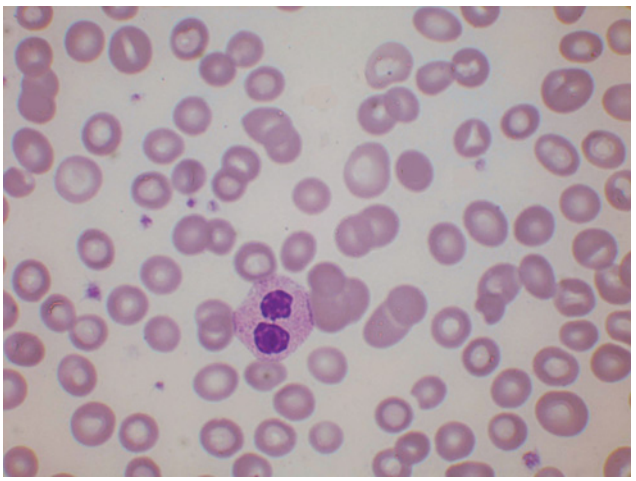


Figure 19.12 A pseudo-Pelger–Huët anomaly in a neutrophil of a patient with myelodysplastic syndrome. Peripheral blood film, MGG.

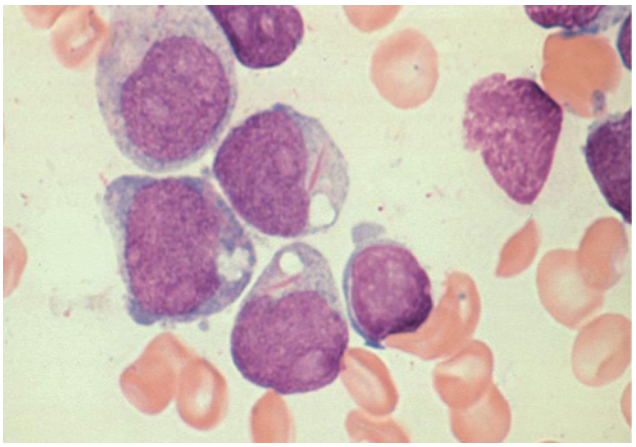


Figure 19.13 Acute myeloid leukaemia showing leukaemic blast cells, several with Auer rods. Bone marrow film, MGG.

(NHL), a trephine biopsy is always indicated since an aspirate may be normal cytologically and immunophenotypically despite there being focal infiltration readily detectable on a core biopsy. Hodgkin lymphoma is very rarely detected in an aspirate, whereas a trephine biopsy specimen may provide the primary diagnostic material, particularly in patients with HIV infection who often present with widespread disease. A trephine biopsy

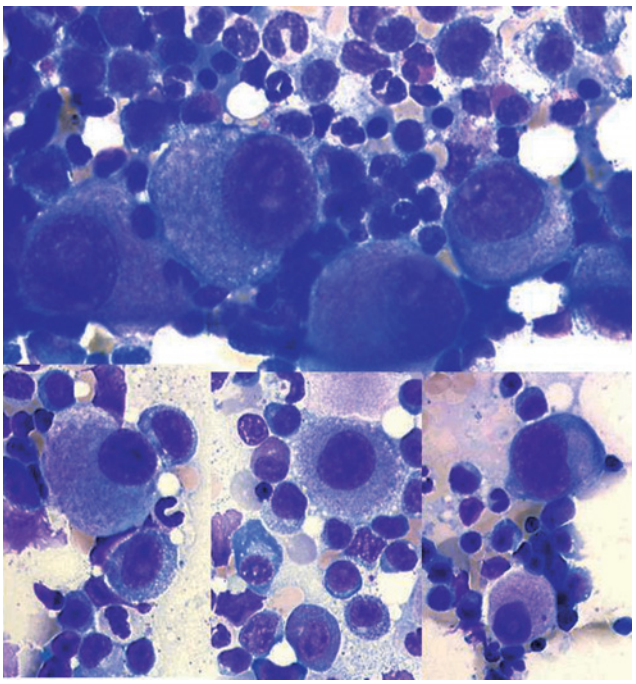


Figure 19.14 Composite image of myelodysplastic syndrome showing the hypolobulated megakaryocytes characteristic of the 5q– syndrome; myelodysplastic syndrome with isolated del(5q) is a specific category in the WHO classification. Bone marrow film, MGG.

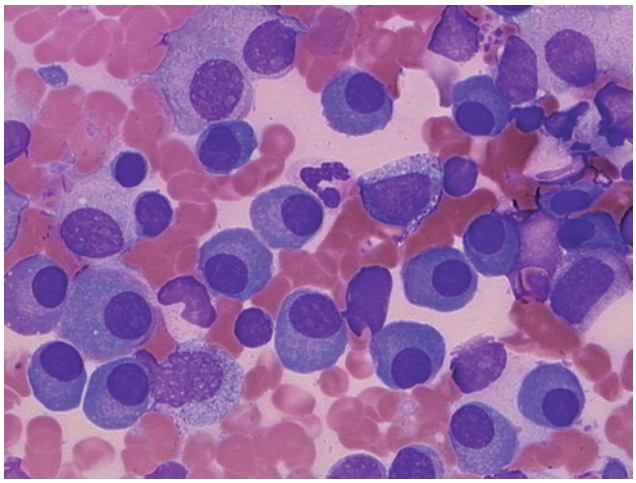


Figure 19.15 Multiple myeloma (plasma cell myeloma) showing virtual effacement of the marrow by myeloma cells. Bone marrow film, MGG.

should be part of the initial investigation of suspected multiple myeloma; sometimes this confirms the diagnosis when the aspirate findings are equivocal and, even when this is not so, it is useful to have a baseline biopsy to compare with post-treatment investigations. A trephine biopsy is not generally necessary in CML or CLL, whereas it is diagnostically important in suspected essential thrombocythaemia and polycythaemia vera and is essential in primary myelofibrosis and systemic mastocytosis.

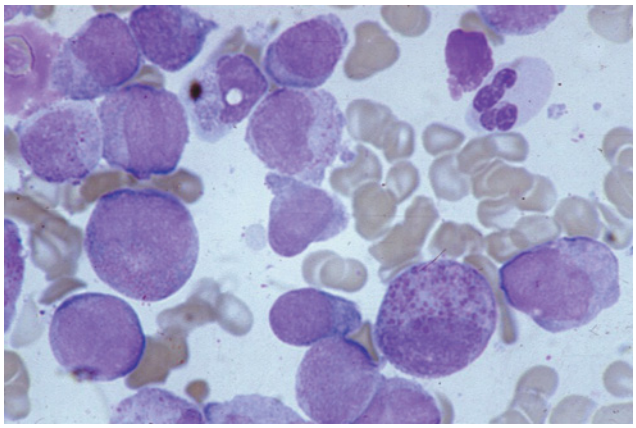


Figure 19.16 Acute myeloid leukaemia (AML) associated with inv(16)(p13.1q22) showing myelomonocytic leukaemia with an abnormal eosinophil precursor that has proeosinophilic (purple-staining) granules. This subtype of AML is often referred to, using an adaptation of FAB terminology, as M4Eo. In the WHO classification this subtype of AML, together with similar cases with t(16;16), constitutes a specific entity within the category 'AML with recurrent genetic abnormalities'. Bone marrow film, MGG.

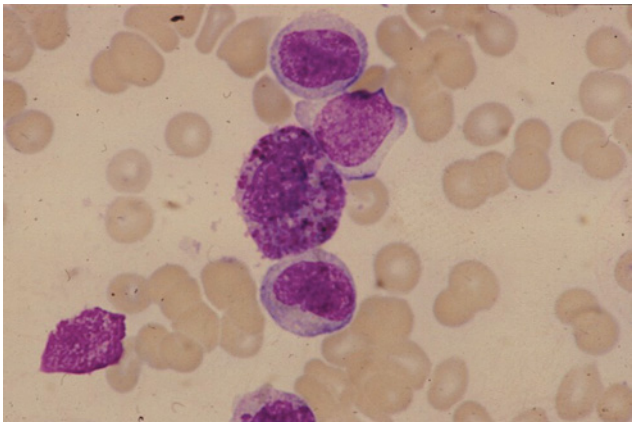


Figure 19.17 Acute myeloid leukaemia (AML) associated with t(16;16)(p13.1;q22) showing blast cells and an abnormal eosinophil precursor with pro-eosinophilic granules. In the WHO classification this subtype of AML, together with similar cases with inv(16), constitutes a specific entity within the category 'AML with recurrent genetic abnormalities'. Bone marrow film, MGG.

There are other indications for trephine biopsy that are more controversial. This procedure may be necessary in ALL or AML if no aspirate or a poor aspirate is obtained. However, if there are numerous circulating neoplastic cells and if a cellular aspirate is obtained, it does not yield any important extra information. In suspected MDS useful extra information that may contribute to the diagnosis is more likely since marrow architecture is often abnormal and immature ($CD34^+$) cells may be detected only on histology or immunohistochemistry.

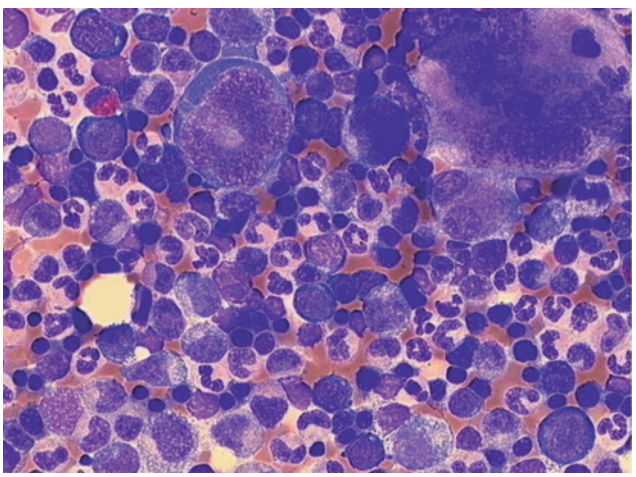


Figure 19.18 Polycythaemia vera with a *JAK2* V617F mutation showing panmyelosis; the large hyperlobulated megakaryocyte (top right) shows emperipolesis. Bone marrow film, MGG.

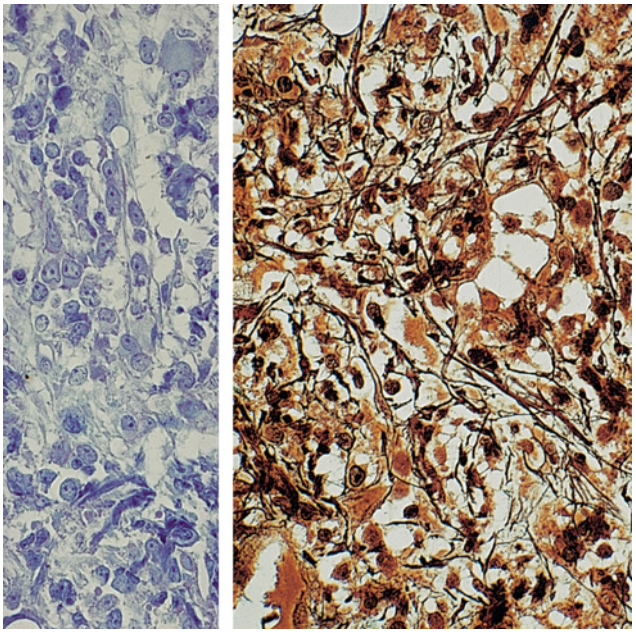


Figure 19.19 Acute megakaryoblastic leukaemia FAB M7 subtype. Bone marrow trephine biopsy shows blast cells (left) and increased reticulin deposition (right).

Cytochemistry

With advances in immunophenotyping and other techniques, the role of cytochemistry in haematological diagnosis has declined considerably. A limited role remains. In acute leukaemia, cytochemistry is indicated if there is limited access to immunophenotyping or if delay in obtaining results is expected. For example, FAB M1 AML (and the 2008 WHO category of AML without maturation) can be distinguished from ALL by means of a Sudan black B or myeloperoxidase stain (Figure 19.20) and acute monoblastic/monocytic leukaemia can be distinguished from large-cell lymphoma by means of a non-specific esterase stain (Figure 19.21). Careful attention to cytological detail usually permits the distinction between FAB M3 variant and M5b AML, but if there is any doubt the former diagnosis is supported by strong reactions for Sudan black B, myeloperoxidase and chloroacetate esterase and the latter by a positive non-specific esterase reaction. Rarely, a specific diagnosis of a subtype of AML is aided by demonstration of blast cells of basophil lineage showing metachromatic staining with a toluidine blue stain (Figure 19.22). Perls stain for iron remains a very important cytochemical stain in suspected MDS (Figure 19.23).

Periodic acid–Schiff and acid phosphatase reactions can help in the diagnosis of ALL (Figures 19.24 and 19.25), but reactions are not specific and we do not recommend their continued use in a routine diagnostic setting if immunophenotyping is available.

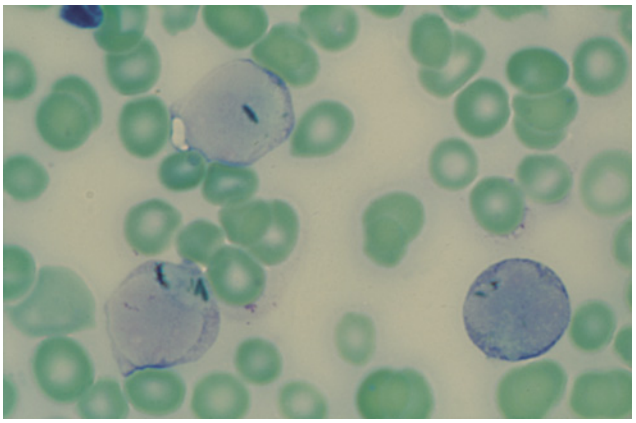


Figure 19.20 Acute myeloid leukaemia (AML) FAB M1 subtype showing Auer rods. This patient had the translocation t(8;21)(q22;q22), identifying a specific entity within the WHO category 'AML with recurrent genetic abnormalities'. Peripheral blood film, myeloperoxidase reaction.

A neutrophil alkaline phosphatase score is likewise redundant in the diagnosis of CML, which now depends on cytogenetic and molecular genetic analysis. A tartrate-resistant acid phosphatase reaction can confirm a diagnosis of hairy cell leukaemia, but is no longer necessary, as long as the immunophenotyping laboratory has an appropriate range of antibodies for this purpose, immunohistochemistry can be performed in the case of a 'dry tap' or *BRAF* mutation analysis can be performed.

Histology

Microscopic examination of sections of tissues other than bone marrow remains an essential technique for the diagnosis of

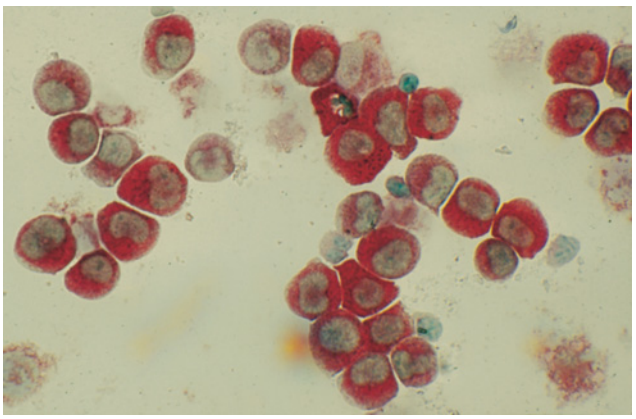
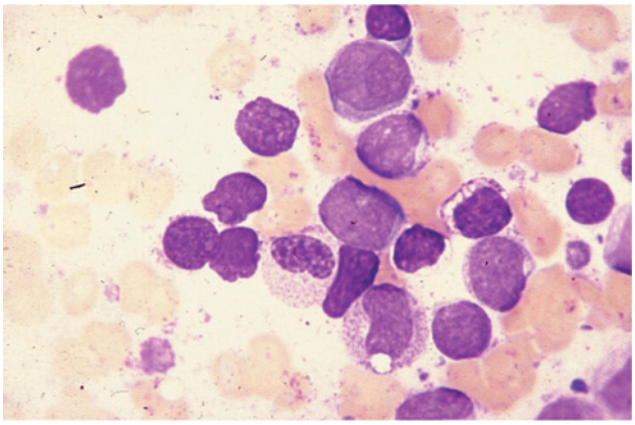
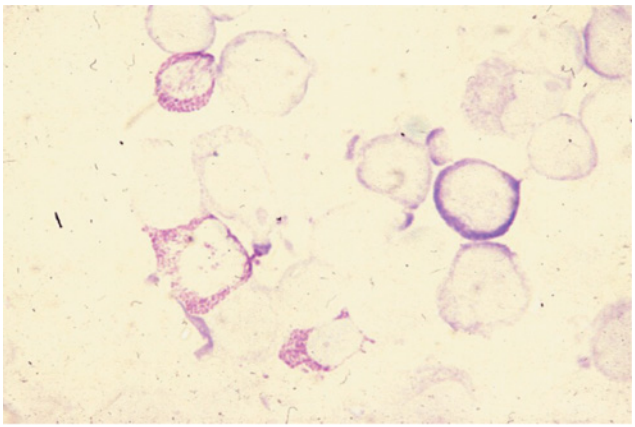


Figure 19.21 Acute monoblastic leukaemia FAB M5a subtype showing strong reaction for non-specific esterase, which confirms the monocytic differentiation. Cytospin preparation, α-naphthyl-acetate esterase.



(a)



(b)

Figure 19.22 Acute myeloid leukaemia (AML) associated with t(6;9)(p23;q34): (a) MGG-stained bone marrow film shows blast cells and two degranulated basophil precursors; (b) toluidine blue-stained bone marrow film demonstrates metachromatic granules and confirms basophilic differentiation. In the 2008 WHO classification, this subtype of AML constitutes a specific entity within the category 'AML with recurrent genetic abnormalities'.

many lymphomas. It may also be necessary for the diagnosis of myeloid sarcoma. Immunohistochemistry (see below) is an essential part of histological assessment for haematological neoplasms.

Flow cytometric immunophenotyping

Immunophenotyping can be based on flow cytometry, immunohistochemistry (on tissue sections) or immunocytochemistry (on cytospin preparations or films). The first two of these techniques are of major importance in haematological diagnosis. Flow cytometry is now carried out using multichannel

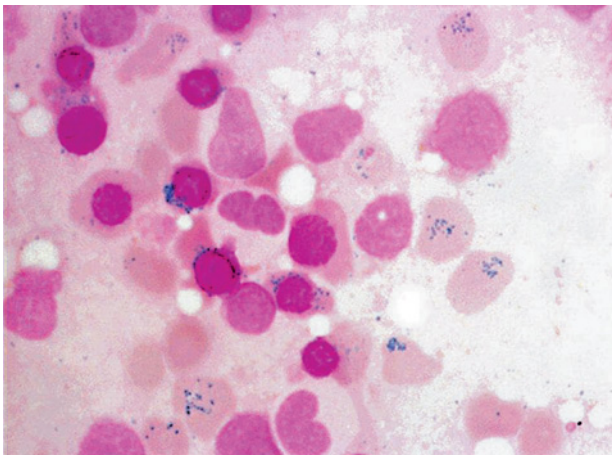


Figure 19.23 Ring sideroblasts and Pappenheimer bodies in erythrocytes in myelodysplastic syndrome. Bone marrow aspirate, Perls stain.

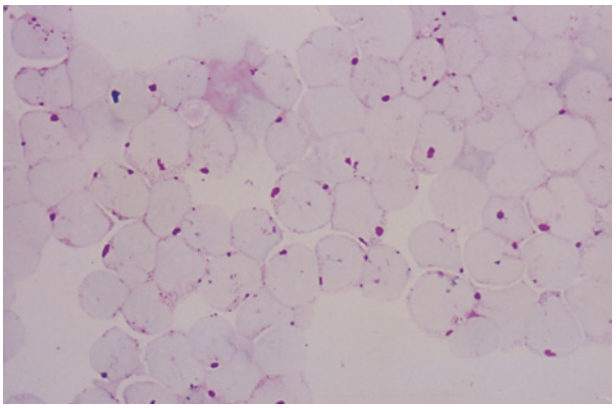


Figure 19.24 Acute lymphoblastic leukaemia showing block positivity to periodic acid–Schiff (PAS) stain. Cytospin preparation of cerebrospinal fluid, PAS stain.

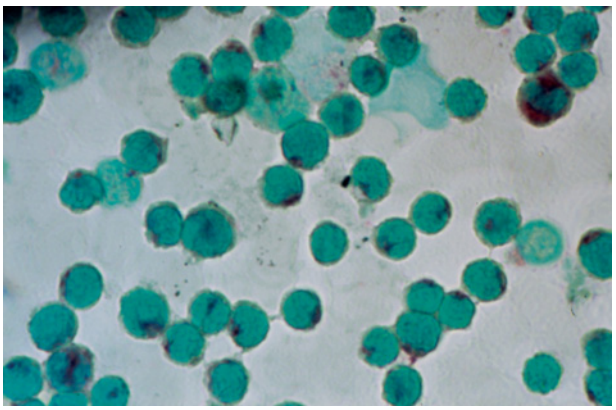


Figure 19.25 T-lineage acute lymphoblastic leukaemia showing focal (Golgi zone) positivity for acid phosphatase in lymphoblasts. Cytospin preparation, acid phosphatase reaction.

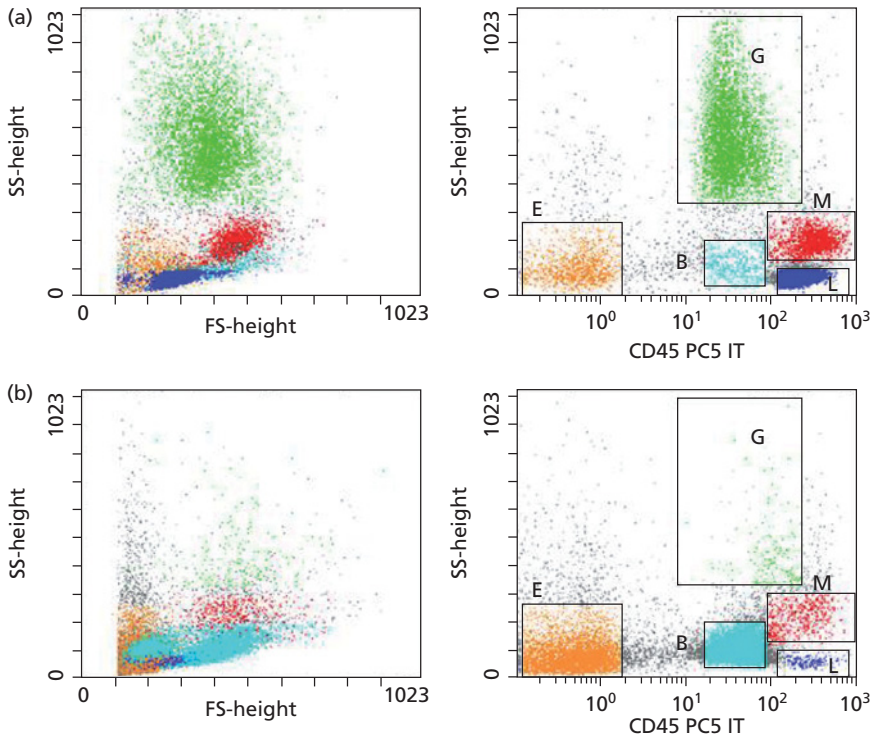


Figure 19.26 Flow cytometry immunophenotyping showing improvement of separation of populations by CD45 and sideways light scatter (SSC) gating. Forward light scatter (FSC) is also shown. (a) Normal bone marrow (left, SSC–FSC plot; right, SSC–CD45 plot); (b) acute myeloid leukaemia bone marrow (left, SSC–FSC plot; right, SSC–CD45 plot). G, granulocytes; M, monocytes; L, lymphocytes; E, erythrocytes; B, blasts. SSC–CD45 gating permits isolation of bone marrow blasts from all other populations, which is not possible by SSC–FSC gating.

instruments that permit the simultaneous assessment of forward light scatter (indicative of cell size), sideways light scatter (indicative of granularity and cell complexity) and the expression of up to 10 surface membrane antigens, all simultaneously in one vial (Figure 19.26). With the addition of techniques to 'permeabilize' the cells, cytoplasmic and nuclear antigens can also be studied. When there are numerous circulating neoplastic cells, immunophenotyping is conveniently done on a peripheral blood sample. Otherwise it can be done on a bone marrow aspirate or on a cell suspension from any infiltrated tissue. The main indications for immunophenotyping are as follows:

1 Confirmation of a diagnosis of ALL and its subclassification and identification of an aberrant immunophenotype (leukaemia-associated immunophenotype) that can be used to monitor minimal residual disease (MRD).

2 Confirmation of a diagnosis of FAB M0 and M7 AML (WHO 2008 classification: AML with minimal differentiation and acute megakaryoblastic leukaemia) and, if not already confirmed by cytochemistry, confirmation of a diagnosis of FAB M5a AML (acute monoblastic leukaemia); support for a diagnosis of acute promyelocytic leukaemia by demonstration of lack of expression of CD34 and especially HLA-DR and by non-specific background autofluorescence observed in the isotype controls; identification of leukaemia-associated immunophenotype that can be used to monitor MRD.

3 Identification of mixed-lineage acute leukaemia.

4 Demonstration of light chain restriction, providing presumptive evidence of monoclonality, in B-lineage disorders.

5 Confirmation of a diagnosis of CLL, assessment of prognosis by quantification of CD38 and ZAP70 expression (both prognostically adverse) and monitoring of MRD.

6 Confirmation of a diagnosis of NHL rather than CLL, and support for specific diagnoses such as follicular lymphoma (CD10 often positive) or mantle cell lymphoma (CD5 and nuclear cyclin D1 positive) (see also Chapter 33).

7 Confirmation of a diagnosis of hairy cell leukaemia by demonstration of expression of all or most of CD11c, CD25, CD103 and CD123 (see also Chapter 27).

8 Confirmation of a diagnosis of multiple myeloma or monoclonal gammopathy of undetermined significance (MGUS) by demonstration that plasma cells (identified by expression of CD138 and strong CD38) have light-chain-restricted cytoplasmic immunoglobulin and often aberrant expression of CD56, together with failure to express CD19 and CD45 (see also Chapter 29).

9 Support for monoclonality of a T-lymphocyte population by demonstration of a uniform aberrant immunophenotype and restricted expression of CD158 (KIR) epitopes and support for specific diagnoses such as T-prolymphocytic leukaemia (CD7 often positive), adult T-cell leukaemia lymphoma (CD25 positive), large granular lymphocyte leukaemia (usually $CD8^+CD4^-$), and hepatosplenic T-cell lymphoma,

which is usually T-cell receptor (TCR)-γδ positive and TCR-αβ negative.

10 Confirmation of a diagnosis of blastic plasmacytoid dendritic cell neoplasm by demonstration of expression of CD4, CD56, CD123 and possibly CD7 in the absence of expression of other lineage-specific markers.

11 Demonstration of an antigen that is a potential target of monoclonal antibody therapy, such as CD20, CD33 or CD52.

12 Diagnosis of paroxysmal nocturnal haemoglobinuria by demonstrating that a subpopulation of neutrophils, monocytes or erythrocytes fails to express antigens, such as CD55 and CD59, that are bound to glycosylphosphatidylinositol, or with a new approach such as proaerolysin conjugated with fluorescein (FLAER).

13 Support for a diagnosis of MDS based on quantification of blasts and detection of aberrant antigen expression.

14 Assessment of infiltration of body fluids (cerebrospinal fluid, pleural effusion) by leukaemia and/or lymphoma.

Immunohistochemistry

Immunohistochemistry for haematological diagnosis is carried out on tissue sections, particularly from lymph nodes or bone marrow trephine biopsy specimens. It is an essential part of the diagnostic procedure in haematological neoplasms, particularly lymphoma, but also myeloid neoplasms. Examples that illustrate its value in lymphoma diagnosis include the following:

1 Confirmation of follicular lymphoma by demonstration that cells forming neoplastic follicles express CD10 and BCL2 (whereas follicles of reactive follicular hyperplasia are BCL2 negative).

2 Confirmation of mantle cell lymphoma by demonstration of CD5-positive B cells with nuclear expression of cyclin D1.

3 Confirmation of a diagnosis of classical Hodgkin lymphoma (CD15 usually positive, CD30 positive) and its distinction from nodular lymphocyte-predominant Hodgkin lymphoma (CD15 and CD30 usually negative and more likely to express B cell-associated antigens such as CD20 and CD79a).

4 Confirmation of Burkitt lymphoma by demonstration of BCL2-negative BCL6-positive B cells with a very high proliferative fraction (approaching 100%) shown using Ki-67 or an equivalent monoclonal antibody.

5 Division of diffuse large B-cell lymphoma (DLBCL) into prognostic groups (see below).

6 Confirmation of hairy cell leukaemia by showing that B cells express CD25, CD72 (DBA44) and annexin A1.

7 Confirmation of clonality in suspected MGUS or multiple myeloma by demonstration of light chain restriction of cytoplasmic immunoglobulin.

8 Confirmation of infiltration in a trephine biopsy specimen in non-Hodgkin and Hodgkin lymphoma (particularly important in patients in whom no neoplastic cells are detected in the bone marrow aspirate and for detection of subtle interstitial or intravascular infiltration in NHL).

9 Confirmation of a diagnosis of hairy cell leukaemia by demonstration of expression of mutant *BRAF*.

DLBCL has been demonstrated by microarray analysis (see below) to be divisible into three major groups with different patterns of gene expression: (i) germinal centre cell; (ii) activated B-cell-like and (iii) other or indeterminate (see also Chapters 31 and 34). Lymphomas showing the gene expression pattern of germinal centre B cells were found to have a better prognosis than activated B-cell-like DLBCL (Chapter 34). It is now possible to assign DLBCL to these two major groups on the basis of immunohistochemistry. The germinal centre group express CD10 or, if they do not, they express BCL6 in the absence of MUM1 expression. DLBCL of the activated B-cell-like group are either negative for both CD10 and BCL6 or are CD10 negative, but express both BCL6 and MUM1. The prognostic significance is seen clearly in patients treated with CHOP (cyclophosphamide, doxorubicin, vincristine and prednisolone), but may no longer reach statistical significance in those treated with R-CHOP (CHOP plus rituximab). In R-CHOP-treated patients, prognostic significance is shown by the demonstration of BCL2 expression, the best prognosis being in those with a germinal centre phenotype, but without BCL2 expression, whose survival curves separate from the other three groups. However, it must be noted that scoring of these immunohistochemical markers is inconsistent between laboratories unless many technical factors are standardized; use of the assignment to germinal centre or non-germinal centre phenotypes to aid treatment decisions would be premature.

In ALL, it is usually possible to perform immunophenotyping on either peripheral blood or a bone marrow aspirate; immunohistochemistry is therefore usually redundant. In AML, immunohistochemistry should be viewed as supplementary to flow cytometric immunophenotyping. It is often unnecessary. However, in the absence of peripheral blood blast cells and if a bone marrow aspirate cannot be obtained, it becomes essential for diagnosis. It can thus be important for the diagnosis of acute megakaryoblastic leukaemias, acute panmyelosis with myelofibrosis and some cases of acute erythroleukaemia, when bone marrow fibrosis may prevent an adequate aspirate being obtained. It is needed for the diagnosis of myeloid sarcoma and blastic plasmacytoid dendritic cell neoplasm (which shows expression of CD4, CD56 and CD123). Immunohistochemistry can be a useful way to demonstrate the presence of an *NPM1* mutation in AML, cytoplasmic rather than nuclear expression being a surrogate marker for these mutations.

Immunohistochemistry for mast cell tryptase is important in confirming a diagnosis of systemic mastocytosis (Chapter 26). Aberrant expression of CD2 and CD25 by the mast cells is also useful (although this pattern of reactivity is not specific for systemic mastocytosis). Immunohistochemistry is much less often needed in other myeloproliferative neoplasms.

Cytogenetic analysis

Cytogenetic analysis refers to study of chromosomes by recognition of their morphology and their characteristic banding pattern, demonstrated by staining with a Giemsa stain (for viewing by light microscopy) (Figure 19.27) or a quinacrine stain (for viewing as fluorescent bands under UV illumination). Chromosomes or specific genes of interest in either interphase or metaphase can also be studied by FISH analysis; this technique is both a cytogenetic and molecular genetic technique since the morphology of the chromosomes can be recognized to some extent, at least in metaphase cells, but the probes used are dependent on hybridization to specific DNA sequences. Cytogenetic, FISH and molecular techniques are complementary and which of these is most useful in an individual patient depends on the suspected diagnosis and the aberrations likely to be present. Classical cytogenetic analysis (chromosome banding analysis) requires cells in metaphase so that for conditions with a low mitotic rate, such as CLL and low-grade NHL, FISH or molecular analysis may be preferred. However, with the newer techniques it is possible to achieve metaphases in more than 90% of cases of CLL using specific culture conditions, including DSP30. Cytogenetic analysis can: (i) provide evidence of clonality and thus confirm that a condition is neoplastic when other evidence is absent or equivocal (e.g. in chronic eosinophilic leukaemia, myeloproliferative neoplasms or NK cell leukaemia/lymphoma), (ii) confirm a specific diagnosis (e.g. CML, AML with recurrent translocations or Burkitt lymphoma) and (iii) give prognostic information (e.g. in CLL, AML, ALL, MDS and multiple myeloma) or indicate that a neoplasm is therapy-induced (t-AML and t-MDS).

It is particularly important that cytogenetic or molecular genetic analysis is performed whenever there is a specific treatment available for patients with a particular genetic abnormality. This includes the following:

1 Conditions sensitive to tyrosine kinase inhibitors (with either *BCR–ABL1,* rearrangement of *PDGFRA* or *PDGFRB,* or *PCM1-JAK2*).

2 Acute promyelocytic leukaemia with *PML–RARA* (sensitive to all-*trans* retinoic acid and As_2O_3).

3 MDS with 5q– (sensitive to lenalidomide).

4 MDS with monosomy 7 (sensitive to azacitidine).

5 Burkitt lymphoma with juxtaposition of *MYC* to the immunoglobulin heavy chain, κ or λ locus (responsive to intensive chemotherapy incorporating certain specific agents plus monoclonal antibody treatment).

Cytogenetic analysis can be considered appropriate whenever it contributes to diagnosis, indicates the prognosis or guides the selection of treatment. It can therefore be considered indicated in the following:

1 All cases of AML: essential for application of the WHO classification, indicates prognosis and is thus relevant to treatment choice (Table 19.1).

2 All cases of ALL: essential for application of the WHO classification, indicates prognosis and is thus relevant to treatment choice; necessary for detection of *BCR–ABL1*-positive cases, although for this purpose FISH and molecular analyses are alternatives (Table 19.2).

3 All cases of mixed-phenotype acute leukaemia: essential for detecting *BCR–ABL1*-positive cases.

4 All cases of suspected CML: essential to confirm *BCR–ABL1*, although FISH or molecular analyses are alternatives.

5 All cases of suspected MDS: necessary for application of the WHO classification, including identification of the 5q– syndrome (Figure 19.28) and essential for application of the revised International Prognostic Scoring System (IPSS-R) and WHO classification-based Prognostic Scoring System (WPSS).

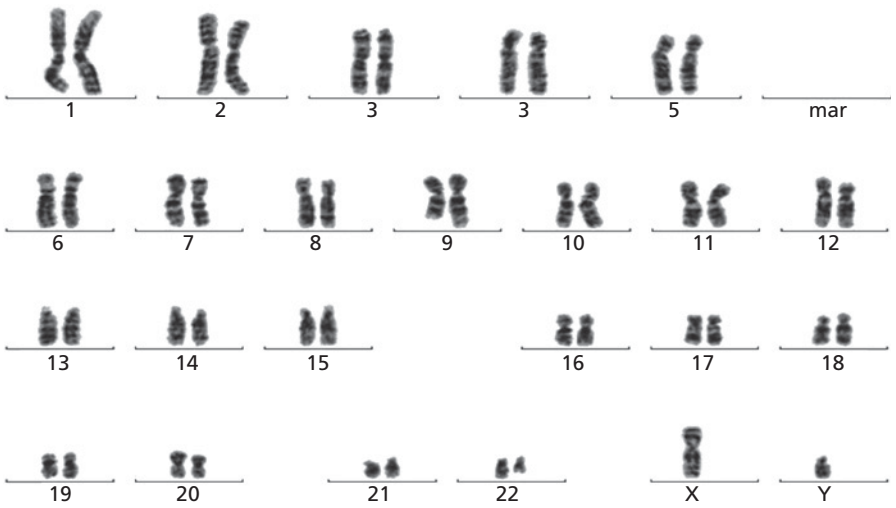


Figure 19.27 A karyogram of Giemsa-stained chromosomes showing t(9;22)(q34;q11.2). The derivative chromosome 22 is the Philadelphia chromosome.

Table 19.1 The 2008 WHO classification of *de novo* acute myeloid leukaemia (AML), including categories where genetic analysis is essential for categorization shown in bold.

AML with recurrent genetic abnormalities*
AML with t(8;21)(q22;q22); *RUNX1–RUNX1T1*
AML with inv(16)(p13.1q22) or t(16;16)(p13.1;q22); *CBFB–MYH11*
AML with t(15;17)(q22;q12); *PML–RARA*
AML with t(9;11)(p22;q23); *MLLT3–MLL*(*KMT2A*)
AML with t(6;9)(p23;q34); *DEK–NUP214*
AML with inv(3)(q21q26.2) or t(3;3)(q21;q26.2); *RPN1–EVI1*
AML with t(1;22)(p13;q13); *RBM15–MKL1*
Provisional entity: AML with mutated *NPM1*
Provisional entity: AML with mutated *CEBPA*
AML with myelodysplasia-related changes†
AML not otherwise categorized
Myeloid sarcoma
Myeloid proliferation related to Down syndrome‡
Transient abnormal myelopoiesis
Myeloid leukaemia associated with Down syndrome
Blastic plasmacytoid dendritic cell neoplasm

*Either cytogenetic or molecular genetic analysis essential for classification.
†Meeting cytogenetic criteria for myelodysplastic syndrome (MDS)-associated cytogenetic abnormalities permits assignment of a case to this category (assignment can also be on the basis of multilineage dysplasia or a previous diagnosis of MDS).
‡Cytogenetic analysis may be necessary to confirm the diagnosis of Down syndrome, particularly in mosaic Down syndrome; an acquired *GATA1* mutation is also present.

6 All cases of suspected chronic eosinophilic leukaemia: necessary for detection of translocations that lead to *PDGFRB* rearrangement (imatinib sensitive) and *FGFR1* rearrangement.
7 All cases of suspected NK leukaemia/lymphoma: can demonstrate clonality and thus confirm the diagnosis.

If resources permit, cytogenetic analysis can also be useful in other myeloproliferative and myelodysplastic/myeloproliferative neoplasms. Cytogenetic analysis can be used for the confirmation of a diagnosis of Burkitt lymphoma, follicular lymphoma or mantle cell lymphoma, but FISH may be preferred. Use of the two techniques may give complementary information. Prognostically relevant information is also obtained in multiple myeloma. However, this is best done with preliminary magnetic activation cell sorting using CD138, CD269 or CD319 monoclonal antibodies, as plasma cells numbers may be low and FISH performed on direct bone marrow films may result in insufficient cells for analysis and give misleading results.

Table 19.2 The WHO classification of acute lymphoblastic leukaemia, with categories where genetic analysis is essential for categorization shown in bold (see also Chapter 21 and 22).

B lymphoblastic leukaemia/lymphoma
B lymphoblastic leukaemia/lymphoma, not otherwise specified
B lymphoblastic leukaemia/lymphoma with recurrent genetic abnormalities
With t(9;22)(q34;q11.2) and *BCR–ABL1*
With t(4;11)(q21;q23) and *MLL*(*KMT2A*)*–MLLT2* or other 11q23 and *MLL* rearrangement
With t(12;21)(p13;q22) and *ETV6–RUNX1**
With hyperdiploidy (>50 chromosomes)
With hypodiploidy (<46 chromosomes)
With t(5;14)(q31;q32) and *IL3–IGH*
With t(1;19)(q23;p13.3) and *TCF3–PBX1*
T lymphoblastic leukaemia/lymphoma

*Molecular genetic analysis is needed since the translocation is cryptic.

Fluorescence *in situ* hybridization

FISH uses labelled oligonucleotide probes that bind to specific DNA sequences. These may be locus-specific probes (including those detecting oncogenes and tumour-suppressor genes), centromeric probes, telomeric probes and whole chromosome paints. Centromeric probes can be used for the detection of monosomies and trisomies. Otherwise, locus-specific probes are those most widely used. By labelling two probes with different fluorochromes it is possible to study two genes that are involved in a specific translocation or other rearrangement. If the probes span the breakpoint, both will be disrupted by the translocation and signals will be adjacent or optically fused on the derivative chromosomes, a double-colour double-fusion technique. Often one probe gives a green signal and the other a red signal so that the fusion signal is yellow. Alternatively, two probes that span the 5′ locus and 3′ locus and overlap can be used in a double-colour break-apart technique; this can be useful when a single gene, such as *MLL* (*KMT2A*), is involved in rearrangements with a large number of other genes.

FISH is applicable to both metaphase cells and interphase cells (Figure 19.29). FISH is particularly useful in the following circumstances:
1 Screening for *BCR–ABL1* fusion (peripheral blood cells are suitable, whereas cytogenetic analysis is better done on bone marrow) (Figure 19.30).
2 Confirming a specific lymphoma diagnosis by demonstration of *BCL2–IGH* in follicular lymphoma, *CCND1–IGHG1* in mantle cell lymphoma or *MYC–IGHG1* in Burkitt lymphoma (Table 19.3).
3 Detection of translocations or inversions or aneuploidy that was not detected on metaphase cytogenetic analyses because

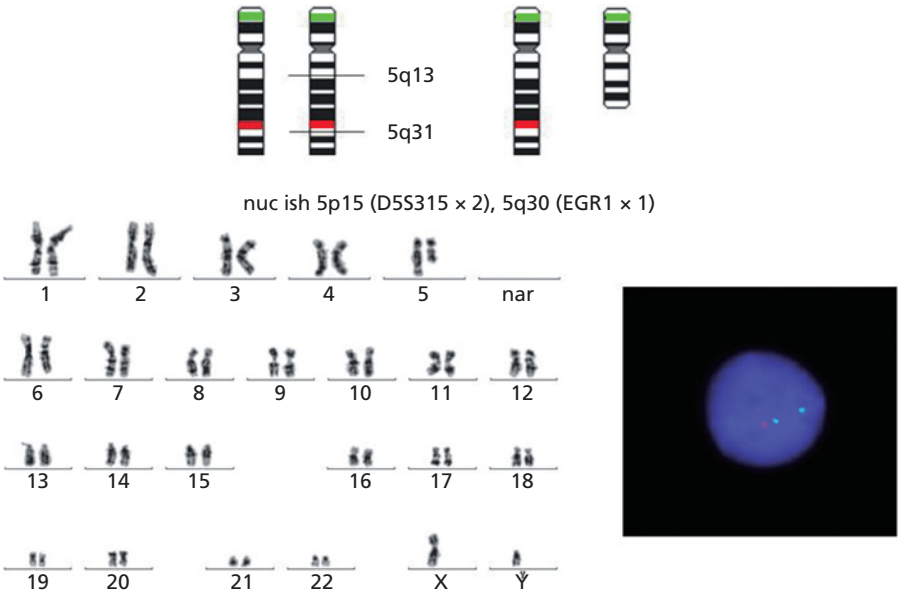


Figure 19.28 An explanatory diagram (top), karyogram of Giemsa-stained chromosomes showing del(5)(q13q31) (bottom left), and FISH showing retention of signal for 5p15 and loss of signal for 5q31 (bottom right) in a patient with 5q– syndrome.

the leukaemic cells did not enter mitosis or the metaphases were of too poor quality, e.g. detection of t(8;21)(q22;q22), t(15;17)(q22;q12), inv(16)(p13.1q22) or *MLL* rearrangement in AML (Figure 19.31) or detection, in ALL, of iAMP21 (using a *RUNX1* probe) or high hyperdiploidy (by using centromeric probes for the chromosomes that are most often triplicated).

4 Screening for a cryptic *FIP1L1–PDGFRA* fusion (by detection of loss of *CHIC2*) in chronic eosinophilic leukaemia.

5 Producing prognostically relevant information in CLL by investigation for trisomy 12, del(6)(q21), del(11)(q22–23), del(13)(q14.3) (Figure 19.32) and del(17)(p13).

Spectral karyotyping, also known as 24-colour FISH, is a modification of FISH that permits visualization of each pair of chromosomes in a unique colour. It can be used as a supplement to metaphase cytogenetics to clarify the nature of complex rearrangements (Figure 19.33). Inversions are not detected.

Molecular genetic analysis

Molecular genetic analysis includes Southern blotting (now little used in routine diagnosis), the polymerase chain reaction (PCR) to study genomic DNA (Figures 19.34 and 19.35) and reverse transcriptase polymerase chain reaction (RT-PCR) to study RNA after its reverse transcription. A PCR reaction can

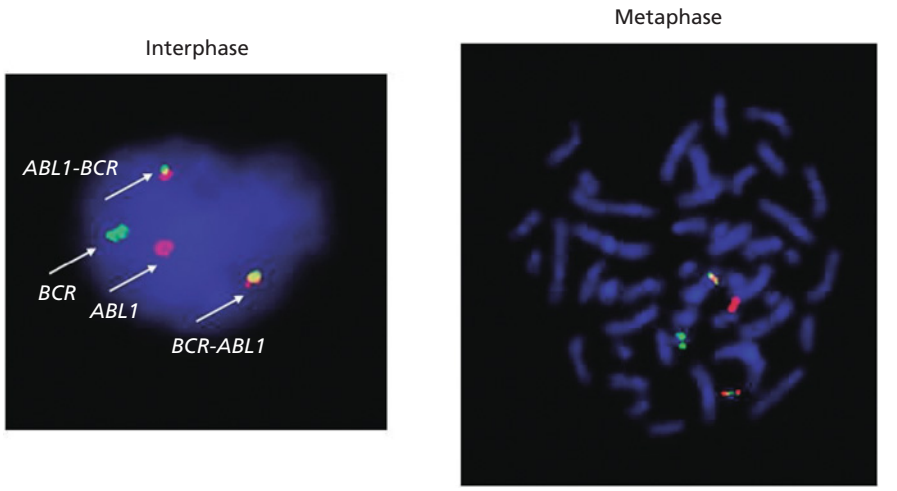


Figure 19.29 Locus-specific probes for *BCR* (green) and *ABL1* (red) applied to an interphase cell and to a metaphase of a patient with chronic myelogenous leukaemia after hybridization with probes for *ABL1* (red) and *BCR* (green) showing *BCR–ABL1* co-localization signals (yellow) on the Philadelphia chromosome and *ABL1–BCR* co-localization signals on derivative chromosome 9.

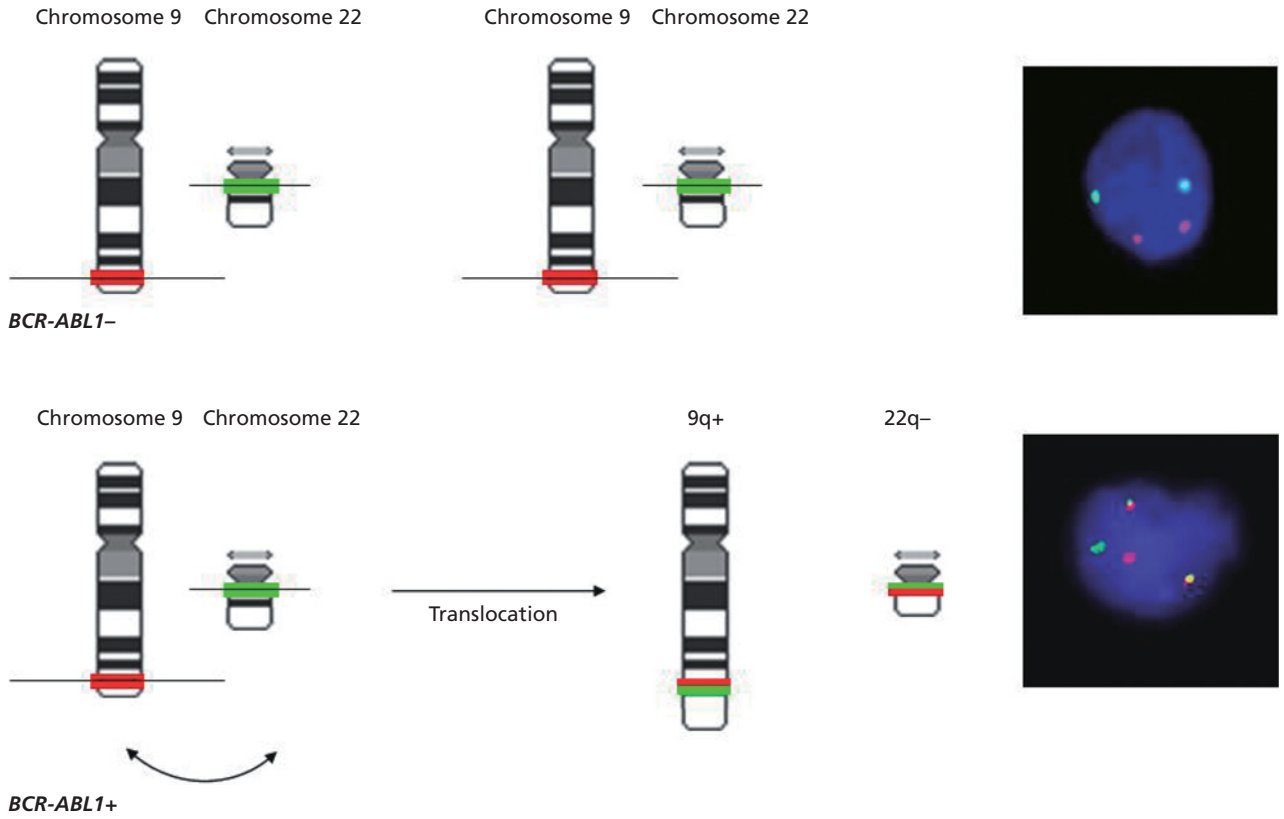


Figure 19.30 Diagram showing the principle of fluorescence *in situ* hybridization for detection of *BCR–ABL1*. Locus-specific probes for *BCR* (green) and *ABL1* (red) have been applied to interphase cells. The normal cell (top) shows two red signals and two green signals, whereas the cell with *BCR–ABL1* shows two normal signals and two fusion signals (*BCR–ABL1* and *ABL1–BCR*).

Table 19.3 Cytogenetic/molecular genetic abnormalities that are incorporated into the 2008 WHO classification of neoplasms of mature T and B cells.

Genetic abnormality	Gene dysregulated	Diagnosis*
t(14;18)(q32;q21) or t(2;18)(p12;q21) or t(18;22)(q21;q11.2)	*BCL2*	Follicular lymphoma
t(11;14)(q13;q32) or, rarely, t(11;22)(q13;q11)	*CCND1*	Mantle cell lymphoma
t(8;14)(q24;q32) or t(2;8)(p12;q24) or t(8;22)(q24;q11.2)	*MYC*	Burkitt lymphoma
inv(14)(q11q32) or t(14;14)(q11;q32) or, less often, t(X;14)(q28;q11)	*TCL1*	T-prolymphocytic leukaemia
t(2;5)(p23;q35) or one of at least five variant translocations with a 2p23 breakpoint	*NPM1–ALK* fusion or other fusion gene incorporating part of *ALK*	Anaplastic large-cell lymphoma, ALK positive

*The genetic abnormality confirms the diagnosis only in an appropriate cytological/histological and immunophenotypic setting.

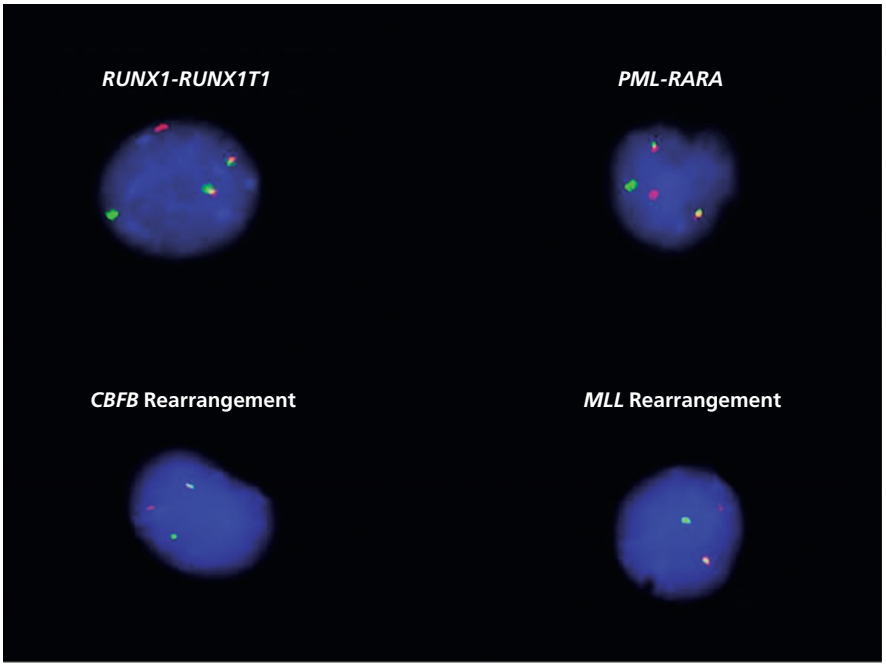


Figure 19.31 Fluorescence *in situ* hybridization in four different patients with acute myeloid leukaemia showing four major cytogenetic/genetic categories of AML: *RUNX1–RUNX1T1* fusion indicative of t(8;21) (top left); *PML–RARA* fusion indicative of t(15;17) (top right); *CBFB* rearrangement, which is likely to indicate *CBFB–MYH11* (bottom left); *MLL* rearrangement (bottom right).

also be made quantitative, as in real-time quantitative PCR, an important technique for monitoring MRD (Figure 19.36). PCR and RT-PCR can be alternatives to metaphase cytogenetic analysis in several subgroups of leukaemia and lymphoma. However, unlike metaphase cytogenetic analysis, but similar to FISH, this is a targeted investigation so that only the abnormality that is specifically sought will be detected. Thus, in some diseases additional diagnostic or prognostic information will be found only by using metaphase cytogenetics. On the other hand, PCR has the advantage that only a small amount of DNA is needed and there is no need for dividing cells. Molecular analysis permits the detection of prognostically and therapeutically relevant cryptic rearrangements, such as the *ETV6–RUNX1* rearrangement associated with a cryptic t(12;21)(p13;q22) in ALL. Genomic PCR can be applied to stored samples.

The roles of genomic PCR and RT-PCR include the following:
1 Detection of rearrangement of immunoglobulin heavy and light chain loci and TCR loci, providing evidence of a clonal disorder if a monoclonal rather than an oligoclonal or polyclonal pattern is detected.

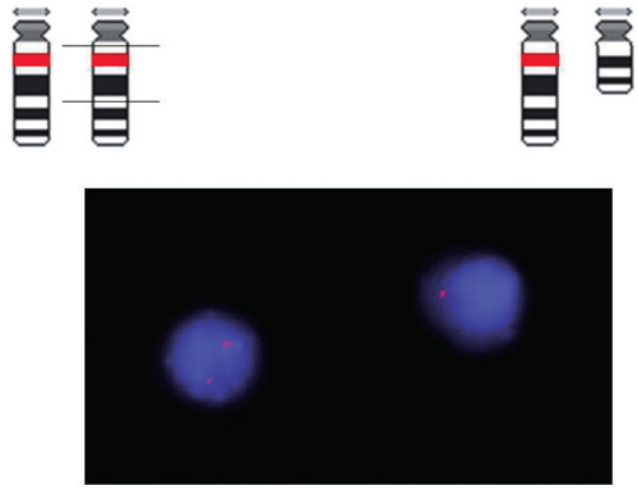


Figure 19.32 Explanatory diagram and fluorescence *in situ* hybridization showing interstitial deletion of 13q14.3 in a patient with chronic lymphocytic leukaemia.

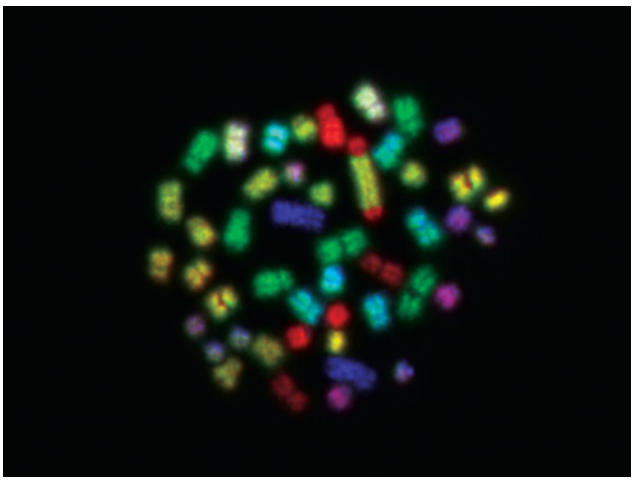


Figure 19.33 Metaphase of a patient with acute myeloid leukaemia and a complex aberrant karyotype after 24-colour fluorescence *in situ* hybridization.

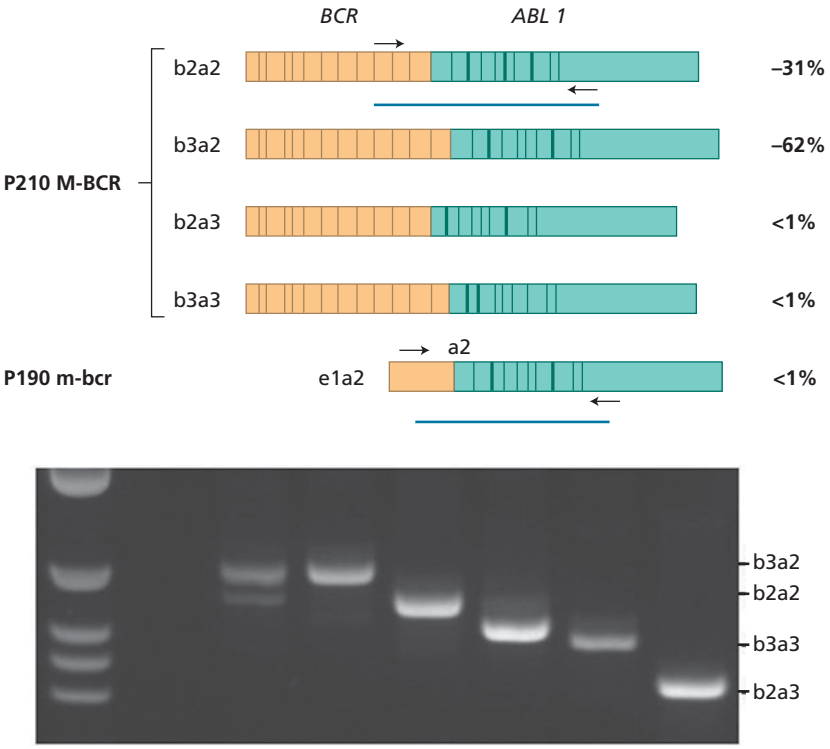


Figure 19.34 An explanatory diagram and a PCR gel showing the various *BCR–ABL1* transcripts that can be identified by PCR.

2 Detection of leukaemia-related and lymphoma-related chromosomal rearrangements by demonstration of gene juxtaposition or fusion in order to confirm a diagnosis, for example detection of *CCND1–IGHG1* indicative of t(11;14)(q13;q32) to confirm suspected mantle cell lymphoma or detection of *BCL2–IGH* indicative of t(14;18)(q32;q21) to confirm suspected follicular lymphoma. Note that neither of these rearrangements is absolutely specific for the disease in question since *CCND1–IGHG1* also occurs in multiple myeloma and *BCL2–IGH* also occurs in DLBCL, but in the context of appropriate cytology or histology they do permit confirmation of a diagnosis.

3 Detection of gene mutations relevant to diagnosis, for example *KIT* D816V in systemic mastocytosis or either *JAK2* V617F or *JAK2* exon 12 mutation in polycythaemia vera. *JAK2* V617F also occurs in many patients with essential thrombocythaemia or primary myelofibrosis, but its detection in a patient with a high haemoglobin concentration confirms the diagnosis of polycythaemia vera and since it is present in the great majority of patients with this diagnosis, measurement of the total red cell mass and plasma volume is generally rendered unnecessary. Patients with suspected essential thrombocythaemia or primary myelofibrosis who are *JAK2* V617F-negative should be further investigated for *CALR* or *MPL* mutation.

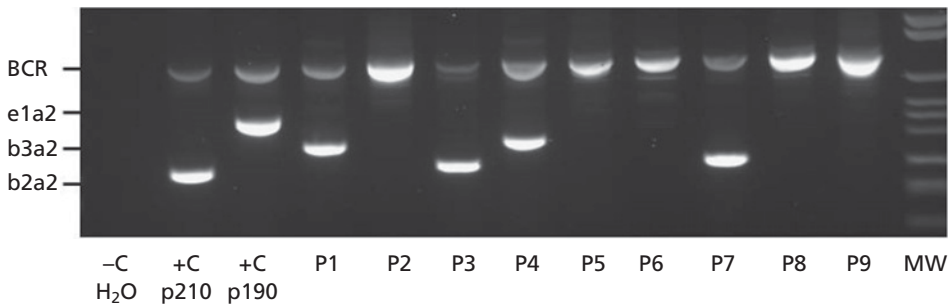


Figure 19.35 *BCR–ABL1* multiplex PCR containing different oligonucleotide ('oligo') combinations. –C: negative control sample containing water instead of a nucleic acid template; +C: positive controls for p210 and p190 fusion types; P: patient samples, P1 and P4 are positive for b3a2 fusion types, P3 and P7 are positive for b2a2 fusion types, P2, P5, P6, P8 and P9 are *BCR–ABL1*negative patient samples; MW, molecular weight standard; BCR, amplification of the non-rearranged *BCR* gene as internal quality control reaction.

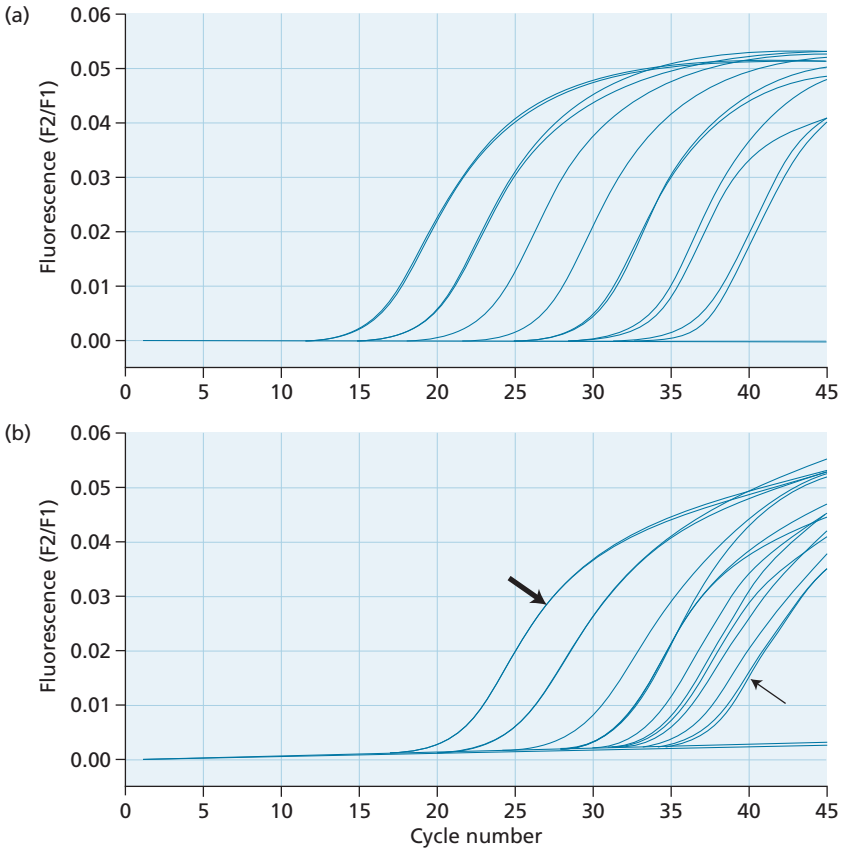


Figure 19.36 Real-time quantitative PCR. Fluorescence is plotted against PCR cycle number. (a) A 10-fold limited dilution series of a *BCR–ABL1*-containing plasmid showing high sensitivity of the PCR. It also serves as a calibration curve for subsequent *BCR–ABL1* calculations. (b) Run with chronic myeloid leukaemia follow-up samples: the large arrow shows a sample with high *BCR–ABL1* load, the small arrow low *BCR–ABL1* load.

4 Prognostic stratification, for example in ALL detection of *ETV6–RUNX1* associated with cryptic t(12;21)(p13;q22) (good prognosis), *TCF3–PBX1* associated with t(1;19)(q23;p13.3) (good prognosis), *MLL–MLLT2* associated with t(4;11)(q21;q23) (poor prognosis) and *BCR–ABL1* associated with t(9;22)(q34;q11.2) (poor prognosis); multiplex PCR, combining a number of primer pairs in a single reaction, is useful for screening simultaneously for more than one rearrangement.

5 Prognostic stratification in AML with normal karyotype, for example *FLT3* internal tandem duplication (ITD) (poor prognosis), *MLL* partial tandem duplication (PTD) (poor prognosis), *NPM1* mutation (good prognosis if not associated with *FLT3* ITD), *CEBPA* mutation (good prognosis if biallelic and not associated with *FLT3* ITD) (see also Chapter 20).

6 Prognostic stratification in CLL: somatic hypermutation, defined as *IGHV* showing less than 98% homology with germline, is found in 50–60% of all cases and correlates with a better prognosis than is found in patients with unmutated *IGHV* genes. *TP53* mutations are prognostically important and guide treatment.

7 MRD detection using *RUNX1–RUNX1T1*, *CBFB–MYH11*, *PML–RARA*, *MLL* rearrangement, *MLL*-PTD, *FLT3*-ITD, *NPM1* mutation or *CEBPA* mutation.

Whole-genome scanning

There are a number of molecular techniques available for whole-genome scanning. Thus far, these have mainly been used to address research topics, but they are now increasingly finding a place in routine diagnostic work. These techniques include comparative genomic hybridization (CGH), in which labelled test DNA is hybridized to normal metaphase preparations, and microarray analysis, in which there is hybridization of the labelled test DNA of interest to probes positioned in high density on a microarray surface. Microarray analysis includes CGH arrays and, more frequently used, genome-wide single-nucleotide polymorphism (SNP) arrays. CGH is labour-intensive, less sensitive for the detection of small regions harbouring copy number alterations and likely to remain a research technique. Microarray analysis has a greater potential for diagnostic application since it has a much higher resolution for the detection of genomic aberrations, requires a low amount of starting material and has the potential to be developed as a fully automated molecular laboratory assay.

In CGH assays, test and reference DNAs are labelled with different fluorochromes before being hybridized to microarrays of either oligonucleotides or bacterial artificial chromosome (BAC)

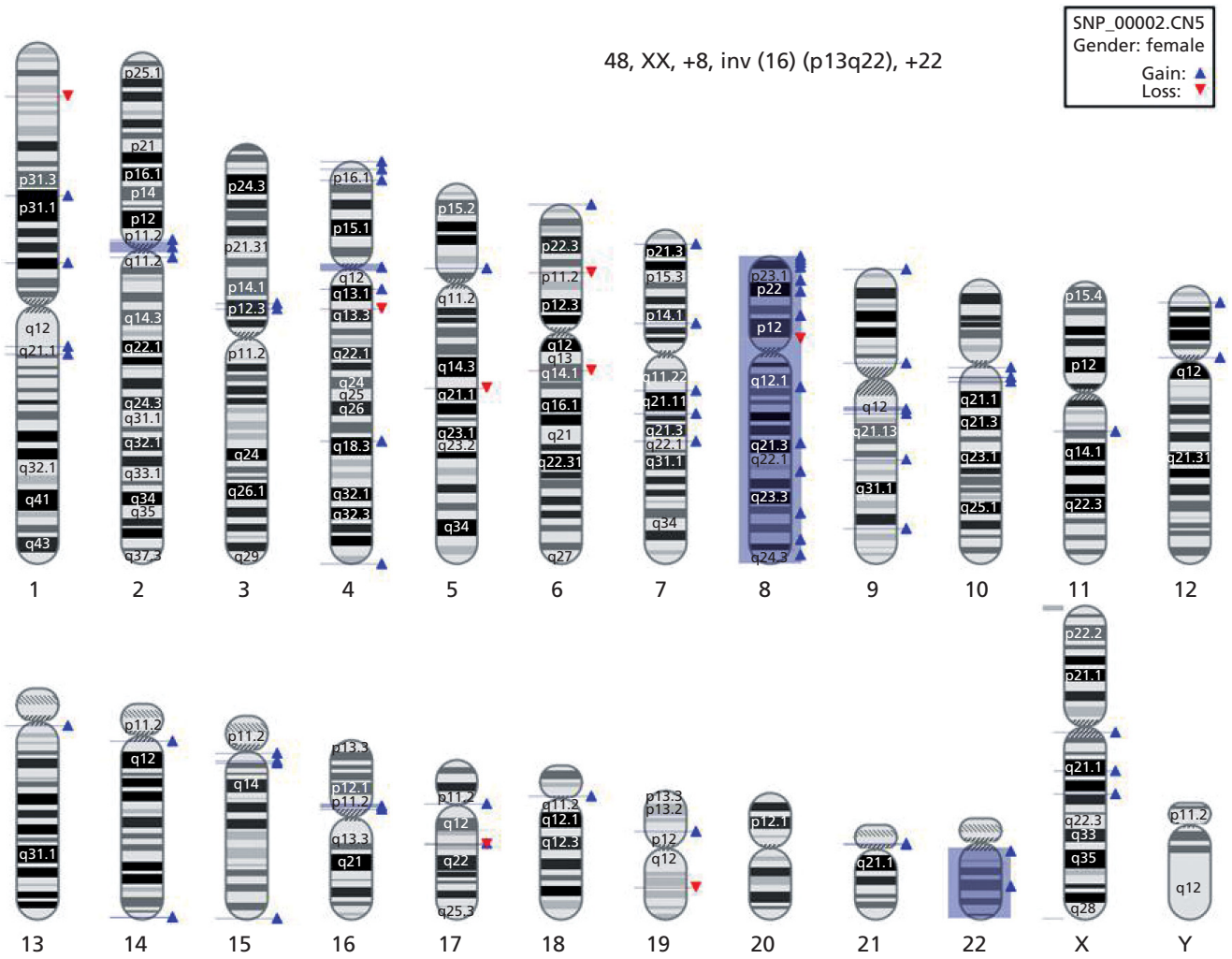


Figure 19.37 Molecular karyotype based on high-density SNP microarray analysis. Alterations are given by arrows indicating either gains or losses of chromosomal material. The karyotype using conventional chromosome banding techniques is given for comparison.

probes. CGH microarray analysis is used for the detection of under-representation or over-representation of specific genomic regions in the test DNA. Deletions and amplifications can thus be detected, for example deletion of 13q14 and 11q22 in CLL or iAMP21 in ALL.

In SNP array analysis, genomic DNA of a sample of interest is digested by restriction enzymes, amplified, labelled and hybridized to oligonucleotide probes. Probes are either bound to microbeads or are synthesized *in situ* in high density on a microarray surface (Figure 19.37). SNP analysis can be applied to genotyping and is able to distinguish heterozygosity from homozygosity. It can thus be used to detect both loss of heterozygosity and alterations of copy numbers of a genomic region (i.e. gains and losses). Acquired somatic uniparental disomy can be detected by comparison of tumour DNA with constitutional DNA. Sensitivity for the detection of a clonal population is similar to that of metaphase cytogenetics. With SNP array technology it is possible to identify uniparental disomy for chromosomal regions that include mutated genes such as *JAK2*, *CEBPA*, *FLT3* and *RUNX1*.

Microarray analysis of gene expression

Microarrayanalysis was initially applied to the study of cellular gene expression by hybridization of test and control RNA specimens, processed and labelled with different fluorochromes, to single-stranded DNA. The DNA probes were either complementary (c)DNA or synthetic oligonucleotides, arrayed in multiple rows of dots on a microarray surface with each dot interrogating a fragment of a single gene. High-density gene expression profiling (GEP) is a well-established molecular method and has been demonstrated to yield information of both diagnostic and prognostic value.

Primary mediastinal B-cell lymphoma can be distinguished from both DLBCL and Hodgkin lymphoma by its molecular signature. Burkitt lymphoma has a distinctive pattern of gene expression, which can help in making a distinction from DLBCL with *MYC* over-expression. Similarly, small lymphocytic lymphoma, marginal zone B-cell lymphoma and mantle cell lymphoma can be characterized by their differential gene expression signatures.

Prognostic information is provided in DLBCL. Microarray studies have shown that this diagnostic category encompasses at least two molecularly distinct diseases, differing in differentiation stage (i.e. cell of origin), oncogenic mechanisms and clinical outcome. Microarray analysis can also divide mantle cell lymphoma into groups with very different prognoses.

Next-generation sequencing (NGS)

The constantly increasing spectrum of molecular mutations is a tremendous challenge for haematological laboratories. Thus, the introduction of high-throughput sequencing technologies, which allow the massive parallel analysis of hundreds of thousands of alleles in a short time, provides new options for molecular mutation analyses and for follow-up diagnostics in myeloid neoplasms. In contrast to whole-genome or exome analyses, which might today not be feasible in a routine setting, amplicon deep-sequencing focuses on distinct genomic loci and their mutation patterns and enables a comprehensive biomarker analysis in a multitude of patients per analysis. It is foreseeable that clearly defined algorithms for molecular investigations will revolutionize diagnosis by including NGS as a routine method in the near future.

Laboratory techniques and the WHO classification of tumours of haemopoietic and lymphoid tissues

The 2008 WHO classification is based on the integration of clinical and pathological features. Among the pathological features, morphology (cytology and histology) is fundamental, but there is an increasing use of immunophenotyping for classification of lymphoid, mixed lineage and some myeloid neoplasms and of cytogenetic and molecular genetic information for classification of both lymphoid and myeloid neoplasms. Tables 19.1 and 19.2 illustrate how cytogenetic and molecular genetic information is integrated into the classifications of AML and B-lineage ALL respectively. Two categories of acute leukaemia of ambiguous lineage are also defined by the presence, in addition to the required immunophenotypic characteristics, of t(9;22)(q34;q11.2)/*BCR–ABL1* and 11q23/*MLL* rearrangement respectively. The 2008 WHO classification has also introduced a new genetically defined categorization for lymphoid and myeloid neoplasms associated with rearrangement of *PDGFRA*, *PDGFRB* or *FGFR1*; these categories encompass cases of *BCR–ABL1*-negative chronic myeloid leukaemia (often with eosinophilic differentiation), AML, ALL and lymphoid or myeloid transformation of a myeloproliferative neoplasm.

Otherwise, among the myeloproliferative neoplasms, only the demonstration of t(9;22)(q34;q11.2)/*BCR–ABL1* to confirm or exclude a diagnosis of CML is essential for classification. However, the demonstration of the *JAK2* V617F mutation or a somatic *MPL* or *CALR* mutation is important for making a diagnosis of a myeloproliferative neoplasm rather than a secondary or familial disorder. A *JAK2* exon 12 mutation similarly confirms that polycythaemia is polycythaemia vera rather than a familial or secondary condition.

Cytogenetic and molecular genetic abnormalities are critical in the diagnosis of certain neoplasms of mature lymphocytes (Table 19.3). However, for the majority of cases of CLL and NHL, such analysis is not needed for diagnosis and classification, although it may provide information of prognostic value.

Conclusions

Modern diagnosis, classification and monitoring of haematological neoplasms require integration of multiple diagnostic tools in a systematic manner in order to provide all the information necessary for optimal management of the patient. In addition, application of new diagnostic tools is constantly increasing our understanding of these disorders. New research tools not only give new information, but are often rapidly integrated into diagnostic practice. Classifications of lymphoid and myeloid neoplasms increasingly incorporate, and depend on, immunophenotypic and genetic analysis. High-throughput sequencing technologies investigating gene panels by next-generation sequencing are likely to be introduced into routine diagnostic algorithms shortly.

Selected bibliography

Alcalay M, Tiacci E, Bergomas R *et al*. (2005) Acute myeloid leukaemia bearing cytoplasmic nucleophosmin (NPMc+ AML) shows a distinct gene expression profile characterized by up-regulation of genes involved in stem-cell maintenance. *Blood* **106**: 899–902.

Bain BJ (2010) *Leukaemia Diagnosis*, 4th edition. Wiley-Blackwell, Oxford.

Bain BJ, Clark DM, Wilkins BS (2009) *Bone Marrow Pathology*, 4th edition. Wiley-Blackwell, Oxford.

Cools J, DeAngelo DJ, Gotlib J *et al.* (2003) A tyrosine kinase created by the fusion of the PDGFRA and FIP1L1 genes as a therapeutic target of imatinib in idiopathic hypereosinophilic syndrome. *New England Journal of Medicine* **348**: 1201–14.

deJong D, Xie W, Rosenwald A *et al*. (2009) Immunohistochemical prognostic markers in diffuse large B-cell lymphoma: validation of

tissue microarray as a prerequisite for broad clinical applications. *Journal of Clinical Pathology* **62**: 128–38.

Falini B, Mecucci C, Tiacci E *et al*. (2005) Cytoplasmic nucleophosmin in acute myelogenous leukaemia with a normal karyotype. *New England Journal of Medicine* **352**: 254–66.

Gleissner B, Küppers R, Siebert R *et al*. (2008) Report of a workshop on malignant lymphoma: a review of molecular and clinical risk profiling. *British Journal of Haematology* **142**: 166–78.

Grossmann V, Schnittger S, Kohlmann A *et al*. (2012) A novel hierarchical prognostic model of AML solely based on molecular mutations. *Blood* **120**: 2963–72.

Haferlach C, Dicker F, Schnittger S, Kern W, Haferlach T (2007) Comprehensive genetic characterization of CLL: a study on 506 cases analysed with chromosome banding analysis, interphase FISH, IgV(H) status and immunophenotyping. *Leukemia* **21**: 2442–51.

James C, Ugo V, Le Couédic JP *et al*. (2005) A unique clonal JAK2 mutation leading to constitutive signalling causes polycythaemia vera. *Nature* **434**: 1144–8.

Kohlmann A, Grossmann V, Nadarajah N, Haferlach T (2013) Next-generation sequencing: feasibility and practicality in haematology. *British Journal of Haematology* **160**: 736–53.

Lenz G, Wright G, Dave SS *et al.* (2008) Stromal gene signatures in large B-cell lymphomas. *New England Journal of Medicine* **359**: 2313–23.

Maciejewski JP, Mufti GJ (2008) Whole genome scanning as a cytogenetic tool in hematologic malignancies. *Blood* **112**: 965–74.

Rosenwald A, Wright G, Chan WC *et al*. (2002) The use of molecular profiling to predict survival after chemotherapy for diffuse large B-cell lymphoma. *New England Journal of Medicine* **346**: 1937–47.

Schlenk RF, Döhner K, Krauter J *et al*. (2008) Mutations and treatment outcome in cytogenetically normal acute myeloid leukaemia. *New England Journal of Medicine* **358**: 1909–18.

Song MK, Chung J-S, Shin D-H *et al*. (2009) Prognostic significance of the Bcl-2 negative germinal centre in patients with diffuse large B lymphoma treated with R-CHOP. *Leukemia and Lymphoma* **50**: 54–61.

Vardiman JW, Thiele J, Arber DA *et al*. (2009) The 2008 revision of the World Health Organization (WHO) classification of myeloid neoplasms and acute leukaemia: rationale and important changes. *Blood* **114**: 937–51.

CHAPTER 20

Acute myeloid leukaemia

20

Alan K Burnett[1] and David Grimwade[2]

[1]Formerly of Department of Haematology, School of Medicine, Cardiff University Cardiff, UK.
[2]Department of Medical and Molecular Genetics, King's College London, London, UK.

Disease epidemiology

Acute myeloid leukaemia (AML) has an incidence of 2–3 per 100,000 per annum in children, rising to 15 per 100,000 in older adults. It can occur at all ages. but has its peak incidence in the seventh decade (Figure 20.1). The incidence does not appear to be increasing beyond that expected in an ageing population, which in itself will considerably increase the burden of the disease to healthcare systems in future years. The fact that most cases occur in older patients has important implications for treatment strategies, in that biological variation associated with chemoresistance and comorbidity, which limits treatment options, increases with age.

Pathophysiology and clinical features

The more carefully AML is studied, the clearer it becomes that there is considerable heterogeneity between cases, with respect to morphology, immunological phenotype, associated cytogenetic and molecular abnormalities and, more recently, methylation profile and pattern of gene and microRNA (miR) expression. This is reflected in the substantially different responses to treatment. Some entities are becoming so distinct that they are regarded as different diseases with specific approaches to treatment. It turns out that AML is genetically much less complex than solid tumours, with cases typically harbouring mutations in over 10 genes. While there remains much to learn about the molecular pathogenesis, in a significant proportion of AML (~40%), there is evidence to suggest that the initiating event is acquisition of a balanced chromosomal abnormality (i.e. translocation, inversion) in an early bone marrow haemopoietic stem/progenitor cell generating chimaeric oncoproteins, which induce leukaemic transformation with the accumulation of additional cooperating mutations.

The tests recommended in the initial work up of a newly admitted patient with AML are summarized in Table 20.1. The observed clinical heterogeneity amongst patients with AML may not only reflect differences in the cytogenetic and molecular genetic characteristics of the leukaemic population, but also the cell of origin. The leukaemic clone has a competitive advantage and can impair normal haemopoiesis leading to marrow failure. Adenopathy or organomegaly can occur, but are not usual features. Disseminated intravascular coagulation, usually accompanied by skin and mucosal haemorrhage due to consumption of platelets and clotting factors, is a frequent presenting feature of acute promyelocytic leukaemia. Gum and skin infiltration are particularly seen in the monocytic variants. Other extramedullary disease can arise, including cerebrospinal infiltration, but this is not usual, unlike in acute lymphoblastic leukaemia. Most patients present with low peripheral blood counts with the consequent clinical features. While in the majority of cases no direct cause is found, there is an association with irradiation, smoking, some rare congenital abnormalities, chemical exposure and obesity. Perhaps the most frequently identified cause is progression from another haematological disorder, e.g. myelodysplastic syndrome (MDS) or myeloproliferative neoplasm (MPN), or as a consequence of prior chemotherapy for another malignancy (therapy-related AML, t-AML).

Non-random chromosome abnormalities are found in the majority of cases (See Chapter 19 and Figure 20.2). These may be structural (gain or loss of material) or reciprocal balanced

Postgraduate Haematology, Seventh Edition. Edited by A Victor Hoffbrand, Douglas R Higgs, David M Keeling and Atul B Mehta.

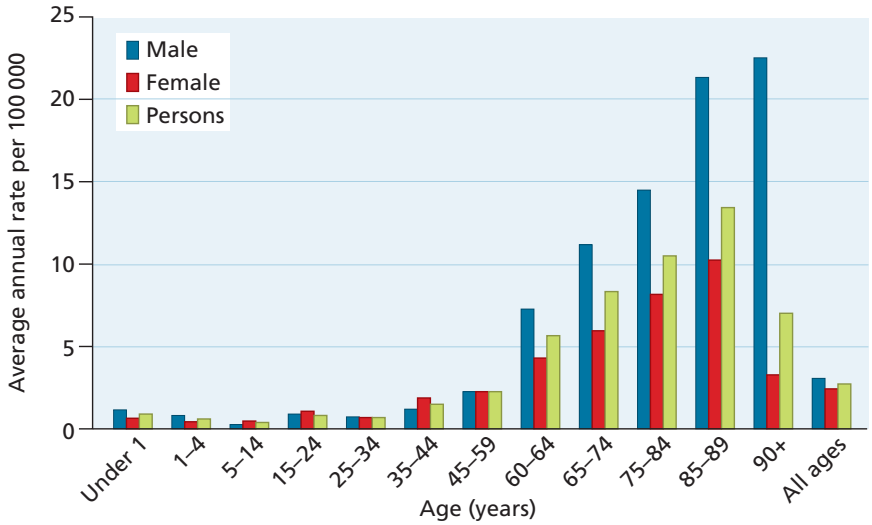


Figure 20.1 Age distribution of AML in the population of Wales: a population study.

rearrangements. The significance of these changes is developing at a steady pace and is discussed below as far as the clinical implications are concerned.

Disease classification

The disease is confirmed by an excess of primitive 'blast' cells in the bone marrow, originally in the FAB classification required to be at least 30%, but 20% is the current threshold. The French–American–British (FAB) morphological classification has been useful in developing a common vocabulary, but has little predictive value since the widespread use of genetic markers. Quality cytochemistry can provide valuable additional diagnostic information. Further precision can be added by immunophenotyping (Chapter 19). Guidelines suggest that panels can be formulated to depict precursor, granulocytic, monocytic, megakaryocytic or erythroid features (Table 20.2). Although widely used, in many cases, where high-quality morphology and cytochemistry is available, immunophenotyping is not strictly required to confirm the diagnosis as AML, except for cases classified as M0 or M7 (see Chapter 19). Although still in research mode, and requiring considerable skill and standardization, the identification of aberrant marker expression at diagnosis is emerging as a useful way to monitor individual patients' response to treatment. It is also essential in diagnosing mixed-phenotype acute leukaemia (Table 20.2).

In recent years, the FAB classification has been superseded by the revised scheme devised under the auspices of the World Health Organization (WHO), which recognizes accumulating knowledge of the cytogenetic and the molecular characteristics, as described in Chapter 19. As discussed later, the leukaemic blasts may demonstrate an 'aberrant' immunophenotype, which together with molecular characterization of cloned breakpoints or mutations has potential to better classify distinct entities with varying prognoses, as well as allow more precise definition of response to treatment.

Cytogenetics and molecular genetics

Over the last four decades there have been major advances in deciphering the genetic changes underlying AML. Approximately 60% of cases are characterized by acquired chromosomal abnormalities which define biologically and prognostically distinct subsets of disease. These include balanced chromosomal rearrangements (translocations and inversions) which frequently target genes encoding haemopoietic transcription factors (e.g. Retinoic Acid Receptor Alpha due to t(15;17)(q22;q21), Core Binding Factor [CBF] complex due to t(8;21)(q22;q22) or inv(16)(p13q22)/t(16;16)(p13;q22)) or epigenetic regulators (e.g. KMT2A [MLL] due to rearrangements of 11q23) (see Chapter 19, and Figure 20.2). Balanced chromosomal rearrangements, which lead to the formation of chimeric oncoproteins, are most prevalent in AML arising in infancy (where *MLL* fusions predominate), children and younger adults (Figure 20.2), but are infrequent in older adults where complex karyotypes are more common. Over the last 20 years there has been dramatic progress in understanding the molecular basis of AML lacking balanced chromosomal rearrangements, including the 40% with normal karyotype, as described in more detail further on (Figure 20.2).

Consequently, it has become clear that the original morphology-based classification of AML is no longer suitable, with an increasing number of disease entities being recognized on the basis of cytogenetic and molecular genetic characteristics. The previous WHO classification introduced

Table 20.1 Test/procedures in the initial work-up of a patient with AML.

Test/procedure	General practice	Clinical trial
Tests to establish the diagnosis		
Complete blood counts and differential count	Yes	Yes
Bone marrow aspirate	Yes	Yes
Bone marrow trephine biopsy	Optional[f]	Optional[f]
Immunophenotyping	Yes	Yes
Cytogenetics	Yes	Yes
RUNX1-RUNX1T1, CBFB-MYH11, PML-RARA, or other gene fusion screening	Optional[g]	Optional[g]
Additional tests/procedures at diagnosis		
Demographics and medical history[a]	Yes	Yes
Performance status (ECOG/WHO score)	Yes	Yes
Analysis of comorbidities	Yes	Yes
Biochemistry, coagulation tests, urine analysis[b]	Yes	Yes
Serum pregnancy test[c]	Yes	Yes
Information on oocyte and sperm cryopreservation	Optional[h]	Optional[h]
Eligibility assessment for allogeneic HSCT	Yes[i]	Yes[i]
Hepatitis A, B, C; HIV-1 testing	Yes	Yes
Chest X-ray, 12-lead ECG: echocardiography (on indication)	Yes	Yes
Lumbar puncture[d]	No	No
Biobanking[e]	Optional[j]	Yes
Prognostic/predictive marker assessment		
NPM1, CEBPA, FLT3 gene mutation	Optional[k]	Yes
WT1, RUNX1, MLL, KIT, RAS, TP53, TET2, IDH1 gene mutation	No	Investigational
ERG, MN1, EVI1, BAALC gene expression	No	Investigational
Detection of minimal residual disease	No	Investigational

[a] Including race or ethnicity, family history, prior exposure to toxic agents, prior malignancy, therapy for prior malignancy, information on smoking.
[b] *Biochemistry*: glucose, sodium, potassium, calcium, creatinine, aspartate amino transferase (AST), alanine amino transferase (ALT), alkaline phosphatase, lactate dehydrogenase, bilirubin, urea, total protein, uric acid, total cholesterol, total triglycerides, creatinine phosphokinase (CPK). *Coagulation tests*: prothrombin time (PTT), international normalized ratio (INR) where indicated, activated partial thromboplastin time (aPTT). *Urine analysis*: pH, glucose, erythrocytes, leucocytes, protein, nitrite.
[c] In women with childbearing potential.
[d] Required in patients with clinical symptoms suspicious of central nervous system involvement; patient should be evaluated by imaging study for intracranial bleeding, leptomeningeal disease, and mass lesion; lumbar puncture considered optional in other settings (eg, high WBC).
[e] Pretreatment leukemic bone marrow and blood sample.
[f] Mandatory in patients with a dry tap (punctio sicca).
[g] Should be performed if chromosome morphology is of poor quality, or if there is typical morphology but the suspected cytogenetic abnormality is not present.
[h] Cryopreservation to be done in accordance with the wish of the patient.
[i] HLA typing and CMV testing should be performed in those patients eligible for allogeneic stem cell transplantation.
[j] Biobanking should also be performed in general practice if at all possible.
[k] Strongly encouraged in AML with normal karyotype.
Source: Döhner *et al*., 2010. Reproduced with permission of American Society of Hematology.

a lower marrow blast threshold of 20% to define the disease, but this is relatively arbitrary and many patients enter treatment with 10% marrow blasts, i.e. high-risk myelodysplastic syndrome. Moreover, in the WHO classification, patients with an established recurrent chromosomal rearrangement, e.g. t(15;17)/*PML-RARA*, t(8;21)/*RUNX1-RUNX1T1*, can be diagnosed as AML, irrespective of marrow blast percentage.

Treatment

Aspirations for treatment

Given the age distribution of patients who will present with the disease, it must first be decided what the goals of treatment are in an individual patient. In young people, there is little doubt that there is the prospect of significant benefit to be gained from an

Table 20.2 Expression of cell-surface and cytoplasmic markers for the diagnosis of acute myeloid leukaemia and mixed-phenotype acute leukaemia.

Expression of markers for diagnoses	
Diagnosis of acute myeloid leukaemia (AML)	
Precursor stage	CD34, CD38, CD117, CD133, HLA-DR
Granulocytic markers	CD13, CD15, CD16, CD33, CD65, cytoplasmic myeloperoxidase (cMPO)
Monocytic markers	Nonspecific esterase (NSE), CD11c, CD14, CD64, lysozyme, CD4, CD11b, CD36, NG2 homologue
Megakaryocyte markers	CD41 (glycoprotein IIb/IIIa), CD61 (glycoprotein IIIa), CD42 (glycophorin 1b)
Erythroid marker	CD235a (glycophorin A)
Diagnosis of mixed phenotype acute leukaemia (MPAL)	
Myeloid lineage	MPO or evidence of monocytic differentiation (at least 2 of the following: NSE, CD11c, CD14, CD64, lysozyme)
B-lineage	CD19 (strong) with at least one of the following: CD79a, cCD22, CD10, or CD19 (weak) with at least 2 of the following: CD79a, cCD22, CD10
T-lineage	cCD3, or surface CD3

Source: Döhner *et al*., 2010. Reproduced with permission of American Society of Hematology.

intensive approach. With increasing age, which is often associated with comorbidity and less responsive disease, the balance of benefit changes to a more palliative approach. Much of what is known about the prospects of successful treatment is derived from large clinical trials. In young patients, these results are usually representative of what can be expected for any age-matched patients. However, only a selected minority of older patients (>60 years) will enter trials, so trial-derived information may be less transferable to the whole population in this age group. A substantial majority of older patients are not considered fit for the usual intensive approach and the priority for such patients has until recently been palliative care, optimization of quality of life and minimal hospitalization. Examination of survival in population studies illustrates the dominant effect of age (Figure 20.3); indeed, a substantial proportion of patients over the age of 60 years do not receive more than a palliative treatment approach. However, newer drugs are being intensively tested in this population that indicate that all patients can be offered treatment with a chance of prolonging life.

Treatment strategy

The initial clinical priority is to apply chemotherapy to improve marrow function by inducing complete remission (CR). Conceptually, this means an approximate 2-log reduction in tumour burden (from the estimated 10^{12} leukaemic cells at disease diagnosis). This becomes compatible with a bone marrow that appears normal morphologically and is functionally able to produce normal numbers of circulating cells. The traditional consensus definition of CR is based on these premises: less than 5% blast cells in a cellular marrow durable for at least 28 days with a peripheral neutrophil count of 1.5×10^9/L and platelet count above 100×10^9/L, and absence of extramedullary disease. In some cases, these criteria may be met, but the morphology is dysplastic. It is not clear whether this is associated with a greater risk of relapse. Similarly, some patients meet the marrow criteria, but do not achieve full peripheral count regeneration, now called CRi (CR with incomplete count recovery), i.e. either the neutrophil or platelet count has failed to reach the level required for the definition of CR. This subgroup tends to have a poorer overall survival. In some circumstances, these features may represent a pre-existing dysplastic state; in others it may represent the effects of over-treatment of that particular individual. In the former case, this may have adverse connotations, while in the latter it represents optimum treatment.

The increased understanding of molecular mechanisms (as described by Bolli and Vassiliou in Chapter 18) has resulted in more sophisticated molecular techniques becoming available; it is clear that it is still possible to detect residual disease when all conventional morphological and functional criteria of complete remission are met. Techniques such as real-time quantitative polymerase chain reaction (RT-qPCR) and multiparameter flow cytometry (MFC) are capable of detection at a level of 1 in 10^4 or 1 in 10^5 residual cells. While the molecular markers do not cover all cases, flow cytometric detection of aberrant phenotype can be characterized in most patients. This is highly skilled technology that requires considerable expertise and lab quality assurance, but may well re-define 'remission' and responses in the near future. Interestingly, in older studies on X-linked markers of clonality, it has been noted that in some cases marrow remissions are clonal. It has recently become clear from work from the Dick and Majeti groups that this

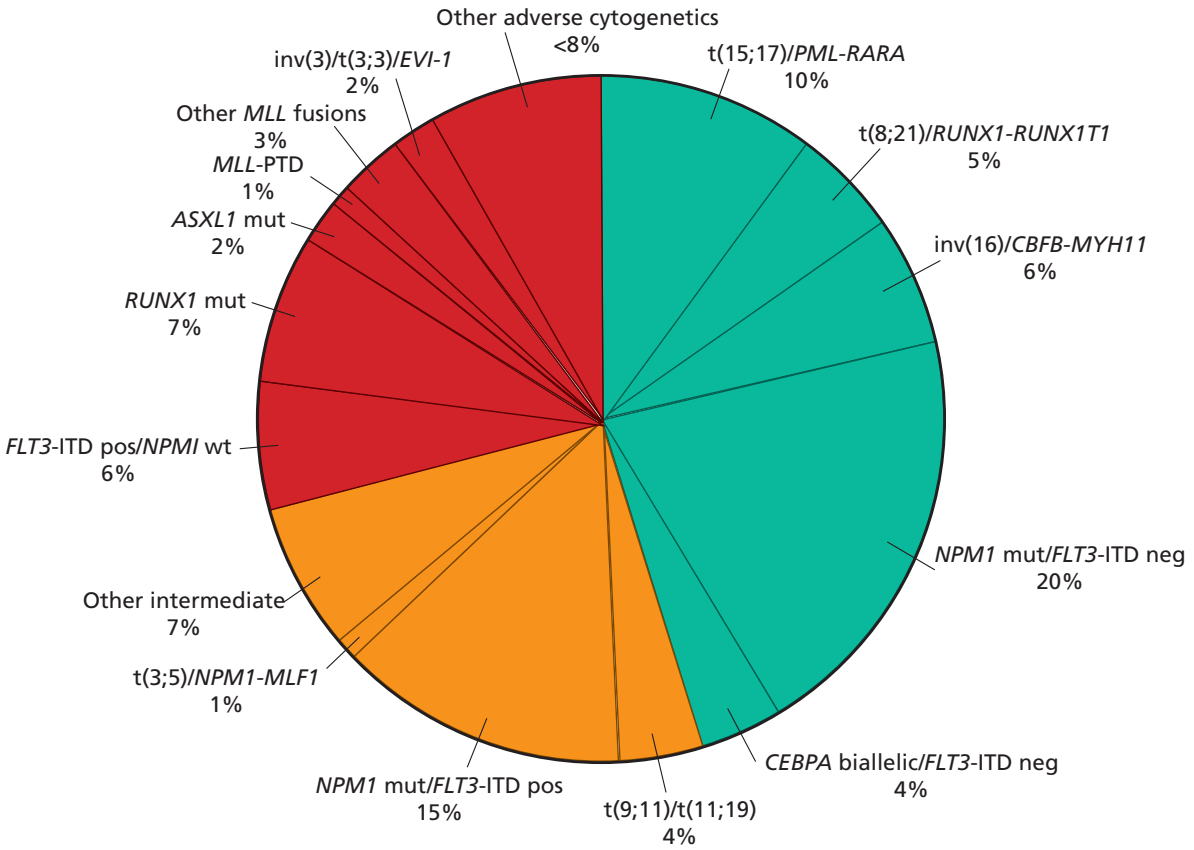


Figure 20.2 Distribution of prognostically relevant cytogenetic and molecular genetic subgroups according to patient age. Key: Favourable risk (green), intermediate (orange), adverse (red). Figure courtesy of Adam Ivey (Medical and Molecular Genetics, Guy's Hospital London), including information from the NCRI AML trials cytogenetic database provided by Anthony Moorman and Christine Harrison (Leukaemia Research Cytogenetics Group, Northern Institute for Cancer Research, Newcastle University) and molecular screening data provided by Tamara Alpermann and Susanne Schnittger (Munich Leukemia Laboratory, Munich, Germany). (Source: Adam Ivey, Guy's Hospital, London. Reproduced with permission.)

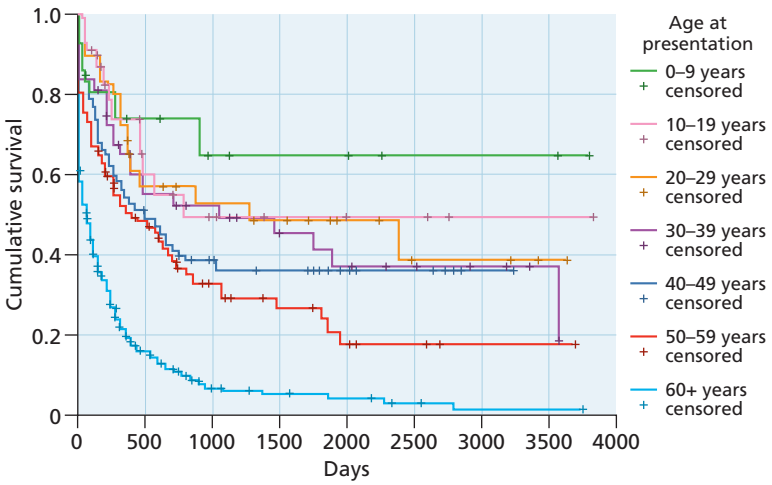


Figure 20.3 Survival of AML in relation to patient age.

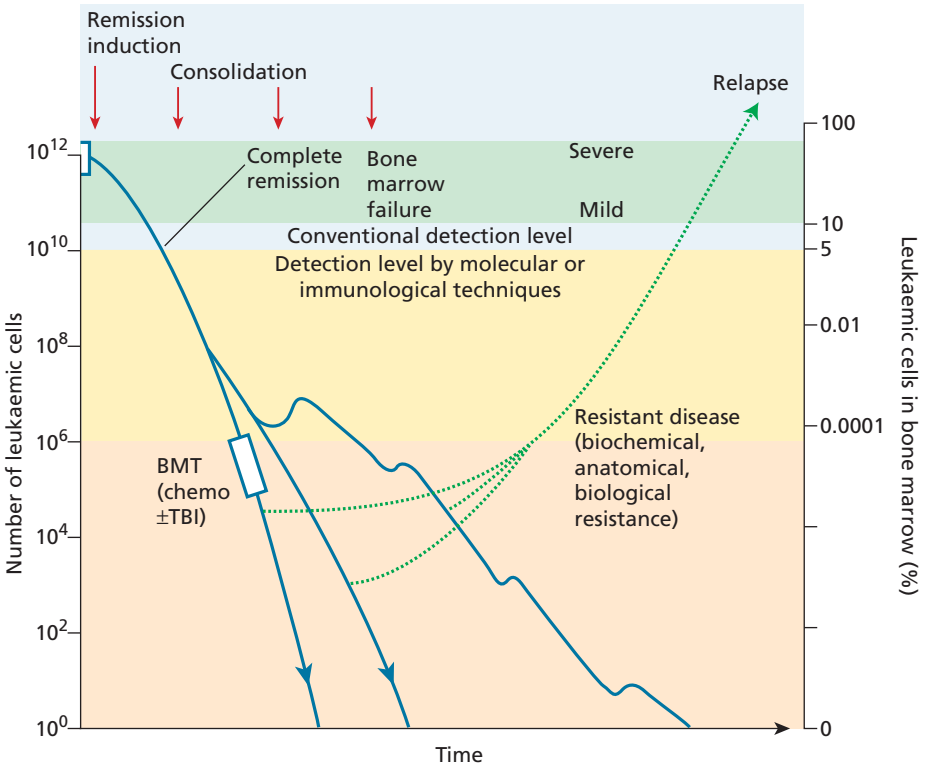


Figure 20.4 Schematic representation of depth of response and kinetics of disease relapse in relation to the treatment pathway in AML.

phenomenon relates to persistence of pre-leukaemic clones harbouring mutations in epigenetic landscaping genes (e.g. *DNMT3A*) which are resistant to chemotherapy and can provide a potential reservoir for disease relapse. As discussed further on, there is now intense interest in using validated molecular markers and leukaemia-associated immunophenotypes to predict imminent relapse before this becomes apparent clinically.

Clinical experience has demonstrated that further intensive postremission treatment is required to 'consolidate' CR. This is delivered at the same intensity as induction in order to achieve further cytoreduction. Under these circumstances, it is possible to achieve disease levels that are beyond the level of molecular detection. It is not clear how many intensive consolidation courses are required, but two or three are generally used in younger patients, and stem cell transplantation may be included. Where intensive induction and consolidation can be given, for example in younger patients, maintenance chemotherapy is not required. A pictorial description of treatment is shown in Figure 20.4.

Treatment details

Induction of remission

The backbone of treatment for 30 years has been the combination of daunorubicin and cytarabine. Historically daunorubicin has been given for 3 days in a dose of 45–50 mg/m^2. Cytarabine is given for 7–10 days as a continuous infusion or by bolus doses of 100–200 mg/m^2 daily. Many clinical trials have been conducted that have tested variations on this standard of care. Alternatives to daunorubicin (doxorubicin, mitoxantrone, idarubicin, aclarubicin) or different doses have not yet been shown to afford consistently superior overall, although recent studies that have explored higher anthracycline doses may improve the remission rate in older patients. Idarubicin may achieve a better quality of remission, as reflected in a reduced relapse risk in younger patients, but it is more myelosuppressive and limits the intensity of consolidation treatment. Intensification of daunorubicin (to 90 mg/m^2) has been reported to be beneficial in younger patients or in a subset of older patients (aged 60 to 65 years), but in the younger patient trial, the control arm (at a daunorubicin dose of 45 mg/m^2) was inferior to what is now routinely expected. Dosing for 5 days instead of 3 also suggested that this method of intensification was effective. Surprisingly, 90 mg has never been compared with the more 'standard' dose of 60 mg/m^2. This question has been addressed in a large UK trial involving more than 1200 patients with no overall difference being found overall or in any patient subgroup. For many investigators daunorubicin at a dose level of 60 mg/m^2 remains the anthracycline of choice. Higher doses of cytarabine (3 g/m^2) in induction have been tested in

recent years, with mixed results and no convincing evidence of overall benefit. Intermediate doses (400 mg/m^2 daily versus 200 mg/m^2 daily) have been tested in younger patients without demonstrating a difference.

Comparison of induction treatment is not simply measured by the rate of remission. By achieving a greater degree of cytoreduction, without necessarily getting more patients into CR, one treatment may be superior because it results in fewer subsequent relapses. The beneficial effect of the addition of a third drug to the induction combination has some evidence to support it. This will usually be etoposide or thioguanine. A large comparative study did not show any difference between these two drugs when used as the third drug in combination with daunorubicin and cytarabine. More recent data from Poland suggests that the addition of cladrabine to daunorubicin and Ara-C is beneficial overall, but in particular in the higher-risk patients. This interesting option requires additional confirmation. There has been much interest in adding the immunoconjugate gemtuzumab ozogamicin (GO) (mylotarg) to induction treatment. This does not improve the remission rate, but reduces the risk of relapse and thereby improves survival. A recent patient-based meta-analysis of the five randomized trials in adults, where GO was added to induction, confirms the overall reduction in relapse and the survival benefit in patients who are not adverse risk. However mylotarg is unlicensed outside Japan. An alternative induction is the FLAG-Ida combination (fludarabine/Ara-C/G-CSF and idarubicin), which proved to be effective in younger patients, again, not by improving the rate of remission, but by reducing relapse. When two successive courses of FLAG-Ida are used for induction, greater myelosuppression is observed in subsequent consolidation, thus reducing compliance with postinduction treatment, but this does not appear to diminish survival. In fact, recipients of only two induction courses of FLAG-Ida may fare just as well as recipients of two DA inductions followed by Ara-C consolidation.

The majority of patients who are going to enter remission will do so after one course of treatment. If an incomplete response is obtained, then a second course of the same combination is indicated. A further group will enter remission, but these patients have thus shown themselves to have less sensitive disease, and this is reflected in a modestly increased risk of relapse later. If patients fail to achieve a substantial reduction in marrow blasts (to <15% blasts) in the first course or fail to enter CR with a second course, even if they then enter remission, they should be considered refractory to the drugs used up to that point and transferred to an alternative treatment schedule where they can still have a prospect of achieving CR, even though they tend to have a higher risk of rapid relapse. On this basis alone they are candidates for allogeneic transplant.

An emerging area is the possibility of re-defining remission status based on minimal (or measurable) residual disease (MRD) assays. This will be discussed later.

Table 20.3 Relationship between complete remission rate and patient age. Patients were all given intensive chemotherapy.

Age (years)	<35	35–55	55–60	61–65	66–70	71–75	75+
Complete remission (%)	88	82	77	62	63	48	59

Source: Data from the UK Medical Research Council AML Trial database.

Results of induction treatment

With the approaches to treatment outlined above, 75–80% of patients under 60 years and 50–60% of older patients will enter morphological CR. Of those who do, about 70% will have required one course. A number of factors influence the prospects of achieving remission. Age is a dominant and independent risk factor and a continuous variable (Table 20.3). Clinical performance score at diagnosis is also highly predictive. In younger patients, fewer tend to present with poorer performance scores so this prognostic factor does not move the overall remission rate to any great extent. A larger proportion of older patients will have poorer risk biology (including poor cytogenetics, secondary AML, drug-resistant phenotype) or poorer prognostic scores where multiparameter scores are used. The distribution of cytogenetic subtypes is related to age, with more-responsive subtypes accounting for a greater proportion in younger patients and less-responsive subtypes aggregating in older individuals. Tumour burden at diagnosis, as represented by white blood count, serum albumin or lactate dehydrogenase levels, will adversely impact on response to induction treatment. It is now possible to measure a number of proteins involved in drug efflux in leukaemic blasts. These 'resistance' proteins, for example P-glycoprotein (see below), tend to be more frequently expressed in AML arising in older patients and correlate with a lower remission rate. If patients have had an antecedent haematological disorder, e.g. myelodysplasia, the remission rate will be about 20% lower than in age-matched groups. About 10–25% of older patients embarking on intensive induction chemotherapy will die during the aplastic phase from non-leukaemic causes, which is essentially a failure of supportive care. Induction deaths tend to be associated with the adverse features already mentioned.

Supportive care (see also Chapter 23)

It is unusual for induction chemotherapy not to clear most of the leukaemic blasts; however, this is at a cost of 3–4 weeks of severe pancytopenia. Supporting patients through the period of marrow suppression is crucial to treatment outcome; indeed many hold the view that the main reason that treatment has improved is due to improvements in supportive care. In collaborative group studies this is reflected in a decline in induction

deaths and deaths in remission over time. It is therefore important that patients are treated in an environment where all necessary supportive facilities are available. Several components of supportive care have to be in place during this period. Careful monitoring of biochemical parameters of renal and hepatic function and coagulation is required. Central venous access is now considered essential, together with high-quality and readily available blood product support.

A priority is the prevention and management of infection. Most patients will receive prophylactic oral antibiotics and antifungals to minimize the risk of infection during the neutropenic period, although routine use of the latter can still be debated. Since hospital-acquired infections are becoming an increasing problem, it can be safer for the patient to be at home, provided that close monitoring can be undertaken in the day hospital and that rapid re-admission to specialist care is available.

Despite prophylactic measures, most patients will become febrile during neutropenia. This must be considered an indication of a serious, and potentially fatal, infection. The common pathogens are staphylococcal, caused by the use of central catheters, and, increasingly, fungal infections (*Candida* and *Aspergillus*), which are related to the duration of severe neutropenia that results from the more intensive chemotherapy now used. Particular patterns of infection will be determined within individual institutions and will dictate the specific approach to empirical antimicrobial intervention. Fungal infections are a particular problem, not only related to more intensive treatment, but also because of building construction work that is such a feature of hospital environments. Guidance on intervention should not only be based on evidence from the literature, but should also incorporate local microbiological issues. Nursing expertise is an essential component. It seems probable that improvements in remission rates in recent years can largely be attributed to better supportive care and nursing skills, which have enabled more intensive treatment to be given safely.

Recombinant growth factor granulocyte colony-stimulating factor (G-CSF) has potential use in two respects. First, if it could curtail the duration of neutropenia, there would be less risk of death during the aplastic phase following induction chemotherapy; this may increase the rate of remission. Second, as many leukaemia cells exhibit receptors for these growth factors, it may be possible to pretreat the patient with growth factor to bring the leukaemic cells into cycle and thereby make them more susceptible to chemotherapy. Extensive studies using G-CSF to curtail neutropenia have been carried out and some general conclusions can be made. The duration of neutropenia can be reduced by a few days, but it is less easy to demonstrate a reduction in episodes of febrile neutropenia. There is generally no improvement in remission rate. Growth factor use has not increased leukaemic growth or involved relapse.

It is primarily a health economics decision that determines whether growth factors are incorporated into routine practice or not. Their use may enable patients to leave hospital earlier. If the local policy is to hospitalize patients during neutropenia, this may save resources. Fewer studies have been carried out to see whether growth factor 'priming' of the leukaemic population would be advantageous. These have generally been unsuccessful, but a recent positive study may re-kindle interest in this approach, particularly because newer mobilizing agents, such as CXCR4 antagonists, may soon become available.

Consolidation treatment

Having achieved remission, the priority is to prevent relapse. Optimization of induction treatment is still required, as it will influence the quality of remission and thereby the subsequent rate of relapse. Mature data from randomized trials suggests that the risk of relapse can be reduced by the addition of the immunoconjugate gemtuzumab ozogamicin (mylotarg) to induction treatment, but it does not change remission rate. However its use post remission has not been effective. Three options are available for younger patients once remission has been achieved: further chemotherapy at induction level of intensity, chemotherapy with autologous or allogeneic stem cell transplantation with or without prior chemotherapy. Chemotherapy will usually involve a further two or three courses usually with high-dose Ara-C. At this point in the treatment, there is a theoretical logic in using different drugs to minimize the risk of selecting chemoresistant leukaemic clones. Combinations, using cytarabine at increased dosage, amsacrine, etoposide and alternative anthracyclines, are often used (Figure 20.5). Few studies have made direct comparisons between specific combinations, but rather try to work out how many consolidation courses are needed. Two or three courses that are intensive enough to induce 3–4 weeks of neutropenia appears to be achievable, but is reaching the limit of tolerability and compliance. High-dose cytarabine (3 g/m^2 on alternate days over 5 days) has been shown to be superior to lower doses (400 mg/m^2 or 100 mg/m^2), but only in patients with more sensitive disease. Trials have also shown that intermediate cytarabine doses may be just as effective with less toxicity. It has been suggested that high-dose cytarabine is more effective in the most responsive subtypes of disease (i.e. those with lower-risk disease) based on cytogenetic prognostic markers. Overall, 45–50% of younger patients who enter remission will relapse, usually within the first 2 years. In older patients, the risk is much higher (80%).

Allogeneic stem cell transplantation

There is little doubt that the most effective way to prevent leukaemic relapse in younger patients is allogeneic transplantation from an HLA-compatible sibling donor. Most of the extensive data available are derived from patients in whom the graft was of bone marrow. Practice has now changed such that peripheral blood is the usual source, and it is clear that a well-matched unrelated donor can provide equivalent outcomes (Chapter 35). In these circumstances, the relapse risk will be reduced from 45%

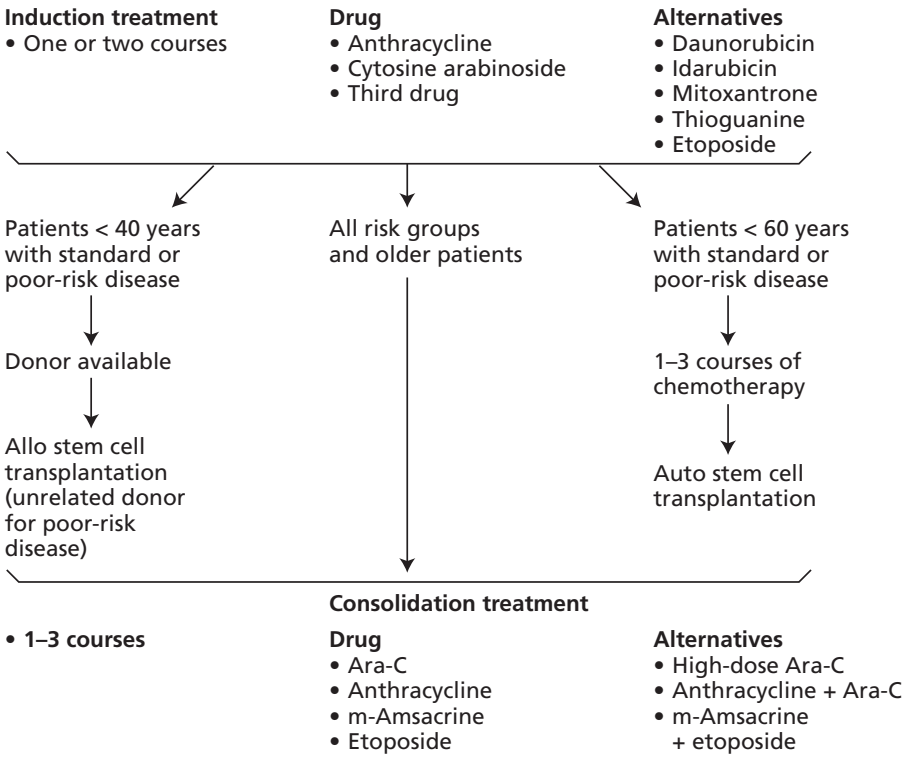


Figure 20.5 Treatment options in AML.

to about 20%. As there are non-leukaemic causes of death, the overall expectation of cure for recipients of allogeneic bone marrow transplantation is around 60% from the time of transplantation. Some of these survivors will have morbidities that survivors of chemotherapy may avoid, e.g. loss of fertility, and this needs to be taken into account when advising patients. As the risk of transplant complications, graft-versus-host disease (GVHD) and infections, in particular, increases with age, this approach is normally limited to patients under 45 years, although the precise age cut-off remains controversial and will be a matter of the relative risk of the transplant and of disease recurrence.

Some transplant-related factors may predict for a more favourable outcome, such as a male donor, a cytomegalovirus (CMV)-negative donor when the host is CMV negative and a higher cell count in the graft, and so influence the decision to undergo the treatment, but these seem to be less powerful than disease characteristics in predicting the overall outcome. The extent to which the high-dose preparative regimen necessary to ensure engraftment or the immunological reactivity of the donor marrow via donor T cells eliminates residual leukaemia has been debated extensively. It is assumed that at least some (and probably a major) contribution comes from the immunological graft-versus-leukaemia (GVL) effect. Experience with donors who are mismatched at more than one HLA locus has not been encouraging, although in specialized centres haplo-identical transplants now appear feasible. Fully matched unrelated donor transplantation has become more reliable, particularly with the development of molecular methods of tissue typing. In expert hands, in carefully selected young patients, this approach is equivalent to having a sibling donor.

Once remission has been achieved, there is probably no definite requirement to administer more than one course of consolidation chemotherapy before the allograft; however, because of the time required to identify a donor and make the necessary arrangements for the transplant, more than one course is usually necessary. It appears that variations in transplantation protocols do not have a major effect on overall survival, for example choice of myeloablative schedule, GVHD prophylaxis or whether bone marrow or peripheral blood is the source of stem cells. However, it may be that the level of immunosuppression with ciclosporine can be manipulated to influence the risk of relapse.

Because the applicability of transplantation is limited, by treatment-related complications, to younger patients and yet has a very powerful antileukaemic effect, there has been considerable interest as to whether non-ablative allogeneic transplantation will have a role in older patients. This approach does not require the traditional intensive treatment to ablate the host marrow, but provides enough immunosuppression to enable the donor stem cells to engraft. Over a period of weeks, the host haemopoiesis becomes donor, i.e. changes from host to mixed to donor chimerism. It is now clear that full chimeric engraftment in older patients can be achieved with treatment

modalities that are not ablative to the bone marrow. The hope is that this provides sufficient GVL effect. In AML, in a conventional allogeneic transplant, it is not clear how important the GVL effect is, so it remains to be seen whether non-ablative transplantation has a role in consolidation of AML in older patients. However, the preliminary results have shown that this is a viable approach and that over the 2–4 year term the outcomes are encouraging.

Preliminary data suggests that non-ablative (reduced-intensity conditioning transplants) can improve survival compared with chemotherapy in intermediate-risk cases, but not in adverse-risk cases. This could be said to be the opposite of the data in younger patients who receive a myeloablative graft. Defining who benefits from a non-ablative transplant is an important ongoing area of clinical investigation.

So there still remains much debate about which patients should be offered an allograft in first remission. At present, for a myeloablative approach, which is validated in younger patients, there is little argument that patients who are identified as high risk should be transplant candidates. Patients with favourable features are unlikely to gain from transplant in first complete remission (CR1), not least because an important proportion can be salvaged if they relapse. The controversy centres around intermediate risk and boils down to how it is defined. The reason why use of this procedure in first remission is still debated is that the risk of relapse is highly variable, and because the encouraging results are distorted by the fact that patients who actually receive the transplant have been selected by surviving in remission until the transplant is performed, i.e. have an unintended selection bias. When the endpoint of published studies is only reported as disease-free survival, the whole issue of using the available transplant as salvage treatment has been ignored. This is inappropriate, so overall survival should be the endpoint on which to judge the data. In advising patients, the risk of relapse, the possibility of salvaging a patient if they do relapse from chemotherapy, and the risk of the transplant procedure itself, need to be weighed in the decision.

Autologous stem cell transplantation

Harvesting of stem cells from the bone marrow or peripheral blood during remission and using them after a period of cryopreservation for haematological rescue after myeloablative chemoradiotherapy has been widely used for younger patients who lack donors. This approach has also been shown to be a more effective way of preserving remission compared with chemotherapy. The treatment-related mortality is lower (5–10%) than with allogeneic transplantation, but it lacks a GVL effect, so the relapse risk is higher (around 35–40%). This results in an overall survival of 50–55% of those who receive this approach. Because the complications are not particularly age related, patients up to their mid-fifties can safely undergo this procedure, but the results of autologous transplantation in older patients (>60 years) are not encouraging.

Patients who receive the autograft early in remission (e.g. within 3 months) do less well than those treated at 3–6 months, because of a higher relapse rate. This may reflect patient selection, but it has also been interpreted to mean that, for an autograft to be successful, consolidation chemotherapy beforehand has an important role in cytoreduction of leukaemia cells before the marrow is harvested, so-called *in vivo* purging of disease. Initially there was concern that returning stem cells to patients would be illogical unless efforts were made to eliminate contaminating leukaemia cells first, so-called '*ex vivo* purging'. Various chemical, cellular and immunologically based techniques were used without clear evidence of benefit. Most clinical experience was gained using 'unpurged' bone marrow supporting myeloablative chemoradiotherapy, which usually comprised cyclophosphamide with total body irradiation or busulfan with cyclophosphamide. One of the problems with the use of autologous bone marrow has been delayed peripheral blood count recovery, particularly of platelets. This seems to be a feature of AML and is less obvious in other disease indications. Current practice uses peripheral blood or combines peripheral blood and marrow stem cells to ameliorate this problem. This has improved haemopoietic recovery but may be associated with an increased risk of relapse, thus giving no overall survival advantage. While autograft has fallen out of fashion as a result of the older randomized trials failing to show a survival benefit, it may well be that applying techniques to more accurately detect submicroscopic levels of disease could select patients who could preferentially benefit. This, however, needs to be prospectively tested.

Comparison of consolidation options

For patients under 55 or 60 years, all three treatment options are available, so the dilemma is which treatment approach to take. About 50% of patients entering remission will survive with chemotherapy alone. Of those who receive an allogeneic or autologous stem cell transplant, 55–60% and 50%, respectively, will survive. Patients who receive a transplant are not equivalent to patients receiving chemotherapy alone. They have survived long enough to receive the transplant, whereas some patients eligible for transplant may not have received one because of early relapse. Some studies demonstrate that 40% of patients with a donor do not receive an allograft.

Another less frequently considered option is to delay the transplant until there is disease recurrence. Primary treatment of relapse with transplantation is associated with a high rate of failure, so it is necessary to establish a second remission first. However, it is possible to salvage overall about 40 to 50% of patients who relapse from chemotherapy. Based on risk factors, it is possible to define those patients who, if they relapse, are likely to enter second remission. In this subgroup the transplant can therefore be delayed beyond CR1 because if first-line treatment does fail, a second CR can be reliably obtained and a transplant delivered. If there is a low chance of second remission, there is a stronger case for transplantation as part of first-line treatment.

Several prospective randomized trials have compared chemotherapy with autologous transplantation and allotransplantation with chemotherapy. In the latter case, these comparisons are not truly randomized, but rather compare patients found to have donors, and assumed to be intended to receive an allograft, with those for whom no donor is found. This 'donor versus no-donor' comparison is a substitute for randomization. Although the conclusions are not universal, and despite the superior ability of allograft to reduce the risk of relapse and in some studies to improve the disease-free survival, there have often been no differences in survival between these approaches. When those with donors are then compared, overall there is only a modest, but statistically significant, survival benefit in favour of allotransplantation; this may, however, be less clinically significant. When the cytogenetic risk of relapse is taken into account (discussed below), it would appear that transplantation is not required for good-risk patients, and in the absence of an emerging non-transplant improvement for poor-risk patients, allogeneic transplantation, including from unrelated donors is the chosen approach. There is a prospect that chemotherapy may continue to improve so the question of whether allotransplantation will continue to be the best option for standard-risk patients remains a matter of considerable debate, particularly with the emergence of several new prognostic markers and continuous re-definition of relapse risk. The evidence continues to evolve and the challenge is to apply it to individual patients in daily clinical practice. Part of the problem is that there is a lack of allograft data in the setting of powerful prognostic information.

Factors influencing the risk of relapse

As treatment of disease in younger patients has improved, so the heterogeneity of disease has also become apparent with respect to differences in relapse risk. On multivariate analysis a number of factors have emerged that can predict the risk of relapse, irrespective of treatment schedules used, including stem cell transplantation.

Cytogenetics

Patients with core binding factor (CBF) leukaemia, defined by the presence of t(8;21)(q22;q22)/*RUNX1-RUNX1T1* and inv(16)(p13q22)/t(16;16)(p13;q22)/*CBFB-MYH11*, have a high remission rate, lower risk of relapse and higher rate of second remission, associated with 5-year survival of 65–75%. These patients tend to have lower expression of the 'resistance proteins' and a low frequency of *FLT3* mutations (described below). Acute promyelocytic leukaemia (APL) is now regarded as a separate entity that is uniquely responsive to retinoic acid. In most cases, the disease is characterized by t(15;17)(q22;q21), which predicts sensitivity to all-*trans*-retinoic acid (ATRA) and arsenic trioxide (ATO), which achieve remission by inducing degradation of the underlying PML-RARα oncoprotein. These abnormalities are frequently associated with additional cytogenetic changes in the leukaemic blasts, such as loss of a sex chromosome and del(9q) in the case of t(8;21), trisomy 22 in inv(16) and trisomy 8 in the case of t(15;17). These additional changes do not adversely affect the favourable prognosis. Similarly, internal tandem duplication (ITD) mutations of the *FLT3* gene (*FLT3*-ITD) occur in about 35% of APL cases, and do not affect prognosis in this group of patients once presenting WBC is taken into account. Together, these good-risk patients comprise about 25% of patients under 60 years (Figure 20.2). CBF leukaemia and APL account for a smaller proportion of AML presenting in older adults (Figure 20.6), but continue to represent a more favourable group, with a survival of 35% compared with an overall survival of 15–20% in that age group in those treated with intensive chemotherapy.

About 15% of younger patients have cytogenetic abnormalities that are associated with a lower remission rate and a relapse risk on conventional chemotherapy of 85%. These include −5/del(5q), −7/del(7q), inv(3)(q21q26)/t(3;3)(q21;q26), t(9;22)

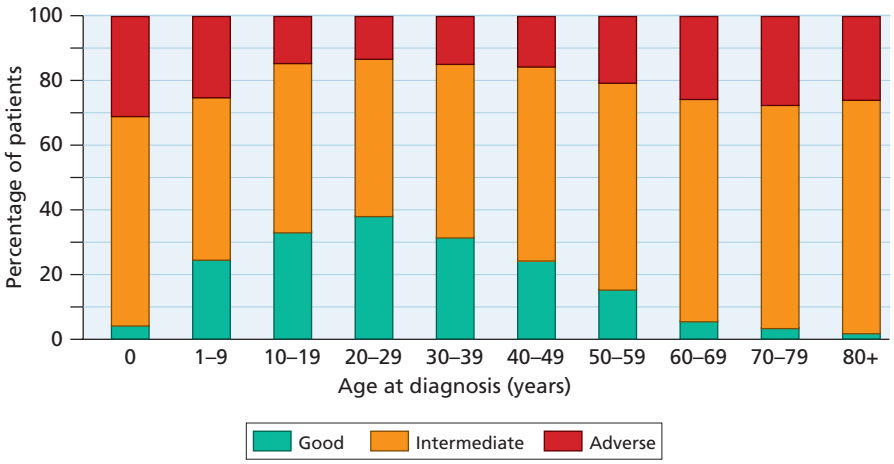


Figure 20.6 Age-specific incidence of cytogenetic risk group in MRC AML trial patients. Cytogenetic risk is defined as per Grimwade *et al.* (Blood 2010). Based on 12,000 patients registered to AML10, AML11, AML12, AML14, AML15 and AML16. (Source: Anthony Moorman and Christine Harrison, Newcastle University, UK. Reproduced with permission.).

and complex karyotype (more than three unrelated changes) in the absence of any favourable cytogenetic abnormality. Such patients need to be identified early because, even if a remission is achieved, it will be short-lived. Currently, transplantation represents the only viable treatment, but even that is associated with a high relapse risk. There is little evidence that the outlook for this high-risk group has improved over the last 20 years.

Patients who do not fall into the categories described are regarded as standard risk. They have a 5-year survival of 40–45% (Figure 20.7a). The impact of this risk stratification is apparent irrespective of chemotherapy used or whether patients receive an allograft or an autograft. There are some minor discrepancies between published series; however, analysis of over 700 children and almost 6000 younger adults from the UK MRC trials has allowed cytogenetic risk stratification to be refined in these age groups. However, risk stratification is an evolving process as the mutational landscape becomes better defined. In older patients, the overall survival of chemotherapy is around 15–20% at 5 years; it is therefore less easy to delineate cytogenetic risk subgroups. Adverse groups are more frequent, and favourable subgroups are less common. This partly accounts for the poorer prognosis of AML in older patients. It is still possible to derive a hierarchical risk stratification in the older patients based on similar criteria already described for younger patients. Some extremely poor subgroups can be identified (Figure 20.7b).

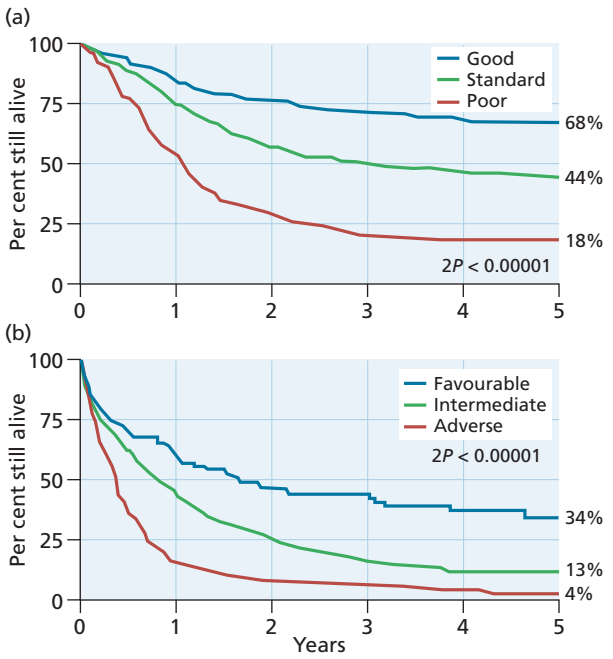


Figure 20.7 The impact of cytogenetic risk groups on treatment outcome. (a) The survival of patients aged under 60 years with good-risk abnormalities – t(8;21), inv(16), t(15;17) – with or without other abnormalities. Poor risk is associated with changes involving chromosome 5 or 7, 3q26 abnormalities or complex karyotype (more than three unrelated abnormalities). The standard-risk group comprises patients who do not have the abnormalities included in the good- or poor-risk categories. (b) The outcome for patients over 60 years old with these abnormalities who were treated with intensive chemotherapy. The relative proportion of the abnormalities in each age group is shown in Figure 20.6.

Age

Increasing age, from children to the elderly who are given intensive chemotherapy, is associated not only with a poorer chance of achieving remission, but also with an increasing risk of relapse, even allowing for obvious differences in comorbidities and distribution of cytogenetic risk groups. Leukaemias in the elderly more frequently express the drug transport proteins associated with chemoresistance (discussed below).

Response to induction chemotherapy

Patients who enter remission with the first course will have a lower relapse risk than those who require a further course. This has been recognized in various ways. For example, the presence of residual blasts in the bone marrow on day 14 can be used as a reason to give additional treatment. The blast percentage in the bone marrow assessed on recovery from the first treatment course has been shown to be highly predictive, i.e. patients who have more than 15% blasts, even though remission is subsequently achieved, will have a high relapse risk. Both cytogenetics and age are related to this response. When the morphological appearances are related to cytogenetic risk group, it is clear that, for good-risk patients, failure to clear the marrow with the first course is not an adverse feature, whereas in poor-risk patients, even those who clear blasts in the first course, will have a poor prognosis. In standard cytogenetic risk patients, it is those who fail to clear the marrow that have an adverse risk. From such data, a risk definition incorporating cytogenetics and marrow response can be obtained and this identifies those in the standard cytogenetic risk group who fail to clear the marrow and are thus poor-risk (Table 20.4). This can be further related to age, which suggests that the impact of older age is clearest in the standard-risk patients who clear marrow blasts.

FLT3-ITD mutation

Not only has *FLT3*-ITD emerged as one of the most frequent mutations in AML but several large series conclude that it is also a major prognostic factor, not for the achievement of remission but particularly for predicting relapse. While there is strong evidence to suggest that *FLT3*-ITD represents a secondary event in disease pathogenesis, its detection has become standard practice

Table 20.4 Overall survival of younger adults according to cytogenetic risk group and response to induction therapy (UK MRC AML10 trial data).

Cytogenetic group	CR (%)	PR (%)	RD (%)
Favourable	77	68	76
Standard	49	41	16
Poor	26	24	4

CR, complete remission; PR, partial remission; RD, resistant disease.

in the diagnostic work-up of AML, providing additional refinement to cytogenetic predictive groups. The question of whether *FLT3*-ITD status should be used to direct patients' treatment depends on the nature of the accompanying cytogenetic/molecular genetic features. Moreover, as discussed below, FLT3 has attracted interest as a potential therapeutic target, with several potential inhibitors under assessment, but thus far none has produced a survival benefit (possibly in part because *FLT3* mutations are secondary events in disease pathogenesis). In contrast to *FLT3*-ITD, mutations involving the activation loop (tyrosine kinase domain, TKD), which occur in 7–10% of patients, have not been shown to be adverse in all studies and may even be favourable, which might be explained by differences in the downstream signalling pathways and the pattern of coexisting mutations. *FLT3* mutation status is complicated by the requirement for rigorous definition of meaningful mutation and a knowledge of the mutant to wild-type ratio, where a high ratio has a more negative impact. Its prognostic relevance is further modified by its frequent association with other mutations, particularly NPM1, which appear to neutralize its negative impact. About 70% of *FLT3* mutated patients have other features that make them high risk anyway, and thus should be directed to stem cell transplantation on that basis alone, although it is far from clear that this provides a satisfactory cure rate.

Other molecular abnormalities

The discovery that the nucleophosmin 1 (*NPM1*) gene was mutated in a third of AML cases, including approximately half with normal karyotype, stimulated considerable effort to further subdivide the prognostic groups using molecular information. Overall, the presence of an *NPM1* mutation confers a favourable outcome. However, as already stated, it frequently coexists with a *FLT3*-ITD, where it modulates the negative impact of this mutation. In a large UK study on patients in MRC trials, the prognostic implications of the four possible combined genotypes is clear (Figure 20.8 and Table 20.5). While it is tempting to suggest that the poor outcome for patients with a *FLT3*-ITD positive/*NPM1* wild-type genotype could be improved by transplantation, so far there are scanty data to confirm that this is the case.

About 10% of cases, usually those with a normal karyotype, have a mutation of the CCAAT/enhancer-binding protein α (*CEBPA*) gene. This can occur as a single mutation in a minority of cases, which does not influence prognosis, or as a biallelic mutation, which confers a favourable prognosis in the absence of *FLT3*-ITD, equivalent to that of patients with favourable cytogenetics. Mutations of c-*KIT* have been reported in 20–30% of patients within the CBF subset of favourable cases, where it identifies a subset who have a higher risk of relapse. It is not clear that this always has a negative impact on survival since such patients remain responsive to chemotherapy. Studies are underway to find out if the addition of KIT inhibitors are of value in this population.

Mutations of the Wilms' tumour (*WT1*) gene that are also likely to represent a secondary event in disease pathogenesis have been identified, and it is generally agreed that they predict for an increased risk of treatment failure. Other mutations that

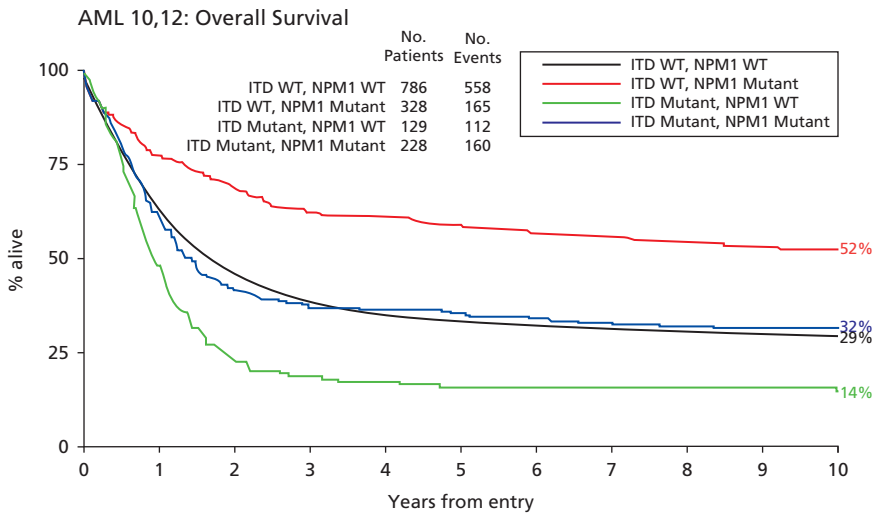


Figure 20.8 Prognostic Impact of *FLT3*-ITD and *NPM1* mutations. Source: Data from UK MRC AML10 and 12 Trials. Abbreviations: ITD, *FLT3* internal tandem duplication (*FLT3*-ITD); WT, wild type.

Table 20.5 Cytogenetic and molecular risk groups.

Risk group	Cytogenetic/molecular genetic abnormality
Favourable	t(15;17)(q22;q21)/*PML-RARA* t(8;21)(q22;q22)/*RUNX1-RUNX1T1* inv(16)(p13q22)/t(16;16)(p13;q22)/*CBFB-MYH11* *NPM1* mutation (in absence of *FLT3*-ITD or *DNMT3A* mutation) Biallelic *CEBPA* mutation
Intermediate	Cytogenetic/molecular genetic abnormalities not classified as favourable or adverse
Adverse	*In the absence of favourable risk cytogenetic/molecular genetic abnormalities*: abn(3q) [excluding t(3;5)(q21~25;q31~35)/*NPM1-MLF1*], inv(3)(q21q26)/t(3;3)(q21;q26)/ *GATA2/EVI1* add(5q)/del(5q), −5 t(5;11)(q35;p15.5)/*NUP98-NSD1* t(6;9)(p23;q34)/*DEK-NUP214* add(7q)/del(7q), −7 t(11q23) [excluding t(9;11)(p21~22;q23) and t(11;19)(q23;p13)] t(9;22)(q34;q11)/*BCR-ABL* −17/abn(17p)/*TP53* mutation Complex karyotype (≥4 unrelated abnormalities) *ASXL1* mutation *DNMT3A* mutation *FLT3*-ITD *MLL*-PTD *RUNX1* mutation

have been identified as independent predictors of adverse outcome include those involving *ASXL1, DNMT3A* and *RUNX1*, as well as partial tandem duplications of the *MLL* gene (*MLL*-PTD). The use of gene over-expression is a little more problematic because of the need to standardize methodology, but a number of studies have reported that higher expression of *EVI1, ERG, MN1, CXXC5, INPP4B* and *BAALC* is associated with poorer survival.

The addition of the molecular subdivision to the prognostic assessment adds considerable complexity for prospective validation and for assessing whether any specific treatment is beneficial, or indeed whether the molecular characteristics add prognostic information to that which is already available, or simply identifies the same patient groups. While current diagnostic algorithms take into consideration only a very limited number of molecular markers, the seminal TCGA consortium study in which the genomes of 200 AML cases were analysed by high-throughput sequencing highlights the challenge ahead. In this study, the average case of AML harboured over 10 mutations and 23 genes were found to be recurrent mutation targets, with over 1000 genes involved overall. Apart from involvement of myeloid transcription factors, signaling molecules, *NPM1*, tumour suppressors (e.g. *TP53*) and epigenetic modifiers, this study highlighted that genes encoding components of the splicing machinery and cohesin complex are recurrent mutation targets in AML. The 'holy grail' is to match up the many molecular subgroups with particular therapies, but this will need huge numbers of fully characterized patients to be statistically robust.

Performance score

In older patients, standard assessments of performance vary considerably. These scores are highly predictive for induction treatment success, and will also inevitably relate to which patients are considered candidates for an intensive approach to treatment.

White cell count

High tumour load is an adverse feature for both induction and relapse risk. The threshold for risk is not definitive. A white cell count of 50×10^9/L is often quoted. In subgroups, a prognostic influence is apparent at much lower counts, for example in APL a white cell count at diagnosis of greater than 10×10^9/L is usually agreed as predictive of a higher relapse risk. The impact of a high white cell count is less than that of cytogenetics, but may have isolated value when cytogenetics is not available.

Resistance proteins

One of the important biological differences between AML presenting in older and younger patients, is the increased frequency in the former of expression of proteins involved in drug transport. These are associated with chemoresistance to some of the drugs used in AML, such as anthracyclines and etoposide. Expression also tends to be associated with a stem cell phenotype and adverse cytogenetics. The most widely studied is P-glycoprotein (P-gp), an energy-dependent transporter protein product of the *MDR1* gene on chromosome 7 and which belongs to the ATP-binding cassette (ABC) transporter family. Older patients with AML frequently over-express this protein, which has been shown to be predictive for inferior rates of remission and remission duration. The expression levels reported may vary in different series because of differences between measurement techniques. Quantitative flow cytometry has brought a degree of consistency to measurement, but a functional assessment (of dye efflux) and blockade with a P-gp inhibitor is also recommended. *In vitro* preclinical studies have demonstrated that agents such as ciclosporine or its analogue PSC-833 can block P-gp function; however, only one trial to date has so far shown that using such agents has clinical benefit. This may be because P-gp is

not the only mechanism of chemoresistance. Multidrug resistance protein 1 is another member of the ABC transporter family and the gene (*MRP1*) is located on chromosome 16. Lung resistance protein (LRP) is a subunit of the major vault protein, which has been identified in some anthracycline-resistant cell lines and appears to be involved in drug transport. LRP expression has been reported in 30–50% of AML cases in different series, but the majority of studies have been unable to show a correlation with response. Above-normal MRP expression can be found in 50% of patients with untreated AML, but there is no consensus about its impact on survival.

Detection of measurable/minimal residual disease (MRD)

Modern treatments of AML ever more frequently apply the concepts of risk-adapted intensification. An additional, equally crucial, objective is to deliver effective therapies that avoid under- or over-treatment; achieving this has prompted precise measurements of persisting disease, or measurable/minimal residual disease (MRD), beyond the threshold of light microscopy. A significant correlation has been shown between MRD quantified by RT-qPCR or multiparameter flow cytometry (MFC-MRD) and outcome. Therefore, the assumption behind the assessment of MRD is that it will add greater definition to risk prediction by the prognostic factors already recognized. In particular, four major areas of application to the post-remission decision-making process have been identified and are being evaluated: (i) assessment of the quality of response to improve risk stratification and enhance postconsolidation choice (e.g. transplant versus no transplant); (ii) post-treatment serial monitoring to anticipate overt relapse and plan pre-emptive therapy; (iii) pre-stem cell transplant assessment and post-transplant surveillance monitoring to assess the risk of relapse, inform immunosuppression and decide the use of donor lymphocyte infusions; and (iv) identification of markers to track treatment response in the context of targeted-therapy approaches.

Because of the relative insensitivity of techniques such as cytogenetics and fluorescence in situ hybridization, molecular and flow cytometry approaches have become the most popular laboratory tests for detecting residual disease. Approximately 60% of children and younger adults have an informative leukaemia-specific marker (i.e. fusion gene, *NPM1* mutation) amenable to MRD detection by real-time quantitative PCR (RT-qPCR). This technology is more robust, generally affords superior sensitivity and is more readily standardized than the original qualitative end-point RT-PCR assays that were used in early MRD studies and which have now been superseded. Before follow-up samples can be assessed, analysis of diagnostic material is essential to establish the appropriate assay to use in each patient (due to variability in chromosomal breakpoints/type of *NPM1* mutation). The level of the leukaemic transcript is routinely normalized to the level of a housekeeping gene (e.g. Abelson, *ABL1*) in each sample. This allows the trend in leukaemic transcripts to be determined at different time points, while controlling for variation in sample quality and processing, including RNA extraction and complementary DNA synthesis. Since these assays depend upon RNA, which degrades rapidly, to maximize the capacity for MRD detection, steps should be taken to ensure that samples reach the laboratory rapidly (ideally within 24 hours). The maximal sensitivity that can be achieved is not only influenced by sample quality, but is also dependent upon the relative expression of the MRD target as compared to the housekeeping gene in leukaemic blasts. This varies according to subtype of AML and between patients, with RT-qPCR assays typically being capable of detecting 1 AML cell in 10^4 to 10^5 normal bone marrow cells. Analysis of peripheral blood is typically associated with 1-log poorer sensitivity.

Molecular MRD monitoring has been most extensively studied in APL, in which assays for the *PML-RARA* fusion are relatively insensitive (typically 1 in 10^4). In this subtype of AML, MRD assessment is included within the standard response criteria, since achievement of molecular remission in the bone marrow by the end of consolidation is a prerequisite for cure. While the majority of patients (~95%) achieve this milestone, approximately 10–25% will ultimately relapse following ATRA and anthracycline-based protocols. The majority of relapses can be predicted by re-emergence of *PML-RARA* transcripts detected in sequential bone marrow MRD monitoring samples, which are usually performed 3-monthly, taking into account assay sensitivity and kinetics of relapse. Early studies that predated availability of arsenic trioxide (ATO) suggested a survival advantage for early treatment intervention for molecular (subclinical) relapse as compared to salvage at frank relapse, due to decreased risk of bleeding and other complications. However, treatment protocols for APL have steadily improved, so for patients with low-risk disease with rapid achievement of molecular remission there is little benefit in further monitoring. Whereas, currently, sequential MRD monitoring until 3 years post treatment is recommended in patients who present with 'high-risk' disease (defined by presenting WBC $>10 \times 10^9$/l), due to the 25% risk of relapse.

A number of studies using RT-qPCR have been conducted in patients with CBF leukaemia and NPM1 mutant AML. These have consistently shown that patients who have a slow kinetic of response, as measured by log-reduction in disease transcripts are at significantly increased risk of disease relapse. Moreover, this information provides a more powerful predictor of outcome, as compared to the mutational profile defined at diagnosis. On this basis, MRD assessment is increasingly being used within clinical trials to refine risk stratification and inform consolidation therapy. In addition, just as in APL, impending relapse can be predicted in AML with a leukaemic fusion gene (e.g. *RUNX1-RUNX1T1*, *CBFB-MYH11*, *DEK-NUP214* etc.) or *NPM1* mutation, based upon persistence of a significant level

of MRD following frontline therapy or by recurrence of PCR positivity with a rising transcript level after an initial molecular response. Studies are currently in progress to determine whether early intervention for molecular relapse is beneficial as compared to treating patients once they are in overt clinical relapse.

Quantification of MRD by flow cytometry may be based on the identification at diagnosis of leukaemia-associated aberrant immunophenotypes (LAIPs). These are absent or very infrequent in normal bone marrow but, using a large panel of monoclonal antibodies, can be described in at least 85% of AML cases at diagnosis. These LAIPs include over-expression of an antigen, coexpression of antigens normally associated with different stages of maturation, but which does not occur in normal haemopoiesis, the absence of myeloid antigen expression or the expression of non-myeloid antigens. An alternative strategy for detection of MRD that can be used when no diagnostic sample is available is a 'different from normal' approach, using a defined panel of antibodies applied to distinguish leukaemic from normal cells, even in the context of marrow regeneration. This methodology is less prone to 'false negative' results, which can be encountered on sequential tracking of LAIPs due to phenotypic switches that can relate to outgrowth of minor subclones giving rise to disease relapse.

The need to test as comprehensive a panel of monoclonal antibodies as possible makes flow-cytometry based MRD monitoring a quite expensive technology. However, the major impulse for using it derives from its applicability to the vast majority of AML cases, with a sensitivity of 1 in 10^4 or 1 in 10^5. Although the sensitivity generally falls below that of molecular techniques, improvements are expected as more colours become available. Since most published data have been based in single laboratories, additional concerns pertain to lack of common standard operating procedures in order to generate comparable results. From a clinical point of view, several published studies have demonstrated that the immunophenotypic detection of MRD, post induction or post consolidation, is independently associated with the risk of relapse. There is also evidence that in MRD-positive patients, the use of allogeneic stem cell transplantation confers a superior outcome, whereas autologous stem cell transplantation does not alter the unfavourable course dictated by MRD positivity. MFC-MRD may also be useful to predict outcome following allogeneic transplant.

The contribution of molecular biology and flow cytometry to the delivery of MRD-directed therapy in AML may be relevant. In fact, given the broad applicability of flow cytometry and the ever rising number of molecular targets identified, one can expect that virtually every patient with AML will be suitable for MRD monitoring and then for individualized management of disease. There is also demand for establishing whether determination of MRD can enhance current risk-stratification strategies based on pretreatment parameters. All these issues can be addressed in the context of large prospective and cooperative studies including parallel determination of MRD by RT-qPCR and flow cytometry. Further refinement may be possible as leukaemic stem cell populations become better defined, which can be tracked by flow cytometry. With increasing availability of mutational profiling, suitable molecular targets can be readily identified. High throughput sequencing technologies and digital PCR provide promising platforms to assess changes in mutational burden. However, optimal application of such approaches will require more detailed understanding of which mutations are initiating lesions and best track the leukaemic clone in order to reliably distinguish which patients are likely to be cured following frontline therapy from those destined to relapse.

Impact of prognostic factors on treatment choice

It is becoming routine to take into account the risk of relapse, as defined by some of the factors described, in order to target treatment. The most obvious example is the growing acceptance that transplantation is not required for patients with good-risk disease. Poor-risk patients must be identified promptly and offered either transplantation or some experimental approach, since currently available chemotherapy is inadequate. Further data regarding whether transplantation (sibling or unrelated) significantly benefits high-risk patients are needed. Similar information is also required about the significance of *FLT3* status and associated genotypes with respect to the impact of transplantation. Children respond very well to intensive chemotherapy, with the majority enjoying prolonged remissions. Only the small number of children with high-risk features require first-line transplantation.

In older patients (>60 years), either patients or doctors make a judgement at diagnosis as to whether intensive chemotherapy will be beneficial. Prognostic factors such as cytogenetics and performance score can inform this choice. Very poor-risk cytogenetics (e.g. complex changes) carries such a low prospect of success that the question arises as to whether such patients should receive intensive treatment, even if considered sufficiently fit.

Acute promyelocytic leukaemia

APL is a special case in which the presence of t(15;17)/*PML-RARA* predicts sensitivity to treatment with all-*trans* retinoic acid (ATRA) and arsenic trioxide (ATO). It has been recognized for more than 20 years that this leukaemia subtype is particularly sensitive to anthracyclines. Recent experience has clearly demonstrated that the combination of ATRA and chemotherapy has made a dramatic improvement, with survival now expected to exceed 80%. Even better prospects are becoming apparent when the combination of an anthracycline

(idarubicin) and ATRA form the backbone of treatment. Simple maintenance with courses of ATRA and orally available agents such as methotrexate and 6-mercaptopurine has been shown in some studies to provide additional benefit, but the role of maintenance, particularly in patients who are molecularly negative after consolidation (who account for 95% of cases) is now in doubt. Maintenance is associated with toxicity and there are concerns that it may contribute to an increased risk of therapy-related MDS/AML, so is becoming less widely used. Since the molecular consequences of t(15;17) are known, evaluation of the role of molecular monitoring is most developed in APL and provides a model for disease monitoring in AML. It has recently been demonstrated that in low- and intermediate-risk patients, the combination of ATRA and arsenic trioxide alone can produce very impressive survivals of over 95%. This has been confirmed in a similar UK trial which used less frequent arsenic dosing and included high risk patients. This has been confirmed in a similar UK trial which used less frequent arsenic dosing and included high risk patients. This changes the landscape of treatment. This approach is feasible in high-risk patients (defined as those with presenting white counts above 10×10^9/L), particularly if combined with a brief exposure to gemtuzumab ozogamicin or idarubicin. To date, arsenic trioxide is not licensed for first-line treatment.

The likelihood is that standard care, at least for the 75% of low-risk patients, will be the 'chemotherapy-free' combination of arsenic trioxide and ATRA. This will become even more attractive if an oral formulation of arsenic becomes available. However, this will be associated with the differentiation syndrome in some cases. In current standard of care this is not frequent because the chemotherapy reduces the risk when combined with ATRA. The syndrome is accurately defined by the presence of: unexplained fever, weight gain, respiratory distress, interstitial pulmonary infiltrates, and pleural or pericardial effusion, with or without hyperleucocytosis. No single sign or symptom itself may be considered diagnostic of the syndrome. However, the earliest manifestations of suspected differentiation syndrome (e.g. unexplained respiratory distress), and prior to the development of a fulminant syndrome, treatment with dexamethasone 10 mg i.v. 12-hourly should be initiated and continued until disappearance of symptoms and signs, and for a minimum of 3 days.

Treatment in the older patient

Improvement in survival in older patients over the last 20 years has been much more elusive. With better supportive care, intensive chemotherapy can expect to achieve remission in 50–60% of cases. However, 80% of cases will relapse by 2 years. This result has been achieved with various combinations of induction and consolidation schedules and is not improved by maintenance to any great extent. There will be a greater interest in maintenance in future studies. Because the outcome is poor, two issues arise. First are there prognostic factors that confirm which patients will benefit from an intensive treatment approach? The data are less convincing than in younger patients, but younger age (60–70 years), higher performance score and favourable cytogenetic risk group can identify a minority of patients with a better than average outcome. However, a significant proportion of patients have adverse factors: older age, poorer performance score, complex cytogenetics or a chemoresistant phenotype. These patients will have a worse prognosis, which raises the second issue of whether they should receive palliative care from the start. One modestly sized study compared a palliative treatment approach with conventional chemotherapy and demonstrated that the use of intensive chemotherapy, because it achieved remission, was more beneficial in older patients; however, no study has yet been large enough to ask that question within the risk groups.

One strategy for improvement is to target the function of P-gp. Only one of several studies using ciclosporine or its analogue has managed to improve survival in relapsed disease, but other studies combining it with first-line treatment have been unsuccessful. This may be because P-gp is not the only resistance mechanism present in leukaemic cells, and once a cell has become resistant by one mechanism there are already other resistance routes. As previously mentioned, the meta-analysis data examining the addition of gemtuzumab ozogamicin to induction treatment provides a small, but significant, survival benefit.

There is an important population of older patients who are not considered fit for a conventional therapy approach. While most physicians will readily recognize such patients, it is more difficult to define them based on patient/disease characteristics. The patients will be older, may have comorbidity, secondary disease or other adverse features. Short of a randomized comparison, which has not been undertaken in the modern era, it is not possible to be sure that such patients would not be better off with conventional chemotherapy. In any event, these patients exist, and previously they would have been offered best supportive care. A key randomized trial showed that low-dose Ara-C (LDAC) given as 20 mg twice a day for 10 days, was superior to best supportive care, importantly without increased toxicity or supportive care. This is, however, inadequate treatment and only appears to benefit the 15–20% of patients who enter CR. More modern alternatives are demethylation agents (azacitidine or decitibine). These are less likely to deliver remission, but can stabilize disease in some patients; however, randomized trials have not shown statistical survival superiority over the LDAC schedule as described.

Management of relapse

The majority of patients will relapse. If this happens after stem cell transplantation, the benefit of further therapy is questionable. However, this is dependent on when the relapse occurs.

Within 1 year, further treatment is unlikely to have sustained benefit and re-transplantation is usually associated with a very high complication rate. If the relapse occurs later, further chemotherapy with re-transplantation may save a few patients. The development of donor lymphocyte infusions (DLI) has been a very effective approach for the treatment of post-allograft relapse in chronic myeloid leukaemia (CML), but has a low rate of success in AML. However, DLI may be more effective if given in the context of MRD rather than full-blown relapse, as has already been shown to be the case in CML.

For patients who relapse after chemotherapy, three factors dictate the clinical outcome: duration of first remission, age and molecular/cytogenetic risk group. Patients with good-risk disease have a high (75–80%) rate of second remission. Patients who are young with a long CR1 will have a reasonable survival, whereas older patients with a short CR1 will do poorly (Table 20.6). Since the second remission rate in good-risk disease defined by cytogenetics is relatively good, transplantation is usually delayed until second remission. There is no clear 'best choice' chemotherapy for other risk groups, so this is often the setting for experimental therapy development.

APL is again a special case. In patients with suspected relapse of APL it is important to confirm *PML-RARA* positivity, to exclude the possibility of therapy-related MDS/AML and to identify patients likely to benefit from molecularly targeted therapy. Patients can respond again to retinoic acid and chemotherapy, but recently arsenic trioxide and the CD33-directed immunoconjugate gemtuzumab ozogamicin have been found to be effective. Remission rates above 80% have been reported; interestingly, following consolidation therapy, a similar proportion can be returned to RT-qPCR negativity. This permits the opportunity for autologous transplantation, the successful outcome of which depends on the graft and, preferably, the patient being molecularly negative. For patients who remain molecularly positive following salvage, allogeneic transplantation is indicated. Arsenic trioxide is now the standard first-line salvage approach in *PML-RARA+ APL*; indeed some relapsed patients can be cured with ATRA + ATO protocols, raising questions as to whether consolidation with an autograft is necessary in all patients who go back into molecular remission.

For non-APL patients, whatever treatment is used to re-establish remission, it is unlikely to be durable without a transplant. Although prospective studies are rare, transplant registry data suggest that about 30% of patients can be salvaged with a transplant with little overall difference whether the source of stem cells is allogeneic or autologous.

Table 20.6 Outcome of relapse in patients over 60 years receiving re-induction treatment ($N = 1529$).

	CR1 duration (months)		
	<6	6–12	>12
Remission rate			
15–59 years (%)	15	33	56
60–69 years (%)	11	29	67
>70 years (%)	13	26	53
Survival from relapse at 2 years			
<35 years (%)	10	16	41
35–60 years (%)	7	14	27
60+ years (%)	4	8	16

Future developments

Classification

The classification of this disease will no doubt continue the trend of taking into account molecular, genetic and clinical features, as well as morphology. Gene expression-based diagnostics are now available to distinguish prognostically relevant molecular subtypes of AML. Moreover, advances in sequencing technology (see Chapter 19) provide increasing opportunity for mutational profiling of large panels of genes to improve disease categorization and inform patient management. The bioinformatics challenges are considerable given the enormous amount of information that these methods produce.

Therapeutics

There is general acceptance that little further progress will be made by simply shuffling currently available drugs with respect to either scheduling or dosage. There may still need to be refinements with respect to toxicity. There is much interest in targeting treatment, either by matching the treatment approach to the patient, based on prognostic factors, or by immunologically directing treatment to leukaemic cells and thereby enhancing the selectivity of treatment. Gemtuzumab ozogamicin is an immunotoxin that is being extensively explored in this respect. This is an immunoconjugate combining an IgG4 anti-CD33 humanized monoclonal antibody with the highly potent antitumour antibiotic calicheamicin. The key to its utility is that when the antibody combines with CD33 antigen, the complex is rapidly internalized to the cell, where the chemical linker between drug and antibody is lysed. A crucial property is that the linker is lysed only intracellularly and not in the circulation. Although expressed in 90% of cases of AML, CD33 is not leukaemia specific. There is expression on haemopoietic precursors, but not on stem cells or, as far as is known, other tissues. The conjugate is clearly active. As a single agent, it can achieve CR in relapse or as first-line treatment in older patients. Pancytopenia is not avoided and transient hepatotoxicity will be seen in some patients. It does not result in the alopecia or mucositis usually

associated with chemotherapy. Its use in AML is under investigation in a number of areas, for example in induction in older patients before chemotherapy, for induction in patients unfit for chemotherapy; as maintenance of remission; as first-line and relapse treatment in APL, as part of transplant conditioning and in simultaneous combination with conventional chemotherapy. All these approaches are still experimental.

Allogeneic transplantation has proved over the years to be a highly effective immunological approach. However, as previously pointed out, the overall survival advantage is not always clear. Part of the reason is that because it is only safely applicable to younger patients, it is competing with the group of patients with the most favourable responses to chemotherapy. Non-intensive transplants have demonstrated that it is feasible to achieve full chimeric status, i.e. 100% donor cells in the bone marrow, without using intensive chemoradiotherapy. This can be done safely in older patients, but there remain concerns about the balance between avoiding relapse on the one hand and GVHD on the other. This could represent a consolidation option for older patients for whom conventional chemotherapy is less successful and where there is an antileukaemic effect of standard allograft. The approach presumes that there will be a significant GVL effect operating in AML. There are only preliminary data on this approach in AML, which is still experimental; assessment in a prospective clinical trial is required.

Many small molecules, particularly tyrosine kinase inhibitors, are becoming available for cancer treatment. So far none has matched the impact of imatinib in CML. The recognition that *FLT3* is commonly mutated in AML has led to the discovery of several powerful, but not specific, inhibitors of the receptor tyrosine kinase. Preclinical models provide considerable encouragement for efficacy. Initial clinical studies show a response in about 50% of patients with relapsed disease, with only occasional CRs; the duration of any responses tended to be short. Because of lack of specificity, it is likely that these agents will need to be used in combination with each other, or with chemotherapy. Inhibitors of RAS pathway molecules have also undergone preliminary assessments. Some responses have been seen, but it is also clear that the agents tested are not specific. Clinical trials are in progress with a range of inhibitors of relevance to *MLL*-rearranged leukaemias, targeting bromodomain and extra-terminal (BET), hDOT1L and CDK6. In addition, trials are underway with inhibitors of mutant isocitrate dehydrogenase 1 and 2 (IDH1, IDH2, collectively mutated in 15% of AML), which are showing promising early results.

Newer, more conventional agents (e.g. clofarabine, vosaroxin, sapacitibine, cladribine and demethylation agents) hold promise from non-randomized trials, but none has been successful in randomized trials. Since many new treatments will be available, novel approaches to clinical trial design will need to develop to make more rapid progress. Much greater international collaboration is needed to provide sufficient numbers of the patient subgroups, or different statistical methods will be required.

Selected bibliography

Burnett AK (2012) Treatment of acute myeloid leukemia: are we making progress? *Hematology American Society of Hematology Education Program* **2012**: 1–6.

Burnett A, Wetzler M, Löwenberg B. (2011) Therapeutic advances in acute myeloid leukemia. *Journal of Clinical Oncology* **29**(5):487–94.

Cancer Genome Atlas Research Network (2013) Genomic and epigenomic landscapes of adult de novo acute myeloid leukemia. *N Engl J Med* **368**(22): 2059–74.

Dohner H, Estey EH, Amadori S *et al*. (2010) Diagnosis and management of acute myeloid leukemia in adults: recommendations from an international expert panel, on behalf of the European LeukemiaNet. *Blood* **115**: 453–74.

Grimwade D, Freeman SD (2014) Defining minimal residual disease in acute myeloid leukemia: which platforms are ready for 'Prime Time'? *Blood* **124**(23): 3345–55.

Grunwald MR, Levis MJ (2013) FLT3 inhibitors for acute myeloid leukemia: a review of their efficacy and mechanisms of resistance. *International Journal of Hematology* **97**: 683–94.

Grimwade D, Hills RK, Moorman AV *et al*. (2010) Refinement of cytogenetic classification in acute myeloid leukemia: determination of prognostic significance of rare recurring chromosomal abnormalities amongst 5,876 younger adult patients treated in the UK Medical Research Council trials. *Blood* **116**(3): 354–65.

Grimwade D, Ivey A, Huntly BJP (2015) Molecular landscape of acute myeloid leukemia in younger adults and its clinical relevance. *Blood*, in press.

Hills RK, Castaigne S, Appelbaum FR *et al*. (2014) Addition of gemtuzumab ozogamicin to induction chemotherapy in adult patients with acute myeloid leukaemia: a meta-analysis of individual patient data from randomised controlled trials. *Lancet Oncology* **15**(9): 986–96.

Ley TJ, Ding L, Walter MJ *et al*. (2010) DNMT3A mutations in acute myeloid leukemia. *New England Journal of Medicine* **363**: 2424–33.

Lo-Coco F, Avvisati G, Vignetti M *et al*. (2013) Retinoic acid and arsenic trioxide for acute promyelocytic leukemia. *N Engl J Med* **369**:(2) 111–21.

Sanz MA, Tallman MS, Lo-Coco F (2005) Tricks of the trade for the appropriate management of newly diagnosed acute promyelocytic leukemia. *Blood* **105**: 3019–25.

Terwijn M, Zeijlemaker W, Kelder A I*et al*. (2014) Leukemic stem cell frequency: a strong biomarker for clinical outcome in acute myeloid leukemia. *PLoS One* **9**(9): e107587.

CHAPTER 21

Adult acute lymphoblastic leukaemia

21

Clare J Rowntree[1] and Adele K Fielding[2]

[1]Department of Haematology, University Hospital of Wales, Heath Park, Cardiff, UK
[2]Department of Haematology, Cancer Institute, UCL, London, UK

Introduction

Acute lymphoblastic leukaemia (ALL) is an aggressive malignancy of T or B lymphocytes. The onset of ALL is usually rapid, with patients presenting with signs and symptoms of bone marrow failure. Occasionally patients will present with primary involvement of nodal or extranodal sites, with minimal evidence of bone marrow involvement. In these cases the diagnosis is one of lymphoblastic lymphoma (LL). The distinction between ALL and LL is arbitrary. If a patient presents with a mass lesion and less than 25% blasts in the marrow then the diagnosis is one of LL. Where marrow blasts total more than 25% the diagnosis is of leukaemia.

ALL is primarily a disease of childhood, with 75% of cases occurring in children under the age of 10 years at diagnosis. However, although ALL is rare in adults, there is an increasing incidence with age after the age of 40 years. The aetiology of ALL is currently unknown.

A new presentation of ALL is a clinical emergency and should be diagnosed and treated without delay. Most patients present with a combination of symptoms resulting from bone marrow failure, including fatigue and weakness due to anaemia, haemorrhagic complications of thrombocytopenia and/or signs and symptoms of recurrent infection due to neutropenia. Approximately 10% of patients will have obvious signs of organ involvement at diagnosis with splenomegaly, liver impairment, lymphadenopathy, testicular swelling or meningeal syndrome with neurological symptoms. Patients with T-cell disease may have large mediastinal masses and can present with features of superior vena cava obstruction.

Diagnosis of adult ALL

Morphology

In adults the blood count will usually be very abnormal, with features of bone marrow infiltration. There is frequently a normochromic normocytic anaemia. The white cell count may be increased (59% of cases), normal (14%) or decreased (27%), and there is usually a thrombocytopenia, with one-third of patients having a platelet count of less than 25×10^9/L at diagnosis. Examination of a blood film reveals circulating lymphoblasts in over 90% of cases. Approximately 15% of patients will have a very elevated white cell count above 100×10^9/L and very occasionally the white cell count can be over 500×10^9/L. It is commoner to see very high white cell counts in T-cell disease rather than B-cell disease. Most adults will be anaemic at diagnosis, although severe anaemia < 80 g/L is unusual.

A bone marrow aspiration is mandatory for the diagnosis of ALL. The bone marrow will be hypercellular with a marked infiltration of leukaemic blasts, which amount to over 25% of total nucleated cells. The majority of cases will have more than 50% infiltration by leukaemic blasts. In some cases the marrow may be difficult to aspirate due to increased reticulin and a trephine biopsy should be performed. Blastic transformation of chronic myeloid leukaemia should be ruled out by morphologic examination where possible (presence of basophilia, prominent granulocyte component and dwarf megakaryocytes in the bone marrow). However, the distinction between de novo Ph+ ALL and CML in lymphoid blast phase by morphology can be very challenging. Molecular studies may be helpful, i.e. the presence of

Postgraduate Haematology, Seventh Edition. Edited by A Victor Hoffbrand, Douglas R Higgs, David M Keeling and Atul B Mehta.

the p190 *BCR-ABL1* transcript strongly suggests de novo ALL, although the finding of the p210 *BCR-ABL1* transcript is less helpful as it can occur in both scenarios.

Lymphoblasts in both B-cell and T-cell ALL vary from small blasts with scant cytoplasm, condensed nuclear chromatin and indistinct nucleoli to larger cells with moderate amounts of light blue to blue-grey cytoplasm, occasionally vacuolated, dispersed nuclear chromatin and multiple variably prominent nucleoli. Historically, the FAB classification was used to classify ALL blasts based on morphological appearances and size of blasts. This classification was of no relevance prognostically and included mature B-cell disease (Burkitt lymphoma), which has now been removed from classifications of ALL. Burkitt lymphoma in leukaemic phase can clearly be defined by its different immunophenotype (see below) and is now treated on different protocols to ALL.

The current World Health Organisation (WHO) classification, shown in Table 21.1, defines ALL and acute lymphoblastic lymphoma together, grouped by their cell of origin, i.e. B cell or T cell. B-cell disease is then further subclassified, based on the presence of recurrent cytogenetic abnormalities that are known to have prognostic value clinically (Table 21.1).

Cell-surface marker analysis

In diagnosing ALL, the bone marrow – or peripheral blood if the count is high – should be examined using a panel of monoclonal antibodies to T-cell-associated and B-cell-associated antigens, which will identify most cases of ALL. Aberrant expression of myeloid antigens is not uncommon and should not detract from the correct diagnosis. Both B-cell and T-cell ALL are divided into subtypes based on the pattern of cell-surface markers present on the leukaemic blast cells. These subtypes are defined according to their stage of differentiation, although aberrant or asynchronous antigen expression is common. The European Group for the Immunological Characterization of Acute Leukaemia (EGIL) has defined a unified classification for ALL based only on phenotypes. The World Health Organisation (WHO) also classifies ALL according to genetic entity – see Cytogenetics below.

Table 21.1 WHO classification of ALL.

B-lymphoblastic leukaemia/lymphoma
B-lymphoblastic leukaemia/lymphoma NOS
• B-lymphoblastic leukaemia/lymphoma with recurrent genetic abnormalities
• B-lymphoblastic leukaemia/lymphoma with t(9;22)(q34;q11.2); *BCR–ABL*1
• B-lymphoblastic leukaemia/lymphoma with t(v;11q23);*MLL* re-arranged
• B- lymphoblastic leukaemia/lymphoma with t(12;21)(p13;q22) *TEL–AML1 (ETV6–RUNX1)*
• B-lymphoblastic leukaemia/lymphoma with hyperdiploidy
• B-lymphoblastic leukaemia/lymphoma with hypodiploidy
• B-lymphoblastic leukaemia/lymphoma with t(5;14)(q31;q32) *IL3–IGH*
• B-lymphoblastic leukaemia/lymphoma with t(1;19)(q23;p13.3); *TCF3–PBX1*
T-lymphoblastic leukaemia/lymphoma

B-lineage ALL

The vast majority of cases of B-ALL express HLA-DR, terminal deoxynucleotidyltransferase (TdT) and CD19. CD79a, CD22 and CD34 are frequently – but not always – expressed and levels can vary. B-ALL accounts for approximately 75% of adult cases of ALL and is subdivided into the following groups:

- *Pro-B-ALL* represents approximately 10% of adult ALL; CD10 is negative, as is cytoplasmic immunoglobulin.
- *Common ALL* (c-ALL) is the major immunological subtype in adult ALL, comprising more than 50% of all cases; c-ALL is characterized by the presence of CD10 and absence of cytoplasmic immunoglobulin.
- *Pre-B-ALL* is characterized by the expression of cytoplasmic immunoglobulin, which is absent in c-ALL, but is identical to c-ALL with respect to the expression of all other cell markers.
- *Mature B-cell ALL* is found in approximately 4% of adult ALL patients and is now classified as Burkitt lymphoma rather than ALL. The blast cells are negative for TdT and express surface antigens of mature B cells, including surface membrane immunoglobulin.

T-lineage ALL

T-cell ALL accounts for approximately 25% of adult ALL. Cells are TdT+, in addition to cytoplasmic CD3+ and CD34+. All cases express the T-cell antigen gp40 (CD7) and they may, according to their degree of T-cell differentiation, express other T-cell antigens, for example the E rosette receptor (CD2) or the cortical thymocyte antigen T6 (CD1). A minority of T-cell ALL blast cells express CD10 together with T-cell antigens. In most cases of T-cell ALL, one or more of the T-cell receptor (TCR) genes are re-arranged. These properties make it possible to classify T-cell ALL according to their stage of intrathymic differentiation:

- *Early T-precursor ALL* accounts for 6% of adult ALL and is associated with inferior outcomes compared to other types of T-cell ALL; it shows characteristic T-cell markers (cyCD3 and CD7), but no further differentiation markers.
- *Thymic (cortical) T-ALL* is the most frequent subtype of T-ALL (10% of ALL). It is characterized by the expression of CD1a; surface CD3 may be present. Since this subtype is associated with a better prognosis, its identification is of particular importance.
- *Mature T-ALL* has a frequency of 6%; the blast cells do not express CD1a, but they are positive for surface CD3.

Cytogenetics

As with AML, cytogenetic abnormalities in ALL are independent prognostic variables and have now been incorporated into the classification of B-cell ALL by the WHO. Cytogenetic analysis is a mandatory part of the clinical diagnosis of ALL and should include both examination of metaphases and fluorescence *in situ* hybridization (FISH) with specific probes, e.g. for *BCR-ABL1* and *MLL-AF4*. Screening by reverse transcriptase polymerase chain reaction (RT-PCR) for the potential *BCR-ABL1* transcripts, p190 and p210, should also be performed. Studies report detection of clonal cytogenetic abnormalities in up to 85% of cases of adult ALL.

Up to 70% of children between 1 and 10 years presenting with ALL will have good-risk cytogenetics with the presence of either hyperdiploidy (between 51 and 65 chromosomes) or the translocation t (12;21)(p12;q22), resulting in the fusion product TEL-AML1 (see Chapter 23). 'Good-risk' cytogenetic abnormalities are rare in adult ALL and decrease with increasing age. However, when good-risk aberrations are detected at diagnosis in adults they carry the same prognostic value as they do in children. Conversely, the incidence of poor-risk cytogenetics, such as hypodiploidy and Philadelphia-positive ALL generally rises with age, although there appears to be a plateau in incidence above the age of 50 years.

The poor prognostic impact of the finding of a t(9;22), the Philadelphia chromosome (Ph+ ALL) – approximately 25% of adults ALL – has long been recognized. The Philadelphia chromosome t(9;22)(q34;q11), results from a translocation involving the BCR gene on chromosome 22 and the ABL1 gene on chromosome 9. One-third of adult patients with Ph+ ALL will have the M-BCR re-arrangement (resulting in a 210-kDa protein), as seen in chronic myeloid leukaemia, detectable by PCR. The remaining two-thirds will have the m-BCR re-arrangement (resulting in a 190-kDa protein). Patients with Ph+ ALL were – in the past – less likely to enter complete remission (CR) with CR rates of 70–80% such patients had an overall survival of 15–20%. ALL treatment protocols have typically assigned patients with Ph+ ALL to 'very high-risk' protocols and a myeloablative transplant has typically been offered, even when sibling donors have not been available.

The translocation t(4;11)(q21;q23), involving the *MLL* gene on chromosome 11q23, has also been recognized as carrying a poor prognosis in adult ALL. In an analysis of data from 1522 adult patients participating in the UKALLXII/ECOG2993 study, the following four groups of patients with distinct karyotypes were shown to have markedly inferior rates of event free survival and overall survival when compared to the whole cohort: t(9;22), t(4;11), complex karyotype (defined as five or more chromosomal abnormalities) and low hypodiploidy/near triploidy. Among patients with Ph-negative ALL, the prognostic relevance of these abnormalities was independent of other previously known poor prognostic factors, such as age and presenting white cell count. This was the first demonstration that cytogenetic subgroups other than Ph chromosome can be used for risk stratification of adults with ALL.

New cytogenetic groups are continuing to be defined in both B-cell and T-cell disease, although their clinical relevance needs to be explored within the context of large clinical trials. It is likely that further risk stratifications based on cytogenetic abnormalities at diagnosis will be part of routine clinical practice within the next 5–10 years.

Minimal residual disease (MRD)

Morphological detection of remission post treatment has very limited sensitivity of 1–5% and a patient in a morphological complete remission may still bear a considerable disease burden. It is now possible to quantitate treatment response very accurately and reproducibly, to the level of 1 leukaemic cell in 10,000 using various methods, as shown in Table 21.2. The most standardized method identifies patient-specific immunoglobulin heavy chain and T-cell receptor gene rearrangements, which

Table 21.2 Comparison of MRD assays in ALL.

Method	Applicability	Sensitivity of leukaemia cell detection	Specimens needed
Flow cytometry	Approx. 95%	1 in 10^4	Fresh cells
RQ-PCR for Ig/TCR re-arrangements	Approx. 90%	1 in 10^5	DNA
RQ-PCR for fusion gene transcripts such as *BCR-ABL1*	Depends on frequency of abnormality; *BCR–ABL1* is 20–25%	1 in 10^5	RNA
High throughput sequencing to quantify Ig/TCR re-arrangements	Approx. 90%	1 in 10^{5-6}	DNA

can then be quantified by real-time polymerase chain reaction (PCR) in a patient specific-assay with a sensitivity of 10^{-5} (1 in 100,000 cells). Flow cytometry can also be used to determine a disease-associated immunophenotype at diagnosis, which can also form the basis of an MRD assay. A diagnostic sample is required in both cases. Where there is a known molecular abnormality such as *BCR-ABL1*, this can also be quantified, using a housekeeping gene as a reference standard for PCR, to determine the level of residual ALL with similar sensitivity. Both molecular and flow cytometry techniques for quantification of MRD are labour intensive when performed properly and should be carried out in reference laboratories, according to standardized protocols. Studies in both childhood and adult ALL have shown a significant correlation between MRD levels and subsequent relapse risk (see Chapter 23). However, it should be noted that interpretation of MRD results is always protocol-dependent. It is not yet known whether high-risk interventions such as stem cell transplant can overcome the poor prognostic significance of residual MRD, and this question is the current focus of several ongoing large Phase III clinical trials.

It is also possible to use MRD analysis to monitor the progress of patients in remission in order to detect impending relapse – the time-window between molecular and haematological relapse is on the order of a few months, possibly yielding an opportunity for early intervention. However, the clinical potential of treatment prior to haematological relapse is not yet clear. Table 20.2 shows a comparison of different methods for detection of MRD with advantages and disadvantages.

Prognostic and predictive factors in ALL

A prognostic factor is a clinical or genetic characteristic that is objectively measurable and that provides information on the likely outcome of the ALL in an untreated person. A predictive factor is a clinical or genetic characteristic that provides information on the likely benefit from treatment. Consequently, prognostic factors define the effects of pre-existing patient or leukaemia characteristics on outcome, whereas predictive factors are used to define the effect of treatment.

A number of clinical and genetic factors have been clearly identified as poor prognostic factors in ALL; advancing age, higher presenting white cell count at diagnosis, presence of CNS disease, certain cytogenetic abnormalities, specific immunophenotypes and specific genetic abnormalities such as *IKZF1* (Ikaros gene) deletions. Prognostic factors are summarized in Table 21.3. Response to initial therapy, early steroid response and quantification of minimal residual disease at protocol-defined time points are validated predictive factors. Most current treatment strategies and trials take some or all of these factors into account in defining a risk status in ALL. There is no universally accepted algorithm by which risk status is measured in all cases. Additionally, whilst there is sometimes a relationship between these poor risk factors (for example, patients with T-ALL often have very high presenting white cell counts and CNS disease), the poor outcome portended by initial presence of a poor prognostic factors can be superseded by a good treatment response,

Table 21.3 Factors commonly used to predict outcome or stratify therapy.

Factor	Poor risk	Neutral or good risk
Presenting WCC	30×10^9/l B ALL $>100 \times 10^9$/l T-ALL	Low presenting WCC
Age	Worse outcome with advancing age	Younger age
Immunophenotype	T-cell phenotype better outcome than B-cell phenotype Early T precursor ALL (ETP) Expression of CD20 worse outcome (some studies)	Expression of myeloid antigens generally not associated with worse outcome
Cytogenetics	t(9;22) (*BCR-ABL1*) t(4;11) (*MLL/AF4*) Complex karyotype (five or more chromosomal abnormalities) Low hypodiploidy/near triploidy	t(12;21) (*ETV6/RUNX1*) Hyperdiploidy
Specific genetic abnormalities	*IKZF1* deletions, particularly in Ph+ ALL	*NOTCH1* and *FBXW7* mutations in T-ALL
Early response to steroid	Poor initial response to steroid therapy	
MRD	Positive signal at protocol-relevant time point	Negative signal at protocol-relevant time point

as documented by having a negative MRD status after induction. Hence, how we vary our therapeutic strategy in response to prognostic and predictive factors will change as studies identify new markers and define their inter-relationships. Figure 21.1 (kindly provided by Anthony Moorman and Christine Harrison, Leukaemia Research Cytogenetics Group, Northern Institute for Cancer Research, Newcastle University, UK) is based on an analysis of adults entered into UK trials of ALL. It shows the genetic heterogeneity of adult ALL, detailing some of the relationships between risk groups and further, their relationships with outcome. Table 21.3 highlights some of the factors that are currently taken into account internationally when making decisions about prognosis and approach to treatment.

Initial approach to a patient with ALL

Clinical presentation

The majority of adult patients with ALL will present with clinical symptoms of bone marrow failure. The history is often very short, i.e. a few weeks and patients frequently report a very rapid progression of their symptoms. The commonest symptoms that cause the patient to seek medical advice are recurrent infection or petechiae or other haemorrhagic problems. Weakness and fatigue due to anaemia are also common presenting features. Approximately half of all patients will have findings of lymphadenopathy, splenomegaly and/or hepatomegaly at presentation. According to data from two large German trials, 14% of patients will have a mediastinal mass at presentation and the majority of these (85%) will subsequently be shown to have T-cell disease. Examination of the cerebral spinal fluid (CSF) at diagnosis reveals leukaemic blasts in 7% of patients, although only 4% have central nervous system (CNS) symptoms such as headache, vomiting, lethargy and cranial or peripheral nerve dysfunction.

Almost any organ can be infiltrated by leukaemic blasts at diagnosis and approximately 10% of patients will have demonstrable involvement clinically. Pleural effusions are most common and tend to occur in association with mediastinal masses in T-cell disease. Bone pain at presentation is rare in adults compared to children, with bone lesions being demonstrated in 1% of adult patients. Involvement of the testes is very rare (<1%) in adult male patients. Leukaemic infiltration of other organs such as lung, kidney and skin is rarely observed and tends to be associated with poorer outcomes.

Diagnostic approach

A new diagnosis of ALL is a medical emergency and should be investigated and treated without delay. All patients require a detailed history to be taken alongside a thorough medical examination. The admitting clinician needs to pay particular attention to signs and symptoms of pre-existing infection, including fungal infection. Patients should be asked specifically about the presence of symptoms of sinusitis and haemoptysis. A careful neurological examination is also mandatory.

Investigations at diagnosis must include a full blood count, blood film examination and bone marrow (see also Chapters 19 and 20). Urgent biochemical testing, including analysis of urate levels, calcium, renal function, liver function and lactate dehydrogenase (LDH) are required. Coagulation abnormalities in the absence of liver failure due to leukaemic infiltration are rare, but a coagulation screen should also be performed.

Many modern protocols include the use of monoclonal antibodies such as rituximab in induction. Testing of hepatitis B and C serology prior to the start of therapy is mandatory if rituximab is to be used. It is also good practice to check the HIV status of a patient before starting induction chemotherapy.

A lumbar puncture should be performed in patients at diagnosis, if there is clinical evidence of CNS involvement with leukaemia, to assess whether or not there is leukaemic infiltration of the CSF. Ideally the CSF should be examined by flow cytometry. If there is no clinical suspicion of CNS involvement then many protocols start systemic chemotherapy to lower the disease burden before the first lumbar puncture is performed to reduce the risk of introducing disease into the CSF. Intrathecal methotrexate should always be given at the first lumbar puncture.

A chest X-ray (CXR) is mandatory for all newly diagnosed patients with ALL, but computerised tomography (CT) scans are not routinely required. However, where there is clinical evidence of lymphadenopathy, mediastinal widening or organ infiltration, a CT scan should be performed.

Supportive care (see also Chapter 23)

Before starting treatment, patients should be well hydrated and receive appropriate medication to prevent urate nephropathy. ALL can be exquisitely sensitive to small amounts of treatment, even in the steroid prephase, and tumour lysis syndrome is not uncommon. Allopurinol should be started 24 hours before induction therapy with steroids or chemotherapy and should be continued for a minimum of 5 days. Rasburicase should be considered as an alternative to allopurinol if the white cell count is high ($> 100 \times 10^9$/L), if there is bulky nodal or extranodal disease or if the patient already has impaired renal function. It should be noted that rasburicase can cause haemolytic anaemia in patients with glucose-6-phosphate dehyrogenase deficiency (G6PD). Consideration should be given to G6PD testing in patients prior to commencing rasburicase.

Due to the combination of myelosuppressive chemotherapy drugs and steroids that patients with ALL receive, they

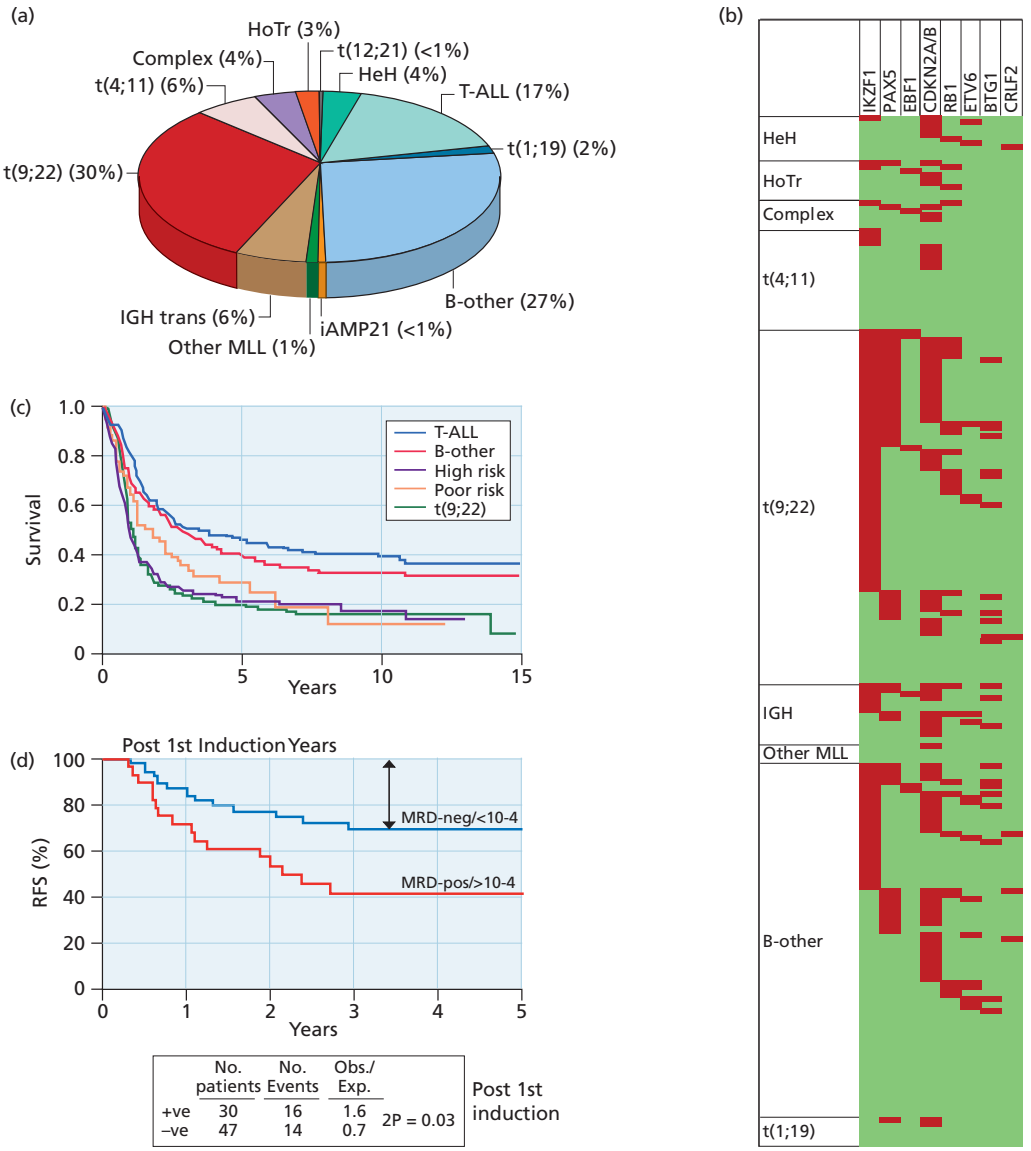


Figure 21.1 Genetic and outcome heterogeneity within adult ALL. (a) Distribution of 1014 UKALLXII/UKALL14 patients (25–60 years) according to primary genetic subtype. (b) Correlation of eight key secondary copy number alterations with the main primary genetic subtype in 257 adults (25–60 years). (c) Outcome of 1065 UKALLXII/ECOG2993 adults (25–60 years) by genetic subgroup. The high-risk group comprises those abnormalities which are now used to direct high-risk treatment on UKALL14: t(4;11), HoTr, complex. The poor-risk group comprises those aberrations which have recently been identified as being associated with a poor outcome, but are not currently used to direct treatment: *IKZF1* deletions, *CRLF2* re-arrangements and *IGH* translocations. The B-other group comprises all cases with successful cytogenetics but that do not harbour an established abnormality. NB: Patients with rarer subtypes e.g. t(12;21), HeH, t(1;19) have been excluded for clarity. (d) Disease-free survival of t(9;22)-negative B-cell precursor UKALLXII patients according to Ig/TCR-based MRD levels post first induction. HeH, High hyperdiploidy (51–65 chromosomes), t(12;21), t(12;21)(p13;q21)/*ETV6-RUNX1*; HoTr, Low hypodiploidy/near-triploidy (30–39/60–78 chromosomes; complex karyotype (five or more unrelated aberrations in the absence of another primary abnormality), t(4;11), t(4;11)(q21;q23)/*MLL-AF4*; t(9;22), t(9;22)(q34;q11)/*BCR-ABL1*; *IGH* trans, translocations involving the *IGH*/14q32 locus; Other *MLL*, translocations involving *MLL* except t(4;11); iAMP21, intrachromosomal amplification of chromosome 21, t(1;19), t(1;19)(q23;p13)/*TCF3-PBX1*. (Source: Anthony Moorman and Christine Harrison, Leukaemia Research Cytogenetics Group, Northern Institute for Cancer Research, Newcastle University, UK. Reproduced with permission.)

are at high risk of infectious complications from treatment, particularly during the induction phase of therapy. Adherence to strict anti-infective strategies is important. All patients should receive prophylaxis against herpes simplex virus and varicella zoster re-activation with aciclovir or an equivalent treatment. Patients also require prophylaxis against *Pneumocystis jirovecii* from the start of therapy. The recommended prophylaxis is cotrimoxazole. It should be noted that allergy to cotrimoxazole is not uncommon and alternative agents such as nebulized pentamidine or dapsone should be used for these patients.

Antifungal prophylaxis is strongly recommended for all patients on ALL therapy from induction until the start of maintenance. Azoles should be avoided when a patient is concurrently receiving vincristine due to the high risk of potentiation of the toxicity of vincristine, particularly neurotoxicity. There is no clear evidence currently about which antifungal prophylaxis regimen should be followed and local policies should apply.

The use of granulocyte-colony stimulating factor (G-CSF), given in parallel with induction chemotherapy, has been shown to reduce the incidence and severity of infections during the first 4 weeks of therapy. Its use has been associated with reduced induction mortality due to infection in several studies. Consequently the addition of G-CSF in induction is also strongly recommended for all adult patients on ALL therapy.

A checklist of procedures required prior to commencing antileukaemia therapy in a patient with newly diagnosed ALL is given in Table 21.4. General aspects of supportive care in haematology are also discussed in Chapter 23.

Specific treatment of ALL

The standard treatment for adult ALL has so far not been defined. It is strongly recommended that, where available, all patients with newly diagnosed ALL are offered treatment within the context of a large Phase III clinical trial. Many countries have well-organized study groups, which focus on developing, running and analysing data from clinical trials. Increasingly, there is stronger, regular communication with practical collaborations between clinicians and scientists to use patient specimens coupled to clinical outcome data to ask questions about the genetics of ALL, using, for example, next-generation sequencing approaches. Treatments have been developed on regimens that have evolved over time, with the gradual introduction of novel agents and new approaches, such that current interventions are now based on a strong literature with clinical and scientific contributions from numerous large international clinical trials.

The aim of initial treatment is to achieve complete remission (CR), which is currently defined on a morphological basis of less than 5% blasts in the bone marrow in the presence of overall haematological recovery (neutrophils $> 1.0 \times 10^9$/L, platelets $> 100 \times 10^9$/L).

Corticosteroids are among the most important drugs in the treatment of ALL. Recent paediatric trials have shown improved outcomes when dexamethasone is used as opposed to prednisolone. Dexamethasone does have better antileukaemic activity than prednisolone *in vitro*; it has better penetration of the CNS and causes fewer thromboembolic events. Although there are less data in adult ALL showing a survival advantage with dexamethasone compared to prednisolone, many adult regimens now have dexamethasone as the preferred corticosteroid of choice in induction. A generalized schema for treatment is shown in Figure 21.2.

Prephase

A steroid prephase for 5–7 days prior to commencing chemotherapy can be very beneficial in the treatment of ALL. Patients will often have a good partial response to steroids. A prephase is particularly useful when a patient has a poor performance status due to burden of disease. Early steroids can reduce disease burden and improve performance status allowing the patient to be in a better position to receive induction chemotherapy. Steroids are also very useful when there is liver infiltration with ALL and vincristine and anthracyclines cannot be administered as planned due to high levels of serum bilirubin. There are rare situations where a steroid prephase is not recommended, due to the need for emergency anthracycline-based chemotherapy, such as superior vena cava obstruction due to a large mediastinal mass. The risk of tumour lysis is always high in these situations and rasburicase should be given prior to chemotherapy.

Induction therapy

The primary goal of induction therapy is complete eradication of ALL cells from the blood, bone marrow, CNS and other extramedullary sites when initially involved. Ideally, CR should be achieved as early as possible with minimal toxicity. Nonetheless, toxicity is often high, especially by contrast to that seen when treating children and many adult patients are hospitalized for weeks to months. There is no universally agreed induction protocol, but most adult treatment regimens are broadly similar in the drug dosing and scheduling. Many protocols use two phases of induction, with the second part being applied regardless of the initial response. Induction regimens in adult ALL are composed of steroid, vincristine, anthracycline (usually daunorubicin or doxorubicin) and L-asparaginase. Cyclophosphamide and cytarabine may also be included, usually in part 2. With modern protocols, CR rates following induction therapy are high with over 90% of unselected adult patients with ALL achieving an initial good response to treatment.

Table 21.4 Suggested checklist for the initial assessment of an adult patient with ALL.

Procedure	Mandatory (M)/ Recommended (R)
Clinical assessment:	
• Full history	M
• Clinical examination of the CNS	M
• Height, weight and body surface area	M
• Assessment of performance status	R
Investigations	
• Full blood count and film	M
• Coagulation screen	M
• Biochemical profile, LDH and urate levels	M
• Blood cultures if patient is febrile	M
• G-6-PD status if rasburicase to be used	R
• Pregnancy test for women of childbearing age	M
• HIV testing	R
• Hepatitis B and C serology if monoclonal antibody therapy, e.g. rituximab, is part of the planned therapeutic regimen	M
• Bone marrow aspirate:	M
◦ Morphology	M
◦ Flow cytometry	M
• Cytogenetics (G-banding and FISH analysis)	M
• Molecular diagnostics	M
• Evaluation for MRD marker	M
• Lumbar puncture if evidence of CNS involvement	M
• Examination of CSF for blasts by cytospin	M
• Flow cytometry of CSF	R
• CXR	M
• CT of neck, thorax, abdomen and pelvis if evidence of nodal/ extramedullary disease	R
• CT head if clinical evidence of CNS/ophthalmic disease	M
• Tissue typing of patient and siblings (unless a young adult being treated on a paediatric protocol)	R
Pretreatment	
• Gain IV access. Avoid central line in phase I induction if possible	R
• Adequate hydration. Use IV fluids if WCC is high or significant extramedullary disease	M
• IV antibiotics if patient is febrile	M
• Prevention of urate nephropathy with allopurinol or rasburicase	M
• Infection prophylaxis against:	
◦ Pneumocystis jrovecii	M
◦ *Varicella* zoster	M
◦ Fungi	R
• Growth factor support during induction	R
• Thromboprophylaxis if platelets $> 50 \times 10^9$/L	R
Treatment	
• Steroid prephase for 5–7 days	R
• Consider entry into clinical trial if available	R

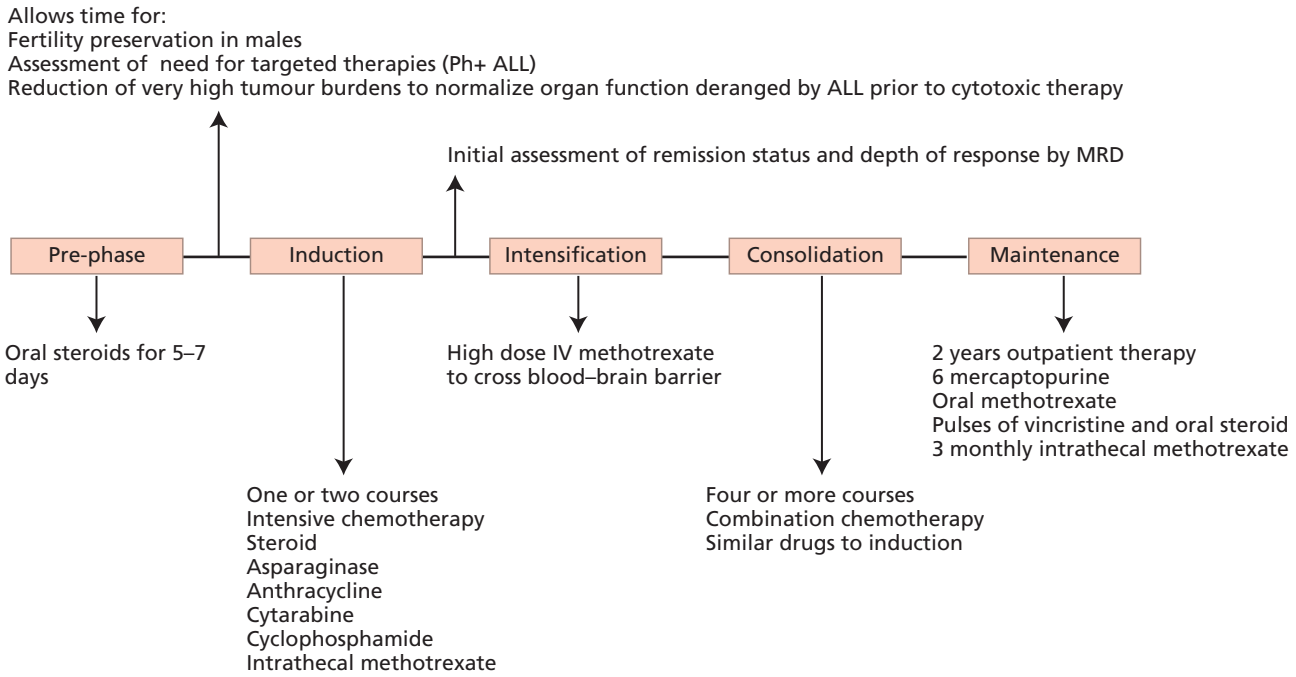


Figure 21.2 A generalized schema for treatment of adult acute lymphoblastic leukaemia.

Toxicity during induction

Induction is generally the most risky time during a patient's treatment and they should be monitored closely and seen regularly for clinical assessment. Induction death rates range from 2–20%, depending on the protocol used and the patient's age. Most induction deaths are due to either severe bacterial sepsis, with fungal infection also posing a significant risk. Steroids often mask signs and symptoms of infection, so astute clinical judgement is required. All patients should have access to a specialist haematology unit 24 hours a day, and they must be treated promptly with intravenous antibiotics for signs of neutropenic sepsis. Any patient with a persisting fever should be investigated for invasive fungal infection.

In addition to the risk of infection, patients are also at risk of problems relating to specific drugs used in induction. L-asparaginase is one of the most effective and important drugs in the treatment of ALL, but it is also one of the most difficult drugs to dose and manage. A major problem with L-asparaginase is the occurrence of coagulation abnormalities during treatment.

L-asparaginase therapy results in reduced levels of fibrinogen and antithrombin levels, but there is little or no evidence that infusions of fresh frozen plasma or fibrinogen can correct the coagulation abnormalities and prevent thrombosis or bleeding. Conversely there is also some evidence that giving fresh frozen plasma replenishes the asparaginase pool that the drug is being used to deplete, negating its effectiveness. Consequently the use of plasma-derived products to correct coagulation abnormalities is not recommended unless the patient is actively bleeding.

Despite deranged coagulation being the abnormality measured in the laboratory, an increased risk of venous thrombosis is the main clinical problem with L-asparaginase. Data indicate that 10–20% of adults receiving L-asparaginase will develop a thrombosis. As a consequence, strategies to reduce the risk of thrombosis are important. It should be noted that with modern adult ALL protocols, patients might not be thrombocytopenic throughout the whole of induction. Hospitalized patients with platelet counts above 50×10^9/L should be given antithrombotic prophylaxis as per UK National Institute of Clinical Excellence (NICE) thromboprophylaxis guidelines. The presence of a central venous catheter increases the risk of thrombosis and many centres will avoid placing a central venous catheter in the first phase of induction until the patient has achieved a remission. There are little data available on how to manage thrombosis in this setting. When a catheter-related thrombosis occurs then the treatment of choice is low-molecular-weight heparin. Treatment of a central venous thrombosis is less clear. Patients should be treated with either low-molecular-weight or unfractionated heparin. Adequate monitoring of the heparin dosing using the anti-Xa assay should be performed as patients may be resistant to heparin due to depletion of antithrombin. In this situation replacement of antithrombin is likely to be valuable, and specialist coagulation advice should be sought.

There are data showing inferior outcomes for patients who develop a thrombosis during induction therapy for ALL despite the fact that the thromboses that occur with L-asparaginase therapy are not usually fatal. It is assumed that the occurrence of a thrombosis results in treatment delays, which are then detrimental to the patient in the longer term.

Patients with previous L-asparaginase related thrombosis can be given further drug with anticoagulation cover during consolidation therapy.

It is common for patients on L-asparaginase to develop transaminitis and hyperbilirubinaemia during induction therapy. If the patient develops acute sepsis then this will exacerbate the problem. All non-essential drugs that are liver metabolized should be withheld. With supportive care the liver abnormalities usually fully resolve, but severe treatment delays often result.

Pancreatitis is a rare, but life-threatening, complication of L-asparaginase, which is more common in children than in adults. Serum amylase levels should always be checked in a patient post L-asparaginase therapy who is complaining of abdominal pain. The development of pancreatitis is a contraindication to further use of the drug.

Consolidation therapy

Consolidation therapy is given to patients in CR with the aim of eliminating residual leukaemia cells and to prevent the re-emergence of drug-resistant disease. Consolidation therapy typically consists of several cycles of treatment with multiagent chemotherapy, including CNS-directed therapy and further doses of L-asparaginase. Modern adult protocols frequently contain at least one delayed intensification block of therapy, which constitutes a modified re-induction schedule mirroring phases 1 and 2 of induction therapy. Smooth progression between courses of therapy to minimize delays is likely to be important, although there are usually specific requirements for peripheral blood count recovery between courses. Slow recovery between courses should be taken seriously as it may indicate actual or impending relapse. Consolidation is generally better tolerated than induction, partly because the patient will have achieved CR and will begin therapy with adequate bone marrow function. Many patients can receive at least some elements of their consolidation therapy as outpatients.

Maintenance therapy

Maintenance therapy remains obligatory for all adult patients who have not undergone an allogeneic stem cell transplant as part of their treatment for ALL. Studies evaluating the omission of maintenance therapy have concluded that this results in inferior outcomes. The standard approach to maintenance includes daily 6-mercaptopurine and weekly methotrexate with pulses of vincristine and steroids for a period of approximately 2 years post consolidation. However, individual protocols vary. The aim of maintenance is to eliminate minimal residual disease and dosing of the chemotherapeutic agents needs to be strictly adjusted to induce moderate myelosuppression, whilst not putting the patient at risk of neutropenic sepsis or haemorrhage. CNS-directed prophylaxis with intrathecal chemotherapy should continue during maintenance. Pneumocystis prophylaxis and herpes zoster prophylaxis should also continue throughout the maintenance period, although antifungal prophylaxis is no longer required.

CNS-directed therapy

The presence of lymphoid blasts in the cerebral spinal fluid is relatively common (7%) in adults with newly diagnosed ALL. Risk factors include T-cell disease, high white cell count and other extramedullary sites of disease. Consequently it is important to target treatment to prevent development of progression of ALL in the CNS. Traditionally intracranial radiotherapy was used, but this resulted in long-term sequelae, particularly in children, with neurodevelopmental delay, short-term memory loss and progressive deterioration of intellectual function all being reported. In adults, several trials have reported CNS recurrence rates of less than 10% when a combination of high-dose methotrexate and intrathecal chemotherapy is used without additional cranial radiotherapy. In most adult and paediatric protocols CNS irradiation is no longer used in the prophylactic setting.

Methotrexate is the mainstay of therapy targeted at preventing CNS relapse in patients with ALL. It is commonly given via the lumbar puncture route throughout all phases of treatment, including maintenance therapy. Cytarabine and steroid can also be given intrathecally and are sometimes given with methotrexate as 'triple therapy', although there are no data suggesting the superiority of adding the additional drugs. The precise number of lumbar punctures needed during therapy is not defined. In childhood ALL this is sometimes modified by risk of CNS disease. In adult ALL 10–20 doses of IT therapy are given, dependent upon treatment protocol.

In addition, many adult protocols include an intensification block of high-dose methotrexate given at doses ≥ 3 g/m^2.

One of the most dramatic complications of ALL therapy is methotrexate (MTX)-induced encephalopathy. Although this side effect is rare, it is very frightening for patients and their relatives when it does occur. MTX-induced encephalopathy is more frequently seen in young adults than older adults and usually occurs when intrathecal methotrexate is being given during a phase of intravenous cytarabine. It may present with fits, or focal neurological signs with or without impaired consciousness. It occurs within 1 day to 3 weeks after exposure to intrathecal methotrexate and hence it is not always the most evident explanation for the signs and symptoms. Fortunately, patients usually make a full recovery, although there have been case reports

where a degree of neurological damage has persisted. It is important to differentiate MTX-induced encephalopathy from cerebral sinus thrombosis with the appropriate imaging. The presence of ALL within the CNS also needs to be ruled out in any patient presenting with a sudden onset of neurological symptoms. In the case of MTX-induced encephalopathy, intrathecal methotrexate should not be given whilst the patient is receiving systemic cytarabine. It can often safely be re-introduced later in the protocol, but this should be approached on a case-by-case basis. Alternatively the patient can be given cytarabine intrathecally, in association with hydrocortisone.

Treatment of CNS disease at diagnosis

For those patients where there is evidence of CNS disease at diagnosis the current advice is to give once- to twice-weekly intrathecal methotrexate alongside induction therapy until the CSF is clear of lymphoblasts on two consecutive samples. Cranial radiotherapy post induction remains the standard treatment of CNS disease for this small group of patients. However, if the patient is due to receive allogeneic stem cell transplant with total body irradiation, CNS irradiation can be withheld. Further intrathecal methotrexate should be given in consolidation and maintenance post radiotherapy.

Stem cell transplantation

Myeloablative allogeneic haemopoietic stem cell transplantation (alloHSCT) is undoubtedly one of the most effective anti-ALL therapies available and there is a clear graft-versus-ALL effect. One of the biggest ever studies of adult ALL (UKALL12/E2993) demonstrated a survival advantage for myeloablative sibling alloHSCT over chemotherapy in a 'donor versus no donor', intention-to-treat analysis. Unfortunately, it is also the most toxic approach within our therapeutic scope. It has played a significant role in the treatment of ALL over many years, in some country's approaches, being reserved for those thought to be at high risk of poor outcome by dint of having poor prognostic factors and in others, being more widely applied wherever a matched sibling donor is available. As alloHSCT practice changes, with unrelated donor stem cells being widely available and the use of haploidentical donor cells, or umbilical cord blood becoming more routine, and with the widespread introduction of reduced-intensity conditioning regimens, so its role, which has never been clearly defined by clinical trial, will also change. Its role will also be affected by changes in ALL therapy. As we learn more about predictive factors, develop new predictive biomarkers and have a larger set of therapeutic tools at our disposal, including T-cell immunotherapies that don't require myeloablation, so we will need to alter our focus of who should receive alloHSCT. Application of alloHSCT is one of the most controversial areas of clinical practice in ALL.

Specific approaches to defined clinical populations in the treatment of ALL

Targeted therapies for ALL: Philadelphia chromosome positive (Ph+) disease

Ph+ ALL is identified by the presence of the fusion between the Abelson (*ABL1*) tyrosine kinase gene on chromosome 9 and the break point cluster region (*BCR*) gene on chromosome 22, resulting in a *BCR–ABL1* fusion protein that is a constitutively active tyrosine kinase. This acquired genetic abnormality is necessary and sufficient for the development of CML (see Chapter 24), but in Ph+ ALL, other abnormalities, such as deletions within the Ikaros gene, are common and an acute leukaemia occurs *de novo*. In the past, this genetic abnormality conferred the poorest prognosis and alloHSCT was recommend for all patients. The treatment of Ph+ ALL has been revolutionized by the addition of selective inhibitors of the action of this fusion protein – tyrosine kinase inhibitors, (TKI) – to treatment. These non-chemotherapy agents, examples of which include imatinib and dasatinib, bind to the kinase domain of *ABL1* and disrupt the ATP binding site. By so doing, they block the catalytic activity of the kinase and subsequent downstream phosphorylation, essentially 'switching off' a growth signal that is otherwise permanently and erroneously switched on. Although there have been no randomized controlled trials comparing chemotherapy with chemotherapy plus imatinib in Ph+ ALL, there is little doubt that the overall outcome of therapy is superior with the inclusion of TKIs, although a higher rate of alloHSCT is also seen when TKIs are used. Although there is some evidence that Ph+ ALL in children can be treated successfully with chemotherapy and TKI combinations alone, this conclusion is based on a small study that was not designed to ask this specific question. In adults, for whom alloHSCT remains the mainstay of therapy, there is no clear evidence to date in the adult setting that Ph+ ALL can be optimally managed without alloHSCT because, even in the presence of TKIs, most patients eventually relapse without the definitive therapy of alloHSCT. Most studies to date have been conducted using imatinib. The TKI dasatinib is less specific for *BCR–ABL1* and also blocks *SRC* kinases, which, unlike in CML, are known to play a role in Ph+ ALL. ALL can become resistant to TKIs, most often by acquisition of mutations in the kinase domain. Mutational analysis can be obtained and is a relevant clinical test in patients with Ph+ ALL who relapse or whose disease is resistant to TKI therapy. Many mutations have been mapped and the implications for therapy with various TKI are known. As an example, the occurrence of the T315I mutation confers resistance to many TKIs – only ponatinib is of clinical utility in this scenario.

Teenagers and young adults (TYA)

Age is known to be a prognostic factor in ALL. Even in paediatric ALL it is well known that children under the age of 10

years have better outcomes than children over the age of 10 years. These differences are at least partly due to the reduced incidence of good-risk cytogenetics in children over 10 years of age (see Chapter 23). However, it was noticed in the late 1990s that there was a further sharp drop in outcomes for adolescents over 16 years of age. Whilst changing disease biology and patient biology play a role, many groups showed a significant impact of the protocol used on long-term outcomes for teenagers and young adults on retrospective data analysis. Traditionally adult protocols have been very transplant focused, whereas paediatric trials only use first remission transplantation in very high-risk disease. In addition, paediatric protocols historically contained significantly more doses of vincristine and L-asparaginase in induction and consolidation, compared to adult protocols.

In the UK patients up to the age of 25 years at diagnosis have been treated on national paediatric trials since 2007 (UKALL 2003 and more recently UKALL 2011). With this approach outcomes for teenagers and young adults have significantly improved with patients aged 16–25 years now having a predicted long-term survival of 75%, with very few patients requiring allogeneic stem cell transplantation in first response. However, analysis of outcomes from the UKALL 2003 trial has continued to show an impact of age on outcomes, with adolescents having inferior long-term survival compared to younger children. MRD status post induction and consolidation is the most important predictor of long-term outcomes in teenagers and young adults. As with older adults, new approaches with the use of novel agents and the selected application of first remission allogeneic transplantation within the context of clinical trials, are required for 16–25 year old patients with persistence of high levels of MRD post consolidation therapy.

ALL in older adults

The mean age of adults diagnosed with ALL is 60 years, with an age-specific annual incidence in the UK rising from 0.45/100,000 for adults aged 35–39 to 0.78/100,000 for adults aged 60–64 and 1.2/100,000 for those over 85 years. Despite these figures, patients over the age of 65 years have typically not been included in most large Phase III national trials. The lack of well-designed trials for the older patient with ALL may be one of the reasons why survival in this group has not improved over time.

Importantly, studies have shown that when the effect of cytogenetics on overall survival was accounted for, age was not a significant prognostic factor. This suggests that the worsening prognosis with advancing age in adult ALL could at least in part be a manifestation of the age-related increase in unfavourable cytogenetics.

There is no standard regimen for the older patient worldwide. Data from intensive transplant-based trials show that older patients fared poorly, with 15% long-term survival. Where patients have been deemed ineligible for national trial protocols, local practices have varied widely and treatment is often given with palliative rather than curative intent. Data from the National UKALL 12 trial reported significantly more infections during induction for patients over the age of 55 years, with chemotherapy doses having to be reduced more commonly in this age group in order to manage toxicity. Infection during induction in this trial was related to poorer outcomes overall. Published data indicate a 3–10 times higher rate of induction deaths in older individuals after treatment for ALL.

Ironically, there has been some progress in improving short-term outcomes for older patients with Ph+ disease. It is possible to achieve CR with minimal toxicity with the use of a tyrosine kinase inhibitor alone. CR rates of over 90% are reported with imatinib in patients over 55 years of age. When combined with steroid therapy, median survival rates of 20 months from diagnosis are reported, with most patients not requiring any admissions to hospital. The long-term follow-up of these patients suggests ongoing relapses and long-term cures are unlikely with this approach.

Balancing toxicity of treatment against potential long-term disease control in older patients is an enormous challenge. Emerging novel agents such as monoclonal antibodies have the potential of adding an independent antileukaemic effect without the addition of substantial toxicity. The challenge will be how to integrate these novel agents into ALL protocols for older individuals, where a current standard of care is not defined.

Treatment of relapsed/refractory disease

If the ALL fails to respond to therapy or relapses in adults, survival is generally very poor and alloHSCT is the only known curative therapy. Several large studies of patients who relapsed following initial therapy have recently been published, which show that once relapse has occurred, it is generally the features of the relapsed disease, rather than the disease at original diagnosis which determine outcome. Salvage therapy with conventional chemotherapy agents is often attempted and gives a second remission rate of about 40–50%, but often with considerable toxicity. Hence, it is very worthwhile to know which patients are most likely to benefit from conventional salvage therapy and which are likely to fare so poorly that they should be directed, if possible, towards clinical trials of novel therapies. Factors predicting poor outcome from relapse include older age, a short duration of first remission and relapse in the CNS. Treatment received in first remission does not predict for outcome after relapse.

Non-chemotherapy approaches to treatment

At the time of writing, TKIs are routinely used in the treatment of Ph+ ALL. Other non-chemotherapy approaches are currently in clinical trial for relapsed or resistant ALL and some are yielding

interesting data with potential promise for the future. A long-term goal of chemotherapy-free treatment for ALL is a worthy aim. Immunotherapy with engineered CD19 chimeric antigen receptor T cells is particularly exciting. Patients' T cells are genetically modified using retroviral vectors to produce the chimeric receptors for a single-chain variable fragment (scFv), targeted against CD19, which is linked to costimulatory molecules that help provide the necessary signals for the T cell to proliferate and lyse ALL cells upon contact with the CD19 antigen. Another approach, which also targets CD19, universally expressed on ALL cells, is the bispecific, antigen-engaging antibody blinatumomab (currently in Phase III trial) in which CD19 and CD3 are concurrently bound, activating an autologous T-cell response against the ALL cells. Antibodies bound to toxins such as the conjugated anti-CD22 antibody inotuzomab are also in Phase III trial. Such therapies have very different modes of action and spectra of potential toxicities to cytotoxic chemotherapy and may rely on other factors yet to be identified for their success. Hence, the predictive biomarkers for response to immunotherapies will require close study to allow us to find the best time within treatment to administer them.

Selected bibliography

Vardiman JW, Thiele J, Arber DA *et al*. (2009) The 2008 revision of the World Health Organisation (WHO) classification of myeloid neoplasms and acute leukaemia: rationale and important changes. *Blood* **114**(5): 937–51.

Bene MC, Castoldi G, Knapp W *et al*. (1995) Proposals for the immunological classification of acute leukaemias. European Group for the Immunological Characterization of Leukaemias (EGIL). *Leukaemia* **10**: 1783–6.

Rowe J, Buck G, Burnett A *et al*. (2005) Induction therapy for adults with acute lymphoblastic leukaemia; results of more than 1500 patients from the international ALL trial: MRC UKALL XII/ECOG E2993. *Blood* **106**(12): 3760–7.

Ludwig W, Rieder H, Bartram CR *et al*. (1998) Immunophenotype and genotypic features, clinical characteristics, and treatment outcome of adult pro-B acute lymphoblastic leukemia: results of the German multicentre trials GMALL 03/87 and 04/89. *Blood* **92**(6): 1898–909.

National Institute for Health and Care Excellence (2010) Clinical Guideline 92. Venous thromboembolism: reducing the risk.

Goldstone AH, Richards S, Lazarus HM *et al*. (2008) In adults with standard-risk acute lymphoblastic leukemia, the greatest benefit is achieved from a matched sibling allogeneic transplantation in first complete remission, and an autologous transplantation is less effective than conventional consolidation/maintenance chemotherapy in all patients: final results of the International ALL Trial (MRC UKALL XII/ECOG E2993). *Blood* **111**(4): 1827–33.

Rowntree C, Hough R, Wade R *et al*. (2013) Outcomes of teenagers and young adults on the UKALL 2003 paediatric trial for children and young people with acute lymphoblastic leukaemia. *Abstract*. *Blood* **122**(21): 57.

CHAPTER 22

Childhood acute lymphoblastic leukaemia

22

Ajay Vora

Department of Haematology, Sheffield Children's Hospital, Sheffield, UK

Introduction

Acute lymphoblastic leukaemia is the most common form of cancer (25–30%) and predominant subtype of leukaemia (75–80%) in children. However, childhood ALL is not a single disease and has considerable phenotypic and genotypic heterogeneity, which is of diagnostic and prognostic importance. Modern protocols take account of this heterogeneity by incorporating complex risk stratification models to determine content and intensity of therapy received by an individual patient. Despite no new drugs being available during that period, treatment outcome has improved substantially in the last four decades such that over 90% of patients can expect to survive without disease in the long term. Translation of recent advances in understanding of the molecular biology of ALL and its influence on phenotype and clinical outcome will help define specific subgroups that might benefit from better targeted therapy.

Epidemiology

The median age at diagnosis for ALL is 13 years and nearly 60% of cases are diagnosed under the age of 25. There is a sharp peak in ALL incidence among 2–3 year olds (>80 per million), which decreases to a rate of 20 per million for 8–10 year olds. The incidence of ALL among 2–3 year olds is approximately fourfold greater than that for infants and is nearly 10-fold greater than that for 19 year olds. The incidence is higher in boys than in girls (four times for T-cell ALL), except that girls have a slightly higher (1.5 times) incidence of leukaemia in the first year of life. The reported incidence of ALL is higher in Northern and Western Europe, North America and Oceania, and lower in Asia, South America and Africa. In industrialized countries, the incidence is higher among children of European descent than among those of African descent. For example, the annual rate (per million population) of childhood ALL from birth to 19 years of age is 29.7 in the UK and 11.0 in India; in the USA, it is 32.9 for white people and 14.8 for African-Americans. T-cell ALL and pre-B leukaemia with t(1;19)/*TCF3–PBX1* (also known as *E2A–PBX1*) fusion are more prevalent among African-American children, who are less likely to have ALL with high hyperdiploidy. The incidence of ALL among children younger than 15 years of age has shown a moderate increase in the past 20 years.

Aetiology

Genetic factors

Only a small proportion (<5%) of patients with childhood ALL have underlying hereditary genetic abnormalities. Children with Down syndrome have a 10- to 30-fold increased risk of developing ALL. Other genetic disorders associated with an increased incidence of ALL include ataxia telangiectasia, Schwachman–Diamond syndrome and Bloom syndrome (see also Chapter 18). There is no definite evidence of an association between leukaemia and congenital immunodeficiencies such as X-linked agammaglobulinaemia. Germline TP53 mutations, a cause of the cancer-predisposing Li–Fraumeni syndrome, are present in

Postgraduate Haematology, Seventh Edition. Edited by A Victor Hoffbrand, Douglas R Higgs, David M Keeling and Atul B Mehta.

40% of patients with low hypodiploid (32–39 chromosomes) ALL.

Fraternal twins and siblings of affected children are at a twofold to fourfold greater risk of leukaemia during the first decade of life than are unrelated children. In the case of identical twins, when leukaemia occurs in one twin, there is a 20% probability that it will also occur in the other twin, owing to ALL transfer in utero via the shared placental circulation. When leukaemia is diagnosed before 1 year of age it almost invariably develops in the other twin, generally within a few months. In identical twins with t(4;11)/MLL–AF4, the concordance rate is nearly 100%, with a short latency period (weeks to a few months). In contrast, the concordance rate in twins with the ETV6–RUNX1 fusion or T-cell phenotype is lower and the postnatal latency period longer, in keeping with the requirement for additional genetic events for leukaemic transformation in these subtypes of ALL. Hyperdiploid ALL also appears to arise before birth, but requires postnatal events for full malignant transformation. In contrast, t(1;19)/TCF3–PBX1 ALL appears to have a postnatal origin in most cases.

Genome-wide association studies (GWAS) have discovered polymorphisms at six genomic loci associated with a two- to fourfold increased genetic susceptibility to childhood ALL in general or a specific genetic sub-type: *ARID5B*, *IKZF1*, *CEBPE*, *CDKN2A*, *BMI1-PIP4K2A*, *GATA3*, which has an association with general susceptibility and Philadelphia-like ALL (GATA3), and *TP63,* which has an association with *ETV6/RUNX1*ALL. There is evidence implicating several of these genes in the pathogenesis of ALL. For example, germline variants in *ARID5B* have the strongest association with ALL susceptibility across the genome, and the loss of Arid5b in mouse leads to significant defects in lymphoid cell development. *IKZF1*, an important transcription factor in all lymphoid lineages, is frequently targeted by copy number alterations in ALL blast cells, and *IKZF1* deletion is associated with a poor prognosis. Loss of *CDKN2A/CDKN2B* occurs in up to 40% of B-precursor ALL and is likely to contribute to cell cycle deregulation in leukaemia. *CEBPE* is related specifically to myeloid cell maturation and terminal differentiation, but intrachromosomal translocations involving the immunoglobulin heavy locus (*IGH*) and *CEBPE* genes also have been described in childhood ALL. *ARID5B, IKZF1, CEBPE* and *BMI1-PIP4K2A* variants cumulatively confer strong predisposition to ALL, with children carrying six to eight copies of risk alleles at a 9-fold higher ALL risk relative to those carrying zero to one risk alleles at these four single nucleotide polymorphisms. The lower incidence of B-ALL with high hyperdiploidy in African-American children in the US may at least partially reflect lower prevalence of the ARID5B risk allele. The convergence of germline ALL susceptibility loci and somatic aberrations on genes involved in lymphoid cell development, cell cycle control, and tumour suppression reinforces the contribution of these key pathways to leukaemogenesis and also points to the possibility that inherited and acquired genetic variations act synergistically in the development of childhood ALL.

Environmental factors

Ionizing radiation and chemical mutagens have been implicated in the induction of leukaemia, but clear aetiological factors for ALL cannot be identified in a majority of cases. *In utero* exposure to diagnostic X-rays is associated with a slightly increased risk of ALL, proportional to the number of exposures. The association between leukaemia and maternal exposure to potential mutagens, neonatal administration of vitamin K, parental use of medications and drugs, proximity to electromagnetic fields and exposure to other potential mutagens have been proposed, but not proven. In the UK, it has been demonstrated that the geographical incidence of childhood leukaemia shows more clustering than would be expected if it were due to chance variations, which could be due to the geographical distribution of environmental risk factors, including infectious agents.

Infection was the first suggested causal exposure for childhood ALL and remains the strongest candidate. Two specific hypotheses have been proposed, both postulating that ALL results from an abnormal response to a common infection. There is no evidence to date of a unique or single transforming virus in ALL, as is the case of leukaemia in some animal species. Instead, it is likely that ALL is promoted indirectly by an abnormal or dysregulated immune response to one or more common infections (viral or bacterial) in a susceptible individual. Susceptible children would be defined as having minimal prior exposure to infection during infancy and have a persistent in utero-generated preleukaemic clone plus a variable degree of genetic susceptibility, as described above.

Pathogenesis (Figure 22.1)

Recent studies have revealed the genomic landscape of childhood ALL at diagnosis and its evolution prior to diagnosis, during therapy and at relapse. Except for MLL rearranged infant ALL, in which a single genetic mutation is likely to be an initiating and driving event, all other subtypes of childhood ALL are due to a series of genetic events within a cell that has acquired an initiating mutation, often *in utero*. A recurring leukaemia-associated chromosomal abnormality is detected by karyotyping, FISH or molecular techniques in around 80% of childhood ALL cases. High hyperdiploidy and ETV6-RUNX1-carrying cells are present at low levels in nearly 1% of cord blood samples from normal neonates, of whom only 1% develop ALL, often after several years; thus these rearrangements are by themselves insufficient to cause leukaemia, as confirmed in experimental models. These initiating mutations drive transcriptional and epigenetic dysregulation and aberrant self-renewal. They also

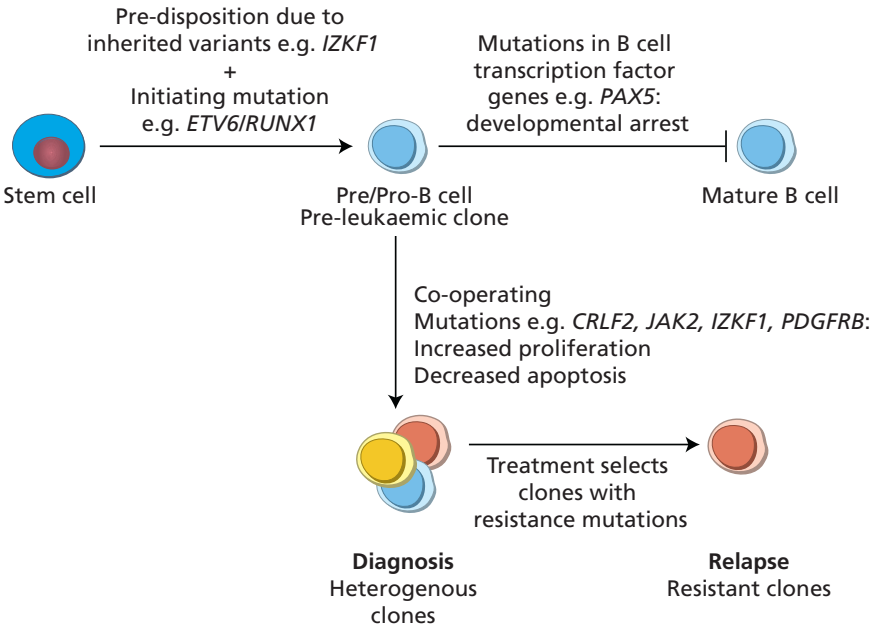


Figure 22.1 Pathogenesis of ALL.

disrupt lymphoid development and result in an arrest in maturation. Additional genetic alterations target multiple cellular pathways, including cell-cycle regulation, apoptosis and chromatin modification. In a subset of cases (BCR-ABL1+ and BCR-ABL1—likeALL), genetic alterations drive aberrant cytokine receptor and kinase signalling. Together, these events result in the proliferation and establishment of the leukaemic clone. Each mutation or combination of mutations gives rise to subclones such that, at diagnosis, the leukaemic cell population is clonally heterogeneous and genetic alterations that confer resistance to treatment may be present in minor clones. These resistant minor clones are selected by treatment to become the dominant population at relapse and are the reason for salvage therapy being less effective than first-line treatment. There is some evidence to suggest that very late 'relapse' of ETV/RUNX1 ALL is in fact a second leukaemia arising from the leukaemic 'stem cell' acquiring a different set of secondary mutations compared with the clone at first presentation.

Clinical features

Children with ALL frequently present with symptoms and signs of pancytopenia due to marrow failure. Fever, often due to leukaemia rather than infection, is present in 60% of patients. Nearly a third of patients, though, may have minimal or no symptoms and signs of marrow failure, but instead present with bone pain and arthralgia due to medullary expansion or periosteal/cortical infiltration as dominant symptoms. To add to the diagnostic confusion, these patients often have mild anaemia or neutropenia and very few circulating blasts, which may not be detected without careful examination of the blood film. Thus, they may be misdiagnosed and treated as juvenile idiopathic arthritis with transient benefit from steroids and/or methotrexate, resulting in a delay in diagnosis. Bulky extramedullary disease, especially overt CNS disease, mediastinal mass and renal infiltrates are more common in T-lineage ALL. CNS involvement is usually leptomeningeal rather than parenchymal and may present with symptoms of meningism or cranial nerve palsies. Ocular involvement is more common at relapse than first presentation and can affect the retina, ocular nerve, orbit, cornea or anterior chamber with hypopyon. Painless enlargement of the scrotum may be due to a testicular infiltrate or hydrocele resulting from lymphatic obstruction. Rarely, the diagnosis is made in a child who has a blood count for minor, non-specific, symptoms.

Differential diagnosis

Immune thrombocytopenic purpura (ITP) is the most common benign condition to be associated with symptoms that raise concern about leukaemia. Severe skin haemorrhage in an otherwise well child with isolated severe thrombocytopenia and normal blood film without circulating blasts makes it possible to distinguish the condition from leukaemia in a vast majority of cases. The presence of mild anaemia due to nose or gum bleeding and activated lymphocytes on blood film due to a recent viral illness, may create diagnostic uncertainty in some cases, but it should be possible for an experienced morphologist to exclude the presence of blasts on a blood film without resorting to a bone marrow aspirate.

Pancytopenia due to aplastic anaemia is difficult to distinguish from 'aleukaemic leukaemia' without marrow

examination. Rare cases of ALL present with an aplastic marrow without blasts, often accompanying severe sepsis, with spontaneous reconstitution of normal haemopoiesis on recovery from the infection. They re-present several weeks to months later with low counts and circulating lymphoblasts. Clonal rearrangements are invariably detected on retrospective examination of the initial marrow using sensitive genetic techniques and it's likely that the 'aplastic' presentation of ALL in such cases is due to severe sepsis causing marrow aplasia, followed by spontaneous remission due to a cytokine release phenomenon.

As described above, bone pain, arthralgia and occasionally arthritis may mimic juvenile arthritis in a minority of patients. Therefore, a marrow examination is mandatory prior to commencing steroids or methotrexate in patients suspected of juvenile 'idiopathic' arthritis.

Paediatric small round cell tumours that involve the bone marrow, including neuroblastoma, rhabdomyosarcoma and retinoblastoma, may be confused with ALL on marrow morphology, but can be easily distinguished on flow cytometry or immunohistochemistry.

Laboratory features

A blood count at diagnosis shows varying degrees of pancytopenia, which is more severe in B-lineage than T-lineage ALL. Initial leucocyte counts (Table 22.1) range from 0.5 to >1000 × 10^9/L (median 12 × 10^9/L); they are greater than 10 × 10^9/L in slightly over half of patients and greater than 100 × 10^9/L in 10–15% of patients. Most patients have circulating blasts, but they may be difficult to detect in those with leucopenia without careful scrutiny of a blood film ('aleukaemic leukaemia'). Reactive hypereosinophilia with pulmonary infiltration and cardiomyopathy is a feature of a rare subtype of B-lineage ALL with t(5;14)(q31;q32), which results in activation of the IL-3 gene on chromosome 5 by the enhancer of the IG gene on chromosome 14. Coagulopathy is less common than in acute myeloid leukaemia, usually mild and not associated with severe haemorrhage, except in patients with very high white cell counts who are at risk of CNS haemorrhage due to leucostasis and tumour lysis. Elevated uric acid and phosphorus levels may be seen due to rapid tumour proliferation in patients with high tumour burden. An enlarged kidney can be detected in 30–50% of patients, but has no prognostic or therapeutic implications. Raised liver enzymes due to leukaemic infiltration are present in 10–20% of patients, are usually mild and have no important clinical or prognostic consequences. A minority of patients may have cholestatic jaundice due to nodes at the porta hepatis or portal tract infiltration. Abnormalities of the bone, such as metaphyseal banding, periosteal reactions, osteolysis, osteosclerosis or osteopenia, are present in half of the patients. As these changes do not affect treatment and outcome, routine diagnostic imaging studies (except for chest radiography to rule out mediastinal mass) is not necessary.

Table 22.1 Characteristics of patients in UKALL 2003 (10/2003 – 06/2011).

	Number (%)
Total	3126
Sex	
Male	1776 (57)
Female	1350 (43)
Age (years)	
Median (range)	5 (1–24)
<2	210 (7)
2–9	2077 (66)
10–15	610 (20)
≥16	229 (7)
Immunophenotype	
B-lineage	2733 (88)
T-lineage	386 (12)
Not known	7
NCI risk group	
High	1310 (42)
Standard	1816 (58)
WCC × 10^9/L	
Median (range)	12 (0.4–881)
<10	1407 (45)
10+	502 (16)
20+	526 (17)
50+	315 (10)
≥100	376 (12)

NCI, National Cancer Institute; WCC, white cell count.

Leukaemic blast cells are identified morphologically at diagnosis in the cerebrospinal fluid (CSF) of approximately one-third of cases, most of whom have no neurological symptoms. Only a minority of such cases (2–3%) fulfil the criteria for CNS leukaemia (CNS 3) which is defined by the presence of at least five leucocytes per microlitre of CSF and the detection of leukaemic blast cells, or by the presence of cranial nerve palsy. However, some studies have suggested that the presence of any leukaemic cells in CSF (even from iatrogenic introduction due to a traumatic lumbar puncture) predicts for an increased risk of relapse.

A dense infiltrate of lymphoblasts is usually detected in a marrow aspirate. Occasionally, the aspirate may be difficult to obtain and appear hypocellular with sparse blasts due to haemodilution. In such cases, a trephine invariably shows a dense infiltrate, but in rare cases with 'aplastic' presentation (see differential diagnosis) may also be severely hypocellular. Morphological

analysis of leukaemic cells distinguishes three subtypes of ALL (L1, L2 and L3), according to the the French–American–British (FAB) classification. Unlike in AML, the FAB classification has no prognostic or therapeutic relevance in ALL and, along with cytochemical staining, has been replaced by immunophenotypic and genotypic diagnostic investigation and classification.

Immunophenotypic classification (see also Chapter 19)

Early pre-B-ALL

Leukaemic blast cells of early pre-B-ALL resemble normal marrow B-cell precursors. The leukaemic cells always express CD19 and almost all cases have cytoplasmic CD22 and CD79α; surface CD22 expression is also evident in most cases. CD10 and terminal deoxynucleotidyltransferase (TdT) are expressed in 90% of cases, and CD34 in more than 75% of cases. The CD20 antigen is present on a minor proportion of blast cells in half of cases, but its intensity might increase during treatment. Early pre-B-ALL cells lack expression of surface and cytoplasmic immunoglobulins.

ALL with rearrangement of the *MLL* gene typically has an early pre-B-ALL phenotype, with distinctive features such as expression of CD15, CD65 and chondroitin protoglycan sulfate, and absence of CD10/22. Hyperdiploidy (chromosome number >50) is typically associated with weak or undetectable CD45 expression.

Pre-B-ALL

The pre-B immunophenotype is defined by the accumulation of cytoplasmic immunoglobulin μ heavy chains with no detectable surface immunoglobulins and is found in approximately 20–25% of cases. Leukaemic cells that express both cytoplasmic and surface immunoglobulin μ heavy chains without κ or λ light chains, a rare finding, have been designated transitional pre-B-ALL. Pre-B-ALL expresses CD19, CD22 and CD79α and, usually, CD10 and TdT, but only two-thirds express CD34. In many cases of pre-B-ALL, surface CD20 is absent or is weakly expressed. ZAP-70 expression is more prevalent in cases with a pre-B-phenotype. Between 20 and 25% of pre-B-ALL cases have either t(1;19)(q23;p13) or der(19)t(1;19)(q23;p13).

(Mature) B-cell ALL

In 2–4% of childhood ALL cases, cells express surface immunoglobulin μ heavy chains plus either κ or λ light chains. Commonly, cells have L3 morphology according to the FAB classification and express CD19, CD22, CD20 and frequently CD10; CD34 is negative. The less common subtype of B-cell ALL is characterized by blast cells with L1 or L2 morphology, and expression of TdT and/or CD34. These patients should be treated as for stage IV B non-Hodgkin lymphoma (Burkitt lymphoma) if they have rearrangement of the *C-MYC* gene.

T-lineage ALL

Around 15% of children and young people with the disease have T-lineage ALL, but the incidence is lower in younger children and increases with age. T-lineage ALL cells express CD7 and CD3, the latter most commonly only in the cytoplasm. Other markers commonly expressed include CD2, CD5 and TdT; CD1a, surface CD3, CD4 and CD8 are detected in fewer than 45% of cases. HLA-DR expression is uncommon, and 40–45% of cases are positive for CD10 and/or CD21. CD79α is also weakly expressed in approximately one-third of cases.

T-lineage ALL has been divided into three stages of immunophenotypic differentiation: early ($CD7^+$, $cCD3^+$, surface $CD3^-$, $CD4^-$ and $CD8^-$), mid or common ($cCD3^+$, surface $CD3^-$, $CD4^+$, $CD8^+$ and $CD1a^+$) and late (surface $CD3^+$, $CD1a^-$ and either $CD4^+$ or $CD8^+$). However, many cases have immunophenotypic patterns that do not fit these thymic maturation stages. T-cell receptor (TCR) proteins are heterogeneously expressed in T-lineage ALL. In approximately two-thirds of cases, membrane CD3 and TCR proteins are absent. In half of these cases, however, TCR proteins (TCR-β, TCR-α, or both) are present in the cell cytoplasm. Most cases with membrane CD3 and TCR chains express the αβ form of the TCR, whereas a minority express TCR γδ proteins.

A subtype of T-ALL, named early thymic precursor (ETP) ALL has been recently identified, which exhibits the gene expression profile of normal ETP cells, a population of recent immigrants from the bone marrow to the thymus that retains multilineage differentiation potential. The phenotype is characterized by positive CD7 and cCD3, negative CD1a and CD8 expression, weak or absent CD5 expression and expression of at least one stem cell or myeloid-associated antigen (e.g. CD34, CD117, CD13, CD33, CD11b). ETP ALL was reported as having a dismal outcome in initial reports, but a more recent UK report shows that its outcome is only slightly worse than for non-ETP ALL

Cytogenetic and molecular classification

Hyperdiploid and hypodiploid ALL

High hyperdiploidy (51–65 chromosomes) (HeH) is the most common genetic abnormality observed in ALL and is present in up to 35% of children and ~10% of adults with the disease. Numerous paediatric and adult studies have shown HeH to be associated with a favourable outcome. *In vitro* studies have suggested that this favourable response is related to sensitivity to methotrexate. Some studies have found that only a subgroup of HeH with trisomies of chromosomes 4, 10 and 17 have a

favourable outcome, others that trisomy 18 is associated with a low relapse risk.

In contrast to the favourable prognosis associated with HeH, patients with near-triploidy (69–81 chromosomes) have a response to therapy similar to that of non-hyperdiploid ALL; cases with near-tetraploidy (82–94 chromosomes) have a high frequency of T-cell immunophenotype. Hypodiploidy (<45 chromosomes) occurs in less than 2% of ALL cases and is associated with a poor outcome, especially for patients with a near haploid (<30 chromosomes) or low hypodiploid (30–39 chromosomes) karyotype. Some cases of near-triploidy and near-tetraploidy are due to a 'doubling up' of chromosome number in an evolved near-haploid or low hypodiploid clone. These patients also have a poor outcome.

ALL with *TEL–AML1* (*ETV6–RUNX1*) rearrangements

The chromosomal translocation, t(12;21)(p13;q22), results in the chimeric fusion product *ETV6–RUNX1* (formerly *TEL–AML1*) and is the most prevalent translocation in paediatric ALL (~25% BCP-ALL), but it is rare among adults. Unlike many chromosomal translocations, it is cytogenetically cryptic and was discovered by FISH in the mid-1990s. Virtually all major clinical trial groups around the world have reported that children with *ETV6–RUNX1* fusion enjoy excellent overall survival with very low rates of relapses. There is some evidence to suggest that the favourable outcome of *ETV6–RUNX1* ALL is related to optimal asparaginase therapy.

ALL with *E2A–PBX1* (*TCF3–PBX1*) rearrangements

Approximately 20–25% of pre-B-ALL cases have the t(1;19)(q23;p13) translocation abnormality that juxtaposes the E2A (TCF3) gene on chromosome 19 and the PBX1 gene on chromosome 1. The resulting E2A–PBX1 fusion protein contains the transcriptional activation domains of E2A linked to the DNA-binding domain of PBX1 and the encoded protein inappropriately activates the transcription of genes normally regulated by PBX1. With contemporary treatment protocols, patients with this leukaemia subtype have excellent response to initial therapy, and a favourable overall outcome.

Another *E2A* fusion gene is created by t(17;19)(q22;p13), in which *E2A* is fused to the gene that encodes the transcription factor hepatic leukaemia factor (*HLF*). Patients with ALL and this gene fusion frequently present with hypercalcaemia and coagulopathy, and have poor prognosis.

ALL with *MLL* gene rearrangements

Structural alterations involving band 11q23 of chromosome 11 are the most frequent cytogenetic abnormalities in infant ALL and occur in 1–2% of older children. The target is the *MLL* gene (for mixed-lineage leukaemia; also known as *HRX*, *ALL-1* and *TRX1*). The most common 11q23 abnormality in ALL is t(4;11)(q21;q23), which produces a chimeric protein that contains the N-terminal portion of MLL linked to the C-terminal portion of AFF1, but *MLL* has been reported to be linked to more than 50 other genes. The *MLL* gene encodes a DNA-binding protein that regulates the expression of many genes, including multiple *HOX* genes. *MLL* is crucial for embryonic development and haemopoiesis, and MLL fusion proteins can transform haemopoietic cells into leukaemia-initiating cells.

Treatment outcome of ALL with an *MLL* gene rearrangement differs by age group, with infants having the worst outcome. Infant ALL with the *MLL–AFF1* gene fusion also has a high prevalence of immature, non-productive and/or oligoclonal antigen-receptor gene rearrangements.

ALL with *BCR–ABL1* rearrangements

The t(9;22)(q34;q11) translocation encodes a chimeric gene consisting of the 5′ portion of *BCR* fused to the 3′ portion of *ABL1*, whereby N-terminal sequences of *ABL1* are replaced by *BCR* sequences. In ALL, breaks tend to occur in the minor breakpoint cluster regions, forming a 190-kDa BCR–ABL. This alteration results in a constitutively active ABL tyrosine kinase that induces aberrant signalling and activates multiple cellular pathways. Genome-wide analysis of *BCR–ABL1* ALL samples revealed deletions of *IKZF1* in 84% of cases, whereas deletions were not found in chronic-phase CML samples. Moreover, deletion of *IKZF1* appeared to be a lesion acquired at the time of transformation from CML to lymphoid blast crisis.

BCR–ABL1 ALL is refractory to standard chemotherapy, but outcomes have improved significantly in recent studies in which a tyrosine kinase inhibitor (TKI) such as imatinib was given in combination with chemotherapy throughout a 2-year treatment programme. Thus, first remission allogeneic stem cell transplant is no longer indicated for the vast majority of patients (nearly 90%) who obtain a good MRD response after 12 weeks of chemotherapy plus TKI. Patients with resistance to first-generation TKI agents may respond to second- and third-generation TKIs such as nilotinib, dasatinib and bosutinib.

ALL with *iAMP21*

A novel, recurrent chromosomal abnormality has been defined in a distinct subgroup of childhood ALL. It is characterized by the presence of an intrachromosomal amplification of chromosome 21 (iAMP21). They have a common/pre-B-immunophenotype, are significantly older (median 9 versus 5 years) and have a lower white cell count (median 3.9 versus 12.4) compared to children without this abnormality. In a previous UK trial, ALL 97, iAMP21 patients had a significantly inferior event-free and overall survival at 5 years compared with all other

patients: 28% (95% CI 11–47%) versus 78% (75–80%) and 71% (50–84%) versus 86% (85–88%), respectively. As a result of this threefold increase in relapse risk, new iAMP21 patients were treated as cytogenetic high risk in the succeeding trial, UKALL 2003, with a significant improvement in 5-year EFS to 85%.

Genetic abnormalities in T-cell ALL

Genes that are dysregulated in T-cell ALL include *SCL* (*TAL-1*), *LMO1* (*TTG-1*), *LMO2* (*TTG-2*) and *HOX11*. An additional alteration found in T-cell ALL is the deletion from chromosome 9p21 of the *CDKN2A* (*INK4A*) and *CDKN2B* (*INK4B*) genes, which encode p16INK4a and p15INK4b, inhibitors of the Cdk4 cyclin D-dependent kinase. Activating mutations of *NOTCH1* are frequently found in T-ALL and have been implicated in its pathogenesis. None of the T-ALL genetic abnormalities predict outcome to the extent that treatment should be stratified on the basis of specific abnormalities.

Novel genetic subtypes of ALL

Rearrangement of CRLF2 and JAK2

Approximately 8% of B-lineage ALL cases have CRLF2 rearrangement at the pseudoautosomal region 1 (PAR1) of Xp/Yp. CRLF2 encodes cytokine receptor-like factor 2, the receptor for thymic stromal lymphopoietin (TSLP).The arrangement occurs either as a rearrangement of CRLF2 into the immunoglobulin heavy chain locus at 14q32, or as a focal deletion immediately upstream of CRLF2 that results in expression of a novel fusion, P2RY8-CRLF2. Both events dysregulate CRLF2 expression, resulting in increased expression by lymphoblasts that may be detected by diagnostic immunophenotyping. CRLF2 rearrangement is particularly common (>50% of cases) in ALL associated with Down syndrome (DS-ALL). Half of cases with CRLF2-rearrangement harbour concomitant activating mutations in the Janus kinase genes JAK1 or JAK2. CRLF2 and JAK alterations are also associated with deleterious IKZF1 alterations and poor outcome. Importantly, JAK2 mutations common in ALL are distinct from those observed in myeloproliferative neoplasms. Coexpression of CRLF2 and JAK mutant alleles transforms model cell lines and activates downstream Jak-Stat signalling, suggesting that these alterations are cotransforming in B-ALL. In DS-ALL, alterations of IKZF1, but not of CRLF2 or JAK2, are associated with a worse prognosis.

BCR–ABL1-like ALL

Of B-ALL cases, 10–20% exhibit a gene expression profile similar to that of BCR–ABL1 ALL but do not have a BCR–ABL1-rearrangement. They commonly show IKZF1 alteration, and have a worse outcome than patients without a BCR–ABL1-like gene expression profile. Nearly 50% of BCR–ABL1-like cases have CRLF2 rearrangements and JAK mutations. Next-generation sequencing, including transcriptome and whole-genome sequencing, has shown that the remaining cases harbour a diverse range of rearrangements, deletions, and sequence mutations that activate cytokine receptor and kinase signalling (e.g, those involving ABL1, EPOR, IL7R, JAK2, and PDGFRB). Several of these were shown to be transforming *in vitro* and to activate kinase signalling in primary leukaemic cells which are sensitive to tyrosine kinase or JAK2 inhibitor. There are also case reports of responses to TKIs in patients with activating kinase mutations and refractory leukaemia.

Prognostic factors

Presenting features

Age, presenting white cell count and immunophenotype (Tables 22.2 and 22.3) are independent predictors of relapse risk in childhood ALL. Although the association of outcome with age

Table 22.2 UKALL 2003 5-year event-free survival by risk groups (% with 95% confidence intervals).

Risk group	n	EFS
Overall	3126	87.3 (86.1–88.5)
NCI standard	1816	90.8 (89.4–92.2)
NCI high	1310	82.5 (80.3–84.7)
MRD high	1037	80.2 (77.7–82.7)
MRD low	1090	94.9 (93.5–96.3)
MRD other	999	86.5 (84.3–88.7)
NCI standard		
MRD high	543	86.5 (83.6–89.4)
MRD low	680	94.7 (92.9–96.5)
MRD other	593	90.1 (87.6–92.6)
NCI high		
MRD high	494	73.1 (69.0–77.2)
MRD low	410	95.0 (92.8–97.2)
MRD other	406	81.2 (77.3–85.1)
Age <10	2287	89.8 (88.4–91.2)
Age 10+	839	80.6 (77.9–83.3)
Immunophenotype		
T	388	81.2 (77.3–85.1)
non-T	2738	88.2 (87.0–89.4)
Early response		
Rapid	2762	89.3 (88.1–90.5)
Slow	364	72.0 (67.1–76.9)
NCI high		
Age <10	471	86.0 (82.9–89.1)
Age 10+	839	80.6 (77.9–83.3)

Table 22.3 UKALL 2003 Prognostic factors in UKALL 2003.

	Univariate log rank		Univariate Cox regression		Multivariate (Cox regression)	
Variable	**Hazard ratio (95% CI)**	**p**	**Hazard ratio (95% CI)**	**p**	**Hazard ratio (95% CI)**	**p**
Sex (F v M)	0.66 (0.52–0.84)	0.0007	0.65 (0.50–0.84)	0.0009	0.80 (0.59–1.09)	0.16
Age*	1.78 (1.48–2.14)	<0.0001	1.07 (1.05–1.10)	<0.0001	1.08 (1.05–1.12)	<0.0001
WCC*	1.36 (1.25–1.48)	<0.0001	1.30 (1.21–1.41)	<0.0001	1.37 (1.21–1.54)	<0.0001
MRD risk (high v low)	3.51 (2.63–4.68)	<0.0001	4.03 (2.85–5.69)	<0.0001	3.51 (2.47–4.99)	<0.0001
NCI risk group (high v std)	2.05 (1.60–2.62)	<0.0001	2.00 (1.57–2.55)	<0.0001	0.88 (0.57–1.36)	0.56
Immunophenotype (T v B/N)	2.11 (1.45–3.08)	0.0003	1.82 (1.34–2.47)	0.0001	0.60 (0.38–0.94)	0.02

*Age and log (WCC) as continuous variables in Cox analyses, grouped in log rank as: Age <2, 2–9, 10–15, ≥16; WCC <10, 10+, 20+, 50+, ≥100.

and presenting white cell count is linear, current risk stratification models use a threshold for both variables to define standard- and high-risk groups. For example, the National Cancer Institute (NCI) criteria employed by the UKALL and US Children's Oncology groups categorizes patients below 10 years with white cell count less than 50×10^9/L as standard risk (60% of patients) and older patients or those with a higher white cell count as high risk.

Historically, the outcome for boys was worse compared with girls, but the difference is not evident in some contemporary studies. In some, but not all, US studies, patients of African-American and Hispanic origin had a significantly worse outcome.

Patients with *BCR/ABL1*, low hypodiploidy, *MLL* rearrangement, *E2A/HLF* and *iAMP 21* cytogenetic abnormalities are classified as high risk and receive more intensive therapy (but not first remission allogeneic transplant in the absence of induction failure) in most protocols. Some groups stratify patients with the good-risk ETV6/RUNX1 abnormality to receive less intensive therapy. As cytogenetic alterations alone do not accurately predict the risk of relapse, there are attempts at incorporating novel genetic alterations into risk stratification models. However, at present, no group is stratifying patients solely on the basis of *IKZF1, BCR/ABL1-like* or *ETP* status to high-risk therapy.

Early response

Risk stratification based on clinical criteria, though, is relatively non-specific. For example, a high-risk group with a 5-year EFS of around 50%, defined by age, gender and presenting WCC, identifies only 20% of patients destined for relapse, with the majority of relapses still arising out of the remaining, apparently, low-risk patients. Speed of response in the first few weeks of treatment is a powerful discriminant of outcome. Morphological slow early response, measured as slow clearance of blood or marrow blasts after a week or two of induction therapy, is associated with a relatively high risk of relapse and around 1% of patients who are not in complete remission at the end of induction (induction failures) have a poor outcome. However, over 60% of relapses arise from the group with a rapid morphological response in the first 2 weeks of therapy. Measurement of minimal residual disease (MRD) offers a much more sensitive and specific means of distinguishing between patients who will and will not relapse.

Minimal residual disease

The bone marrow of patients in morphological complete remission may contain up to 10^{10} residual leukaemic cells. This submicroscopic level of leukaemia (minimal residual disease, MRD) can be detected using sensitive molecular or flow cytometric methods. Molecular quantitation of MRD using real-time quantitative PCR of antigen receptor gene rearrangement (immunoglobulin and T-cell receptor) is the method of choice in the UK and Europe. Each leukaemic clone carries a unique molecular signature of rearranged antigen receptor genes which can be detected at diagnosis using appropriate PCR primers and sequenced. A patient-specific probe designed against the junctional region of the rearranged gene is then used to track the leukaemic clone in remission samples. Flow cytometric methods utilize combinations of cell-surface antigens uniquely expressed on leukaemic cells (leukaemia-associated immunophenotype, LAIP) to detect MRD.

Several large studies have revealed that the positive and negative predictive value of a particular MRD result depends on the sensitivity of the technique used to measure it, the time point at which it is measured, and the treatment received by the patient before and after the point of assessment. Despite differences in these variables amongst published studies, there is sufficient consensus in the results relating to certain clinical end-points. Around 50% of patients become MRD negative for two clonal markers by the end of induction (EOI) and have a <5% risk

of relapse. Patients with detectable MRD at the end of induction subdivide into two groups on follow-up. In around 95% of these patients (45% of all patients) the MRD level falls below 0.5% by the end of consolidation therapy. This group has a 20% relapse risk and benefits from further intensification of therapy, as demonstrated in UKALL 2003. The remaining 5% (1% of all patients) have persistent MRD ≥0.5% at postinduction follow-up, and an associated >80% risk of relapse. The optimal therapy for this MRD very high-risk group is uncertain and many groups consider them eligible for experimental therapy followed by first remission allogeneic stem cell transplant. The kinetics of MRD clearance is slower in T-lineage compared with B-lineage ALL and the prognostic outlook for a given detectable MRD level is worse at earlier time points for B lineage and later time points for T lineage.

Pharmacogenetic variables

Mutations and polymorphisms of genes encoding several drug-metabolizing enzymes have been associated with response to therapy. The best-known is the thiopurine *S*-methyltransferase (TPMT) gene, which encodes the enzyme that catalyses mercaptopurine to non-cytotoxic methyl mercaptopurine metabolites. Approximately 1 in 300 patients have an inherited homozygous deficiency of TPMT, which results in all the administered dose of mercaptopurine (a prodrug) being converted to cytotoxic thioguanine nucleotides (TGNs). Such patients should be treated with 5–10% of standard doses to avoid marrow aplasia. Patients who are heterozygous for this mutation (approximately 10%) have intermediate levels of enzyme activity. Both homozygous and heterozygous deficiencies are associated with better EFS, likely due to higher dose intensity of mercaptopurine.

Treatment

Drugs and protocols

Eight categories of chemotherapeutic drugs form the mainstay of childhood ALL therapy (Table 22.4). Current front-line treatment protocols contain varying combinations of these eight drugs given over a 2–3 year period. A typical treatment regimen consists of induction, pre-emptive CNS-directed, intensification and maintenance (also called continuing) therapy courses. Induction and intensification therapy are designed for rapid debulking of tumour load during the first few months of therapy, while maintenance therapy, administered over the remaining 2–3 years, is aimed at clearing MRD. Systemic therapy has variable penetration across the blood–brain barrier and, therefore, is insufficient to prevent isolated CNS relapse, the risk of which is substantially reduced by pre-emptive CNS-directed therapy.

Historical background

Studies conducted by collaborative groups in many different countries have contributed to the current generally gratifying outcome of treatment of childhood ALL. The first attempt at a cure was pioneered in the mid-sixties by the St. Jude's group in Memphis, USA, who showed that durable remissions could be achieved in roughly 30–40% of patients with a combination chemotherapy protocol, which they called total therapy, containing remission induction, pre-emptive CNS therapy and prolonged continuation therapy. While other groups tried to optimize this basic template by doing randomized studies of various components within it, the Berlin–Frankfurt–Munster (BFM) group took the second major step forward in the late 1970s, documenting that long-term remission rates could be improved to 60–70% by intensified induction and consolidation therapy. Others, including the MRC in the UK, subsequently confirmed the benefits of intensified therapy, even using different combinations and schedules of intensification from the original BFM model. Subsequent UK studies have investigated alternatives to cranial radiotherapy for CNS-directed treatment (UKALL XI), and compared the efficacy and toxicity of dexamethasone with prednisolone, and thioguanine with mercaptopurine (ALL97). Cranial radiotherapy and HD MTX reduce the risk of CNS relapse compared with prolonged IT MTX, but are associated with equivalent overall EFS due to an increase in non-CNS relapses. Hence, presently in the UK, cranial radiotherapy is restricted to patients with overt CNS disease (CNS3) at presentation who have persistent CSF blasts after two intrathecal methotrexate injections. Dexamethasone is more effective than prednisolone in reducing the risk of systemic and CNS relapses in several studies. Thioguanine is more effective than mercaptopurine at preventing CNS relapses, especially in younger boys, but is associated with an increased risk of death in remission and sinusoidal occlusion syndrome or veno-occlusive disease (SOS/VOD) of the liver. A proportion of patients with the latter toxicity have chronic portal hypertension due to periportal liver fibrosis.

Current UK strategy

Risk stratification

The risk stratification approach currently used in the UK is shown in Figure 22.2. At diagnosis, patients <10 years old with WCC $<50 \times 10^9$/L are classified as clinical standard risk, patients ≥10 years of age and those with WCC $\geq 50 \times 10^9$/L as clinical intermediate risk and patients with a cytogenetic abnormality involving rearrangement of the *MLL* gene or hypodiploidy <40 chromosomes or *iAMP21* or *E2A/HLF* abnormality or failure to remit at day 29 or induction, are classified as clinical high risk (HR). The clinically defined standard- and intermediate-risk groups are stratified by measurement of MRD at the end of induction (time point 1) and recovery from consolidation

Table 22.4 Drugs used in the treatment of childhood ALL.

Drug class	Mechanism of cytotoxicity	Route(s) of administration	Part in treatment protocol	Significant toxicities	
				Acute	Late
Vinca alkaloids	Disruption of mitotic spindle	Intravenous bolus	Throughout	Neurotoxicity Constipation Vesicant	Neurotoxicity
Glucocorticoids	Apoptosis by binding to intracellular steroid receptors	Oral and intravenous	Throughout	Hyperglycaemia Weight gain Mood alteration Hypertension	Osteoporosis Obesity Avascular necrosis of bone
L-asparaginase	Depletion of L-asparagine	Subcutaneous or intramuscular	Induction and intensification	Hypersensitivity Thrombosis Pancreatitis Hyperlipidaemia Neurotoxicity	None?
Anthracyclines	DNA intercalation Inhibition of topoisomerase II function	Intravenous infusion	Induction and intensification	Vesicant Myelosuppression	Cardiotoxicity
Alkylating agents (cyclophosphamide)	Interstrand DNA cross-links	Intravenous bolus or infusion	Intensification	Myelosuppression Haemorrhagic cystitis	Secondary leukaemia Infertility
Methotrexate	Anti-folate	Oral, intravenous and intrathecal	Maintenance CNS-directed therapy	Mucositis Nephrotoxicity (high dose) Hepatotoxicity	Osteopenia Neurotoxicity
Thiopurines (mercaptopurine and thioguanine)	Incorporation of thioguanine nucleotides into newly synthesized DNA	Oral	Maintenance	Hepatotoxicity Myelosuppression	Hepatic?
Epipodophyllotoxins (etoposide and t)	Interferes with topo II religation of double-stranded DNA breaks	Intravenous infusion	Intensification	Myelosuppression	Secondary leukaemia
Cytarabine	Inhibits DNA polymerase	Intravenous, subcutaneous or intrathecal	Intensification	Hypersensitivity Myelosuppression Central neurotoxicity	Neurotoxicity

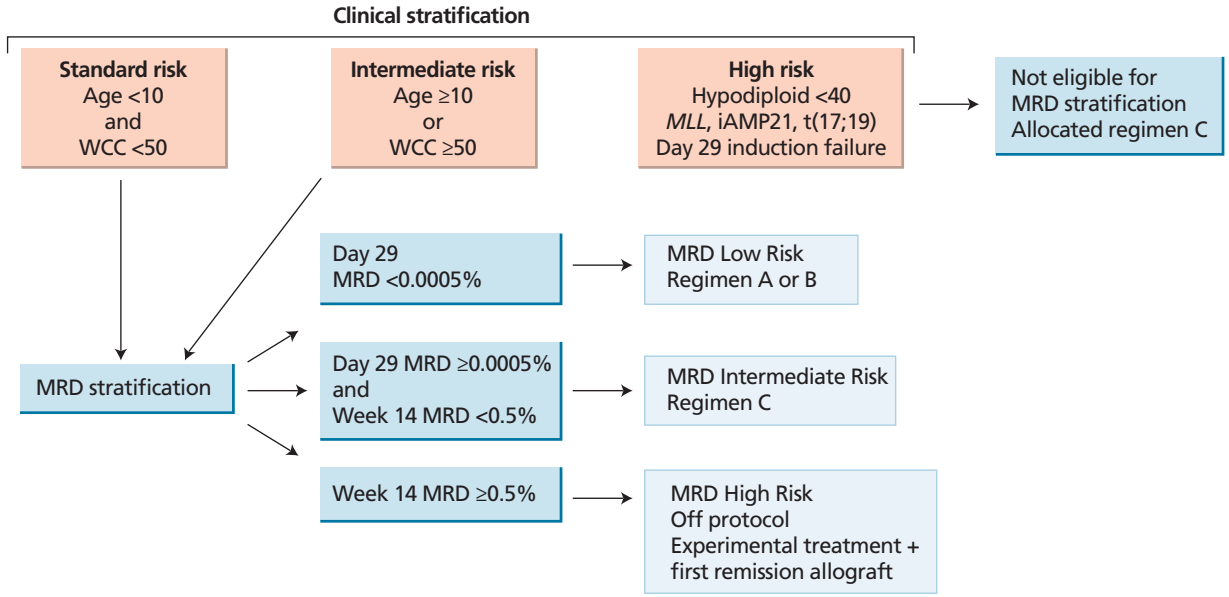


Figure 22.2 UK risk stratification for children and young people with ALL. Age in years; WCC: white cell count × 10^9/L.

(prior to start of interim maintenance, time point 2). Patients with an MRD level <0.005% at TP1 are classified as MRD low risk, whereas patients with MRD above that level at TP1, which has fallen to below 0.5% at time point 2 (TP2) are classified as MRD intermediate risk. Patients with persistent MRD ≥0.5% are classified as MRD high risk and receive experimental therapy to reduce MRD prior to a first remission allogeneic stem cell transplant.

Treatment

Patients receive one of three escalating intensity treatment regimens depending on their clinical and MRD risk group (Figure 22.3). Initial treatment allocation is on the basis of clinical risk criteria with treatment post induction being determined by MRD response. The treatment regimens described below are the standard treatment arms of the current randomized UK trial, UKALL 2011.

Induction and consolidation

Clinical standard-risk patients receive a three-drug induction containing vincristine, steroids and asparaginase for 4 weeks, and intermediate- and high-risk patients receive in addition daunorubicin. All patients receive three doses of intrathecal methotrexate (IT MTX) in induction, with patients who have blasts in their cerebrospinal fluid at diagnosis receiving an additional two doses.

For consolidation, clinical standard-risk patients who are MRD low risk receive daily oral mercaptopurine and three doses of weekly intrathecal methotrexate. Patients who are clinical intermediate-risk patients and MRD low risk receive in addition 4 weeks of cyclophosphamide and cytarabine (BFM consolidation). Clinical high-risk patients and MRD intermediate-risk patients receive an additional four doses of vincristine and two doses of pegylated asparaginase during the BFM consolidation course.

Interim maintenance and delayed intensification

Following consolidation, all patients receive 2 months of interim maintenance prior to a single delayed intensification course. MRD low-risk patients receive oral mercaptopurine and methotrexate with monthly vincristine and steroid pulses during interim maintenance. Interim maintenance for clinical high-risk and MRD intermediate-risk patients consists of escalating doses of intravenous methotrexate without folinic acid rescue, vincristine and pegylated asparaginase.

All patients receive a single delayed intensification. For MRD low-risk patients this consists of a single dose of pegylated asparaginase at day 4 and vincristine, dexamethasone, doxorubicin for 3 weeks, followed by cyclophosphamide and cytarabine, as given during the BFM consolidation course. MRD intermediate- and clinical high-risk patients receive in addition two doses of vincristine and one dose of pegylated asparaginase.

Continuation therapy

Regardless of clinical and MRD risk group, all patients receive oral mercaptopurine and methotrexate, monthly vincristine and steroid pulses and 3-monthly intrathecal methotrexate. Boys receive treatment for 3 years and girls for 2 years from the start of interim maintenance.

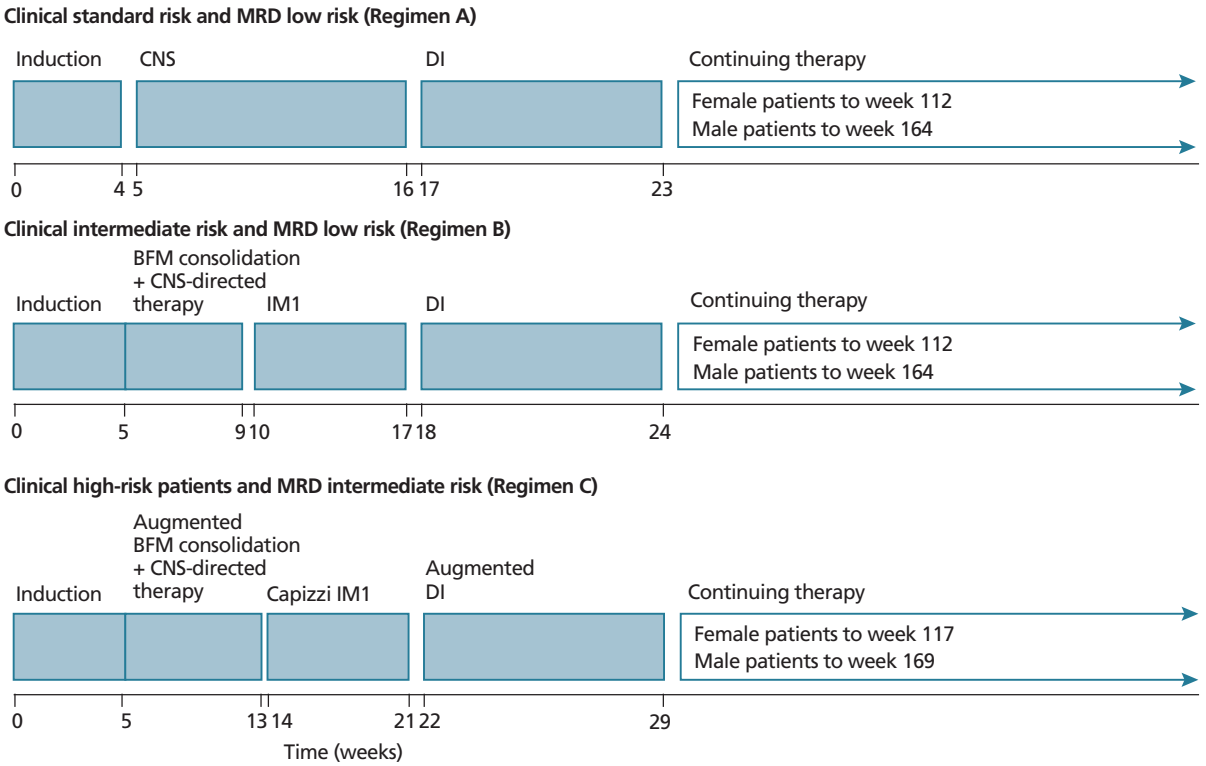


Figure 22.3 Outline of current UK treatment regimens. IM, interim maintenance; DI, delayed intensification; BFM, Berlin–Frankfurt–Munster; MRD, minimal residual disease; MRD, intermediate risk patients assigned to augmented therapy transferred to Regimen C after induction.

Steroid and asparaginase formulations, doses and schedules

All patients receive dexamethasone 6 mg/m^2 with a 10 mg ceiling dose during induction and maintenance courses. In delayed intensification courses, all patients receive dexamethasone at 10 mg/m^2 (without a cap) for 14 days in a week-on, week-off schedule. All patients receive pegylated asparaginase 1000 units/m^2/dose given intramuscularly throughout treatment. Standard- and intermediate-risk patients receive three doses (two in induction and one in delayed intensification course) and high-risk patients eight doses (additional two in interim maintenance and delayed intensification courses).

CNS-directed therapy

Patients with ≥5 leucocytes/μL and blasts in a diagnostic CSF sample with <10 red cells/μL (CNS-3) receive an extra two IT MTXs in induction. Patients with traumatic lumbar puncture and blasts in the CSF, as well as those with <5 leucocytes/μL which were blasts (CNS-2) also receive an extra two IT MTXs during induction. All other patients received intrathecal methotrexate as described above. High-risk patients receive Capizzi intravenous methotrexate at doses <500 mg/m^2 without folinic acid rescue. Thus, no patients receive cranial radiotherapy during first-line treatment

Allogeneic haemopoietic stem cell transplantation (HSCT)

Less than 5% of patients are eligible for an allogeneic transplant in first remission. These include patients with >25% blasts in their marrow at day 29 of induction or with a high-risk karyotype and >5% blasts at that time point. In addition, patients with MRD >0.5% at TP2 are eligible for first remission allogeneic SCT after experimental therapy to reduce the MRD level. Autologous stem cell transplantation is of no benefit in ALL.

Current outcomes

Given the above treatment, the 5-year EFS of over 3000 children and young people (aged 1–25 years) recruited to the recently concluded UK trial, UKALL 2003 (2003–2011) was 87.2%, which compares favourably with outcomes reported from contemporary trials by other groups (Table 22.5). A low incidence of isolated CNS relapse (1.9%) was observed, equivalent to that observed in studies in which a significantly higher proportion of patients received cranial irradiation. Randomized interventions within the trial demonstrated that treatment can be de-escalated without compromising survival in MRD low-risk patients, and augmentation of postremission therapy reduces relapse risk for patients with detectable MRD at the end of induction. Thus, in

Table 22.5 Outcomes with contemporary childhood ALL protocols.

Trial	Group	Region	Years	Subgroup (n)	EFS (years)	OS (years)
Several	COG	US, Canada, Australia, New Zealand	2000–2005	All patients (6994) B-ALL (5845) T-ALL (457)	N/A N/A N/A	91.3% (5) 92.0% (5) 81.5% (5)
Total XV	SJCRH	US	2000–2007	All patients (498) B-ALL (422) T-ALL (76)	85.6% (5) 86.9% (5) 78.4% (5)	93.5% (5) 94.6% (5) 87.6% (5)
00–01	DFCI	US, Canada	2000–2004	All patients (492) B-ALL (443) T-ALL (49)	80.0% (5) 82.0% (5) 69.0% (5)	91.0% (5) N/A N/A
AIEOP-BFM 2000	BFM	Western Europe	2000–2006	All patients B-ALL (4016) T-ALL (464)	N/A 80.4% (7) 75.9% (7)	N/A 91.8% (7) 80.7% (7)
ALL-9	DCOG	Netherlands	1997–2004	All patients (859) B-ALL (701) T-ALL (90)	81% (5) 82% (5) 72% (5)	86% (5)
UKALL 2003 (age 1–25)	MRC/NCRI	UK	2003–2011	All B T	87.3% (5) 88% (5) 82% (5)	91.6% 92.3% 86.4%

AIEOP–BFM, Association of Italian Paediatric Oncology and Berlin Frankfurt–Munster; COG, Children's Oncology Group; SJCRH, St. Jude Children's Research Hospital; DFCI, Dana Farber Cancer Institute Consortium; MRC, Medical Research Council; NCRI, National Cancer Research Institute.
*Infants <1 year old excluded where possible.

the current UK trial, postremission treatment intensity is stratified by MRD response at the end of induction and recovery from consolidation.

The improved outcomes reported in UKALL 2003 and other recent trials have been obtained in the absence of new drugs for the treatment of ALL for over 40 years. The use of dexamethasone and pegylated asparaginase throughout treatment is likely to be an important contributor to the improvement seen in UKALL 2003. Several large randomized clinical trials have demonstrated the improved efficacy of dexamethasone in preventing systemic and CNS relapse compared with prednisolone. Pegylated asparaginase has better pharmacokinetic and pharmacodynamic properties than the native formulation and a lower risk of hypersensitivity reactions on re-exposure.

Treatment of distinct sub-groups

Young people (age 16–25 years)

A decade ago, retrospective comparisons demonstrated a consistent and large EFS or OS advantage for young people with ALL treated according to paediatric protocols compared with adult protocols. The reasons for this were unclear, but possibly included physician experience and compliance, patient compliance, supportive care and specific aspects of protocol design. In particular, early dose intensification of chemotherapy, higher cumulative doses of steroids, vincristine and L-asparaginase and less frequent use of alkylating agents, anthracyclines, high-dose cytarabine and allogeneic stem cell transplant (with the associated higher treatment-related mortality) in paediatric protocols. The results of these retrospective comparisons have been validated in prospective trials that recruited 16–25-year-old patients to a paediatric protocol, such as UKALL 2003, in which that age group had a 5-year EFS of 75%.

Infants

Acute lymphoblastic leukemia (ALL) in infants under 1 year of age is rare and biologically different from ALL in childhood. Infant ALL is characterized by a high frequency of MLL gene rearrangements, a very immature B-cell phenotype (proB ALL), expression of myeloid markers, lack of CD10/22 expression, presentation with a high tumour load and a poor outcome. MLL rearrangement status, presenting white blood cell count (WBC), age at diagnosis and prednisone response are independent prognostic factors and a model comprising age, MLL status and WBC stratifies patients into three risk groups with distinct 4-year EFS (low: MLL germline, EFS 74%; intermediate: MLL rearranged, WBC <300 or age <6 months, EFS 43% and high: MLL rearranged and age <6 months and WBC >300, EFS

18%). The current treatment approach within an international trial, Interfant 06, combines this risk stratification algorithm and MRD response to select patients for first remission allogeneic transplant. Unfortunately, despite these international efforts, the outcome for this subgroup remains poor. Molecular investigations and preclinical studies indicate that epigenetic modifiers and immune-based approaches might be effective in MLL rearranged cases and are to be tested in future trials.

Down syndrome

Children and young people with DS have a 10–20-fold increased risk of developing ALL compared to those without DS. This increased risk is limited to the first three decades of life, but with a notable absence of ALL in those under a year of age. T-cell ALL is rare in DS, as is the presence of CNS disease at diagnosis. ALL associated with DS (DS-ALL) has a distinct frequency of genetic changes from ALL seen in the non-DS population. Recurring cytogenetic abnormalities conferring either a favourable or a poor prognosis are less common. Favourable risk cytogenetics; high hyperdiploidy and ETV6/RUNX1 fusions, occur in 50% of children with non-DS-ALL compared to 10–20% in DS-ALL. Activating CRLF2, JAK2 and IZKF1 deletions are found more commonly in DS-ALL. DS-ALL has an inferior survival due to a combination of increased relapse risk and high treatment-related mortality (TRM). Many groups treat DS-ALL with reduced intensity treatment and recommend additional supportive care measures. As in non-DS patients, good-risk karyotype and MRD low-risk status are associated with a significantly lower risk of relapse in DS-ALL. These subgroups may benefit from treatment de-escalation to reduce the risk of treatment-related mortality and morbidity.

Relapse

Relapse may occur during or after treatment, usually within 2 years after cessation of therapy, but occasionally later (relapses occurring as late as 10 years after diagnosis have been reported). The bone marrow remains the most common site of relapse, but the proportion of patients with relapse involving an extramedullary site, particularly the CNS, has steadily increased as systemic first-line therapy has become more effective. Even when extramedullary relapse appears to be an isolated event, it is often associated with submicroscopic residual disease in the bone marrow. Although the clonal populations at relapse may not genetically resemble the dominant population at diagnosis, sensitive PCR techniques reveal that they are present as minor subclones, which have been selected due to their relative resistance to chemotherapy.

The chances of cure after relapse are related primarily to timing of relapse in relation to first diagnosis, immunophenotype and site. Adverse risk factors include short initial remission, T-cell immunophenotype and marrow involvement. MRD response remains predictive of prognosis after relapse in patients who relapse after completion of first-line therapy, and such patients may be cured by second-line chemotherapy alone. Patients with on-treatment relapse and those with later relapse and a slow MRD response to re-induction require a second remission allogeneic stem cell transplant for cure. The probability of EFS after transplantation is significantly associated with MRD burden before transplant and the kinetics of MRD following transplant.

Early and late toxicity (Table 22.4)

Modern intensive therapy is associated with a small risk of mortality (1% in induction, 2% in CR) and significant acute and late morbidity. A majority of the latter are attributable to cranial radiotherapy and there has been a gratifying reduction in the number of patients with endocrine, growth and neuropsychological problems since it was replaced by alternative approaches to prevention of CNS relapse. Dexamethasone is more toxic than prednisolone and likely to be a significant contributor to the risk of serious bacterial and fungal infection. In a recently concluded European study, AIEOP-BFM 2000, there were twice as many induction deaths among patients in the dexamethasone arm compared with the prednisolone arm and other studies have reported a similar experience. There is also concern that dexamethasone is associated with a higher risk of osteonecrosis than prednisolone, but that has not been a universal experience. Dose, scheduling and interaction with anthracyclines are important determinants of the risk of steroid-associated toxicity and the current UK trial, UKALL 2011, is testing novel schedules of dexamethasone in an attempt to reduce the toxicity, while retaining the efficacy of this drug. Toxicity attributable to asparaginase occurs in a significant proportion: thrombosis in 3% of patients with 1% involving the central nervous system. Hypersensitivity reactions occur in 1.5% of patients overall, with a majority occurring in high-risk patients primarily on re-exposure to the drug during the consolidation or first high-risk interim maintenance course.

Treatment in a resource-poor setting

The drugs, staffing and infrastructure required to deliver the treatment described above is very expensive, with an estimated UK cost of £50,000/case. Over 80% of new cases of ALL will arise in countries that cannot afford this cost. Thus, there is an urgent need for studies to test more affordable treatment protocols that can be delivered within a resource-poor setting. The challenge of delivering optimal care to such a large number of patients in a resource-poor setting has been met by clinicians in those countries through a combination of adaptation and innovation. For a number of reasons, apart from lack of resources, outcomes are worse compared with the West, even within a tertiary care setting. A lower incidence of good-risk karyotypes (*ETV6/RUNX1*, high hyperdiploidy),

malnutrition, partial compliance and poor general hygiene conspire to increase the risk of relapse and infection-related deaths. There are no easy solutions to these problems. Simple flow cytometric algorithms for measurement of MRD may be applicable in this setting and could provide a relatively cheap method to identify low-risk patients who can be cured without post-remission intensification therapy, thus avoiding infection-related deaths in this group. Learning from the experience of tuberculosis, pilot programmes should test the feasibility of directly observed treatment to improve compliance with maintenance chemotherapy. Universal access to treatment, though, remains the single most difficult challenge in this setting. As these nations' overall wealth increases, only a massive investment in public health infrastructure will address this challenge.

Future strategies and conclusions

Given the good results of current protocols, future studies of childhood ALL therapy are faced with the law of diminishing returns. However, there remains a substantial minority of patients with primary refractory disease (around 2%) or early relapse (10%) who cannot be cured with current treatment, including haemopoietic stem cell transplant. These patients may be identified in first CR as having persistent high-level MRD during the first 20 weeks of treatment. Intervention with novel agents (monoclonal antibodies, tyrosine kinase inhibitors, nelarabine, clofarabine and autologous CAR T cells) followed by HSCT early in first CR (between weeks 12 and 20) might offer some of these patients a cure. In the UK, we plan to test novel agents during the week 12–20 window in these patients, using MRD as a surrogate marker.

As cure rates improve, greater attention should focus on reducing treatment-related deaths, which make up an increasing proportion of treatment failures. Identification of groups at high risk of toxicity (e.g. Down syndrome) and pharmacogenomic expression profiling will guide targeted supportive care and individualized drug dosing to reduce toxic deaths. There is evidence that gene expression profiling of leukaemic blasts can predict *in-vitro* and *in-vivo* chemosensitivity and treatment in future could be customized to a patient's pharmacogenomic and leukaemia gene expression profiles. The current UK trial, UKALL 2011, is testing ways of further reducing the risk of treatment-related toxicity. In future, new drugs designed to target leukaemia-specific receptors and proteins could replace elements of conventional chemotherapy regimens responsible for some of the major toxicities, thereby reducing toxicity, whilst retaining overall efficacy of treatment. Ultimately, translation of recent advances in understanding of the molecular biology of ALL and its influence on phenotype and clinical outcome will help define specific subgroups that might benefit from such an approach. Lastly, international collaboration, as highlighted by the INTERFANT protocol, will need to increase so as to properly investigate new treatment strategies and the biological determinants of treatment response in rare subgroups of patients such as those with the Philadelphia chromosome and near-haploidy abnormalities.

Acknowledgements

This version of the chapter retains some of the content of the previous version written by Dr Ching-Hon Pui and Dr Dario Campana.

Selected bibliography

Bateman CM, Colman SM, Chaplin T *et al.* (2010) Acquisition of genome-wide copy number alterations in monozygotic twins with acute lymphoblastic leukemia. *Blood* **115**(17): 3553–8.

Biondi A, Schrappe M, De Lorenzo P *et al.* (2012) Imatinib after induction for treatment of children and adolescents with Philadelphia-chromosome-positive acute lymphoblastic leukaemia (EsPhALL): a randomised, open-label, intergroup study. *Lancet Oncology* **13**(9): 936–45.

Hunger SP, Mullighan CG (2015) Redefining ALL classification: toward detecting high-risk ALL and implementing precision medicine. *Blood* **125**(26): 3977–87.

Jabbour E, O'Brien S, Ravandi F, Kantarjian H (2015) Monoclonal antibodies in acute lymphoblastic leukemia. *Blood* **125**(26): 4010–16.

Lennard L, Cartwright CS, Wade R, Richards SM, Vora A (2012) Thiopurine methyltransferase genotype-phenotype discordance, and thiopurine active metabolite formation, in childhood acute lymphoblastic leukaemia. *British Journal of Clinical Pharmacology* **76**(1): 125–36.

Maude SL, Teachey DT, Porter DL, Grupp SA CD19-targeted chimeric antigen receptor T-cell therapy for acute lymphoblastic leukemia. *Blood* **125**(26): 4017–23.

Moorman AV, Ensor HM, Richards SM *et al.* (2010) Prognostic effect of chromosomal abnormalities in childhood B-cell precursor acute lymphoblastic leukaemia: results from the UK Medical Research Council ALL97/99 randomised trial. *Lancet Oncology* **11**(5): 429–38.

Mullighan CG (2013) Genomic characterization of childhood acute lymphoblastic leukemia. *Seminars in Hematology* **50**(4): 314–24.

van Dongen JJM, van der Velden VHJ, Bruggemann M, Orfao A (2015) Minimal residual disease diagnostics in acute lymphoblastic leukemia: need for sensitive, fast, and standardized technologies. *Blood* **125**(26): 3996–4009.

Vora A, Goulden N, Wade R *et al.* (2013) Treatment reduction for children and young adults with low-risk acute lymphoblastic leukaemia defined by minimal residual disease (UKALL 2003): a randomised controlled trial. *Lancet Oncology* **14**(3): 199–209.

CHAPTER 23

Supportive care in the management of leukaemia

23

Eliza Gil[1], Vanya Gant[2] and Panagiotis Kottaridis[1,2]

[1]University College London Hospitals NHS Trust
[2]Royal Free London NHS Foundation Trust, London, UK

Introduction

The outlook for patients with leukaemia has tremendously improved within the last 10 years, mainly because of new and more effective treatments, such as targeted therapies, development of new combinations and protocols, safer delivery of higher doses of chemotherapy and improvements in supportive care.

Supportive care requires a multidisciplinary approach by medical and non-medical staff working in dedicated haematological wards with experience in treating patients with haematological malignancies. Complications requiring supportive care can be classified as follows:

- Psychological, social and financial
- Reproductive
- Anaemia
- Thrombocytopenia
- Infections
- Chemotherapy-induced nausea and vomiting (CINV)
- Nutritional
- Metabolic
- Skin, nail, dental, auditory and visual complications
- Pain control
- Palliation.

Psychological

Most patients diagnosed with leukaemia may experience one or more of the following:

- Acute or chronic anxiety
- Uncertainty about the future
- Anger
- Mood disturbances
- Adjustment difficulties
- Distress
- Family communication problems
- Changes in body image
- Depression
- Decision-making difficulties
- Challenges balancing illness and treatment alongside the demands of life.

As there is growing evidence that the provision of effective psychosocial care improves the outcomes of patients with cancer, it is essential for all haematological units to be able to provide such a demanding type of care to patients and their relatives. Central to the implementation of this care is that health professionals have the necessary communication and assessment skills, for example, discuss prognosis and treatment options available with the patient or the move from curative to palliative treatments. Such communication is vital to enable the provision of appropriate, accurate and detailed information to the patient at key stages relating to the pathological process of the disease. Because this information has the potential to affect the patient's decision-making process, the provision of information must be appropriate, accurate and detailed.

This is primarily the task of a senior haematologist accompanied by a specialist nurse and supported by junior doctors and nurses, psychologists or occasionally psychiatrists.

Postgraduate Haematology, Seventh Edition. Edited by A Victor Hoffbrand, Douglas R Higgs, David M Keeling and Atul B Mehta.

Unfortunately, high workload and a lack of available time for health professionals might limit their ability to engage in conversation with patients in order to elicit their specific needs.

Therefore, the development of effective strategies to assist clinicians to dedicate more time to their patients and improve their communication skills is fundamental to achieving optimal psychosocial outcomes for patients.

Social

Social support is a critical, yet under-utilized resource when undergoing cancer care. Under-utilization occurs in two conditions: (a) when patients fail to seek out information, material assistance and emotional support from family and friends or (b) when family and friends fail to meet the individualized needs and preferences of patients.

Patient's supportive groups run by departments, hospitals, organizations or charities might provide major social support to patients and relatives.

Financial

Most cancer patients experience financial problems. This might be more relevant to patients with leukaemia, as profound immunosuppression and delayed immune reconstitution might result to prolonged absence from work with devastating financial losses. Therefore, governmental bodies, health boards, organizations and charities must do more to ensure people living with cancer are signposted to good-quality specialist and timely welfare benefits support to help them cope and manage the financial impact of their diagnosis.

Reproductive

Fertility preservation

Fertility issues should be addressed in all patients of reproductive age before cancer treatment. In men, cryopreservation of sperm should be offered, regardless of the risk of gonadal failure. In women, the recommendation of fertility preservation should be individualized, based on multiple factors, such as the urgency of treatment, the age of the patient, the marital status, the regimen and the dosage of cancer treatment. If the risk of gonadal failure is very low, fertility preservation may not be required. If the patient is a potential transplant candidate, fertility preservation should be considered, as the infertility rate post transplant is excessively high. In principle, embryo cryopreservation or oocyte cryopreservation is recommended as a fertility preservation option, if there is enough time for ovarian stimulation before initiation of cancer therapy. In leukaemias, as opposed to some slow-progressing lymphomas, the urgency of treatment is such that rarely can female patients undergo oocyte stimulation techniques.

However, it is very reassuring that the risk of infertility in patients with ALL or AML who undergo conventional chemotherapy is very low, as contemporary treatment protocols entail lower doses or are devoid of alkylating agents.

GnRH-a

Clinical data support a viewpoint that prepubertal girls have greater tolerance than postpubertal women during gonadotoxic treatments. Ovarian suppression with gonadotropin-releasing hormone agonist (GnRH-a) during chemotherapy has been advocated by various groups. GnRH-a acts by suppressing the pituitary ovarian axis, decreasing ovarian perfusion and preventing germ cell apoptosis. Other collateral benefits include suppression of menses and resulting prevention of anaemia and bleeding during the aplastic phase.

One hypothesis is that the administration of GnRH-a could put the ovary into an endocrine prepubertal state, with less susceptibility to cancer treatments. Studies in monkeys have demonstrated the protective role of GnRH-a against the primordial follicle loss associated with cyclophosphamide; however, they failed to find the same role in ionizing-radiation-induced ovarian injuries. The available clinical research is highly controversial, with some meta-analyses showing a benefit of ovarian suppression by GnRH-a, but others not showing a benefit with this approach. Most of the studies have been conducted in patients with lymphoma and solid tumours and rarely in leukaemia patients. The American Society of Clinical Oncology concludes that there is insufficient evidence that ovarian suppression protects fertility from gonadotoxic therapies and therefore large randomized clinical studies with long-term follow-up should be performed to clear the safety and efficacy of GnRH-a treatment.

Anaemia

Chemotherapeutic agents induce anaemia through directly impairing hemopoiesis, including synthesis of RBC precursors, in the bone marrow. In addition, nephrotoxic effects of particular cytotoxic agents (e.g., platinum-containing agents) can also lead to anaemia through decreased renal production of erythropoietin. The malignancy itself can lead to or exacerbate anaemia in several ways: leukaemia cells may directly suppress haematopoiesis through bone marrow infiltration. They may produce cytokines that lead to iron sequestration, which decreases RBC production and may even shorten survival. Additional indirect effects may include nutritional deficiencies caused by loss of appetite, haemolysis by immune-mediated antibodies or changes in coagulation capability.

There is no good evidence to support a particular red cell transfusion policy in acute leukaemia. Therefore, the decision to conduct PRBC transfusion should not be made strictly based on whether the Hb level has reached a certain threshold or 'trigger'. As a general rule, transfusion is rarely indicated when the Hb level is greater than 90 g/L. However, different transfusion thresholds should be adopted, having taking into consideration three general categories of patients: (i) asymptomatic ones without significant comorbidities, for which observation and periodic re-evaluation are appropriate, (ii) asymptomatic with comorbidities or high risk, for which transfusion should be considered and (iii) symptomatic, for which patients should receive transfusion.

Older age and the presence of pre-existing cardiovascular, pulmonary or cerebral vascular disease may compromise a patient's ability to tolerate anaemia. Hence, decisions related to whether immediate correction of anaemia is needed must be based on an assessment of individual patient characteristics, degree of severity of anaemia, presence and severity of comorbidities, and clinical judgment of the physician.

Use of erythroid-stimulating agents (ESA)

RBC production is normally controlled by erythropoietin, a cytokine produced in the kidneys. First introduced in 1989, ESAs are a synthetic, recombinant human erythropoietin that can stimulate erythropoiesis in patients with low RBC levels.

The use of ESA remains very common in patients treated by chemotherapy for solid tumours and lymphoid malignancies. The American Society of Hematology (ASH) and the American Society of Clinical Oncology (ASCO) recommend a strict benefit/risk evaluation by clinicians for its use compared with RBC transfusions.

In 2007, the FDA made substantial revisions to the label information and regulations regarding epoetin α and darbepoetin α, including the addition of a 'Black Box' label warning and implementation of a risk management programme known as Risk Evaluation and Mitigation Strategy. The strengthened FDA restrictions were mainly based on the results of eight randomized studies that showed a decrease in overall survival and/or locoregional disease control with ESA use for advanced breast, cervical, head and neck, lymphoid, and non-small cell lung cancers. Of the eight studies, three investigated ESA effects in patients who underwent chemotherapy. All eight trials had an off-label target Hb level of greater than 120 g/L.

Worsened health outcomes associated with the use of ESAs have been confirmed in three recent meta-analyses of 51 to 53 randomized controlled trials by Bohlius *et al*., Tonelli *et al*. and Bennett *et al*. However, this association has been refuted by two other meta-analyses reporting no significant effect of ESAs on mortality or progression. In addition, several recent pharmacovigilance trials reported no decrease in survival with ESA use in patients with chemotherapy-related anaemia.

Increased thromboembolic risk has also been associated with ESA treatment in patients with cancer. The cause of VTE is complex, including the underlying malignancy per se and the chemotherapy regime. Other risk factors for thromboembolic disease in patients with cancer include prior history, hypercoagulability due to underlying thrombophilia, elevated pre-chemotherapy platelet counts, recent surgery, hormonal agents, immobilization, steroids and comorbidities such as hypertension.

In patients receiving chemotherapy for acute myeloid leukaemia or undergoing allogeneic hemapoietic stem cell transplantation, the use of ESAs has not been clearly assessed and is still controversial, which explains the need for prospective evaluation of ESAs in these two populations.

Special considerations

All patients in whom allogeneic transplantation may be considered should receive cytomegalovirus (CMV)-negative or leucodepleted products until their CMV status is known. Patients treated with purine analogue-based chemotherapeutic regimens (i.e. FLAG) should be supported with irradiated blood products.

Thrombocytopenia

Prophylactic platelet transfusion should be administered to patients with thrombocytopenia resulting from impaired bone marrow function to reduce the risk of bleeding when the platelet count falls below a predefined threshold level. This threshold level for transfusion varies according to the patient's diagnosis, clinical condition, and treatment modality. The optimal platelet dose has not been identified.

Several recent clinical trials have addressed the question of which peripheral blood platelet count should trigger prophylactic transfusions, i.e. transfusions given to patients with severe thrombocytopenia without clinical haemorrhage. The prophylactic strategy seems beneficial, and available clinical guidelines generally recommend prophylactic transfusions when the peripheral blood platelet count falls to 10×10^9/L. This recommendation is supported by four randomized clinical trials that showed no statistically significant difference between the compared groups with regard to mortality, remission rates of the malignant disease, frequency of severe haemorrhages or erythrocyte transfusion requirements when comparing 10×10^9/L with a higher threshold. The use of 10×10^9/L as the threshold is in addition supported by a retrospective study comparing 224 patients transfused at a threshold of 20 and 256 patients transfused at a threshold of 10; these two groups showed similar bleeding incidence and erythrocyte transfusion requirements.

A recent trial suggested that a policy of giving platelet transfusions only as treatment for bleeding might become a new

standard of care in selected patients. However, higher rates of grade 3 or 4 bleeding events were reported among patients with acute myeloid leukaemia who received no prophylaxis, as compared with those who did receive prophylaxis. On the basis of these findings, and the results of a study by Stanworth *et al*., as published in *NEJM* in 2013, the use of prophylactic platelet transfusions in patients with haematologic cancers treated with chemotherapy or allogeneic haemapoietic stem cell transplantation appears to be important as the proportion of patients who had bleeding events of WHO grade 2, 3, or 4 was reduced by 7% overall in the group that received prophylactic platelet transfusions.

Platelet preparations

Platelets for transfusion can be prepared either by separation of units of platelet concentrates (PCs) from whole blood, which are pooled before administration, or by apheresis from single donors. Comparative studies have shown that the posttransfusion increments, haemostatic benefit and side-effects are similar with either product. Thus, in routine circumstances, they can be used interchangeably. In most centres, pooled PCs are less costly. Single-donor platelets from selected donors are preferred when histocompatible platelet transfusions are needed. Both preparations can be stored for up to 5 days after collection at 20–24 °C with good maintenance of platelet viability.

TPO receptor agonists

Thrombopoietin (TPO) mimetics, such as romiplostim and eltrombopag (EP), have been developed to treat patients with refractory immune thrombocytopenia. TPO is a potent endogenous cytokine that binds to its receptor (c-Mpl) on megakaryocyte progenitors and stimulates a number of signal transduction events promoting megakaryocyte proliferation and maturation, and platelet release. Recent studies have suggested that use of TPO mimetics can be of clinical value in the treatment of refractory severe aplastic anaemia. Multiple studies have also been undertaken to explore the potential utility of eltrombopag in the treatment of thrombocytopenia in patients with AML or MDS. In addition to inducing thrombopoiesis, EP exhibits an antileukemic effect. Given the platelet-increasing effects, the *in vitro* and *in vivo* antileukemia effects and the potential to chelate excess iron in multiply transfused patients, the utility of EP in patients with AML and MDS is actively being explored as a dual thrombopoietic and antineoplastic agent in different types of cancer.

Granulocyte transfusions

Small randomized controlled studies over the past 20 years have shown some efficacy in the use of granulocytes; however, most of the studies lack power and provide inconclusive evidence. Nevertheless, there is increasingly anecdotal evidence that pooled buffy coat or apheresed granulocytes may be used to treat localized, primarily skin infections, unresolving fevers or severe invasive fungal infections in neutropenic patients.

The largest study so far (RING study) was initially presented at the 56th Annual Meeting of the American Society of Hematology in San Francisco, and subsequently was published in *Blood* in September 2015. The primary end points of the study were survival and microbial response at 42 days from randomization. Both chemotherapy and recipients of haemopoietic stem cell transplants were eligible to participate. Surprisingly, there was no statistical difference between the two groups; however, analysis suggested that patients receiving a high dose (more than 0.6×10^9/L) might benefit.

As HLA alloimmunization has been reported in 17% of patients, patients who are transplantation candidates should be re-tested after granulocyte transfusions to ensure the absence of HLA antibodies directed against the potential donor. This might be more relevant in recipients of haploidentical transplants.

Granulocyte colony-stimulating factor (G-CSF)

Myelosuppression due to cytotoxic drugs is a major limiting factor in the treatment of leukaemia. Neutropenia places patients at a high risk of fever, infections, septicaemia and ultimately death.

Three prophylactic strategies that have garnered much attention are: prophylactic antibiotics with antibacterial and antifungal agents, prophylactic G-CSF and mandatory hospitalization during profound neutropenia. G-CSF was licensed in 1991 by the United States Food and Drug Administration (FDA) for use in patients undergoing cytotoxic treatment for decreasing the incidence of febrile neutropenia (FN) and reducing the duration of neutropenia and fever following myeloablative chemotherapy.

The American Society of Clinical Oncology (ASCO), the American National Comprehensive Cancer Network (NCCN) and the European Organisation for Research and Treatment of Cancer (EORTC) have set-up guidelines for the use of G-CSF and GM-CSF in patients receiving chemotherapy with the objective of preventing fever and infections and to maintain chemotherapy dose intensity. They advocate the use of colony-stimulating factors to prevent FN in patients at a >20% risk for fever. Risk factors for FN include prior chemotherapy, abnormal hepatic and renal function, low white blood count, chemotherapy and planned delivery of ≥ 85% of the dose of chemotherapy.

The efficacy of G-CSF administered during and/or after induction chemotherapy in patients with AML or acute lymphoblastic leukaemia (ALL) has been examined in randomized, multicentre trials.

Conflicting results are available concerning the use of G-CSF, although results of a meta-analysis suggest that use of colony-stimulating factors in febrile neutropenia accelerates neutrophil recovery and reduces the duration of hospitalization. ASCO guidelines do not recommend the routine use of G-CSF in

patients with afebrile neutropenia; however, the use of G-CSF should be considered in patients with febrile neutropenia who have risk factors predictive of a poor outcome or who are at high risk of infection-associated complications.

Dose intensification is being investigated in an attempt to improve outcomes in patients with cancer. Administration of G-CSF permits the use of intensified myelotoxic chemotherapy regimens that would otherwise not be tolerated. Indeed, ASCO guidelines state that primary prophylaxis with G-CSF is essential in patients receiving certain dose-dense chemotherapy regimens and EORTC guidelines state that prophylactic G-CSF should be administered as supportive treatment in settings where chemotherapy dose intensification has been shown to be associated with survival benefits.

In conclusion, prophylaxis with G-CSF can shorten the duration of chemotherapy-induced neutropenia in patients with leukaemia, as well as shortening the duration of hospitalization for infection and the duration of intravenous antibacterial therapy. In addition, G-CSF prophylaxis can facilitate the administration of dose intensity.

Infections

Infection in leukaemic patients is often life-threatening, if not lethal, and significantly contributes to morbidity in both community and hospital settings during chemotherapy. Infection risk in leukaemia relates not only to the destruction of formed elements of the blood responsible for immunity, but also to its treatment, which causes additional and profound destruction of innate immunity and damage to the body's physical barriers.

Patients with leukaemia are at risk of invasive infection with many organisms, some typically considered benign – many of which are part of endogenous flora – from the gastrointestinal tract, oropharynx or skin, but also myriad opportunistic pathogens, many of which exist in the environment. This is in addition to infection with the 'professional pathogens' that afflict the non-immunosuppressed, infection with hospital strains of bacteria and simultaneous infection with multiple pathogens. Table 23.1 lists important pathogens in this patient cohort. Table 23.2 lists the locations of access for endogenous bacteria frequently causing invasive infection.

There have been marked changes in the aetiology of infections in leukaemia over recent decades; notably, an increase in Gram-positive bacteraemias and a rise in Gram-negative organism antibiotic resistance. This may relate in part to the generally unwise use of quinolone antibiotic prophylaxis.

Patients with leukaemia are repeatedly exposed to both antibiotics and the hospital environment, increasing the risk of both colonization and infection with antibiotic-resistant organisms.

Table 23.3 lists those host factors responsible for increasing the risk and severity of infection. Most patients will have more than one such risk factor.

These factors – especially chemotherapy – confound 'predictable' patterns of infection, flowing from the primary 'phagocytic failure' of myeloid leukaemia, cell-mediated, lymphocyte-generated immunity, or both. These are inextricably linked, making predictive rules of risk and organism type far from absolute.

Risk of infection always increases with the depth and duration of neutropenia. Bacterial infections can occur at any stage during neutropenia, whilst fungal infections typically occur later – invasive *Candida* infections after 7 days of neutropenia and invasive moulds such as *Aspergillus* spp. after around 14 days, with peak incidence paradoxically around the time of neutrophil recovery.

Invasive fungal disease (IFD) carries considerable morbidity and mortality. High IFD risk is associated with induction chemotherapy for AML and MDS, purine analogue therapy (such as fludarabine and high-dose cytarabine) and Allogeneic stem cell transplant recipients (Allo SCT) with GVHD necessitating corticosteroids. Patients classified as being 'high risk' should be considered for antifungal prophylaxis with agents active against both yeasts and moulds.

Febrile neutropenia

There are many possible causes of fever in neutropenic patients, including infection and infusion of blood products, drug reactions, mucositis, tumour lysis syndrome, total body irradiation and GVHD. Fever, however, remains the main and often sole manifestation of infection; febrile episodes mandate thorough clinical assessment and instant administration of empiric antibiotic therapy. Individual patients' degree and duration of immunosuppression and treatment must be considered when evaluating the most likely aetiology of their symptoms.

A full clinical examination must be performed, including of the skin, oropharynx, vascular access devices and perianal region (Figure 23.1).

Blood cultures must be taken at the onset of fever and prior to antibiotic administration. At least two sets are recommended, taken peripherally and from any vascular access device – ideally from each lumen; 10–15 mL of blood should be inoculated into each bottle (aerobic and anaerobic) – volumes lesser or greater than this will reduce diagnostic yield. Blood cultures are typically incubated for 5 days; most cultures from bacteremic patients are positive within 24 hours. Microscopy of a positive blood culture will guide antimicrobial therapy by Gram reaction and organism morphology. A full identity is normally obtained within 24–48 hours. Recent advances in molecular diagnostics suggest that such delays will soon fall away.

Empirical antibiotic therapy

Febrile neutropenia is only confirmed microbiologically in <30% of cases and the majority of infections in febrile neutropenic patients must therefore be treated empirically. Febrile

Table 23.1 Important pathogens in leukaemia patients.

	Bacteria	Endogenous/exogenous source of infection, notable characteristics
Gram-positive bacteria	Coagulase-negative staphylococci	Endogenous
	Staphylococcus aureus	Endogenous
	Enterococcus spp.	Endogenous
	Viridans group streptococci	Endogenous
	Streptococcus pneumoniae	Endogenous
	Streptococcus pyogenes	Endogenous
Gram-negative bacteria	*Escherichia coli*	Endogenous
	Klebsiella spp.	Endogenous
	Enterobacter cloacae	Endogenous
	Pseudomonas aeruginosa	Endogenous or exogenous; associated with water; colonizes medical devices
	Acinetobacter spp.	Exogenous Typically acquired in the intensive care unit environment. Forms resistant biofilms on devices, often multi-resistant, and survives in the environment
	Stenotrophomonas maltophilia	Exogenous Low virulence organism typically unable to overcome host barriers. Infection associated with carbapenem use, to which it is constitutively resistant
Yeasts and moulds	*Candida* spp.	Responsible for both superficial skin and mucosal infections and fungemia
	Aspergillus spp.	Causes rapidly progressive pneumonia Responsible for ca. 90% of invasive fungal infection in this cohort, of which >90% are pulmonary
	Pneumocystis jirovecii	Rare in this group; but not negligible
	Cryptococcus spp.	Rare cause of pneumonia and meningitis
	C. immitis, H. capsulatum, B. dermatitidis, P. brasiliensis, sporotrichosis, *P. marneffei*	Endemic, geographically limited highly virulent pathogens
	Zygomycetes, *Fusarium, Scedosporium* spp.	Low pathogenicity; tendency to produce skin lesions, disseminated disease and high mortality; difficult to treat

neutropenia with bacteraemia is exceedingly dangerous and is a medical emergency. Mortality varies considerably and can be as high as 40%. Unrecognized or incorrectly treated Gram-negative bacteraemia carries a mortality of over 95%. Accordingly, febrile neutropenic patients must be immediately treated with broad-spectrum antibiotics with antipseudomonal activity, typically a β-lactam/β-lactamase inhibitor combination (such as piperacillin/tazobactam) or carbapenem (meropenem). The latter can also be used in most cases of penicillin allergy, although some cross reaction does exist. Rising global levels of antibiotic resistance prompt the use of more than one empiric antibiotic of a different class to minimize chances of treatment failure. Individual units must appraise themselves of their and their country's resistance patterns and tailor empirical therapy accordingly. Subsequent isolation of a pathogen and its sensitivities will then drive evidence-based, focused de-escalation in most cases. Current patterns of resistance in most countries suggest addition of an aminoglycoside (gentamicin or amikacin) to the initial β-lactam/β-lactamase inhibitor combination or carbapenem (meropenem). Such initial dual therapy should be considered mandatory in severe cases of sepsis, especially in the presence of shock.

MRSA colonization or suspected line infection prompts use of additional therapy with vancomycin or teicoplanin. Previous blood and urine culture/infection results should be sought, to include antibiotic susceptibility data. Previous infection with resistant organisms mandates the choice of empiric therapy remaining effective. This rule is becoming ever more important in view of the inexorable global rise in Gram-negative resistance, including carbapenems. Some countries (such as the UK)

Table 23.2 Locations of access for endogenous bacteria frequently causing invasive infection.

		Typical context of pathogenicity	Important considerations
Skin	Coagulase-negative staphylococci	Line-associated infections	Biofilm formers Often highly antibiotic resistant
	Corynebacterium spp.	Bacteraemia associated with indwelling devices	*C. jeikeium* has a predilection for bacteremia *via* iv devices in this group
	Propionibacterium acnes	Rare cause of endocarditis and septic arthritis	Invasive infection often associated with indwelling prostheses
Oropharynx	Anaerobes	Dental abscesses	Often associated with polymicrobial infections
	Viridans streptococci	Dental plaque and caries Opportunistic cause of infective endocarditis	Pre-existing gum and tooth disease may relapse during chemotherapy
Nasopharynx and upper respiratory tract	*Streptococcus pneumoniae*	Pneumonia, otitis media, bacteraemia, meningitis	Risk markedly increased in asplenia or decreased serum [IgG]. Polyvalent vaccine available but not necessarily effective in leukaemia
GI tract	*Enterobacter cloacae*	Urinary tract infections, bacteraemia	Resistant to beta-lactam antibiotics except carbapenems in some countries
	Enterococcus faecalis	Bacteraemia, endocarditis, urinary tract infections	Typically replaced by *Enterococcus faecium* (which is more resistant) following antibiotic treatment
	Escherichia coli	Urinary tract infections, bacteraemia	Often spontaneously translocates from the gut in neutropenia
	Klebsiella pneumoniae	Urinary tract infections, bacteraemia	Often highly antibiotic resistant. Forms resilient biofilms on plastic.

suggest testing for multidrug resistant (MDR) organism carriage, usually by means of a rectal swab. The prevalence of such MDR organisms in some countries has now reached epidemic proportions; in some cases threatening elective chemotherapeutic treatment, as no antibiotics can be relied on to be effective in the case of bacteraemia – which if untreatable carries a mortality in excess of 95%. Antimicrobial choice should mirror local guidelines.

Immediate use of empirical antifungal agents is discouraged. Antifungal treatment should be considered in neutropenic patients with fever non-responsive to broad-spectrum antibiotic therapy and clinical or radiological signs supporting the diagnosis of invasive fungal disease. There are several important considerations to the choice of antifungal agent, including evidence of pre-existing fungal colonization (such as *Candida* colonization of skin/oropharynx) and local epidemiological surveillance of prevalent fungal pathogens; this is discussed in greater detail below (Figure 23.2).

Investigating febrile neutropenic episodes

Whilst most febrile neutropenic episodes are not associated with localizing symptoms or signs, any focal sign or symptom must be mandatorily investigated, as source control is paramount.

Respiratory symptoms

A chest radiograph must be performed on all patients with respiratory symptoms.

The differential diagnosis involves bacterial, viral and fungal causes; the diagnostic possibilities of the last are very wide. Serum antibody tests are essentially useless. Correctly taken viral

Table 23.3 Host factors responsible for increasing the risk and severity of infection.

- Age
- Metabolic disturbance:
  - Acidosis
  - Diabetes
  - Iron overload
- Neutropenia
- Dys/hypogammagobulinaemia
- Failure of T-cell-mediated acquired immune pathways
- Treatment:
  - Chemotherapeutic agent, particularly:
    - Anthracyclines
    - Cytarabine
    - Purine analogues, e.g. fludarabine
  - Corticosteroids
  - Irradiation
  - Antibiotic prophylaxis – this selects for organisms with innate or acquired resistance to that agent, as well as increasing the risk of candidaemia
  - Gastric acid suppression – associated with an increased predisposition to invasive infection with streptococci as well as *Clostridium difficile*
- Breaches of physical barriers to infection:
  - Skin breaks
  - Mucositis
  - Chemotherapy-induced colitis
  - Perianal disease
- Vascular access devices

swabs of the retropharynx are essential and should be submitted at least for PCR-determined diagnosis of influenza, parainfluenza subtypes, adenovirus, metapneumovirus and respiratory syncytial virus.

Knowledge of these agents' behaviour and epidemiology is key: influenza is inverse in its seasonality with respect to parainfluenza and is related to hemisphere; both are equally dangerous for Allo SCT patients. Respiratory syncytial virus and the coronaviruses show similar seasonal patterns to influenza; adenovirus and metapneumovirus have no seasonality. *Mycoplasma pneumoniae* and *Chlamydia pneumoniae* usually occur in the context of community outbreaks and are therefore unlikely in patients with prolonged hospital admissions.

Prolonged neutropenia and persistent fever in spite of antibiotic therapy mandates high-resolution CT imaging to examine evidence of pulmonary fungal disease. Broncheoalveolar lavage (BAL) and preferably tissue samples are infinitely superior to radiological abnormality *per se*. BAL is useful even though in the absence of biopsy, especially as the patient has usually already been treated with multiple anti-infective agents, it reduces the yield of useful information to approximately 20%.

Aspergillus spp. are responsible for over 90% of respiratory fungal infections in neutropenic patients and may be associated with pleuritic pain and haemoptysis. Detection of galactomannan (a component of the *Aspergillus* cell wall) in serum is considered useful by some; issues of sensitivity and false positivity secondary to some antibiotics and diet, together with the need for repeated serial testing limits performance, which nevertheless improves for bronchial lavage or other body fluids. β-glucan estimation is not mould-specific, but might emerge as an alternative serological diagnostic test. It is most likely useful for exclusion. Most instead rely on strong clinical, microbiological or radiological evidence for initiation of antifungal treatment. Recovery from neutropenia may drive radiographic worsening of fungal lesions; not be misinterpreted as failure, but rather as radiologically more impressive 'hardening' of lesions as these become necrotic and haemorrhagic. There are currently no standardized and validated PCR-based serum diagnostics;

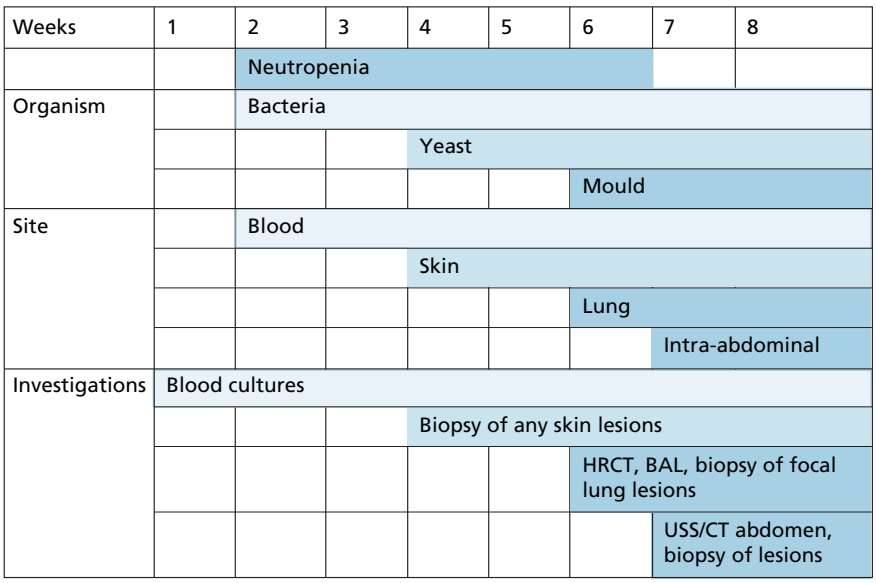


Figure 23.1 Infective causes of fever in chemotherapy-induced neutropenia as these relate to time.
N.B. The above general trends are not absolute as each patient journey will be different for reasons of underlying host constitution, pre-existing comorbidities, chemotherapeutic regimens and interventions.

Anaerobes	Gram positive					Gram negative			
	MRSA	*Staphylococcus aureus*	*Enterococcus*	β-haemolytic streptococci: *S. pyogenes* Group B strep	α-haemolytic streptococci: Viridans group strep *S. pneumoniae*	*E. coli*	Klebsiella	Enterobacter Serratia Citrobacter	Pseudomonas
	Teicoplanin/Vancomycin								
						Gentamicin/amikacin			
				Penicillin					
			Amoxicillin/ampicillin						
		Co-Amoxiclav							
		Piperacillin tazobactam							
		Flucloxacillin							
				Erythromycin/clindamycin					
				Cefuroxime/ceftriaxone					
				Ceftazidime					
						Ciprofloxacin/ofloxacin			
				Levofloxacin					
		Meropenem							
		Ertapenem							
	Linezolid								
Metronidazole									

Figure 23.2 The spectrum of antimicrobial activity of commonly used antibiotics.
N.B. This table refers to *predicted* spectrum of action. In reality, acquired resistance has been described for almost all the bacterial groups listed. The degree to which this has developed varies by country and location, and is in some cases endemic to the point where some antibiotics are no longer predicted to be effective at all.

this may reflect *Aspergillus* spp. tendency to remain fixed in tissues.

Patients with leukaemia are at a greater risk of *Mycobacterium tuberculosis* infection. Neither the tuberculin skin test nor interferon-γ release assays are to be relied on in such patients as the tests require intact cellular immunity. Microscopy showing acid fast bacilli and/or culture positivity for *M. tuberculosis* should prompt both expert advice and an infection control 'look back' for possibly exposed contacts.

Other opportunistic respiratory pathogens, including *Toxoplasma gondii*, atypical mycobacteria and Nocardia must always be considered. In patients with dyspnoea and hypoxia, particularly on exertion, *Pneumocystis jiroveci* pneumonia (previously *Pneumocystis carinii* pneumonia) must be excluded.

The breadth of possible pathogens highlights the importance of obtaining diagnostic sputum, BAL and/or tissue samples and PCR-based diagnostics whenever possible.

Abdominal symptoms

Patients with acute leukaemia frequently suffer from gastrointestinal symptoms, particularly diarrhoea – usually chemotherapy-induced – but sometimes infective in origin. Diarrhoeal stool should be sent for culture and if persistent or accompanied by abdominal pain, imaging evidence of colitis or associated with fever, should also be sent for analysis for *Clostridium difficile* infection.

Neutropenic, acutely unwell patients with abdominal pain, may have neutropenic enterocolitis. This is a life-threatening condition with a high mortality rate and its management remains contentious. Early recognition is essential and most patients respond to prompt treatment with broad-spectrum parenteral antibiotics, typically a carbapenem alongside supportive management. Antidiarrhoeal agents must not be used.

Neutropenic patients with persistent fever despite antibiotic treatment and with no evidence of a focus of infection should have abdominal imaging – with ultrasound in the first instance – even in the absence of abdominal symptoms, to investigate for evidence of hepatosplenic abscesses. Any identified lesions in the liver and/or spleen should be biopsied if at all possible. This is because whilst *Candida* spp. remains the most likely causative pathogen of multiple hepatosplenic lesions, a very broad range of alternative causative pathogens exists, including many species of bacteria, mycobacteria and fungi.

Urine should be sent for culture in patients with urinary symptoms, and in Allo SCT recipients with haematuria for adenovirus and BK virus PCR.

Skin lesions

Skin lesions often represent the metastatic spread of infection and punch biopsies should always be taken (even in thrombocytopenic patients) for histopathology, culture and ideally for PCR analysis. Of particular concern is ecthyma gangrenosum: ulcerative lesions with necrotic centres typically seen in the perineum, axillae and submammary regions, caused by systemic *Pseudomonas aeruginosa* infection and associated with a high mortality rate. Patients with this condition are invariably exceedingly toxic with concurrent *Pseudomonas* bacteraemia. Fungi, particularly *Aspergillus*, *Candida* and *Fusarium* spp., can also disseminate to the skin. The latter organisms also often involve the retina; fundoscopy is mandatory in such cases to interdict permanent blindness if possible. Atypical mycobacteria also present with skin lesions, typically nodules. Some are relatively benign and will extinguish with relatively little therapy in cases where the source of the atypical organisms is the intravenous access device.

Infection of, and via, vascular access devices

If there is clinical evidence of infection at the site of a vascular access device, or if the patient's febrile episodes correlate with line use, the device should be removed if possible. The degree to which line removal becomes mandatory depends on the organism associated with it, as some are far more capable of driving septic shock and/or serious morbidities than others. Accordingly, we consider line-associated *Pseudomonas aeruginosa*, *Staphylococcus aureus*, *Klebsiella* spp. and *Candida* spp. mandatory reasons for immediate removal. This relates to their inherently pathogenic nature, capacity for metastatic spread to heart valves and joints, as well as adherence to plastic and biofilm formation, making medical therapy bound to fail.

It is also very important to differentiate between *intraluminal* colonization (usually with no local evidence of infection) and local infection *around the catheter site*. The former *may* be amenable to antibiotic 'treatment' of the line – so-called 'line locks'; the latter will never be 'cured' by anything but line removal, as the infection is perpetuated at the avascular tissue/plastic interface and therefore not amenable to antibiotic-mediated killing. A protocol for intravascular line management is shown in Figure 23.3.

Suspect intravascular lines should always be removed. *E. coli*, however, is not a biofilm former. Accordingly, *E. coli* bacteraemia should prompt the search for an alternative source (translocation, urinary tract infection, etc.) rather than automatic line removal.

In the context of a febrile neutropenic patient with a suspected line infection, vancomycin or teicoplanin should be added to the empirical antibiotic regimen. Lines infected with coagulase-negative Staphylococci *may* be salvageable with combined systemic antibiotics and antibiotic line locks.

If the patient remains febrile following the removal of the line, perform an ultrasound scan to examine for infected thrombus at the line site, however profoundly thrombocytopenic the patient is.

Invasive fungal infection

This remains one of the most significant challenges to effective treatment and outcomes for this group of patients. Much evidence suggests that length of neutropenia, initial chemotherapy cycles for acute myeloid leukaemia, high-dose steroids and graft-versus-host disease drives susceptibility. Risk of IFD is highest for patients undergoing remission induction chemotherapy for AML and MDS, those treated with purine analogues (fludarabine, high-dose cytarabine) and allogeneic stem cell Allo SCT recipients with severe chronic GVHD needing steroids. There is a hierarchy of risk in such conditions, making clear, calculable definitions of risk difficult.

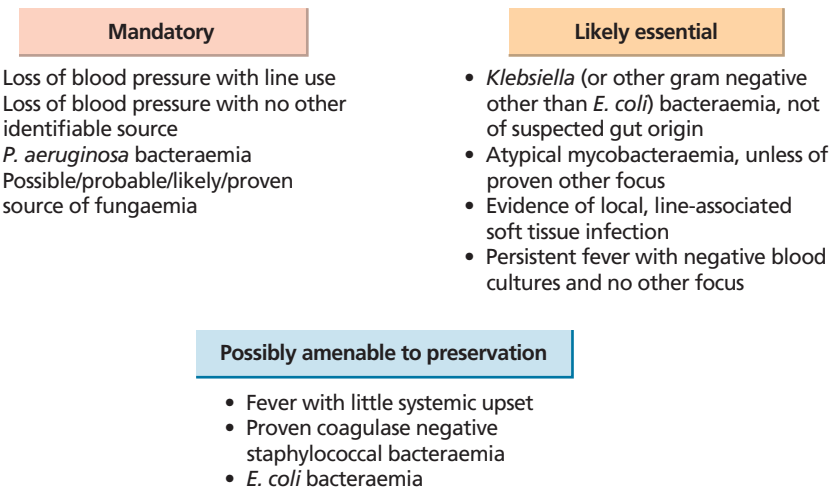


Figure 23.3 Line management options as these relate to clinical and bacteriological factors.

Liposomal amphotericin, echinocandins or voriconazole form the mainstay of antifungal treatment; voriconazole possesses many side-effects, is unpredictably metabolized and interacts with many drugs, making its use often difficult. Liposomal amphotericin is active against almost all fungal pathogens, whilst voriconazole has a similarly broad spectrum of action except for lacking activity against *Mucor* spp. Intrinsic and/or inducible resistance to both agents is rare, but nevertheless exists, such as *Candida lusitaniae* for the former, and azole resistance in *Aspergillus* spp. is increasingly recognized in Europe, probably driven by widespread agricultural use.

In some countries *Candida* spp. are the commonest cause of IFD in neutropenic patients; patients with leukaemia often have additional risk factors for candidaemia including mucosal disruption, vascular access devices and treatment with broad-spectrum antibiotics – in some cases unnecessary and excessive. Fluconazole has good activity against *C. albicans,* most non-albicans *Candida* and *Cryptococcus* spp., but is not active against moulds: *Aspergillus* spp., *Mucor* spp. and *Fusarium* spp. Flucytosine, the echinocandins and itraconazole have a similar spectrum of activity to fluconazole but itraconazole is active against *Aspergillus* spp., and the echinocandins lack activity against *Cryptococcus* spp.

Patients classified as being at 'high risk' of invasive fungal disease should be considered for antifungal prophylaxis with agents active against both yeasts and moulds. Such classifications drive the use of prophylactic antifungal agents, administered during critical periods of chemotherapy-induced neutropenia. Azoles are agents of choice here; data for posaconazole prophylaxis are particularly strong.

Interestingly, many units continue to believe that it is acceptable for such prophylaxis to only be for yeasts, such as *Candida albicans*, using the widely used agent fluconazole. Unfortunately, the really dangerous pathogen in this setting is *Aspergillus* spp. – either *fumigatus* or *flavus* in most cases. These moulds are not susceptible to fluconazole. For this reason the use of mould-active agents in situations of 'high risk' for invasive fungal infection is important if not mandatory. There are country-specific differences in choice of prophylactic agent here; drivers for choice include local preference, marketing pressures and, of course, cost. Mould-active agents include the subsequent generations of triazole antifungals, such as itraconazole, voriconazole (almost exclusively reserved for treatment rather than prophylaxis) and posaconazole. All have some advantage over the others; and all have their own disadvantages. More recent strong evidence indicates that posaconazole is particularly effective in reducing the incidence of invasive mould infection in such patients; the drug, however, is expensive and (like other azoles except fluconazole) has variable and unpredictable oral absorption profile in its syrup form. This relates to its lipophilic properties, shared with the other azoles apart from fluconazole. All rely on an acid pH for ionization and absorption; this is often impeded by the concurrent administration of powerful antacid medication, including H2 blockers and proton pump inhibitors. A different oral tablet formulation of posaconazole with better absorption is now available.

Treatment of proven focal infection

Antimicrobial therapy should be tailored according to any positive microbiological results. What is found has a very close relation to where it is found. For example, rapidly evolving, painful skin lesions in the presence of neutropenia, which progress to necrosis in a matter of hours, represent a clinical emergency. These lesions are almost invariably driven by *Pseudomonas aeruginosa*, and are termed lesions of ecthyma gangrenosum. This condition may spontaneously appear either on the limbs, the face or the genitalia and can be lethal within 12 hours. Some strains of *P. aeruginosa* are known to contain virulence genes, which code for exceedingly cytolytic toxins. These kill adjacent tissue, which then, devoid of blood supply and therefore antibiotics, acts as yet more material for further unhindered growth and penetration into more tissue. This condition can be very rapidly lethal and the smallest necrotic lesion, when associated with neutropenia, must be immediately excised and sent for culture and histology. This is particularly important for genitalia, where large areas of skin and subcutaneous tissue can be lost forever in a matter of hours, leading to the need for subsequent reconstructive plastic surgery; this is in addition to the commonly associated and potentially lethal dissemination of *P. aeruginosa* into the bloodstream.

Respiratory system

Fever and pulmonary symptoms in this patient group are particularly common. Other than that provided by PCR for viruses from a pharyngeal swab, diagnosis is rare: clinical necessity drives prescription of very broad-spectrum antibacterial agents, which result in resolution in most cases. Potentially diagnostic bronchoscopy and bronchoalveolar lavage is often not performed, or performed too late; this relates to its relatively poor diagnostic yield, especially in view of the reluctance to perform transbronchial biopsies in the presence of thrombocytopenia.

High-resolution CT scan appearances, especially if interpreted by those with experience, can be very useful. They can (for example) document diffuse, usually bilateral parenchymal abnormalities, suggesting a diagnosis of either virus or *Pneumocystis*, or 'classical' fungal lesions – peripherally distributed relatively soft nodules of varying size, ignoring anatomical margins and often accompanied by a 'halo', or alternatively larger, cavitating lesions, which might be amenable to CT-guided biopsy. Finally, bacterial infection due to *Pneumococcus* or *Haemophilus* spp. usually appears as more confluent areas of anatomically well-defined consolidation.

In reality, more aggressive efforts to make a diagnosis are often triggered by failure of initial antibacterial therapy – often

followed by a trial of broad-spectrum antifungal therapy, which also produces no clinical benefit, with persisting fever and progression of CT scan appearances. It is at this point that the rarer causes of pulmonary disease in such patients needs to be considered – and the differential diagnosis remains very wide. Of note here is the possibility that *Aspergillus* spp. may still be responsible for antifungal-unresponsive disease, especially in the presence of uncontrolled malignancy. The reasons why an identical fungus, with (presumably) identical tissue concentrations of antifungal agents of proven efficacy should fail to be adequately treated if there is persisting and active malignancy are far from clear. Finally, other agents capable of causing pulmonary disease, such as *Nocardia* spp., *Mycobacterium tuberculosis, Toxoplasma gondii*, and endemic moulds such as *Histoplasma* and *Cryptococcus* spp., as well as rarer moulds should also be borne in mind in this exquisitely susceptible patient group.

Gastrointestinal system

The gastrointestinal tract is not spared concerning infection. The more important issue relates to the simultaneous and predictable serious effects of chemotherapy on the gastrointestinal tract, which leads to ulceration and mucosal defects from mouth to anus. Accordingly, diarrhoea, abdominal pain and fever can be caused by either chemotherapy-induced destruction of mucosae (with consequences for bacterial penetration to deeper tissues and translocation to the blood), or by a primary, local process of invasion. Over and above this predictable process are specific conditions, to be considered in the face of the 'ordinary' therapy-induced problems of mucosal destruction. Accordingly, upper gastrointestinal symptoms may be due to local infection with invasion by *Candida* spp., *Cytomegalovirus* or *Herpes* spp; small intestinal symptoms (usually accompanied by CT-scan evidence of wall thickening) may be due to invasive fungal enteritis, and diarrhoea, fever, toxicity and abdominal pain may be due to colonic inflammation secondary to destruction of local mucosal barriers and local penetration by normal colonic flora (so-called 'typhlitis'). *Clostridium difficile* infection, with sometimes lethal pseudomembranous colitis, should always be excluded with relevant stool antigen and toxin tests; this disease is associated with both neutropenia and disorders of immunoglobulin synthesis, due to disease or treatment. The situation is made far more difficult because many, if not all, of these symptoms are predictably caused by chemotherapy; clinical vigilance and acumen are needed to identify those patients whose presentation seems to fall out of the expected pattern. Finally, the rectum and anal canal are highly susceptible to local invasion by lower gastrointestinal tract bacteria. The development of rectal or anal pain, whether or not associated with local signs, should prompt rapid investigation – with MRI if possible. Pre-existing defects, such as fistulas and haemorrhoids, will predictably lead to local problems of bacterial infection in neutropenia – and mandate immediate institution of rapid, broad-spectrum antibiotic therapy in order to limit the spread of bacteria to deeper pelvic tissues.

Skin

The skin can act as a window to diagnosis in patients with leukaemia. Quite apart from relatively slowly developing lesions of *Leukaemia cutis,* the appearance of lesions (especially in the context of fever and toxicity) should always be considered as not only ominous, but also an opportunity to make a rapid diagnosis of myriad infective conditions, many of which have the skin as part of far more disseminated disease. Immediate biopsy for histology, viral PCR and culture for both bacteria, fungi and mycobacteria should be the rule. Papular lesions are most often associated with disseminated fungal infection with organisms such as *Fusarium* or *Scedosporium* spp., more planar, initially erythematous lesions may represent disseminated *Varicella zoster* and rapidly developing, necrotic lesions may represent the rapidly lethal ecthyma gangrenosum lesions of *P. aeruginosa* and neutropenia. The skin is also the target for other non-infective conditions such as drug allergy (most often antibiotics) and little understood conditions, whose histology reveals inflammatory processes of obscure aetiology.

Vascular access devices

Indwelling vascular access devices, such as Hickman lines and peripherally inserted central catheters (PICC lines) are all susceptible to colonization with bacteria, fungi and mycobacteria. Development of fever in the absence of another clear focus should always prompt the possibility that the device is responsible for the fever, and blood cultures should be taken via the line lumens (each being cultured separately) as well as peripherally. Retention of the line in the presence of proven bloodstream infection with the line as source of infection is contraindicated with the exception of coagulase-negative staphylococcal infection, when it is not unreasonable to attempt to 'cure' the line of infection with antibiotic line locks. Glycopeptide antibiotics infused into the line lumen and retained for several hours are in massively higher concentration than that which might be achieved in tissue and such concentrations overcome the decreased antibiotic susceptibility of organisms which are trapped within their own biofilm. Such line locking is not invariably successful, and persisting fever, bacteraemia, or both would be an indication for removal.

Antibiotic resistance

The rise of antibiotic resistance in Gram-negative bacteria is an increasing, global issue; there are few if any 'new' antibiotics in the Pharma pipeline, and better antibiotic stewardship currently remains the only mitigating strategy. Identification of colonization or infection of a patient with an antibiotic-resistant

Table 23.4 Mechanisms of resistance and important considerations flowing therefrom.

Antibiotic	Bacteria	Comments
Chromosomal *AMP C* β-lactamase	*Citrobacter freundii* *Enterobacter* spp. *Morganella morganii* *Proteus* spp. *Serratia* spp.	Able to break down penicillins and cephalosporins Treatment with third-generation cephalosporins (includes ceftazidime and ceftriaxone) induces hyperproduction of resistance enzymes. Not inhibited by clavulanic acid.Usually require treatment with a carbapenem
Extended spectrum β-lactamases	Most commonly found in: *E. coli* *Klebsiella pneumoniae* Described in other Gram-negative bacilli	Confer resistance to penicillins and cephalosporins Usually inhibited by a β-lactamase inhibitor, such as clavulanic acid
Carbapenem	Enterobacteriaceae, particularly: *E. coli* *Klebsiella pneumoniae* *Pseudomonas aeruginosa*	Multiple known mechanisms of acquired carbapenem resistance. Very challenging management. May remain sensitive to tigecycline, colistin, fosfomycin or aminoglycosides. These infections often require dual or triple therapy combinations
Methicillin	*Staphylococcus aureus*	Resistant to all β-lactams due to altered penicillin binding site. All UK isolates so far sensitive to glycopeptides (vancomycin and teicoplanin). Variable sensitivity to doxycycline, aminoglycosides, trimethoprim and rifampicin. Alternatives: daptomycin, linezolid.
Vancomycin	*Enterococcus* spp.	Vancomycin-resistant enterococci are typically acquired through exposure to glycopeptide treatment and can lead to persistent carriage in the GI tract

organism should immediately prompt review of not only current treatment, but also the empirical antibiotic regimen to be used should they become febrile at a later date. Table 23.4 shows major antibiotic resistance mechanisms; of note is that several may coexist, rendering treatment choices exceedingly limited, if not entirely impossible.

Investigation of persistent fever

Persistent fever with neither positive microbiology, nor an identified focus of infection, remains an extremely common scenario for patients with leukaemia. Many causes for fever other than infection exist in these patients, due to the disease itself, drug-related effects (such as that seen with cytosine arabinoside, amongst others) or allergy. Most, if not all, patients will be given empiric antibiotic therapy, as fevers in the presence of neutropenia must be treated. The antibiotic regimen may not need to be changed unless the patient is unstable, clinically deteriorating or there is strong clinical evidence of a progressive and local infective process, such as spreading cellulitis. However, there should be daily clinical review and consideration as to possible unidentified foci of infection.

It is essential to explore infection risk factors, including the patient's country of origin, travel history and pre-existing serology if available. Diagnoses to be considered include infections due to opportunistic organisms allowed to thrive subsequent to the immunosuppressive effects of their disease, treatment or both. Examples here include extrapulmonary tuberculosis, disseminated atypical mycobacterial infection, toxoplasmosis and, in patients with geographic risk factors, malaria, brucellosis and *Strongyloides*.

A progressive infective process and/or patient deterioration requires consideration of antimicrobial change. If the patient is clinically septic an aminoglycoside (gentamicin or amikacin) should be added. Worsening skin or soft tissue infection may be due to many organisms other than the traditional *S. aureus* and *S. pyogenes*. Gram-negative cellulitis may develop in the face of deep immunosuppression and neutropenia and rapidly spread to become lethal. The choice of antibiotic in these situations remains very location specific, as the likely responsible organisms will carry geographically well-defined sets of resistance mechanisms. Local knowledge is essential for adequate, timely and effective treatment – and should always be carried out in partnership with microbiologists. Table 23.5 provides an initial suggested framework of possibilities and investigation for such situations, and should ideally always be in partnership with microbiologists, infectious diseases physicians or both.

When to stop anti-infective agents

Nowhere else in the field of infectious disease and microbiology does the term 'course of anti-infective agent' apply less.

Table 23.5 Suggested initial framework for investigation of persisting fever in leukaemic patients.

	Review	Of note	Action
History	Any focal symptoms?		ID/micro review or advice on need for further investigations, to include optimal imaging modalities
	Social history:		
	Country of birth?	Born in the tropics	Check if eosinophilia
	Occupation?	Contact with fresh water or animals	Consider:
	Pets?		–Malaria film
	Travel history:	Travel to malarious area	–Serial stool samples for ova/cysts/parasites/ PCR
	Where?	Tick/mosquito bites	
	Insect bites?	Contact with rodents or animal bites	–Serology testing for Brucella, filarial, Lyme disease, strongyloides, schistosomiasis, toxoplasmosis as appropriate
	Animal contacts?		
	Fresh water contact?	Schistosomal areas	–Direct PCR for specific pathogens
Lab results	FBC	Eosinophilia	Review social, travel and drug history
	LFTs	Transaminitis	Abdominal USS
		Raised bilirubin	Serology: hepatitis viruses, EBV and CMV if not already done
			Veno-occlusive disease (transplants)
	Blood cultures	Any previous positive cultures?	Repeat peripheral and line cultures
			Consider mycobacterial cultures
			Review empirical antimicrobial in light of any previous positive results
	Urine culture	Any previous positive cultures?	Send MSU for culture
			If catheterized and specimen of urine reveals organisms in pure growth this *may* be significant and needs discussion with microbiology
	Bone marrow culture	Was bone marrow sent for culture?	Consider sending for prolonged/mycobacterial culture
Lines	Suppuration	Blocked/unable to aspirate?	If clinical evidence of infection, remove line immediately
	Erythema		
	Swollen limb		Blood cultures from line
			Send tip if removed
			Consider USS site for evidence of thrombus
Skin	Focal signs?	Focal lesion(s)	Biopsy for culture and histology if not pathognomonic
		Rash	
			Viral swab
			Fungal culture of scrape
Pulmonary	Focal signs?	Cough or shortness of breath	A chest X-ray is rarely informative.
		Purulent sputum or haemoptysis?	High-resolution CT scanning imperative
			Haemoptysis may reflect thrombocytopenia, *Aspergillus* infection, or both
		Pleuritic pain	
		Hypoxia	Broncheoalveolar lavage may be useful
			Biopsy of focal lesions carries a high rate of diagnosis; at the expense of risk (bleeding, pneumothorax)
Abdomen	Focal signs?	Pain/diarrhoea	Consider perforation; neutropenic colitis; *C. difficile*
			CT scanning useful to examine for bowel thickening

(*Continued*)

Table 23.5 (*Continued*)

	Review	Of note	Action
Oropharynx	Focal signs?	Pain Swelling Odynophagia	Dental X-rays Maxillofacial or dental review. Orthopantogram or skull CT scanning Ensure adequate anaerobic antibiotic cover if fever considered of oropharyngeal origin Consider endoscopy if not neutropenic to exclude HSV ulceration and oesophageal *Candida*
Perianal region	Focal signs?	Pain, erythema Discharge Bleeding	Perianal imaging (typically MRI) Surgical review to exclude/include fistula

There is little enough evidence base for such a concept in other than a handful of very well-defined and organism-specific situations in immunocompetent patients. Leukaemic patients vary in their susceptibility to infection and their ability to combat it, either unaided or with the help of modern anti-infectives. Persisting neutropenia (functional or real) requires a less cavalier approach to antibiotic withdrawal; this, however, does not imply that patients with proven neutropenic sepsis and bacteraemia, whose symptoms and signs rapidly disappear on appropriate antibiotic therapy, should not have their antibiotics withdrawn as soon as is practicable. The concept of 'step-down' therapy here also seems flawed. When such antibiotics are stopped in the safe environment of the hospital, the patient will either quickly relapse and need antibiotic support for longer, or more likely remain well and free of fever. The concept of 'oral step-down' is strange. Antibiotics are either needed to prevent sepsis, or they are not. Once the sepsis has resolved, 'oral step-down' is illogical. It does, however, set the scene for generation of antibiotic resistance and *C. difficile*.

The question of when to stop antifungal agents (either intra-chemotherapy cycle prophylaxis, or treatment for proven or highly likely disease) is equally fraught with issues that preclude 'one size fits all'. General rules apply, however: permanent recovery of neutrophil count, complete remission from leukaemia, evidence of haemorrhagic necrosis of lung lesions on CT all prompt the possibility of stopping these drugs with no relapse. Continuing fever, uncontrolled tumour and/or neutrophil counts below 0.5×10^9/L all make the decision to stop more difficult.

Special considerations in bone marrow Allo SCT recipients

Bone marrow Allo SCT recipients are vulnerable to opportunistic infection for months if not years following transplantation, the nature and depth of which relates to transplant type, conditioning regimen, closeness of HLA match, presence of GVHD and chimaerism. Prior to engraftment during conditioning these patients are at risk of bacteraemia with gut commensals due to mucositis, candidaemia almost invariably selected for by antibiotics, and *Herpes simplex* re-activation if not adequately prophylaxed. Cord transplants have a particularly long period of susceptibility to these agents; extending to perhaps years because these patients are immunologically naïve and will continue to suffer defects in both humoral and cell-mediated immunity with increased susceptibility to both bacteria, viruses and fungi. Until well-developed antibody responses are made, they will also be prone to infection with encapsulated bacteria and it is essential that they are vaccinated against pneumococcus, meningococcus and *Haemophilus influenzae* type B, although efficacy remains far from optimal and absolute. Haemopoietic stem cell Allo SCT recipients should also routinely be re-vaccinated, starting 6 months post vaccination (Table 23.6).

Table 23.6 A typical post-transplant vaccination schedule.

Initiate vaccination course		No. of doses
6 months	Pneumococcal conjugate vaccination	3
	Haemophilus influenzae B	3
	Influenza vaccine	1–2, repeat annually
8–9 months	Meningococcal vaccine	1
12 months	Diphtheria/tetanus/pertussis	3
	Polio	3
24 months	Measles/mumps/rubella (if immunocompetent)	1–2

Conclusions

Infections in patients with leukaemia develop both as a consequence of the disease itself, as well as the chemotherapy necessary for remission and/or cure. Febrile neutropenia and sepsis of bacterial origin represents a clinical emergency and empiric antibiotic choice must be appropriate and sufficient. Recent global changes in Gram-negative resistance patterns make these choices increasingly difficult, and rectal screening for carriage of multidrug-resistant Gram-negative organisms ahead of the period of neutropenia is becoming ever more widespread and essential. Equally, correct management of fungal infection relies on risk stratification and in appropriate cases, prophylaxis. The performance of fungal diagnostics are considered suboptimal by many, making empiric antifungal treatment, usually based on CT scanning and exclusion of other causes of fever, the norm. Response rates and 'cure' in such patients often relates to the prognosis of the underlying leukaemia; persistent malignant disease very significantly impacts on the ability of these drugs to eradicate the fungal infection. Finally, viral infections remain a significant threat for those whose cell-mediated immunity is poor, either subsequent to disease or to chemotherapy. In addition, many can rapidly spread through an inpatient haematology unit should infection control precautions be less than excellent.

Chemotherapy-induced nausea and vomiting (CINV)

Nausea and vomiting is common in patients undergoing treatment for leukaemia.

Both symptoms are distressing, cause significant morbidity and reduce quality of life. Nausea and vomiting contribute significantly to nutritional deterioration and increased pain intensity. The emetogenicity of the chemotherapeutic agents, repeated chemotherapy cycles and patient risk factors significantly influence CINV. The use of a combination of a serotonin 5-HT3 receptor antagonist, dexamethasone and a neurokinin 1 (NK1) receptor antagonist has significantly improved the control of acute and delayed emesis in single-day chemotherapy. Palonosetron, a second-generation 5-HT3 receptor antagonist with a different half-life, a different binding capacity and a different mechanism of action from the first-generation 5-HT3 receptor antagonists appears to be the most effective agent in its class. Aprepitant, the first and only agent clinically available in the NK1 receptor antagonist drug class has been used effectively as an additive agent to the 5-HT3 receptor antagonists and dexamethasone to control CINV. Rolapitant and netupitant are other NK1 receptor antagonists that are currently in Phase III clinical trials. Olanzapine, a US-FDA approved antipsychotic, has emerged in recent trials as an effective preventative agent for CINV, as well as a very effective agent for the treatment of breakthrough emesis and nausea.

Clinical trials using gabapentin, cannabinoids and ginger have not been definitive regarding their efficacy in the prevention of CINV. Additional studies are necessary for the control of nausea and for the control of CINV in the clinical settings of multiple-day chemotherapy and bone marrow transplantation.

Non-pharmacological strategies

Non-pharmacological strategies are very important in the overall management of the patient. Examples include avoiding food with strong tastes and smells; small but frequent meals, behavioural approaches (e.g. relaxation) and acupuncture/reflexology.

Nutritional

Chemotherapy side-effects resulting in the deterioration of dietary intake have been shown to diminish nutritional status and quality of life. Malnutrition has also been found to affect treatment effectiveness. Furthermore, it makes patients more vulnerable to the side-effects of chemotherapy. Past clinical reports have revealed that malnutrition adversely affects the possibility of remission in children undergoing induction chemotherapy for leukaemia. Malnutrition is multifactorial and can be affected by factors other than treatment. Sociodemographic-economic status, inadequate dietary intake, psychological issues prior to and during treatment, and the malignancy itself, play a role in the nutritional status and quality of life of patients.Tube feeding by nasogastric tube can prevent severe weight loss and hypoalbuminaemia in most patients. This is more relevant in patients undergoing a haemopoietic stem cell transplant, as underweight or impaired nutritional status was found to increase the risk of early death in both paediatric and adult recipients. In a more recent study in transplant patients, enteral nutrition was well tolerated and significantly improved patient outcome by reducing early post-transplant mortality and GVHD incidence and by inducing more rapid neutrophil engraftment.

Metabolic complications

Fluid balance

Fluid balance should be monitored carefully in patients treated with intensive chemotherapy or bone marrow transplantation. Daily weight monitoring is also equally important. Fluid input/output monitoring should take place several times within each 24-hour period in order to be able to observe unwanted changes in fluid balance and to be able to intervene, if necessary. Elderly patients are more susceptible to fluid imbalance due to underlying comorbidities, including congestive heart failure and reduced GFR, and therefore should be monitored more closely.

Vomiting is pretty common in patients on chemotherapy and in combination with diarrhoea and sweating due to pyrexia might rapidly lead to dehydration. Potential vasodilatation due to sepsis might exacerbate hypovolaemia and therefore strict monitoring and fluid replacement should be undertaken.

Hydration prior to cytotoxic agents

Hydration, urine alkalinization (pH 7 to 8), and electrolyte monitoring can prevent the hyperuricaemia, hyperphosphataemia, hypocalcaemia and hyperkalaemia. This is more relevant at diagnosis as the disease bulk is increased and the risk appears to be higher. The goal of IV hydration is to improve renal perfusion and glomerular filtration, and induce a high urine output to minimize the likelihood of uric acid or calcium phosphate precipitation in the tubules. However, IV hydration can lead to potentially dangerous fluid overload in patients with underlying acute kidney injury or cardiac dysfunction. Close monitoring is very important to prevent complications associated with tumour lysis syndrome or fluid overload.

Tumour lysis syndrome

Tumour lysis syndrome (TLS) is an oncologic emergency that is caused by massive tumour cell lysis with the release of large amounts of potassium, phosphate and nucleic acids into the systemic circulation. Catabolism of the nucleic acids to uric acid leads to hyperuricaemia, and the marked increase in uric acid excretion can result in the precipitation of uric acid in the renal tubules and can also induce renal vasoconstriction, decreased renal blood flow and inflammation, resulting in acute kidney injury. Hyperphosphataemia with calcium phosphate deposition in the renal tubules can also cause acute kidney injury. TLS is a frequent complication of advanced stage Burkitt lymphoma and acute leukaemia with profound leucocytosis. The symptoms associated with TLS largely reflect the associated metabolic abnormalities (hyperkalaemia, hyperphosphataemia, and hypocalcaemia). They include nausea, vomiting, diarrhoea, anorexia, lethargy, haematuria, heart failure, cardiac dysrhythmias, seizures, muscle cramps, tetany, syncope and possible sudden death.

Aggressive intravenous hydration is the cornerstone of preventing TLS and is recommended prior to therapy in all patients at intermediate or high risk for TLS. Patients at risk for TLS should initially receive 2–3 L/m^2 per day of IV fluid. Urine output should be monitored closely and maintained within a range of 80–100 mL/m^2 per hour. Diuretics can be used to maintain the urine output, if necessary, but should not be required in patients with relatively normal renal and cardiac function. The best diuretic for patients with TLS is unknown; loop diuretics such as furosemide appear preferable because they not only induce diuresis, but may also increase potassium secretion.

Allopurinol is a hypoxanthine analogue that competitively inhibits xanthine oxidase, blocking the metabolism of hypoxanthine and xanthine to uric acid. Allopurinol effectively decreases the formation of new uric acid and reduces the incidence of obstructive uropathy in patients with malignant disease at risk for TLS. However, there are several limitations to its use. Because it acts by decreasing uric acid formation, allopurinol does not reduce the serum uric acid concentration before treatment is initiated. Thus, for patients with pre-existing hyperuricaemia (serum uric acid ≥7.5 mg/dL [446 μmol/L]), rasburicase is the preferred hypouricaemic agent. Allopurinol increases serum levels of the purine precursors hypoxanthine and xanthine, which may lead to xanthinuria, deposition of xanthine crystals in the renal tubules, and acute kidney injury. Since allopurinol also reduces the degradation of other purines, dose reductions of 65–75% are needed in patients being treated with mercaptopurine or azathioprine. Allopurinol has the potential to interact with a number of other drugs, including cyclophosphamide, high-dose methotrexate, ampicillin and thiazide diuretics. The usual allopurinol dose in adults is 100 mg/m^2 every 8 hours (maximum 800 mg per day).

Among patients who are unable to take oral medications, IV allopurinol can be administered at a dose of 200–400 mg/m^2 per day, in one to three divided doses (maximum dose 600 mg per day). Treatment is generally initiated 24 to 48 hours before the start of induction chemotherapy and is continued for up to 3–7 days after until there is normalization of serum uric acid and no other laboratory evidence of tumour lysis.

An alternative approach to allopurinol for lowering serum uric acid levels is to promote the degradation of uric acid by the administration of urate oxidase (uricase), which catalyses oxidation of uric acid to the much more water-soluble compound allantoin. *Rasburicase* is well tolerated, rapidly breaks down serum uric acid, and is effective in preventing and treating hyperuricaemia and tumour lysis syndrome (TLS). This rapid reduction in serum uric acid is in contrast to the effect of allopurinol, which decreases uric acid formation and therefore does not acutely reduce the serum uric acid concentration. The EMA and FDA dosing guidelines both recommend a rasburicase dose of 0.2 mg/kg once daily for up to 7 days. Rasburicase is contraindicated in patients with glucose-6-phosphate dehydrogenase (G6PD) deficiency. Patients being considered for rasburicase (especially males) who have the potential for G6PD deficiency should undergo screening prior to initializing treatment.

Hyperleucocytosis

Hyperleucocytosis is generally defined as an initial white cell count/blast count of more than 100×10^9/L. There are no trials that prove leucapheresis is of benefit in these patients, but this procedure is generally safe and should be considered in

patients with AML, CML and lymphoproliferative disorders with large lymphocytes (i.e. prolymphocytic leukaemia), presenting with symptomatic hyperleucocytosis, including blurred vision, headaches and chest symptoms. Since leucapheresis only temporarily decreases the WBC, early initiation of more definitive therapy should be considered. Anaemic patients with leucocytosis should be transfused with caution as there is a risk of increased plasma viscocity.

Differentiation syndrome (DS)

Differentiation syndrome (DS), also known as retinoic acid syndrome, is the main life-threatening complication of therapy with differentiating agents (all-*trans* retinoic acid or arsenic trioxide) in patients with acute promyelocytic leukaemia (APL). It typically occurs during induction therapy with differentiating agents, especially in patients presenting with high white cell count at diagnosis or relapsed disease. The median day of onset of clinical manifestations of the DS ranges between 7 and 12 days after starting therapy with differentiating agents.

This syndrome is characterized by unexplained fever, weight gain, hypotension, acute respiratory distress with interstitial pulmonary infiltrates and/or a vascular capillary leak syndrome leading to acute renal failure. The diagnosis of DS is mainly based on clinical and radiological features; however, it usually requires the exclusion of alternative causes that could explain the signs and symptoms of the syndrome. The early addition of chemotherapy to ATRA and the administration of high doses of steroids in the form of dexamethasone at the onset of the first sign or symptom has reduced the mortality associated with this syndrome from 30% down to 5% or less. It is recommended to start treatment with intravenous dexamethasone at a dose of 10 mg twice daily promptly at the very earliest suspicion of DS. Most patients will show rapid improvement and complete resolution of signs and symptoms after starting intravenous dexamethasone. Corticosteroid therapy should be continued until the complete disappearance of symptoms, and then tapered. Other supportive measures are also crucial in the correct management of DS. In the PETHEMA protocols, furosemide is usually administered to treat signs or symptoms of fluid overload. Some cases with refractory acute renal failure and/or fluid overload may need renal replacement therapy. Invasive or non-invasive mechanical ventilation is indicated in some patients with severe acute respiratory failure who do not respond to high-flow oxygen therapy. Patients with prerenal failure and hypotension require careful fluid monitoring in conjunction with inotropic support.

Whether the differentiation agents should be discontinued when DS takes place remains controversial. It seems likely that both ATRA and ATO can be safely continued in the majority of patients with DS as long as intravenous dexamethasone has been started. It is reasonable to temporarily discontinue ATRA or ATO when DS progresses rapidly, despite starting dexamethasone, or when DS is presented as a full-blown syndrome (i.e. severe DS) or with severe end-organ dysfunction (e.g. mechanical ventilation or dialysis required). Once the signs and symptoms of severe DS are completely resolved, the differentiating agent should be re-started and continued until CR and/or the achievement of minimum length duration of differentiating therapy.

Fortunately, the vast majority of patients presenting with DS have a dramatic and rapid resolution of the clinical picture when treated early with dexamethasone and the appropriate supportive measures.

Skin, nail and dental problems

Skin reactions to drug therapy are extremely common. Chemotherapy drugs can cause dry skin, hyperpigmentation, discoloration, serpentine hyperpigmentation, photoallergic reactions and acne (mainly treatment with steroids). Radiation treatment can cause dry or red skin in the area being treated. Certain areas are more sensitive than others. Patients should be encouraged to use moisturizing creams on a regular basis. They should also protect skin from the sun by wearing a hat and clothing that covers the arms and legs, and apply sunscreen with a sun protection factor (SPF) of 30 or higher before going outside.

Nail changes include:

- Beau's lines (horizontal depressions of the nail plate). These skin reactions can occur a few weeks after a course of chemotherapy.
- Mee's lines (white horizontal discoloration of the nail plate involving the entire nail width).
- Leuconychia (white horizontal discoloration involving partial nail width).
- Onycholysis (separation or loosening of a fingernail or toenail from its nail bed).
- Onychodystrophy (malformation of the nail).

Nails should be trimmed and kept clean and gloves should be worn for house-cleaning at discharge. Occasionally antibiotics and antifungal agents might be necessary to treat underlying infections.

Radiation recall

Radiation recall is a skin reaction that can occur when certain chemotherapy drugs are given after radiation therapy. It usually appears in the area of skin where the radiation was given. The skin becomes red and tender. It may also peel or blister, resembling sunburn. Radiation recall may be treated with

corticosteroids to lessen inflammation or chemotherapy may be delayed until the skin heals.

Dental problems

Dental care also is very important. Leukaemia and chemotherapy can make the mouth sensitive, easily infected and likely to bleed. Patients should receive a complete dental examination and, if possible, undergo necessary dental care before chemotherapy begins, since any existing lesions that might normally lie dormant can flare up and become life-threatening once the patient is immunosuppressed. Dental extractions should preferably be performed at least 10 to 14 days before the commencement of chemotherapy, so that epithelization of the extraction site has been completed prior to the initiation of treatment. Dental procedures that could trigger a bleeding episode or/and bacteraemia during the aplastic phase should be avoided. The use of chlorhexidine mouthwash is very important, primarily during the phase of aplasia. Aesthetic restorations of the enamel defects are also often necessary, especially after completion of treatment.

Auditory toxicity

The auditory system is vulnerable to many chemical exposures. Antibiotics, such as the aminoglycosides, and chemotherapeutic agents, such as cisplatin, can cause considerable ototoxicity in as many as 40% of patients. Dosage reduction and regular assessment with auditory studies during the course of therapy might reduce toxicity and prevent further damage.

Visual toxicity

Ophthalmic complications induced by cytotoxic chemotherapy are often under-estimated and under-reported due to priority given to other life-threatening conditions; however, most of the ophthalmic complications are reversible if detected early enough. Among the functional complaints are blurred vision, decreased colour vision, diminished visual acuity, diplopia, night blindness and photophobia.

Among the drugs used on a regular basis to treat leukaemia, cytarabine can cause conjunctivitis and methotrexate can cause oedema, ocular pain, blurred vision, photophobia, conjunctivitis and blepharitis. Chlorambucil has been reported to cause keratitis and daunorubicine excessive lacrimation and conjunctivitis. Alemtuzumab can cause optic neuropathy and endophthalmitis. Finally imatinib can cause periorbital oedema in up to 47% of patients.

The possible reversal of some of these side-effects, if discovered in time, emphasizes the need for clinicians to be aware of these ocular reactions and suggest immediate consultation by an ophthalmologist.

Pain

Many leukaemia patients experience bone or joint pain that results from the bone marrow being heavily infiltrated by leukaemia cells. Oral and gut mucocytis are probably the most common site of pain, especially in recipients of high-dose chemotherapy or total body irradiation. Infections can also cause pain, especially in areas such as the skin, mouth, eyes and perianally. Avascular necrosis of bone is another cause of pain in patients with acute lymphoblastic leukaemia and is often severe enough to need relief with opiates. Peripheral neuropathy and associated neuropathic pain is a widespread side-effect of treatment with certain chemotherapeutic agents. The pain associated with these neurotoxic effects can be prolonged, severe and relatively resistant to intervention. The overall incidence is not clearly delineated, although it is documented frequently with vincristine and platinum-based agents.

A variety of modalities to treat and control pain include:

- Pain control drugs with non-opioid and opioid analgesics
- Neuropathic pain control drugs such as carbemazepine, gabapentin, and antidepressants (amitriptyline, nortriptyline and imipramine)
- Continuous pump infusions
- Nerve block therapies
- Psychosocial interventions – stress management, counselling, coping mechanisms
- Acupuncture
- Massage therapy
- Relaxation techniques.

Palliation

Relapse post chemotherapy or following a bone marrow transplant is a common theme in leukaemia. Intensive salvage chemotherapy or other modalities can often achieve prolonged remission; however, refractory unresponsive-to-chemotherapy disease is not uncommon in this setting. Several experimental treatments can be offered; however, it is important to weigh the possible limited benefits of a new treatment against the possible downsides as quite often further treatment might not improve the health or change the outcome or survival. Under these circumstances palliation can be offered to the patient. Palliative care is a proactive approach involving a multiprofessional team. As well as controlling pain and other distressing symptoms, it applies a holistic approach to meeting the physical, practical, functional, social, emotional and spiritual needs of patients and carers facing progressive illness and bereavement. It is not uncommon for patients to ask to remain under the care of the same haematology team; however, the palliative care team should also be involved as the choice of dying at home versus

dying in a hospice requires some expertise and external support better managed by the palliative care services.

Selected bibliography

Karaiskos I, Giamarellou H (2014) Multidrug-resistant and extensively drug-resistant Gram-negative pathogens: current and emerging therapeutic approaches. *Expert Opinion in Pharmacotherapy* **10**: 1351–70.

Montesinos P, Bergua JM, Vellenga E *et al.* (2009) Differentiation syndrome in patients with acute promyelocytic leukemia treated with all-*trans* retinoic acid and anthracycline chemotherapy: characteristics, outcome, and prognostic factors. *Blood* **22**: 775–83.

Pathak S, Roth M, Verma A *et al*. (2013) Eltrombopag for the treatment of thrombocytopenia in patients with malignant and non-malignant hematologic disorders. *Expert Opinion in Drug Metabolism and Toxicology* **9**: 1667–75.

Seguy D, Duhamel A, Rejeb MB *et al*. (2012) Better outcome of patients undergoing enteral tube feeding after myeloablative conditioning for allogeneic stem cell transplantation. *Transplantation* **15**: 287–94.

Tacke D, Buchheidt D, Karthaus M,*et al*. (2014) Primary prophylaxis of invasive fungal infections in patients with haematologic malignancies. 2014 update of the recommendations of the Infectious Diseases Working Party of the German Society for Haematology and Oncology. *Annals of Hematology* **93**(9): 1449–56.

Townsley DM, Desmond R, Dunbar CE *et al*. (2013) Pathophysiology and management of thrombocytopenia in bone marrow failure: possible clinical applications of TPO receptor agonists in aplastic anemia and myelodysplastic syndromes *International Journal of Hematology* **98**: 48–55.

Vehreschild JJ, Böhme A, Cornely OA *et al*. (2014) Prophylaxis of infectious complications with colony-stimulating factors in adult cancer patients undergoing chemotherapy–evidence-based guidelines from the Infectious Diseases Working Party AGIHO of the German Society for Haematology and Medical Oncology (DGHO). *Annals of Oncology* **25**(9): 1709–18.

Villafuerte-Gutierrez P, Villalon L, Losa JE, *et al*. (2014) Treatment of febrile neutropenia and prophylaxis in hematologic malignancies: a critical review and update *Advances in Hematology* **2014**: 986938.

CHAPTER 24

Chronic myeloid leukaemia

24

David TO Yeung and Timothy P Hughes

Department of Haematology, SA Pathology, Adelaide, Australia

Introduction

Chronic myeloid leukaemia (chronic granulocytic leukaemia) is a myeloproliferative neoplasm characterized by rearrangement of the long arms of chromosome 9 and 22, resulting in the Philadelphia (Ph) chromosome, creating the fusion oncogene *BCR–ABL1*. This genetic event that encodes for a constitutively active tyrosine kinase occurs in a haemopoietic progenitor and confers proliferative and antiapoptotic effects. The natural history of untreated CML is triphasic. It most commonly presents in chronic phase (CP), characterized by markedly increased myeloid activity, resulting in increased numbers of morphologically and functionally normal granulocytes, predominantly comprised of neutrophils and myelocytes. Basophilia is a distinct feature of CML, although eosinophilia is also usually present. Thrombocytosis is commonly seen. Without effective treatment, patients may progress from chronic phase into accelerated phase (AP), where myelopoiesis is increasingly ineffective with a rising blast count, culminating in blastic phase (BP) when the blast percentage rises further in the context of marrow failure. In CML-BP (blast crisis, blastic transformation), the disease clinically resembles acute leukaemia and the outcome is dismal. Disease progression used to be almost universal over a period of 5–10 years, but has been dramatically reduced in the era of tyrosine kinase inhibitor (TKI) therapy. For most patients, therapy is expected to be lifelong, though recent trials suggest that a proportion of patients may be able to cease their TKI therapy entirely without losing their remission status.

Epidemiology and aetiology

The incidence of CML in developed nations had been relatively steady over the last decade and remains at 1–2:100,000, with a slight male predominance. The disease is most commonly diagnosed in the fifth and sixth decades of life in developed countries. In the developing world, presentations in the third and fourth decades are more common, possibly reflecting younger population demographics. CML is rare in the paediatric population.

Apart from exposure to ionizing radiation (evidence derived from epidemiological studies of Hiroshima and Nagasaki atomic bomb survivors), there is no other known risk factor leading to the development of CML, whether inherited or acquired.

Clinical features, natural history, laboratory haematology and cytogenetics

The laboratory features of CML in its various phases, as defined in the two most commonly used classifications, are summarized in Table 24.1. The WHO classification is increasingly gaining acceptance, though the European Leukemia Net (ELN) / MD Anderson Cancer Center criteria were used to define CP in many pivotal TKI studies and still have relevance for interpretation of clinical data. An important distinction between the two is the higher myeloblast count for CML-BP, set by the ELN (30% rather than 20% used by WHO), and a role for excess

Postgraduate Haematology, Seventh Edition. Edited by A Victor Hoffbrand, Douglas R Higgs, David M Keeling and Atul B Mehta.

Table 24.1 Definition of advanced phase CML disease.

	ELN	WHO
Accelerated phase	• PB or BM blast 15–29% • PB or BM blast + promyelocytes ≥30% • PB basophils ≥20% • Platelets ≤100 × 10^9/L unrelated to therapy • Clonal evolution	• PB or BM blast 10–19% • PB basophils ≥20% • Platelets ≤100× 10^9/L unrelated to therapy • Platelets >1000 × 10^9/L unresponsive to therapy • Clonal evolution on treatment • Increasing spleen size and increasing WBC count unresponsive to therapy
Blast crisis	• PB or BM blast ≥30% • Extramedullary blast proliferation, apart from spleen	• PB or BM blast 20% • Extramedullary blast proliferation, apart from spleen • Large foci or clusters of blasts in the bone marrow biopsy

BM, bone marrow; PB, peripheral blood

promyelocytes in the determination of advanced phases, in contrast to the WHO criteria (Table 24.1).

Chronic phase (CP)

The majority of patients (around 95%) are diagnosed in CML-CP. Many are asymptomatic and the diagnosis is often incidental to investigation for an unrelated medical problem. The most common presentations include unintentional weight loss, night sweats, fatigue and splenomegaly (which may cause early satiety and anorexia). The most striking haematological abnormality is usually a leucocytosis, with the differential heavily favouring mature neutrophils, band forms and myelocytes. Eosinophilia is common, and absolute basophilia is usually present (Figure 24.1a). Absolute monocytosis is uncommon and, if present, usually associates with the e1a2 breakpoint (see below). Rarely, patients may present with a marked thrombocytosis with only mild leucocytosis. Unlike other myeloproliferative disorders, CML is very rarely associated with arterial or venous thromboses. Dysplasia is uncommon in CML, and if present, should prompt consideration of atypical CML as a differential diagnosis.

The CML-CP marrow is usually grossly hypercellular as a result of markedly increased myeloid activity (Figure 24.1b). Myeloid maturation is normal or slightly left shifted, and blasts are usually less than 5%. The myeloid to erythroid ratio is markedly increased and megakaryocyte numbers may be either decreased or increased compared to the normal marrow; small, hypolobated forms are common. Increased reticulin fibres are seen in 30% of cases and correlate with larger spleen size. Pseudo-Gaucher cells and sea blue histiocytes may be seen secondary to increased numbers of phagocytic cells amidst increased cell turnover. Although now rarely used, the leucocyte alkaline phosphatase reaction using naphthol-ASD-chloroacetate esterase usually results in decreased staining in CML neutrophils, but increased staining with reactive neutrophilia.

Accelerated phase (AP)

Whilst CML-CP is characterized by over-production of myeloid cells, essentially with normal morphology and function, AP disease is characterized by either impaired differentiation or drug resistance. There may be an increased myeloblast percentage in peripheral blood or marrow, as well as increasing leucocytosis, splenomegaly, basophilia or thrombocytosis, despite ongoing therapy. Thrombocytopenia, unrelated to therapy, may herald progressive marrow failure. Cytogenetic abnormalities additional to those found in the baseline Ph^+ clone are often present at progression to AP, a phenomenon termed clonal evolution. These abnormalities frequently include duplication of the Ph chromosome, trisomy 8 or 19 and isochromosome 17q. CML-AP usually progresses to BML-BP within 1–2 years without appropriate treatment.

Blastic phase (BP)

CML-BP is characterized by rapidly rising blast counts and marrow failure as a result of uninhibited growth of myeloid precursors with absence of maturation or differentiation. Morphologically, CML-BP is indistinguishable from other cases of acute leukaemia. In ~70% of cases the blasts are of myeloid lineage by morphology and immunophenotype, and may express antigens associated with granulocytic, monocytic, megakaryocytic

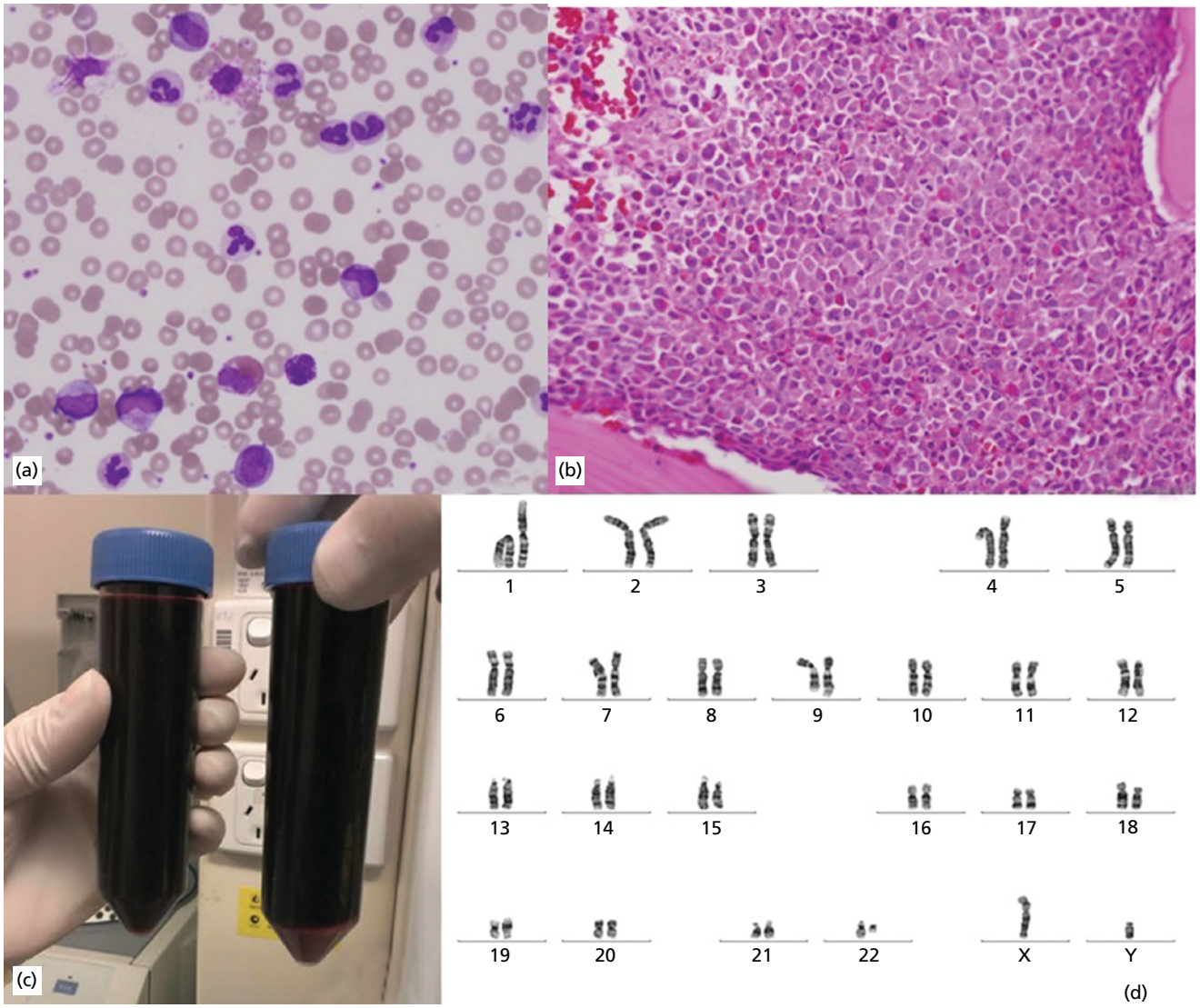


Figure 24.1 (a) Blood film from a newly diagnosed CML-CP patient aged 47 years. Leucocytosis is prominent, with increased neutrophils and myelocytes, but no obvious dysplastic features. Several myelocytes, a basophil and an eosinophil can be seen towards the bottom of the photomicrograph. (b) A grossly hyper-cellular marrow trephine from the same patient. (c) Whole blood is processed prior to RQ-PCR. Red cell lysis agent is added to whole blood and a white cell pellet is obtained after centrifugation. On the right, a newly diagnosed CML patient yielded a sizable white cell pellet at the bottom of the tube, compared to that of a normal control on the left. (d) Karyogram showing the translocation t(9;22). (Source: (d) Ms Sarah Moore, SA Pathology, Adelaide, Australia. Reproduced with permission.)

or erythroid differentiation. Aberrant expression of lymphoid markers (e.g. CD7) is common. Blasts in the remaining 25–30% of cases usually express lymphoid antigens of B lineage. Aberrant expression of myeloid antigens in CML lymphoid BP is also quite common.

Although mixed lineage leukaemia is rare, most cases bear the t(9;22) lesion. These cases are recognized as a separate entity under the 2008 WHO classification of haemopoietic neoplasms and make up <1% of acute leukaemias (see also Chapter 19).

Cytogenetics and molecular biology

The diagnosis of CML rests on demonstration of the Ph chromosome with cytogenetic analysis, or the *BCR–ABL1* fusion gene with RT-QPCR (reverse transcriptase quantitative polymerase chain reaction, also known as RQ-PCR, qRT-PCR or QPCR), in the context of compatible blood and marrow morphology. Cytogenetics and RQ-PCR (see below) offer complementary information and both tests should be performed routinely at diagnosis.

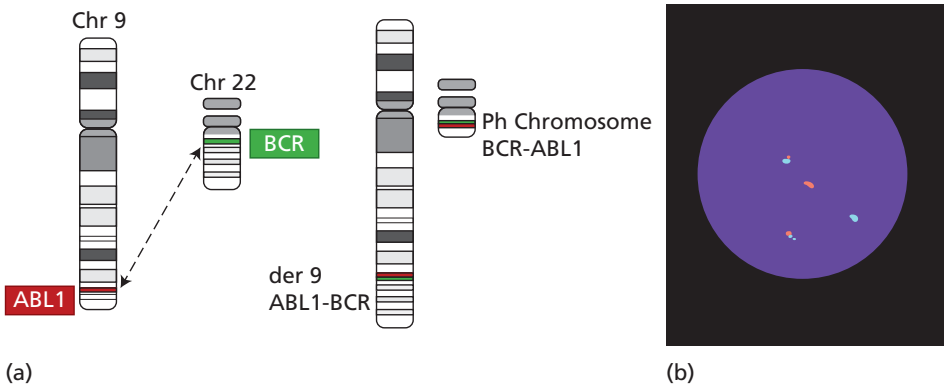


Figure 24.2 (a) The t(9;22) translocation leads to juxtaposition of genetic material from chromosomes 9 and 22, with formation of two fusion genes *BCR–ABL1* and *ABL1–BCR*, the former being crucial to the pathogenesis of CML and Ph+ ALL. (b) Fluorescence *in-situ* hybridization. Hybridization of an oligonucleotide sequence complementary to the *BCR* gene attached to a flurophore (green in this case) leads to its localization in a metaphase spread. The same can be done for *ABL1* (red in this case). A normal cell has two red and two green signals. A t(9;22) translocation event will juxtapose red and green flurophores (with some probe sets seen as two fused yellow signals), one from *BCR–ABL1* and one from *ABL1–BCR* (the two signals to the left of the picture). The remaining red and green signals originate from the normal chromosome, which is not rearranged. (Source: (b) Ms Sarah Moore, SA Pathology, Adelaide, Australia. Reproduced with permission. FISH probe set from Abbott Diagnostics, Vysis LSI BCR/ABL Dual Color, Dual Fusion Translocation Probe.)

The Ph chromosome was first linked to CML in 1960 by Nowell and Hungerford in Philadelphia. Rowley in 1973 showed this to be a result of a reciprocal translocation between the long arms of chromosome 9 and 22 (Figure 24.1d, see also Chapter 19). This event, now denoted as t(9;22)(q34;q11), is known to fuse the *BCR* (breakpoint cluster region) and *ABL1* genes together. The resulting *BCR–ABL1* fusion gene encodes a constitutively active tyrosine kinase that is central to the pathogenesis of Ph^+ leukaemias (Figure 24.2a). Variant translocations, where *BCR–ABL1* resides on a third chromosome, is seen in ~5–10% of patients and may be cryptic, discernible only with fluorescent *in situ* hybridization (FISH) analysis (Figure 24.2b). These variants have little or no bearing on clinical outcome. Amplification of *BCR–ABL1* may result from duplication of the Ph chromosome, which, along with trisomy 8, i(17q) and trisomy 19, leads to increased risk of blastic transformation and together these are referred to as major route cytogenetic lesions. Cytogenetic analysis at diagnosis remains important for prognostication as major route lesions are not discernible on FISH or RQ-PCR. Apart from CML, t(9;22) is also seen in a number of *de novo* high-risk acute lymphoblastic leukaemia cases. The finding of t(9;22) in *de novo* AML is rare and usually denotes CML-BP from an undiagnosed antecedent CML-CP.

The breakpoints of the *BCR* and *ABL1* genes occur within the introns of each gene and are unique to each individual (Figure 24.3). The *ABL1* breakpoint is almost invariably located in the intron between exons 1 and 2, whereas the *BCR* breakpoint usually occurs in an intron between exons 13–15. The resulting fusion transcripts are named e13a2 and e14a2 transcripts (previously known as b2a2 and b3a2); both may coexist in one individual due to alternate splicing. These transcripts encode for a 210 kDa protein product (p210). A breakpoint immediately downstream of the first *BCR* exon produces the e1a2 transcript, which encodes the p190 BCR–ABL1 protein. This is more commonly seen in Ph^+-ALL and is infrequent in CML. Rarely, the *BCR* breakpoint occurs between exons 17–20, leading to the larger p230 protein from the e19a2 transcript, associated with a more prominent neutrophilia and thrombocytosis. Other transcripts such as e6a2, e8a2 and e18a2 are also occasionally reported.

The translocation is carried by the leukaemic stem cell population, such that both myeloid and lymphoid progeny may carry t(9;22) in an affected individual. ABL1 is a tyrosine kinase with multiple substrates. Its activity is tightly regulated by the phosphorylation status of critical tyrosine residues on ABL1 itself. Juxtaposition of BCR and ABL1 promotes dimerization of BCR–ABL1 proteins through a coiled-coil domain in BCR and formation of tetramers through subsequent dimerization of dimers (Figure 24.4). Critical tyrosine residues are brought into close proximity with adjacent ABL1 kinase domains, leading to autophosphorylation and activation. In addition, native ABL1 is usually regulated by autoinhibition mediated by intramolecular binding of an N-terminus myristate moiety into the kinase domain. This myristate moiety is lost through the translocation event. These modifications lead to constitutive activation of ABL1, which through downstream messengers, results in cellular effects such as loss of growth inhibition, decreased

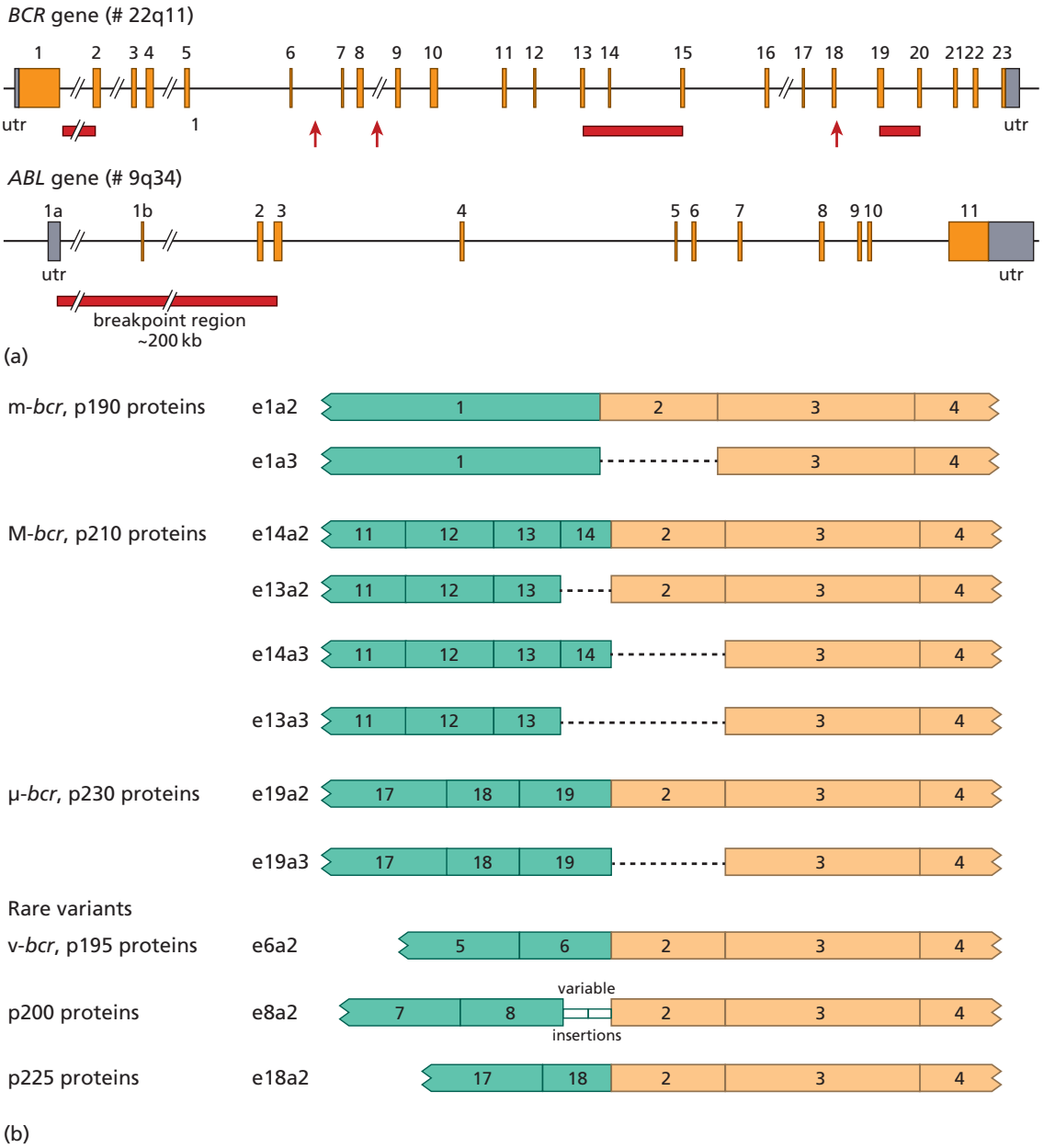


Figure 24.3 (a and b) Formation of the *BCR–ABL1* fusion gene. Breakpoint in the *ABL1* gene almost invariably occurs in the intron between exons 1–2. In contrast, breakpoints in the *BCR* gene may occur in introns between exons 1–2, 13–15 and 19–20. These regions are also known as the m-*BCR*, M-*BCR* and μ-*BCR* regions, respectively, and in turn yield the p190, p210 and p230 BCR–ABL1 proteins. The e13a2 and e14a2 transcripts are most commonly encountered in CML, whereas Ph+ ALL cases are usually associated with the e1a2 transcript. Other breakpoints (v-*BCR*) are also occasionally encountered. (Source: Weerkamp *et al.*, 2009 [*Leukemia*; **23**: 1106–1117]. Reproduced with permission of Nature Publishing.)

apoptosis and decreased adherence to bone marrow stromal cells. Many of the downstream pathways are well studied, and involve proteins such as RAS/RAF/MEK/ERK, PI3K/AKT, STAT and MYC (Figure 24.5). BCR–ABL1 also leads to genomic instability and facilitates disease transformation through accumulation of subsequent genetic lesions. DNA damage, such as double-stranded breaks, are mediated through increased ROS (reactive oxygen species) generated by BCR–ABL1. The repair mechanisms secondary to ROS-associated DNA damage in *BCR–ABL1* positive cells are more error-prone compared to normal mammalian cells, which in turn contributes to genomic instability.

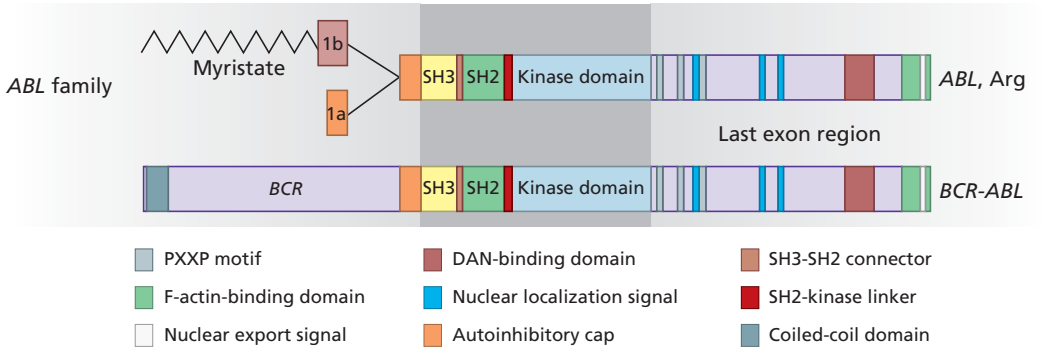


Figure 24.4 The native ABL1 protein possesses either an auto-inhibitory cap region encoded by exon 1a, or a myristate moiety encoded by exon 1b. The myristate moiety can fold over and bind to a pocket on the same ABL1 molecule to exert an auto-inhibitory effect. Replacement of exon 1a/b with BCR not only leads to loss of auto-inhibition, but also leads to tetramerisation through introduction of the coiled-coil domain in BCR, resulting in cross phosphorylation of critical tyrosine residues on neighbouring ABL1 and constitutive activation. (Source: Hantschel & Superti-Furga, 2004 [*Nat Rev Mol Cell Biol* 2004; 5: 33-44]. Reproduced with permission of Nature.)

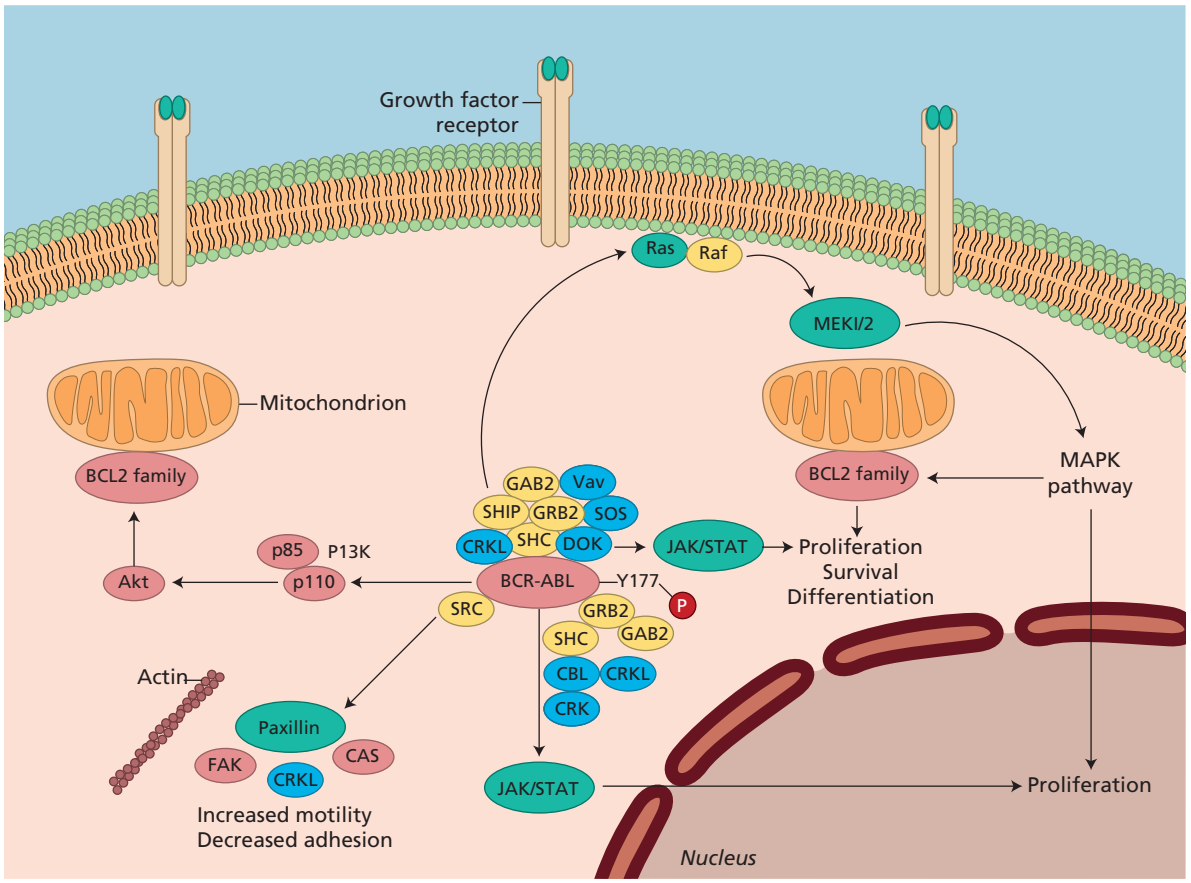


Figure 24.5 BCR–ABL1 is a constitutively active tyrosine kinase with multiple substrates. Interactions between BCR–ABL1 and secondary messengers such as GRB2, GAB2, CRKL lead to inhibition of apoptosis and increased proliferation through activation of the PI3K/AKT, JAK/STAT and RAS/RAF/MEK pathways. (Source: Weisberg *et al.*, 2007 [*Nat Rev Cancer* 2007; B: 345–56]. Reproduced with permission of Nature Publishing.)

Measurement of residual disease and monitoring with cytogenetics and *BCR–ABL1* RQ-PCR

CML patients receiving therapy are commonly monitored with blood counts, marrow cytogenetics and molecular assays. A complete haematological response (CHR), the first treatment milestone, is achieved with normalization of blood counts and morphology, as well as resolution of splenomegaly. Achieving CHR from treatment with cytotoxic agents (such as hydroxycarbamide) has little or no impact on the risk of transformation and can only be regarded as a short-term goal of therapy or palliative treatment.

When patients are successfully treated with tyrosine kinase inhibitors that specifically target the Ph^+ clone, the risk of transformation starts to fall when a reduction in the overall number of leukaemic cells can be achieved. Examination of marrow metaphases will reveal a decreasing proportion of cells with the Ph chromosome as a proportion of total metaphases examined, as normal haemopoietic stem cells and their progeny recover. The proportion of residual Ph^+ cells correlate to different levels within a formal classification of cytogenetic responses. Minimal and minor cytogenetic responses are defined as a reduction in the percentage of residual Ph^+ metaphases to 66–95% and 36–65%, respectively. Partial cytogenetic response (PCyR) equates to 1–35% Ph^+ metaphases and patients without Ph^+ clones detected on karyotyping are said to have achieved complete cytogenetic response (CCyR). The term major cytogenetic response (MCyR) is sometimes used to encompass CCyR and PCyR. Achievement of CCyR is an important milestone and has been associated with excellent long-term progression-free survival in multiple clinical studies.

Although cytogenetic examination is a valuable tool for monitoring disease response, it is labour intensive, and entails significant inconvenience and discomfort for the patient. Furthermore, it is routine in some laboratories to examine only 20 metaphases per patient, resulting in a detection sensitivity for residual disease of ~5%. FISH may be used as an adjunct to cytogenetics when the Ph chromosome results from a cryptic translocation. FISH may also increase the sensitivity of residual disease detection. Fluorescent fusion signals signifying residual Ph chromosomes are easily identified by experienced operators, allowing for 200–500 metaphases to be examined, correlating to a lower threshold of detection to ~1% (Figure 24.2b). However, there is currently insufficient data to correlate specific clinical outcomes with various levels of residual disease as detected by FISH, and this method is not currently recommended for routine disease response monitoring.

The introduction of imatinib as the first widely effective therapy stimulated the development, standardization and widespread implementation of RQ-PCR as a more sensitive disease monitoring tool. Due to its demonstrated clinical significance correlated with specific transcript levels, and the ease of use, RQ-PCR has gained wide acceptance. *BCR–ABL1* RNA transcripts, whether measured from a peripheral blood or a bone marrow sample, correlates well with the number of residual Ph^+ cells. Peripheral blood is more commonly studied as it is convenient to collect and entails less discomfort for the patient. The transcript numbers of a control gene such as *BCR, ABL1* or *GUSB* is usually measured from the same sample. The test result is reported as a ratio of *BCR–ABL1* transcripts numbers to control gene transcript numbers, as a percentage. As RNA is labile, control gene transcript numbers differentiate acceptable specimens from those subjected to degradation from transport and storage.

Reporting results on the International Scale (IS) allow interpatient and intrapatient comparability between different laboratories through the use of a simple conversion factor. By convention, 100% IS is a diagnostic baseline value derived from 30 diagnostic samples from the IRIS (International Randomized Study of Interferon and STI571) study. As well as a ratio, it is also common to describe the residual tumour burden as a $\log_{10}$ reduction from the standardized baseline, with 10%, 1%, 0.1% being a 1, 2 and 3 log disease reduction, respectively (Figure 24.6). A 2 log reduction, or *BCR–ABL1* of ≤1% IS, is a convenient and reliable surrogate for CCyR. A 3 log reduction, or *BCR–ABL1* of ≤0.1% IS, is a milestone response also named major molecular response (MMR, MMolR), associated with significantly superior event-free survival. Deeper responses, such as MR^4, $MR^{4.5}$ and MR^5 have also recently been defined as 4, 4.5 and 5 log tumour reductions, being equivalent to *BCR–ABL1* of

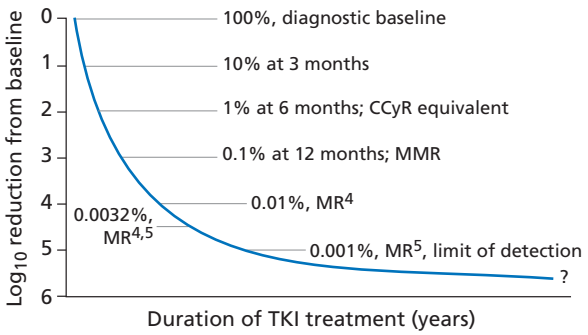


Figure 24.6 Idealized optimal reduction in *BCR–ABL1* over duration of TKI treatment. A RQ-PCR value of 100% on the international scale is derived from the baseline value of a subgroup of patients in the IRIS study. After 3 months of TKI treatment, patients should have a 1 log reduction in their *BCR–ABL1* to 10%. After 6 months, patients should achieve *BCR–ABL1* ≤1%, which is a surrogate for CCyR. After 12 months of therapy, the optimal response is considered MMR or *BCR–ABL1* ≤0.1%. Deeper molecular responses may follow after a variable period. MR^5 or $MR^{4.5}$ act as a platform for patients who wish to join TKI cessation studies. RQ-PCR assays with sensitivity to >5 logs below the diagnostic baseline are not routinely available.

≤0.01% IS, 0.0032% IS and 0.001% IS, respectively. (In some literature, MR^5 may be referred to as complete molecular response or CMR, though this term is no longer favoured.)

In general, patients receiving TKI therapy should have RQ-PCR testing every 3 months for the first 2 years, then every 3–6 months if MMR is achieved. Patients who fail to achieve time-dependent treatment targets (see later) may benefit from more frequent testing. An increase of RQ-PCR value of more than twofold, associated with loss of previously achieved milestones, may be a consequence of treatment resistance and is worthy of further investigation.

Treatment

Monotherapy with a TKI that targets the ABL1 kinase is currently regarded as standard treatment for CML-CP and AP. Patients with CML-BP may be treated either with a TKI alone or in combination with chemotherapy, though duration of response tends to be short without consolidation with an allogeneic stem cell transplant. Oral cytotoxics, such as hydroxycarbamide (hydroxyurea), are used commonly for cytoreduction of leucocytosis before confirmation of diagnosis (leucapheresis can also be used in more extreme cases). During the initial cytoreduction phase, it is common to coadminister allopurinol to blunt any rise in serum urate, and to monitor for electrolyte disturbances carefully. There is no longer any role for oral busulfan for upfront treatment. Although interferon-α results in durable complete cytogenetic response in ~20% of patients, its toxicity precludes generalized use. Pegylated interferon has more favourable toxicity and efficacy profiles. Its role as a low-dose adjunct to TKIs is currently being examined in several clinical studies. The combination of interferon and cytarabine, considered the standard of care prior to the widespread adoption of TKIs, now has little role in CML treatment due to its significant toxicity profile.

The characteristics of the five TKIs currently available are discussed below. Radotinib is encountered rarely outside Asia and will not be discussed here. Haematological toxicities are commonly encountered in the first few months of TKI treatment, and usually improve with time as re-population of the marrow by normal Ph-negative clones occurs. Dose intensity, especially in the initial months of treatment, is critical to the long-term success of TKI treatment. Supportive measures, such as colony growth-stimulating factor (GCSF, EPO) or transfusion support may be used as temporary adjuncts to allow for ongoing TKI treatment.

Imatinib

Imatinib was the first TKI to be licensed for use in CML. Its introduction transformed a previously fatal disease for many into a manageable chronic illness with excellent long-term survival. Imatinib competitively inhibits tyrosine kinases including c-kit, ABL1, ABL2, c-FMS and PDGFRB. It binds to inactivated forms of BCR–ABL1 by occupying the ATP-binding pocket and prevents conformational changes (such as repositioning of the 'activation loop') required for BCR–ABL1 activation.

The Phase III IRIS trial enrolled treatment-naïve CML-CP patients, who were randomized to receive either imatinib 400 mg daily, or interferon-α and low-dose cytarabine. Superiority of the imatinib arm was demonstrated unequivocally, leading to crossover of 89.2% of the control arm patients by 2 years of follow-up. At last report, the overall survival for IRIS patients randomized to the imatinib arm was 85% at 8 years, and 93% when only CML-related deaths were considered.

The IRIS study established the prognostic significance of treatment milestones. For instance, the 5-year analysis showed progression-free survival (defined in IRIS as progression to AP/BC, loss of a CHR or major cytogenetic response) to be 97% for patients who achieved CCyR by 12 months, versus 93% for patients who achieved only PCyR and 81% for patients who achieved less than PCyR. Of the patients who achieved CCyR, those who achieved the even deeper response of MMR had a PFS of 100%. Results from other TKI studies similarly confirmed the prognostic utility of these milestone responses (Figure 24.7).

Even though imatinib treatment has proved highly successful for many, ~40–50% of patients who start imatinib therapy will discontinue – about half for intolerance and half for resistance. The most common resistance mechanism identified is kinase domain mutations (see below). Clonal evolution, such as acquisition of an extra copy of the Ph^+ chromosome or development of isochromosome 17q may also be involved. Non-adherence, whether intentional, due to toxicity or financial pressures, or unintentional, are also significant contributors to therapeutic failure. Pharmacokinetic considerations leading to subtherapeutic serum imatinib concentrations should also be taken into account, such as severe gastrointestinal disturbances or administration of concomitant medications with the potential to induce CYP P450 activity (Figure 24.8). However, the cause of imatinib failure remains undetermined in a number of patients.

The likelihood of a patient achieving the desired therapeutic outcome with imatinib treatment may also be predicted using clinical scoring systems. The Sokal and Hasford scores are both well established, and online calculators are widely available. They both stratify newly diagnosed CML-CP patients into high-, intermediate- and low-risk groups. The Sokal index includes age at diagnosis, spleen size (measured in cm below the costal margin in the midclavicular line), platelet count and peripheral blood blast percentage. The (Euro) Hasford score includes the peripheral blood eosinophil and basophil percentages in addition to the Sokal parameters. These scores correlated well with progression-free survival in the pre-imatinib era. With the dramatic improvements in overall and progression-free survival that accompanied the introduction of TKI treatment for patients in all risk groups, the difference in outcomes is no longer as pronounced. A further refinement of the

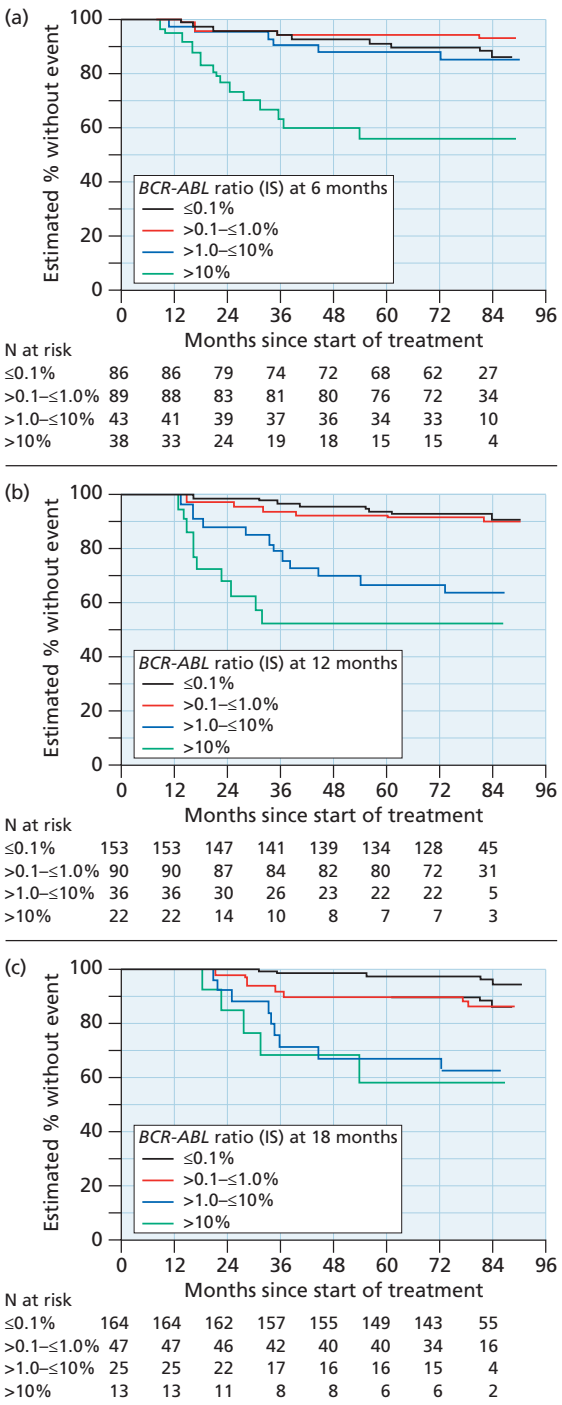


Figure 24.7 Event-free survival (EFS) from the IRIS cohort. The majority of IRIS patients enjoy excellent EFS. However, patients *BCR–ABL1* RQ-PCR of >10%, >1% and >0.1% at 6, 12, and 18 months after starting therapy had inferior EFS. Events in this analysis included (i) loss of complete haematologic response, ii) loss of major cytogenetic response, (iii) progression to AP/BC, or (iv) death due to any cause. (Source: Hughes *et al.*, 2010 [*Blood*;116:3758-65]. Reproduced with permission of American Society of Hematology.)

Hasford score is the EUTOS score, which simplified the calculation by only including the spleen size and basophil percentage, and stratifies patients into two risk groups with respect to the probability of achieving CCyR 18 months after starting a TKI. The validation of this scoring system is ongoing. A number of emerging biomarkers may also provide prognostic information, though interventions based on stratifications using this information have not been prospectively verified. These include serum drug levels, the intracellular transporter OCT1 (organic cation transporter 1; also called SLC22A1), CIP2A (inhibitor of tumour suppressor protein PP2A), gene expression profiles and polymorphisms involved in drug metabolism and apoptosis.

The most common starting dose of imatinib is 400 mg daily. A dose of 600 mg daily may lead to higher molecular responses and is well tolerated in many patients. However, randomized studies comparing imatinib 400 mg daily to 800 mg daily have not demonstrated a clear benefit to justify the uniform use of higher doses. The more common toxicities of imatinib tend to be gastrointestinal disturbances (nausea, vomiting, anorexia, diarrhoea, dyspepsia) and oedema (periorbital, facial and pedal). Judicious use of diuretics, prokinetic agents and taking imatinib with a full meal may ameliorate these symptoms. Haematological toxicities (anaemia, neutropenia and thrombocytopenia) are commonly associated with TKIs and tend to be most severe at treatment initiation. Cardiac toxicity, an initial concern, appears minimal with long-term clinical follow-up.

Nilotinib

Nilotinib is a second-generation TKI effective in imatinib-resistant or -intolerant patients. In addition to BCR–ABL1, it also inhibits PDGFR and c-kit. It is active against a range of imatinib-resistant kinase domain mutations, except for T315I, F359V/C, E255K/V and Y253H/F. It is administered as a twice daily dose on an empty stomach. Treatment is usually well tolerated apart from cytopenias at treatment initiation. Common associated biochemical abnormalities include elevations in serum liver enzymes, lipase and amylase, and hyperbilirubinaemia (thought to be secondary to decreased hepatic bilirubin excretion into bile and is generally reversible). Pancreatitis is an uncommon, but notable idiosyncratic toxicity. Incidence of cross intolerance with imatinib is low with regards to non-haematological side-effects. Although all TKIs can prolong QTc interval, the effect of nilotinib is perhaps the best documented, especially when nilotinib is taken with a fatty meal.

Nilotinib is currently indicated for second-line treatment of CML-CP and CML-AP in patients who have resistance or intolerance to at least one TKI at a dose of 400 mg BID. There is currently inadequate clinical evidence to recommend its use in Ph^+ ALL or CML-BP. The data surrounding its frontline use came from the ENESTnd study, which showed CML-CP

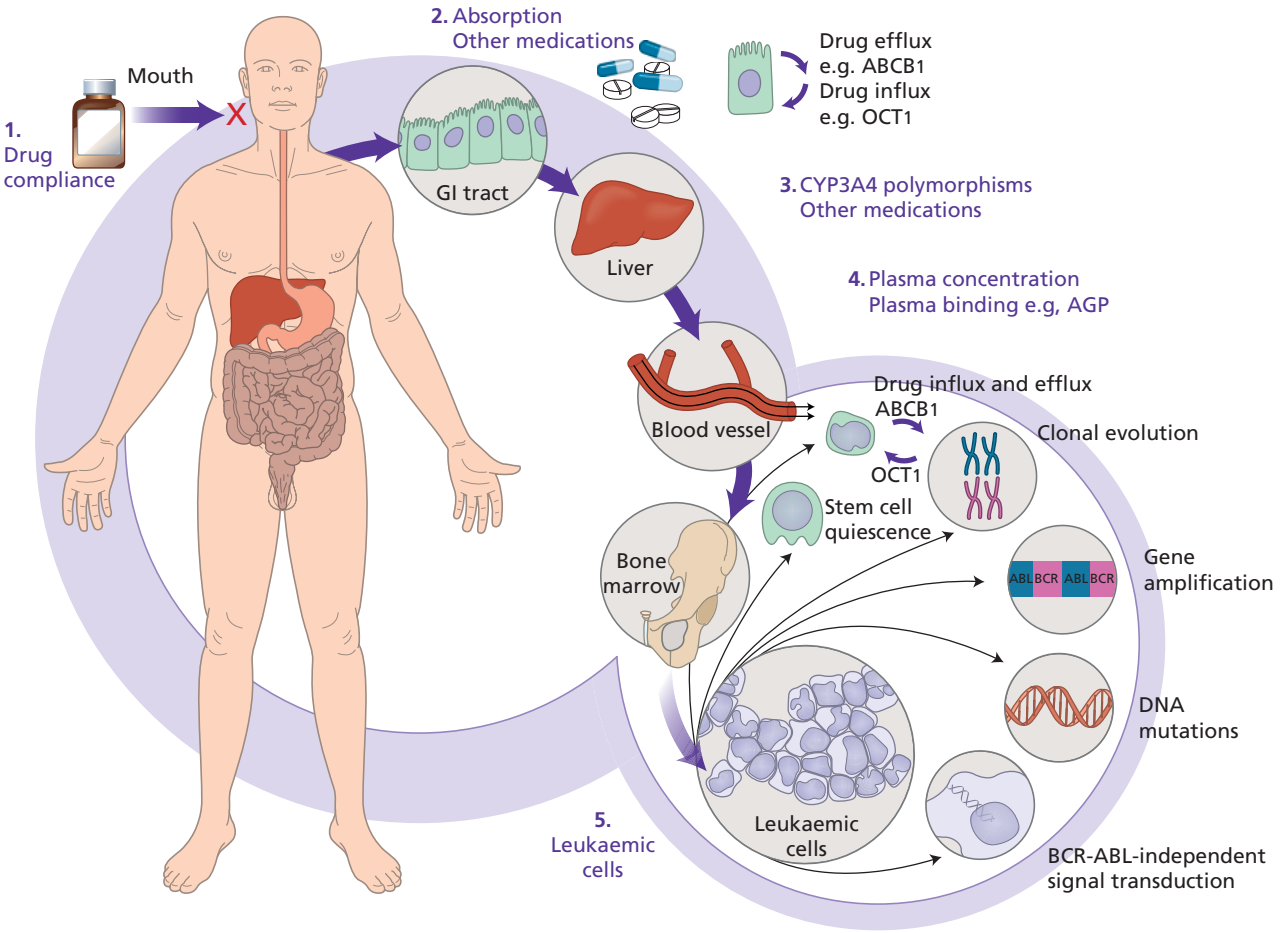


Figure 24.8 A number of mechanisms are implicated in imatinib-related treatment resistance. AGP, α-1 acid glycoprotein. (Source: Apperley, 2007 [*Lancet Oncol*. **8**(11): 1018–29]. Reproduced with permission of Elsevier.)

patients treated with frontline nilotinib (at either 300 mg BID or 400 mg BID) have decreased risk of disease transformation to AP/BP, reduced risk of kinase domain mutation acquisition, and higher achievement of MMR/MR$^{4.5}$, compared to imatinib-treated patients. Nilotinib-treated patients were also more likely to achieve early molecular response (EMR; *BCR–ABL1* <10% at 3 months), a surrogate for PFS (Figure 24.9). However, overall survival differences were marginal between imatinib and nilotinib. Whilst efficacy was similar in the nilotinib 300 mg BID and 400 mg BID arms, toxicity is more common at higher doses. Consequently, first-line patients are dosed at 300 mg BID. Nilotinib-treated patients have higher incidence of hyperglycaemia and dyslipidaemia, and accelerated development of arterial occlusive disease is an emerging concern. This manifests as ischaemic heart disease, strokes and peripheral arterial occlusive disease, and appears to be more common in both nilotinib arms compared to the imatinib arm in the ENESTnd study. However the effect is especially pronounced in patients treated with nilotinib 400 mg BID.

Dasatinib

Dasatinib is a second-generation TKI and a dual Src/ABL inhibitor, though it also inhibits a number of other tyrosine kinases. It is approved for frontline treatment of CML-CP, as well as cases resistant or intolerant to at least one prior TKI in all phases, and for second-line use in Ph$^+$-ALL. It has the greatest *in vitro* activity against BCR–ABL1 amongst the three TKIs approved for frontline treatment and, unlike imatinib, dasatinib can bind to BCR–ABL1 in active conformation. Even though dasatinib has a short serum half life, it is usually given as a daily dose (100 mg for CP and 140 mg for advanced phases), taken with or without food. There is a higher incidence of toxicity without improvement in efficacy with a twice daily regimen. In the Phase III DASISION study, dasatinib-treated patients had superior achievement of MMR and MR$^{4.5}$ compared to imatinib treatment, though no significant difference in overall survival and progression-free survival was demonstrated. Interestingly, the number of patients who developed kinase domain mutations was similar in both arms, with a higher number of T315I cases in

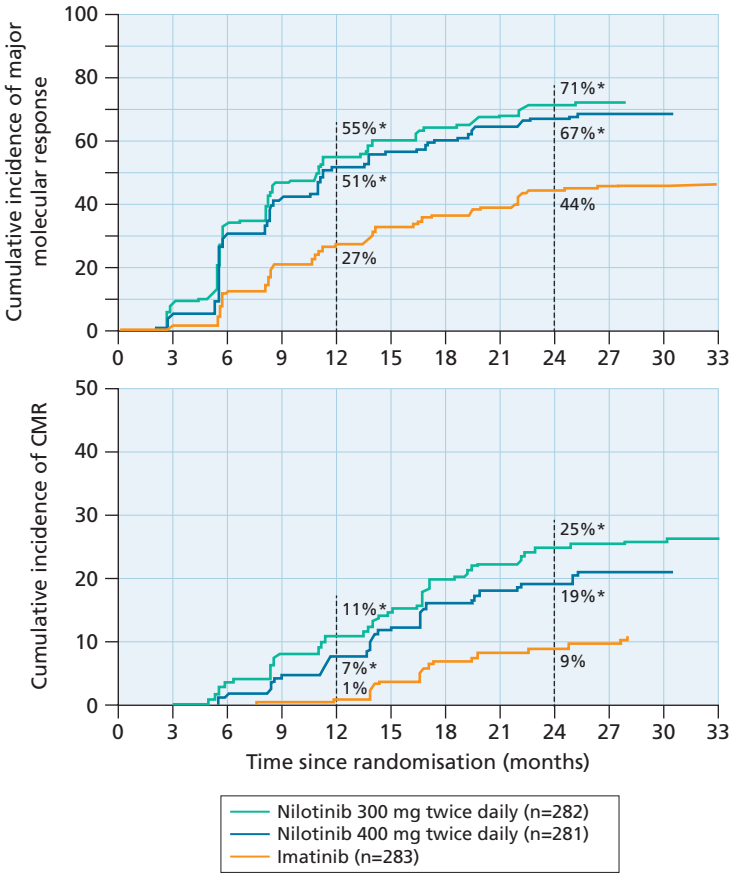


Figure 24.9 Results from the ENESTnd study. Newly diagnosed CML-CP patients were randomized to three arms – nilotinib 300 mg BID, nilotinib 400 mg BID and imatinib 400 mg OD. Patients who received nilotinib had superior probability of achieving MMR and MR[4.5]. Similarly, data from the Phase III DASISION study showed that patients assigned to dasatinib had better molecular responses compared to patients receiving imatinib. (Source: Kantarjian *et al.*, 2011 [*Lancet Oncol.* **12**(9): 841–51]. Reproduced with permission of Elsevier.)

dasatinib-treated patients. In the second-line setting, dasatinib is active against a range of imatinib-resistant mutations, although it is inactive against T315I/A, V299L and F317V/C mutations.

Idiosyncratically, dasatinib causes pleural effusions in up to 40% of some patient groups. Patients with pre-existing cardiac or pulmonary disorders and advanced disease phases are particularly at risk. The most important risk factor appears to be age, with patients over 70 years of age having an incidence of over 50%. Onset of pleural effusions are unpredictable and may first present months to years after treatment commencement. The pleural fluid is usually an exudate, often with a lymphocytosis which may be oligoclonal, suggesting a role of immune dysregulation in its aetiology. Usual treatment of symptomatic, dasatinib-induced, pleural effusions includes a brief treatment interruption, judicious use of diuretics and thoracentesis as necessary. Most experts would favour a short course of prednisolone as immunomodulation, but no firm evidence in favour of this approach is available. Switching to an alternate TKI may be necessary in some patients, if pleural effusions do not resolve with these simple measures. Occasionally, the pleural effusion may persist after switching to another TKI. Pulmonary hypertension (PAH) appears to be a rare complication of dasatinib therapy, though given the difficulty in diagnosing PAH and its inconsistent symptoms, the true incidence is difficult to ascertain. Only small series had been reported, the largest of which is a French cohort of nine cases, from a database that implied an incidence of 0.45%. This is usually, but not always, reversible upon dasatinib cessation, although pulmonary arterial pressures rarely return to normal. Other idiosyncratic side effects associated with dasatinib treatment include pericardial effusions, lymph node follicular hyperplasia and peripheral blood lymphocytosis.

Bosutinib

Bosutinib is a second generation TKI indicated for the treatment of CML patients with intolerance or resistance to at least one other TKI, for all disease phases. It is administered as a single daily dose of 500 mg daily. In the second-line setting, ~50% of CML-CP patients who switch from imatinib to bosutinib will achieve CCyR. This falls to less than one-third for patients who have failed two previous TKIs, and is 25% and 20% in accelerated and blastic phases, respectively. Clinical evidence suggests inferior responses for patients with the V299L, L248V, F359C (and perhaps G250E) mutations. The T315I mutant clone is totally resistant to bosutinib.

Evaluation of bosutinib's efficacy in the frontline setting is ongoing. Gastrointestinal side effects, including diarrhoea, nausea, abdominal pain and vomiting are common and experienced by up to 80% of patients. Other common non-haematological toxicities include rash, headache and liver enzyme elevations.

Ponatinib

Ponatinib is a third-generation TKI, and is the only drug with activity against the T315I mutation. Cases with compound mutations (where more than one kinase domain point mutation is present in the same *BCR–ABL1* transcript), especially those involving the E255K/V mutation, are potentially resistant to this drug. Ponatinib is effective in patients with resistance or intolerance to at least one previous line of treatment in all phases of CML disease, as well as Ph^+-ALL, especially in patients with T315I as the only known mechanism of resistance. Rash, elevated lipase, pancreatitis and headaches seem to be common non-haematological toxicities of the drug. Of significant concern, 27% of patients who participated in early clinical studies reported having had either arterial or venous thrombotic events. In light of emerging toxicities, a Phase III study comparing imatinib to ponatinib as frontline therapy was prematurely terminated, though ponatinib retains marketing approval in the EU and US for Ph^+ leukaemias with the T315I mutation, and in cases where other TKIs have no real prospect of therapeutic success. This drug should be used with extreme caution and at the lowest effective dose in patients with risk factors for vascular disease, and patients with a history of vascular events may not be suitable candidates for this drug.

Choice of TKI up front

Currently, imatinib, nilotinib (300 mg BID) and dasatinib are available in most countries for up-front treatment in CML-CP. All three are reasonable options and although the second-generation TKIs lead to lower rates of progression, all three lead to similar long-term overall survival. The choice of therapy in each individual patient would rest mainly on a patient's comorbidities and their risk of disease transformation, as well as the priority placed on achievement of deep molecular responses and opportunities for subsequent trial of treatment cessation.

Both nilotinib and dasatinb may be reasonable choices in patients with high Sokal/Hasford risk. The daily dosing of dasatinib may be more convenient in younger patients at higher risk of non-compliance, but should be avoided in patients with significant coexisting cardiopulmonary disease who risk decompensation with pleural effusion and pulmonary hypertension. Unlike dasatinib, nilotinib-treated patients appear to be at lower risk of acquiring mutations. Nilotinib also appears to lead to superior achievement of deep molecular responses, and may be ideal in patients for whom the opportunity to cease treatment is a high priority (e.g. women planning to start a family). However, its twice-daily dosing away from food may compound non-adherence. Nilotinib should be avoided in patients with previous vascular events and uncontrolled cardiovascular risk factors, especially at the higher dose of 400 mg BID.

Imatinib, in contrast, appears to have a favourable long-term toxicity profile. Whilst side-effects are common and cause significant discomfort, most patients learn to manage these symptoms. In older patients with pre-existing vascular or respiratory problems, imatinib may be a safer choice. As the patent on imatinib expires in many jurisdictions in the next few years, its use up front may become more cost effective, with more potent agents held in reserve for cases with drug resistance. However, patients who receive imatinib up front at a dose of 400 mg daily are at higher risk of transformation to AP/BP, in comparison to those who receive up-front nilotinib or dasatinib.

Incorporation of both imatinib and a second-generation TKI into a combined treatment strategy may be a reasonable compromise to enhance safety and retain efficacy. For instance, the Australasian TIDEL-II study started newly diagnosed CML-CP patients on imatinib 600 mg daily. Patients were assessed against time-dependent molecular targets (*BCR–ABL1* $\leq$10%, $\leq$1% and $\leq$0.1% at 3, 6 and 12 months, respectively). Patients who failed these targets switched to nilotinib with or without a trial of imatinib 800 mg/day, as did patients with imatinib intolerance or loss of response. The 24-month MMR, MR4.5 and the 3-year overall and progression-free survival from TIDEL-II were similar to that achieved by patients who received a second-generation TKI up front. However, this strategy allowed many patients to stay on imatinib. Over the first 2 years of study treatment, ~50% of patients remained on imatinib and ~30% of patients switched to nilotinib, whilst ~20% had discontinued study treatment.

In the paediatric setting, the majority of evidence and experience concerns the use of imatinib, where the efficacy and side-effect profile is very similar to that seen in adults. One important exception is decelerated growth in height, especially in prepubertal children treated with imatinib. Imatinib is known to cause hypocalcaemia, hypophosphataemia and decreased bone remodelling, all factors that may lead to stunted growth. Imatinib, as with other TKIs, is dosed according to body surface area in children, similar to other chemotherapeutic agents. Data for the use of dasatinib has been accumulating, especially in the second-line setting after imatinib failure, and most clinicians would favour this drug in the frontline setting for advanced phase disease. The experience of nilotinib in the paediatric population remains limited.

Treatment goals

For CML-CP patients treated with TKIs, disease-specific survival is excellent and overall survival is only slightly inferior

to an age-matched population. However, despite the introduction of increasingly potent TKIs, survival for patients with CML-BP has not improved dramatically, and prevention of disease transformation is therefore of paramount importance. There is a reliable correlation between time-dependent reductions in tumour burden (measured by RQ-PCR or cytogenetics) and survival. Milestone responses, such as achievement of CCyR at 6 months and MMR at 12 months, differentiate patients at high risk of disease transformation from patients with optimal treatment response, with the latter being adopted as surrogate end points for many clinical trials.

Consensus statements concerning optimal time-dependent treatment targets are available. The two most widely used are published by the US-based NCCN (National Comprehensive Cancer Network) and the ELN (European Leukemia Net), with subtle differences between the two. The ELN stratify responses at 3, 6 and 12 months after therapy commencement into three categories – Optimal, Warning and Failure. Optimal response is associated with favourable long-term survival. Conversely, patients who fail any of the time-dependent targets are at high risk of disease progression should they continue with the same treatment. Patients who have received a 'warning' are in the intermediate category and these patients should be closely monitored and counselled to adhere to therapy. In contrast, the NCCN dichotomizes patients into those who reach their targets and those who do not, and stratify interventions according to the relative shortfall between response and target. The relevant milestones are summarized in Table 24.2. Treatment recommendations are based on consensus expert opinions, as good-quality evidence demonstrating the benefit of interventions is not currently available. Interventions based on early molecular responses are a particularly contentious issue. Patients who record a *BCR–ABL1* IS > 10%, 3 months after starting therapy have inferior overall survival, a high risk of disease transformation (between 8 and 12%), and inferior achievement of MMR/$MR^{4.5}$ (Figure 24.10). There is a divergence of opinion as to whether a therapeutic intervention should occur for patients failing to achieve this milestone, or whether it might be preferable to wait for the 6-month *BCR–ABL1* to provide further prognostic information prior to acting, there is little data to support either approach. Interestingly, data from the Adelaide cohort of imatinib-treated patients suggest that patients with *BCR–ABL1* >10% at 3 months may be further risk stratified, based on an individual patient's *BCR–ABL1* halving time. This is the approximate duration over which a patient's *BCR–ABL1* decreases by 50%, and patients with a halving time of >76 days were at the highest risk of treatment failure. The German CML IV study group came to the similar conclusion that the velocity of an individual patient's *BCR–ABL1* drop may add additional prognostic information to RQ-PCR measurements at single time points.

Duration of treatment

With the success of TKI therapy, an increasing number of patients achieve deep and durable molecular responses. Although it was originally assumed that patients would need to continue TKI therapy indefinitely to prevent late disease relapses, a number of patients have since participated in treatment cessation studies and have achieved treatment-free remissions (TFR). Two of the earliest treatment cessation studies were the French STIM study and the Australian TWISTER study (Figure 24.11), where imatinib-treated CML-CP patients who had undetectable *BCR–ABL1* (with assays of >5 log sensitivity) for an uninterrupted period of 2 years stopped imatinib and were monitored monthly with RQ-PCR. Rapid molecular relapse (i.e. RQ-PCR moving from undetectable to positive) was noted in 60% of patients, mostly occurring within the first 6 months after stopping. All relapses were sensitive to imatinib re-challenge, with most patients re-establishing deep molecular response, and no disease transformation was reported. The remaining 40% have remained in a deep molecular remission off treatment, many for over 5 years. There have been, as yet, no late relapses reported in this group, raising the possibility that TFRs may truly be durable in the long term. The biological determinants of successful treatment cessation remain elusive, but appear to correlate with total TKI treatment duration. This may relate to senescence of Ph^+ clones over long periods of TKI treatment. Additional emerging data suggest not all patients who experience early molecular relapse require re-treatment. In about a third of these patients, RQ-PCR becomes detectable, but remains below 0.1%, even without treatment. This observation is pending confirmation in larger studies.

Further studies are currently underway to examine whether the criteria for treatment cessation can be relaxed to $MR^{4.5}$ or MR^4 from MR^5 without detrimental effects, whether treatment with a more potent TKI up front will further improve the subsequent success of TKI cessation and whether adjunctive therapy such as pegylated interferon will be of benefit just prior to or in the initial period of therapy cessation. It should be noted that although early results are promising, treatment cessation should not be attempted outside of a clinical study without the ability to perform frequent and highly sensitive molecular monitoring. Furthermore, the current findings should not be extrapolated to patients with advanced phase disease.

Treatment resistance

Treatment resistance may be encountered in ~10–20% of patients receiving TKI in the frontline setting (an additional ~10–20% of patients fail therapy because of varying degrees of toxicity, and some are better classified as failing therapy because of intolerance, though the two frequently overlap). Kinase domain mutations are identified in 30–50% of cases and

Table 24.2 Treatment targets as recommended by NCCN and ELN.

	NCCN 2014			ELN 2013			
		Suboptimal				Failure	
	Optimal	Definition	Management	Optimal	Warning	Definition	Management
3 months	BCR–ABL ≤10% IS or ≤35% Ph^+	>10% IS, >35% Ph^+	If IM is used: increase dose to 800 mg/d or change to 2nd-gen TKI. If frontline NIL or DAS, change to another TKI (not IM). HSCT or clinical trial may also be appropriate	BCR–ABL ≤10% IS or ≤35% Ph^+	>10% IS or Ph^+ 36–95%	No CHR, >95% Ph^+	
6 months	BCR–ABL ≤10% IS or ≤35% Ph^+	>10% IS, >35% Ph^+	Change to another TKI (not IM). HSCT or clinical trial may also be appropriate	BCR–ABL ≤1% IS or 0% Ph^+	1–10% IS or Ph^+ 1–35%	> 10% IS or > 35% Ph^+	If failing IM, switch to 2nd-gen TKI or PON, and do HLA tissue typing. If failing 2nd-gen TKI, switch to an alternative 2nd-gen TKI or PON, do HLA tissue typing and consider HSCT. Consider results of mutation analysis in TKI selection and only use PON in T315I cases.
12 months	BCR–ABL ≤0.1% IS or 0% Ph^+	1–35% Ph^+	Preferably change to another TKI (not IM) ±HSCT depending on response; or increase IM dose to 800 mg/d if taking lower dose IM, or continue same treatment and observe closely.	BCR–ABL ≤0.1% IS	BCR–ABL 0.1–1% IS	BCR–ABL >1% IS or >0% Ph^+	
		>35% Ph^+	Preferably change to another TKI (not IM) ±HSCT deepening on response, or clinical trial may also be appropriate				
18 months	BCR–ABL ≤0.1% IS or 0% Ph^+	>0% Ph^+	Preferably change to another TKI (not IM) ±HSCT depending on response, or clinical trial may also be appropriate				

Ph^+, number of metaphases on cytogenetic examination with the Philadelphia chromosome
Partial 1–35% Ph^+; Complete 0% Ph^+; Major 0–35% (partial + complete)
IM, imatinib; NIL: nilotinib; DAS, dasatinib; PON, ponatinib.

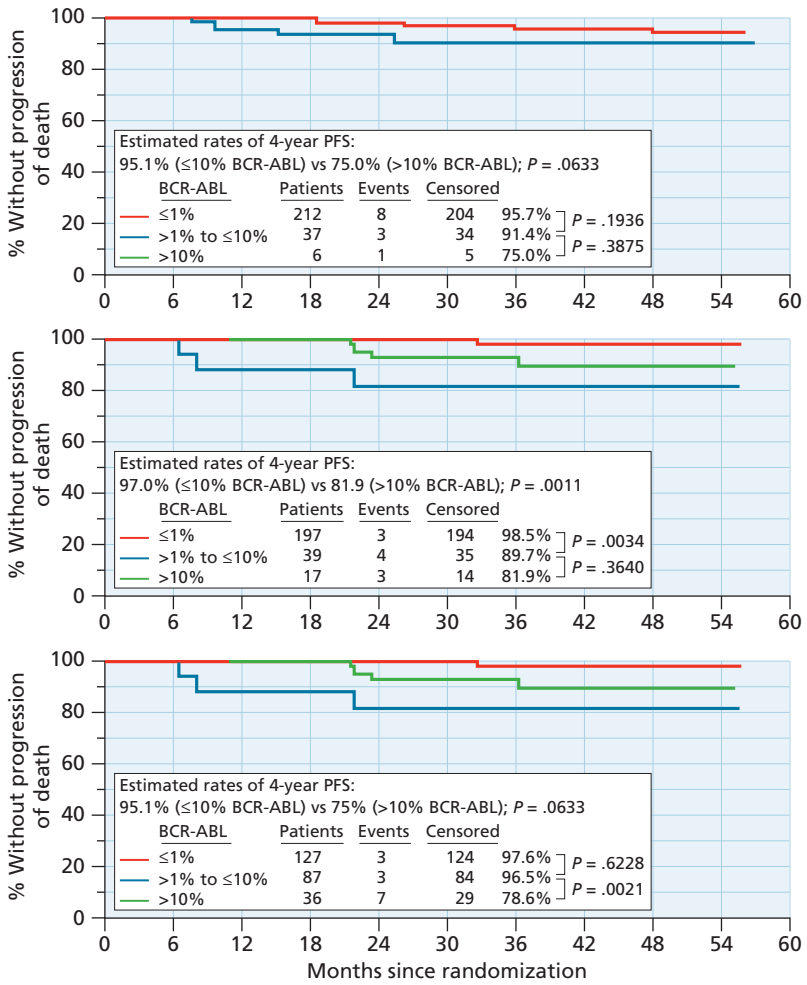


Figure 24.10 Patients failing to achieve an early molecular response experience inferior survival regardless of frontline therapy received. These results from the ENESTnd study showed progression-free survival to be 75.0%, 81.9% and 78.6% in the nilotinib 300 mg BID, nilotinib 400 mg BID and imatinib 400 mg OD arms, respectively, in patients with *BCR–ABL1* >10% at 6 months. (Top panel – nilotinib 300 mg BID, middle panel – nilotinib 400 mg BID, bottom panel – imatinib 400 mg OD). In comparison, patients with *BCR–ABL1* ≤1% 6 months after starting therapy enjoyed 4-year progression-free survival of 95.7–98.5%. The same conclusion can be drawn from the landmark analyses from the DASISION study comparing imatinib to dasatinib. (Source: Hughes *et al.*, 2014 [*Blood* **27**; 123(9): 1353–60]. Reproduced with permission of American Society of Hematology.)

are the most frequently identified mechanism of treatment resistance. More than 100 kinase domain mutations are now known to confer varying degrees of resistance to the TKIs. These point mutations result in amino acid changes, which decrease binding affinity of TKIs, but not the usual substrates (Figure 24.12). The consequence is restoration of constitutive kinase activity. Imatinib appears to be susceptible to the widest range of mutations. The second-generation TKIs are active against most imatinib-resistant mutations, though the T315I mutation confers resistance to all TKIs except ponatinib (Figure 24.13). This threonine residue at position 315 in the ATP binding pocket is critical to the binding of most TKIs and is substituted for isoleucine in ~10–15% of mutations. Other commonly encountered mutations of note include the V299L and F317V/C mutations resistant to dasatinib, and the Y253H, E255K/V and F359C/V mutations resistant to nilotinib. Sanger sequencing is the standard diagnostic method for mutation analysis, and should be performed when a patient fails to reach their time-dependent targets or has loss of a previously achieved response. Most importantly, screening for kinase domain mutations should always be done in patients who are resistant to a TKI, prior to selecting the most suitable TKI for ongoing therapy. Sanger sequencing can only detect mutant clones if they are present in >10–15% of the total *BCR–ABL1*+ population, and only when the total *BCR–ABL1* is >0.1% IS. More sensitive mutation detection techniques using next-generation sequencing and mass spectrometry can reveal subclonal population dynamics; these are currently undergoing clinical evaluation.

Advanced phase diseases

In advanced phase CML, TKI monotherapy is much less effective compared to CML-CP. Although a number of patients can achieve complete haematological responses, these are rarely durable without follow-up combination chemotherapy and allogeneic HSCT. Fortunately, over 90% of CML patients are diagnosed in CP, and with effective TKI treatment, only ~5% of these patients transform to AP/BP.

Evidence regarding the use of TKIs in *de novo* CML-AP is scant. In the STI571 0109 study, imatinib 600 mg daily led to a

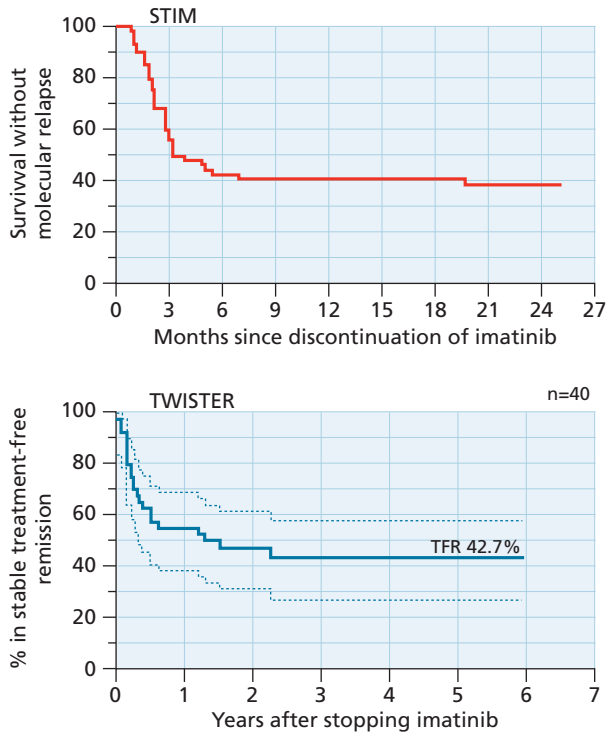


Figure 24.11 For patients with deep molecular responses, a trial of TKI cessation may lead to durable treatment-free remissions (TFR). The STIM (top panel) and TWISTER (bottom panel) studies both showed the rate of TFR in patients who maintained MR^5 for 2 continuous years prior to trial of therapy cessation to be ~40%. (Source: (Top panel) Hughes *et al*., 2014 [*Blood* **123**(9): 1353–60]. (Bottom panel) Ross *et al*., 2013 [*Blood* **122**(4): 515–22]. Reproduced with permission of American Society of Hematology.)

CCyR of 20%. In this study, response rates of AP patients with both clonal evolution and increased blast counts were inferior to those with either feature alone. Extrapolating from data with frontline TKI use in CML-CP patients, the more potent second-generation TKIs should be superior to imatinib in CML-AP. In the MD Anderson Cancer Center cohort of 51 patients, imatinib led to overall survival of 87% at 3 years, with 46% of patients achieving MMR by 18 months. In contrast, patients treated with second-generation TKIs had a 3-year overall survival of 95%, with an 18-month MMR rate of 56%.

Outcomes for CML-BP patients remain dismal, even with potent TKIs, and patients invariably develop resistance to TKI monotherapy. Patients with lymphoid BP are best treated with Ph^+-ALL induction chemotherapy regimens. TKIs are of benefit when used prior to or in combination with chemotherapy, though this may compound marrow toxicity and delay cell count recovery. Whether TKIs have effective CNS penetration is unclear and standard ALL CNS prophylaxis should be offered. Myeloid BP is best treated with AML-type induction regimens, with combinations such as cytarabine and an anthracycline. As with lymphoid BP, TKIs in combination with chemotherapy are generally preferred. Treatment responses in CML-BP are rarely durable and transplant-eligible patients should proceed to allogeneic HSCT once a second chronic phase has been established.

In patients with significant comorbidities or those otherwise unfit for chemotherapy, TKI monotherapy would be a reasonable option. As with CML-AP, there is a paucity of data regarding the use of TKI monotherapy in *de novo* CML-BP. In one Phase II study, 154 treatment-naïve CML myeloid BP patients were treated with imatinib. This led to CHR lasting for ≥4 weeks in 8.8%, and CCyR in 8.1%.

For patients who progressed to AP/BP from CP whilst on treatment, the choice of the subsequent agent should be directed by results of the mutation analysis. About a third of these patients will achieve CCyR with a second-line agent. The T315I mutation is one of the most commonly found mutations in this context and will preclude successful use of all TKIs except for ponatinib. The PACE study included patients with treatment resistance to either nilotinib or dasatinib, or who had the T315I mutation; ~30% of myeloid BP and ~30–50% of lymphoid BP patients achieved complete haematological response and cytogenetic response with ponatinib treatment. However, median duration of response is only ~5 months.

Fertility preservation and pregnancy

The long-term effect of TKIs on prospective fertility is not known. An increasing number of children had been fathered by male patients receiving imatinib, without increased congenital or developmental abnormalities compared to the general population. In contrast, effective contraception is advised for female patients receiving imatinib as it is associated with a higher risk of birth defects. There is currently a lack of human data with regards to the teratogenicity of second- and third-generation TKIs, and effective contraception is advised for both male and female patients. Barrier contraception may be preferred as it does not interfere with CYP450 enzyme activity.

Female patients wishing to conceive electively should ideally achieve and maintain a deep molecular response (e.g. $MR^{4.5}$) for at least 2 years before proceeding to a trial of cessation that will allow for conception. The appropriate washout period between TKI cessation and attempting to conceive is unknown, though pharmacokinetic data would suggest a week should suffice. Regular haematological and molecular monitoring during this period and the subsequent pregnancy is advised. TKIs should be restarted post partum for patients with loss of molecular response. Should patients require treatment whilst still pregnant (e.g. loss of complete haematological response with moderate or marked leucocytosis and/or thrombocytosis)

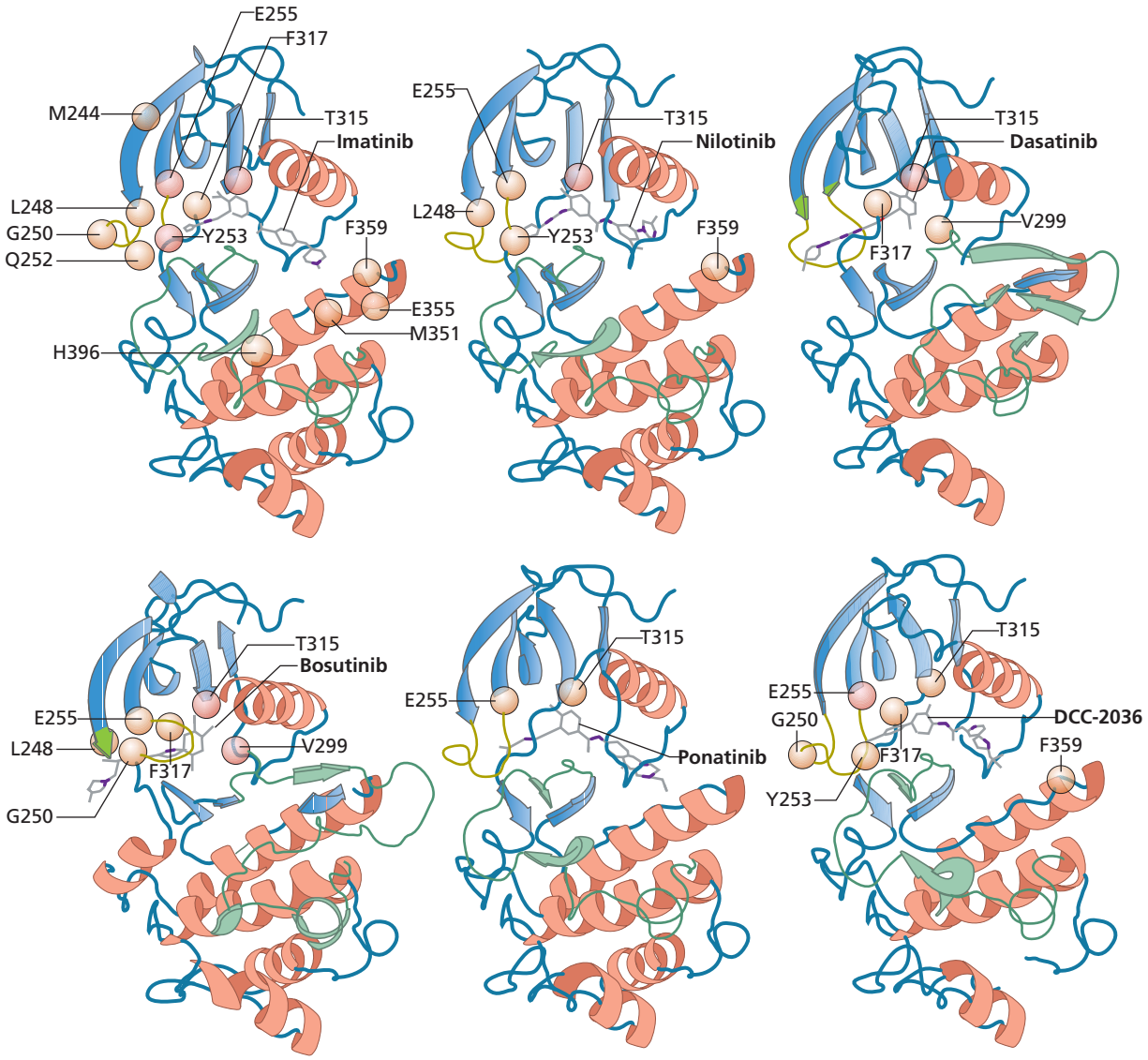


Figure 24.12 The crystal structure of the ABL1 kinase domain and critical amino acid residuals conferring resistance to TKIs. Orange spheres indicate a particular nucleotide exchange confer moderate resistance and red spheres indicate high-level resistance. The spatial relationship of each TKI and the ATP binding pocket is shown. The P-loop is shown in yellow and the activation loop in green. Imatinib is susceptible to the widest range of kinase domain mutations. The threonine residual at position 315 is crucial to the binding of all currently available TKIs, except for ponatinib. The development of DCC-2036 (rebastinib), originally intended as an inhibitor for T315I mutated BCR–ABL1, is currently halted for CML. See Figure 24.13 for a more comprehensive list of point mutations conferring resistance to nilotinib and dasatinib. (Source: O'Hare *et al*., 2012 [*Nature Rev Cancer* **12**: 513–26]. Reproduced with permission of Nature Publishing.)

interferon-α would be a reasonable choice in the second or third trimester. Regular leucapheresis may be an alternative. Data regarding the use of hydroxycarbamide (hydroxycarbamide hydroxyurea) in this setting are scant and busulfan should be avoided. In general, TKIs should be stopped in women who accidentally conceive while taking these drugs, though the course of action should be individualized.

Haemopoietic stem cell transplantation

Prior to the introduction of TKIs, CML (especially CML-CP) was one of the most common indications for allogeneic HSCT. Although optimized protocols for HSCT in CML-CP can result in transplant-related mortality as low as ~5–10%, TKIs are still a better option for the vast majority of these patients. Consequently, HSCT is now usually reserved for patients with

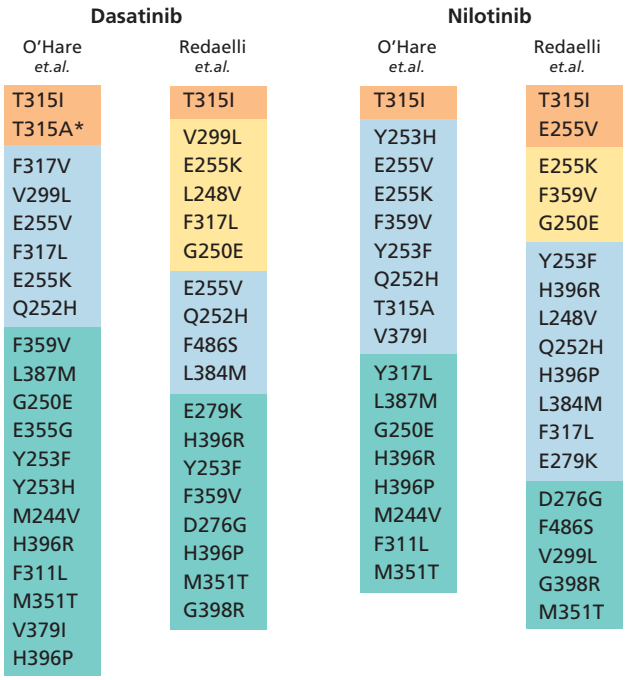


Figure 24.13 Kinase domain mutations are the most commonly identified cause of treatment resistance to TKI. Shading of O'Hare: orange, resistant; blue, intermediate sensitivity; and green, sensitive. Shading of Redaelli: orange, highly resistant; yellow, resistant; blue, moderately resistant; and green, sensitive.
*Although T315A is rated as resistant, it is rarely detected. (Source: O'Hare *et al*., 2009 [*Blood*; 110(7):2242-2249] and Redaelli *et al*., 2009 [*Journal of Clinical Oncology*; 2009; 27(3):469-471].)

advanced phase disease or therapeutic failure to all available TKIs.

Allogeneic HSCT outcomes may be predicted using the EBMT risk score, which includes variables such as patient age, stage of disease, donor type (related versus unrelated), stem cell source (PB versus BM) and length of diagnosis (>12 months versus <12 months). The bulk of transplant outcome data came from patients who underwent myeloablative conditioning, which usually consists of high-dose cyclophosphamide and either busulfan or total body irradiation. The experience with reduced-intensity conditioning in CML had been steadily increasing, and is preferred for older patients.

Problems common to HSCT in other diseases are also applicable in CML. Transplant-related mortality, graft-versus-host disease and relapse-related mortality are significant problems. Many centres will use TKI therapy post allogeneic HSCT and in cases where there is evidence of residual leukaemia by molecular monitoring, though the role for TKIs in this setting has not been fully evaluated. Unlike CML-CP patients, marrow RQ-PCR may be a more sensitive marker for minimal residual disease in CML-AP/BP. Donor lymphocyte infusions may be of benefit in patients with disease relapse post transplant. However, exacerbation of GVHD may be a concern.

Atypical CML/chronic neutrophilic leukaemia (see also Chapters 25 and 26)

Occasionally patients present with leucocytosis with a prominent neutrophilia, though the Philadelphia chromosome and the *BCR–ABL1* fusion gene are both undetected. After exclusion of cryptic variant t(9;22) and atypical *BCR–ABL1* transcripts, a diagnosis of either atypical chronic myeloid leukaemia (aCML) or chronic neutrophilic leukaemia (CNL) may be appropriate. Atypical CML is an overlap syndrome with features of both myeloproliferative and myelodysplastic disorders. Patients usually present with leucocytosis, mainly caused by increased number of neutrophils and their precursors. Maturation is dysplastic, and in contrast to CML and chronic myelomonocytic leukaemia (CMML), basophilia, monocytosis and eosinophilia are rare. Patients with aCML fare poorly and often die from marrow failure or transformation to AML. Like CML and aCML, CNL is characterized by leucocytosis with a predominant neutrophilia. However, unlike aCML, neutrophil morphology is normal in CNL. Neutrophil precursors such as metamyelocytes, myelocytes and promyelocytes are <10% of circulating cells in CNL, which differentiates this disease from CML. The course of CNL is variable, though neutrophilia is usually progressive, accompanied by cytopenia of other lineages resulting in marrow failure, or transformation to AML.

Recurrent genetic lesions have recently been found in the *CSF3R* and *SETBP1* genes in a number of CNL and aCML patients. *CSF3R* is known to be important in the pathogenesis of congenital neutropenia. *CSF3R* mutations in CNL and aCML were found in 59% of patients in one series and are located in two clusters – membrane proximal mutations (T615A and T618I) and truncation-deletions in the cytoplasmic (D771fs, S783fs, Y752X, and W791X). Ruxolitinib (a *JAK2* inhibitor) and dasatinib have *in vitro* activity in some cases, though the clinical utility of these agents is yet to be demonstrated. In another series, *SETBP1* mutations were found in 24% of aCML cases. Most are located in residues 858–871 in the SKI homology region. Neither of the two mutations are specific to CNL/aCML – *CSF3R* can be mutated in AML and T-ALL, and *SETBP1* in CMML cases. The overlap between *SETBP1* and *CSF3R* is yet to be determined, as is their role in the pathogenesis of these diseases.

Prospects

Although the treatment outcome for CML patients has dramatically improved in the last decade, a number of challenges

remain. Early transformation to CML-AP/BP remains the outstanding cause of CML related death. A number of emerging biomarkers with prognostic significance are currently under investigation, and early identification of patients who are at high risk of transformation should become possible. Perhaps more challenging will be the task of finding effective therapies for high-risk patients thus identified, and in patients with *de novo* CML-BP. This may rely on the clinical development of small molecule inhibitors with novel mechanisms of action and synergistic activity against *BCR–ABL1*+ cells. Advances in allogeneic HSCT, such as continuing decreases in transplant-related mortality and improvements in supportive care, may also be important for this group. At the other end of the spectrum, an increasing number of patients achieve stable and deep molecular responses. Selection of patients likely to be able to stop therapy without relapse, as well as strategies to boost the overall TFR rate is an increasingly urgent direction for future research.

Acknowledgement

The authors would like to thank Dr Agnes Yong for her critical review of this chapter.

Selected bibliography

Andolina JR, Neudorf SM, Corey SJ (2012) How I treat childhood CML. *Blood* **119**: 1821–30.

Apperley J (2009) CML in pregnancy and childhood. *Best Practice in Research and Clinical Haematology* **22**: 455–74.

Baccarani M, Deininger MW, Rosti G *et al*. (2013) European LeukemiaNet recommendations for the management of chronic myeloid leukemia: 2013. *Blood* **122**: 872–84.

Branford S, Melo JV and Hughes TP (2009) Selecting optimal second-line tyrosine kinase inhibitor therapy for chronic myeloid leukemia patients after imatinib failure: does the BCR-ABL mutation status really matter? *Blood* **114**: 5426–35.

Gotlib J, Maxson JE, George TI and Tyner JW (2013) The new genetics of chronic neutrophilic leukemia and atypical CML: implications for diagnosis and treatment. *Blood* **122**: 1707–11.

Hughes T, Deininger M, Hochhaus A *et al*. (2006) Monitoring CML patients responding to treatment with tyrosine kinase inhibitors: review and recommendations for harmonizing current methodology for detecting BCR–ABL transcripts and kinase domain mutations and for expressing results. *Blood* **108**: 28–37.

Hughes T, White D (2013) Which TKI? An embarrassment of riches for chronic myeloid leukemia patients. *Hematology American Society of Hematology Education Program* **2013**: 168–75.

Van Etten RA, Mauro M, Radich JP *et al*. (2013) Advances in the biology and therapy of chronic myeloid leukemia: proceedings from the 6th Post-ASH International Chronic Myeloid Leukemia and Myeloproliferative Neoplasms Workshop. *Leukemia and Lymphoma* **54**: 1151–8.

CHAPTER 25

The myelodysplastic syndromes

25

Kavita Raj and Ghulam J Mufti

Department of Haematological Medicine, King's College Hospital, UK

Introduction

The myelodysplastic syndromes are a diverse group of clonal stem cell disorders manifesting commonly with macrocytosis and cytopenias due to impaired blood cell production, morphological dysplasia, a variably cellular bone marrow and an increased risk of myeloid leukaemic transformation. Most patients with a myelodysplastic syndrome (MDS) are over the age of 60 years, with the median onset occurring in the seventh decade of life and increasing with age. The clinical course varies, reflecting the diverse pathobiology of the disease, with some patients having a more indolent course and longer life expectancy, while others present with aggressive disease that evolves rapidly into acute myeloid leukaemia (AML). Overall, approximately 40% of patients will transform to AML during the course of their disease, but a third also succumb to infections. Prognostication is enabled by classification schemes, which take into account the diverse clinical features, and is further supported by scoring systems so that therapy can be both symptomatic and risk directed. The recent discovery of molecular pathways that are commonly disrupted in MDS provides a rational basis for clustering phenotypes, and these are being integrated into risk scoring systems and are likely to influence therapy.

History

The first description of MDS was in 1900 by Von Leube, who used the term leukanamie to describe a patient with severe megaloblastic anaemia that progressed to acute leukaemia. In the 1930s, the term 'refractory anaemia' was coined to refer to a group of patients with a macrocytic anaemia that was unresponsive to iron or other dietary haematinics. While some of these anaemias were apparently attributable to coexisting chronic diseases, others arose *de novo* in otherwise healthy individuals. Bone marrow analysis often revealed a spectrum of abnormalities, notably aberrant maturation of marrow precursors, a variable increase in marrow blasts and, in some patients, the presence of ring sideroblasts. In the 1950s, it was appreciated that AML in the elderly was often preceded by a preleukaemic state characterized by peripheral blood cytopenias and increased numbers of blasts in the marrow. However, the term 'pre-leukaemia' became less popular in the 1970s when it became apparent that patients may never develop leukaemia, but die of complications arising directly from the cytopenias. These include anaemia, infectious problems and bleeding complications. In the 1980s, the term 'myelodysplasia', or 'myelodysplastic syndromes' gained widespread acceptance and reflects the heterogeneity of the disease. Indeed, it was the recognition that MDS is a heterogeneous entity of related disorders of haemopoiesis that provided the rational basis for devising diagnostic classification schemes that have evolved over the last three decades.

Incidence

According to European registry data, the overall annual incidence rate for MDS in adults is 3–4 per 100,000. However, this rate increases markedly with age, exceeding 30 per 100,000 for individuals over the age of 80 years. Over 60% of patients are over the age of 70 at diagnosis, with males more likely to be diagnosed with MDS than females by a ratio of 1.4: 1. MDS in childhood

Postgraduate Haematology, Seventh Edition. Edited by A Victor Hoffbrand, Douglas R Higgs, David M Keeling and Atul B Mehta.

is rare with an annual incidence of 2–3 per million, or less than 5% of paediatric haematological malignancies, tending to occur in children less than 5 years old.

Aetiology

The aetiology of primary MDS is unclear. Exposure to ionizing radiation, benzene, solvents, pesticides and smoking are implicated; however, case-control studies are inconsistent. There is some evidence to suggest that polymorphic variation in certain genes may increase susceptibility to MDS, particularly where the role of the encoded protein is to counter environmental insults to the cell. Examples include enzymes responsible for the metabolism of carcinogens, proteins involved with oxidative stress and DNA damage repair proteins.

In approximately 15% of patients, MDS occurs following prior exposure to cytotoxic chemotherapy and/or wide-field radiotherapy, when it is referred to as therapy-related MDS. These cases of MDS typically arise in patients who have received previous cancer chemo/radiotherapy or prolonged immunosuppressive therapy. The most commonly implicated cytotoxic agents are the alkylating agents, such as cyclophosphamide, melphalan and the topoisomerase II inhibitors, such as etoposide, as well as ionizing radiation and antimetabolite drugs. The latency period between exposure and the onset of disease tends to be 5–6 years for the alkylating agents and for radiation, but is shorter (2–3 years) for the topoisomerase II inhibitors, although the latency can vary from 1 to 10 years and beyond. The risk increases with age and with prolonged exposure to low-dose chemotherapy, such as prescribed for controlling vasculitic and other autoimmune disorders. The risk of developing therapy-related MDS following autologous transplantation for lymphoma varies widely between series, with some reports as high as 12%. Most cases develop within 5 years of transplantation, with the major risk factors being the duration and amount of pretransplant chemotherapy, whether the patient received total body irradiation and the dose of infused stem cells. Cases of therapy-related MDS often show marked trilineage dysplasia, genetic abnormalities such as −5 and/or −7 with alkylating agents and 11q23 *MLL* gene rearrangement with topoisomerase II inhibitors, rapid evolution to AML and poor response to *de novo* AML therapies.

Classification

FAB classification

In 1982, the French–American–British (FAB) group proposed a morphological classification of MDS that divided MDS into five subgroups (Table 25.1). The FAB system was based on the percentage of blasts and ringed sideroblasts in the bone marrow and the presence or absence of a peripheral blood monocytosis.

Table 25.1 The FAB classification of MDS.

Subtype	Blood	Bone marrow
Refractory anaemia (RA)	<1% blasts	Dysplasia, <5% blasts
Refractory anaemia with ringed sideroblasts (RARS)	<1% blasts	As for RA and >15% ringed sideroblasts
Refractory anaemia with excess blasts (RAEB)	<5% blasts	Dysplasia, 5–19% blasts
Refractory anaemia with excess blasts in transformation (RAEBt)	<5% blasts	Dysplasia, 20–29% blasts or Auer rods
Chronic myelomonocytic leukaemia (CMML)	$>1 \times 10^9$/L monocytes	Dysplasia, <30% blasts

Refractory anaemia (RA) is defined as anaemia with a hypercellular marrow exhibiting dyserythropoiesis, but minimal granulocytic or megakaryocytic abnormalities and a marrow blast count of less than 5%. If the proportion of ring sideroblasts exceeds 15% of erythroid cells, the patient is diagnosed as having refractory anaemia with ring sideroblasts (RARS). Patients with increased numbers of marrow blasts are categorized as having refractory anaemia with excess blasts (RAEB) if the blast percentage is between 5 and 19%, or refractory anaemia with excess blasts in transformation (RAEBt) if the blast percentage is between 20 and 29%. Patients with bone marrow blasts greater than 30% are diagnosed as AML. Finally, chronic myelomonocytic leukaemia (CMML) encompasses all cases where the circulating monocyte count exceeds 1×10^9/L. However, this subgroup is particularly heterogeneous, comprising a spectrum of patients whose clinical course varies from more typical MDS to more proliferative states characterized by high monocyte counts and splenomegaly.

WHO classification

The World Health Organization (WHO) classification was first published in 2001 and more recently in a revised form in 2008 (Tables 25.2 and 25.3). It emerged through an international collaboration, the aim of which was to improve the FAB classification by incorporating recent clinical and genetic data. In particular, it recognized the fact that patients with RA with less than 5% marrow blasts had a poorer clinical prognosis if there was evidence of multilineage dysplasia, as opposed to unilineage erythroid dysplasia. New diagnoses of refractory cytopenia with multilineage dysplasia (RCMD) and refractory cytopenia with multilineage dysplasia and ring sideroblasts (RCMD-RS) were

Table 25.2 The Revised WHO classification (2008) of MDS.

Subtype	Blood	Bone marrow
Refractory cytopenias with unilineage dysplasia (RCUD) Refractory anaemia (RA) Refractory neutropenia (RN) Refractory thrombocytopenia (RT)	Unicytopenia or bicytopenia	Dysplasia in >10% of cells of one myeloid lineage only, <5% blasts, <15% ring sideroblasts
Refractory anaemia with ring sideroblasts (RARS)	Anaemia	Erythroid dysplasia only, >15% ring sideroblasts
Refractory cytopenia with multilineage dysplasia (RCMD)	Cytopenia(s)	Dysplasia in >10% of cells in two or more myeloid lineages, <5% blasts, <15% ring sideroblasts
Refractory cytopenia with multilineage dysplasia and ring sideroblasts (RCMD-RS)	Cytopenia(s)	Dysplasia in >10% of cells in two or more myeloid lineages, <5% blasts, >15% ring sideroblasts
Refractory anaemia with excess blasts-1 (RAEB-1)	Cytopenia(s) <5% blasts	Unilineage or multilineage dysplasia, 5–9% blasts, no Auer rods
Refractory anaemia with excess blasts-2 (RAEB-2)	Cytopenia(s) 5–19% blasts	Unilineage or multilineage dysplasia, 10–19% blasts ± Auer rods
Myelodysplastic syndrome – unclassified (MDS-U)	Cytopenia(s) <1% blasts	Dysplasia in <10% of cells in one or more myeloid lineage, cytogenetic abnormality supportive of diagnosis, <5% blasts
MDS associated with isolated del(5q)	Anaemia ± thrombocytosis <1% blasts	Prominent megakaryocytes with hypolobated nuclei Isolated del(5q) cytogenetic abnormality, <5% blasts

introduced for patients with bicytopenia or pancytopenia whose marrows exhibited dysplastic features affecting more than 10% of cells in two or more myeloid lineages. Patients with pure erythroid dysplasia, but otherwise lacking evidence of a clonal karyotypic abnormality should be reassessed after 6 months to ensure persistence of these features before making a diagnosis of MDS.

A subgroup recognized by the WHO classification was MDS with isolated deletion of the long arm of chromosome 5, so-called MDS with isolated del(5q). The clinical and morphological features of this subgroup are distinct, occur more frequently in elderly women, and are characterized by macrocytic anaemia and thrombocytosis that is commonly referred to as 5q– syndrome. The bone marrow shows prominent megakaryocytes that are decreased in size with conspicuously non-lobated or hypolobated nuclei (Figure 25.1). Erythropoiesis is often reduced, with mild to moderate dysplasia, and myeloblasts comprise less than 5% of the myelogram.

Table 25.3 Relative frequency of MDS subtypes.

Subtype of MDS	Approximate percentage of MDS cases	Cytogenetic abnormalities	Median survival (months)
RA	10	25	66
RARS	10	<10	72
RCMD	25	50	33
RCMD-RS	10	50	33
RAEB-1	25	30–40	18
RAEB-2	15	40–50	10
Isolated del(5q)	5	100	116

The RAEB entity was also refined to include two subgroups based on the percentage of blood and bone marrow blasts. The term 'blast cell' here refers to myeloblasts, monoblasts, promonocytes and megakaryoblasts, but not erythroblasts. CD34-positive cells, defined immunophenotypically, are not considered to be synonymous with the blast cells. The two subgroups were denoted as RAEB-1 (blast count <5% in peripheral blood and 5–9% in marrow) and RAEB-2 (blast count 5–19% in peripheral blood and 10–19% in marrow). The morphological finding of Auer rods is also sufficient for making a diagnosis of RAEB-2. The FAB entity RAEBt, defined originally by a blast count in the marrow of over 20%, was abolished in favour of a diagnosis of AML with multilineage dysplasia that often yields poor-risk cytogenetics and refractoriness to chemotherapy. In addition, patients with recurrent leukaemic chromosomal translocations typical of *de novo* AML, namely t(15;17), t(8;21), and inv(16) or t(16;16), are considered to have AML, regardless of the percentage of blasts in the bone marrow.

In the revised WHO classification, there has been some refinement of the system to include patients presenting with

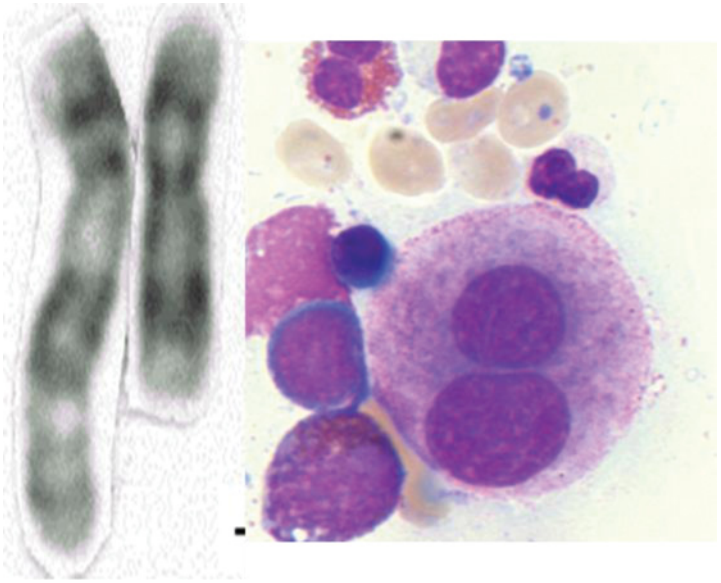


Figure 25.1 Diagnostic findings in 5q– syndrome demonstrating: (a) chromosomal abnormality and (b) typical morphology. Note loss of material from long arm of chromosome 5 and characteristic hypolobated megakaryocytes.

cytopenia affecting a single lineage or bicytopenia and in whom bone marrow dysplasia is restricted to a single lineage. This category is denoted refractory cytopenia with unilineage dysplasia (RCUD), which includes refractory anaemia, refractory neutropenia and refractory thrombocytopenia. The entity of MDS unclassifiable (MDS-U) is reserved for patients with pancytopenia who fail to fit any of the other categories. Finally, childhood MDS has been incorporated as a separate entity to cover all forms of paediatric myelodysplasia, including the subgroup of refractory cytopenia of childhood (RCC) that includes cases that fail to meet the criteria for the RAEB subgroups, specifically where there is evidence of dysplasia, but with fewer than 5% blasts in the marrow.

Finally, the WHO recognizes two additional histological subtypes of MDS, namely hypoplastic MDS and myelofibrotic MDS, each of which account for approximately 10% of cases. Hypoplastic MDS may be difficult to distinguish from aplastic anaemia, but the presence of dysplastic megakaryocytes and reticulin fibrosis favours this diagnosis. Myelofibrotic MDS is differentiated from primary myelofibrosis by the absence of splenomegaly, must be diagnosed from a trephine biopsy that often reveals an excess of blasts and generally follows an aggressive clinical course. Both hypoplastic and myelofibrotic MDS are histological entities based on trephine findings rather than being specific diagnostic entities.

The WHO (2008) also recognizes that some patients have both dysplastic and myeloproliferative features and includes these within the group myelodysplastic/myeloproliferative neoplasms. These include CMML, atypical CML, juvenile myelomonocytic leukaemia and, for those that do not fit into any of these categories, but have evidence of morphological features of myelodysplasia and proliferative features such as a platelet count $\geq 450 \times 10^9$/L or WBC count $>13 \times 10^9$/L, with or without splenomegaly, and no prior history of a myeloproliferative disorder, or of cytotoxic or growth factor treatment and absence of the Philadelphia chromosome, BCR–ABL1 fusion, rearrangement of *PDGFRA*, *PDGFRB*,*FGFR1* and absence of isolated del(5q), t(3;3)(q21;q26) or inv(3)(q21q26). Alternatively if the presentation of *de novo* disease has mixed dysplastic and myeloproliferative features and does not fit into any other MDS, MPN or MDS/MPN category, it is categorized as myelodysplastic/myeloproliferative neoplasm, unclassifiable.

Pathogenesis

Until recently, clues to the pathogenesis of MDS have relied heavily on cytogenetic analysis. However, technological advances such as next-generation sequencing have led to impressive developments in understanding the mutational landscape in MDS and in correlating these with the biological and clinical heterogeneity. In MDS genetic mutations commonly affect RNA splicing, DNA methylation, chromatin modification, transcription regulation, DNA repair, signal transduction and the cohesion complex pathways. Their role alongside the effect of the bone marrow microenvironment and the host immunity in the emergence and progression of the disease are discussed in the following section.

MDS: a clonal expansion of HSCs with ineffective haemopoiesis and leukaemic transformation

MDS is a clonal stem cell disorder which commonly affects all three myeloid lineages (i.e. megakaryocytic, erythroid and

granulocytic/monocytic), with trilineage dysplasia and cytogenetic abnormalities. Flow cytometry reveals an expansion of the long-term HSCs, particularly in high-risk MDS. Furthermore, the common myeloid progenitors (CMPs) are expanded in low-risk MDS, whereas expansion occurs in the granulocyte monocyte progenitors in high-risk MDS. Perhaps unsurprisingly, the situation is complicated and varies from one type of MDS to another. Thus, in 5q– syndrome, the cytogenetic abnormality can be found in both myeloid progenitors and pro-B cells, indicative of a true haemopoietic stem cell origin, while in patients with trisomy 8 the cytogenetic abnormality is often absent from the CD34+CD38− fraction, suggesting that there might be a different initiating event occurring within the haemopoietic stem cell. If so, then this would support the multistep theory of disease progression by which the dysplastic clone evolves over time through acquisition of additional mutations, allowing selection of more proliferative subclones.

Whole-genome sequencing studies of bone marrow from patients with AML that had developed from antecedent MDS shows that 85–90% of bone marrow cells are clonal in these patients in both the MDS and AML phase. Thus almost all cells of the bone marrow myeloid lineage (red cell, granulocytic/megakaryocytic precursors and megakaryocytes) are clonal.

The clonal architecture in MDS is evident from studies of *TET2* mutations, which are detected in a small fraction of the immature HSCs, but in a higher proportion of the mature HSCs, suggesting that the mutation is initiated in the immature cell, but propagated in the mature cell. The mechanisms whereby the MDS HSC achieves clonal dominance is as yet unclear.

Evolution to AML is driven by the acquisition of further driver mutations leading to the formation of subclones of haemopoietic cells with impaired differentiation and/or maturation capacity leading to an accumulation of immature cells and AML. Thus progression to AML is characterized by persistence of the antecedent founder clone and at least one subclone, with new somatic mutations displaying intratumoural heterogeneity. The initial driver mutation may also shape the future trajectory of clonal evolution through constraints on the repertoire of the cooperating genetic events (see also Chapter 18). The bone marrow microenvironment, also referred to as the stem cell niche mechanisms, which comprises the stromal, endosteal and vascular environment, maintains the stem cell numbers through intimate cell–cell contact. Interference with this niche by targeting the molecular interactions between the stem cell and its microenvironment represents a novel therapeutic strategy. Such targets include homing molecules such as CD44 or the ligand–receptor interactions between CXCR4 on stem cells and its chemokine ligand SDF-1 (CXCL12) and between Tie-2 on stem cells and angiopoietin-1 on osteoblasts. In comparison with age-matched controls, the low-risk MDS marrow displays a particularly inflammatory environment, which may provide the milieu for the accumulation of mutated clones. Within this environment myeloid-derived suppressor cells accumulate and inhibit autologous haemopoiesis. These cells are stimulated by the S100a9 ligand via CD33 and Toll-like receptor 4 signalling, leading to ineffective haemopoiesis. The interception of the ligand with soluble receptor or interruption of the signalling cascades mediates reversal in ineffective haemopoiesis and is an area of active study.

Immunological abnormalities in MDS

Immunological abnormalities are commonly encountered in MDS, suggesting that they may play a role in the aetiology and pathogenesis of the disease. This is particularly apparent in cases of hypoplastic MDS that share a number of features in common with aplastic anaemia, notably clinical presentation with macrocytosis and varying levels of dyserythropoiesis. A hypocellular bone marrow is encountered in approximately 10% of cases of MDS and may indeed represent one end of the MDS spectrum that overlaps biologically with aplastic anaemia. Both disorders are characterized by a clonal expansion of T cells that are antigen driven, and there is also an association with HLA-DR15. While the pathogenesis of hypoplastic MDS and aplastic anaemia remains unknown in detail, both these entities appear to be characterized by a pathological immune response triggered by abnormal haemopoietic stem cells that results in the autoimmune destruction of normal stem cells and/or their niche. In addition, acquired mutations in the PIGA gene characteristic of paroxysmal nocturnal haemoglobinuria (PNH) are also encountered, presumably through expansion of small clones in an otherwise hypoplastic marrow, suggesting that PNH similarly is part of this spectrum. Conversely, dysplastic clones carrying cytogenetic abnormalities can emerge in both aplastic anaemia and PNH after treatment and can evolve into frank MDS over time.

A role for the immune system in the pathogenesis of MDS is also borne out by the higher incidence of autoimmune disease in these patients. T-cell-mediated inhibition of haemopoiesis appears to be an important aspect of this mechanism, with oligoclonal CD8+ cytotoxic T cells being found in many patients. However, the antigens produced by the MDS cells that lead to these T-cell responses are largely unknown although, interestingly, WT1-specific CD8+ cells are often detectable in the peripheral blood of MDS patients with trisomy 8. Sometimes, the expanded T cells in MDS can themselves become neoplastic, resulting eventually in the diagnosis of T-cell large granular lymphocytic leukaemia that also falls into this overlap spectrum of bone marrow failure states. Recently, analysis of self-reactive T cells in MDS has focused on the regulatory T cells that play a central role in maintaining immune tolerance. This has revealed a correlation between expanded numbers of CD4+CD25highFoxP3+ regulatory T cells and more aggressive

forms of MDS, as defined by their blast cell percentage, possibly due to suppression of host antitumour mechanisms. However, the most persuasive evidence for immune dysregulation in MDS arguably comes from the recognition that some forms of the disease, notably hypocellular MDS, like aplastic anaemia, can respond well to immunosuppressive therapy with antithymocyte globulin and/or ciclosporin resulting in durable haematological recovery and abrogation of T-cell clones.

Apoptosis in MDS

One of the defining, albeit paradoxical, features of MDS is the presence of cytopenias, despite a typically hypercellular bone marrow. For those patients undergoing leukaemic transformation, the cytopenias arise due to maturation block of the malignant cells. However, in cases of MDS that lack an excess of blasts, the cytopenias are a reflection of the ineffective haemopoiesis that is a hallmark of the disease. The mechanism appears to be one of increased apoptosis of haemopoietic precursors in the marrow, as demonstrated using *in situ* end-labelling of fragmented DNA to reveal cells undergoing programmed cell death. Apoptosis is more prominent in early MDS, such as RA and RARS, than in advanced MDS with excess myeloblasts. Indeed, for the blasts to overcome this apoptotic tendency indicates that they have lost their G2/M checkpoint control that appears to be a necessary requirement for progression to leukaemia. This progression is accompanied by a change in favour of proapoptotic proteins such as c-Myc in CD34-positive precursors at diagnosis to antiapoptotic proteins such as Bcl-2 in leukaemic blasts at time of transformation. Moreover, patients with higher rates of apoptosis have a considerably better overall survival than patients with lower rates of apoptosis, likely reflecting the clonal evolution of the MDS towards AML. This finding is corroborated by flow cytometry analysis of MDS marrow samples to measure relative levels of apoptosis (by annexin V staining) versus proliferation (by Ki67) that demonstrates a shift from apoptosis to proliferation as the disease progresses. These findings underpin the theory that apoptosis plays a major role in the pathophysiology of MDS.

Apoptosis can be initiated by various cytokines, notably tumour necrosis factor (TNF)-α, Fas-ligand and TNF-related apoptosis-inducing ligand (TRAIL). Indeed, all these cytokines and other related ones are typically upregulated in the marrow in MDS, serving as negative regulators of haemopoiesis. Apoptosis can also be triggered by cytotoxic T cells and by signals from marrow stromal cells, probably via activation of similar pathways, and intrinsically following DNA damage and mitochondrial dysfunction that leads in turn to cytochrome c release and caspase activation. This balance of proapoptotic to antiapoptotic signals swings in favour of the latter as the MDS evolves towards AML, with upregulation of NF-κB, altered expression of adapter molecules such as Flice inhibitory protein (FLIP) and enhanced activity of members of the BCL-2 and IAP (inhibitors of apoptosis protein) families. This is borne out by a murine model of MDS/AML in which BCL-2 is conditionally over-expressed alongside an activating NRAS mutant gene, resulting in a disease entity reminiscent of MDS. In this model, using a stronger promoter to drive BCL-2 gene expression leads to more rapid AML progression, providing experimental insight into the process of leukaemogenesis.

Cytogenetic abnormalities in MDS

Recurrent cytogenetic abnormalities occur in over 50% of patients with MDS and in 90% of t-MDS (Table 25.4) and provide evidence of clonality. However, apart from the 5q– syndrome, cytogenetic abnormalities are not specific for particular clinicomorphological subtypes of MDS. Generally, the loss or gain of genetic material, through chromosomal deletions, unbalanced translocations and aneuploidy, is more characteristic of MDS than AML. Conversely, balanced translocations that are typical of de *novo* AML are rarely seen in MDS.

The common chromosomal abnormalities found in MDS include loss of Y, 5q– or monosomy 5, 7q– or monosomy 7, trisomy 8, 20q–, abnormalities of 11q23, and deletions of 17p, 12p, 13q and 11q, among others. None of these is specific for MDS as they can also occur in AML and myeloproliferative states.

The International Prognostic Scoring System (IPSS) classified clones as normal, single recurring, double recurring, complex, with at least three anomalies in the same cell, or miscellaneous. Twelve different cytogenetic categories are identified. Normal karyotype occurred in 60%, –Y in 2%, del(5q) 6%, del(20q) 2%, miscellaneous single 9%, trisomy 8 5%, double chromosome abnormalities 3%, miscellaneous double 3%, chromosome 7 abnormality 1%, miscellaneous complex 2% and complex in

Table 25.4 Cytogenetic abnormalities in MDS with approximate frequency.

Abnormality	Primary MDS (%)	Therapy-related MDS (%)
Complex karyotype	15–20	80–90
del(5q)/monosomy 5	15–20	30–40
del(7q)/monosomy 7	10–15	40–50
Trisomy 8	10–15	10–15
del(20q)	5–10	–
del(17p)	<5	–
del(13q)	<5	–
del(11q)	<5	–
del(12p)	<5	–

8%. These were refined into three risk-based cytogenetic subgroups: good 70% (normal, del (5q), del(20q) and –Y), poor 16% (complex or chromosome 7 abnormalities) or intermediate 14% (all other abnormalities) with median survival times of 3.8, 0.8 and 2.4 years, respectively. The complex group consisted of 66 patients of whom 63 had chromosome 5 and/or 7 abnormalities in addition to other abnormalities (abnormalities of chromosome 5 n = 33, chromosome 7 n = 7 and both 23). However, 14% of patients had cytogenetic abnormalities of unknown prognostic significance.

Analysis of a larger cohort of 2902 patients with primary untreated MDS or oligoblastic AML addressed the significance of less common, as well as combined abnormalities. The median number of metaphases analysed was 20. If a cytogenetic anomaly occurred in at least five patients it was a distinct subgroup. Double abnormalities (i.e. two distinct clonal MDS-related cytogenetic abnormalities within one cell) are classified into three subgroups: del(5q) with one additional clonal aberration, −7/del(7q) with an additional clonal karyotypic abnormality and any other combination of two abnormalities. Complex karyotypes are also further subdivided into those with three unrelated cytogenetic abnormalities and those with more than three karyotypic abnormalities. Independent clones where two subclones occur in parallel form a subgroup provided neither clone is complex. If either clone is complex it counts as a complex karyotype. The incidence of individual abnormalities was 0.4% for inv(3)/t(3q)/del(3q), 6.4% for del(5q), 1.6% for –7, 0.5% del (7q), 4.7% +8, 0.7% del(11q), 0.6% del(12p), 0.4% i(17q), 0.4% +19, 1.7% del(20q), 2.2% –Y. Rare single abnormalities (present in <10 patients) all had an incidence of 0.3% or less, accounted for 9% of all abnormalities and included der(1;7), +1q, –1/1p–, t(5q), +11, t(11q23), –13/13q–, del(16q), del(17p), +21, –21, –X, +Mar and other rare subtypes. Double abnormalities including del(5q) accounted for 1.6%, including –7/del(7q) 1.2% and any other double 3.4%. The complex subgroup with three abnormalities was present in 2.1% and >3 abnormalities 7%. Independent clones were detected in 0.9%. These abnormalities are arranged by impact on overall survival and predilection for transformation to AML into five groups: very good, which includes single del(11q) or –Y (2.9% of patients) with a median overall survival of 60.8 months and a hazard ratio (HR) for AML transformation of 0.47(0.3–0.7); good (65.9%) including normal karyotype, single del(5q), del(12p), del(20q), double including del(5q) and predicts a median OS of 48.5 months, HR 1.0 (0.8–1.3); intermediate (20.7% of patients) including single del(7q), +8, iso(17q), +19, +21 or any other individual clonal or double (any other double) and predicts a median overall survival of 25.0 months, HR 1.59 (1.4–1.9); poor (3.6%) consisting of single –7, der(3)(q21)/der(3)(q26), double including –7/7q– or complex with three abnormalities predicting a median overall survival of 15 months, HR 2.83 (2.2–3.7) and finally the very poor (7%) with complex >3 abnormalities predicting a median OS of 5.7 months, HR 4.37 (3.5–5.5). This classification has been incorporated into the revised IPSS and has a greater weighting in predicting prognosis than blast percentage.

SNP-A karyotyping

Techniques such as gene expression profiling, to identify expressed transcripts, and comparative genomic hybridization (CGH), to identify small regions of DNA gain or loss, are employed to understand the pathogenesis of MDS. High-resolution single nucleotide polymorphism (SNP) genotyping microarrays have been used to detect cytogenetically cryptic genomic aberrations (Figure 25.2). These SNP microarrays use chromosomal markers to identify regions of loss of heterozygosity, which occurs either through deletion or functionally via copy-neutral uniparental disomy (UPD), where the individual acquires a duplicated copy of an entire or part of a chromosome derived from one parent due to mitotic recombination, resulting in LOH without a change in copy number (CN). One study of 119 low-risk MDS patients revealed that UPD occurs in 46% of cases, while deletions and amplifications occur in 10% and 8% of cases, respectively. SNP-A is complementary to metaphase cytogenetics and cyto-FISH.

In studies of patients with myeloid neoplasms, 44% have abnormal metaphase cytogenetics, 49% have a normal karyotype and 7% are uninformative. The addition of SNP-A detects cytogenetic abnormality in 74–75% of patients, including cryptic lesions in those with normal or uninformative karyotypes, submicroscopic lesions and UPD involving chromosomes 1, 5, 7, 11, 17 and 21. A subgroup of patients with both normal karyotype and absence of SNP-A defects have a significantly better outcome (overall survival 16 versus 43 months, $p < 0.0001$) than those with defects detected by either MC or SNP-A. Furthermore, the presence and number of SNP-A lesions 0 versus 1 or 2 versus >2 are independent predictors of inferior OS and EFS in MDS. Areas of UPD suggest homozygous mutations in the region and UPD at chromosome 4q coincides with the *TET2* gene that is frequently mutated in MDS.

Deletion chromosome 5q

Interstitial deletion of the long arm of chromosome 5 del(5q) is the commonest cytogenetic abnormality in MDS, occurring in 10–15% of cases. The 5q– syndrome was originally described in 1974 by Van den Berghe and was the first entity within the group of refractory cytopenias to be associated with a consistent chromosomal aberration, namely the interstitial deletion of 5q. The syndrome is associated with macrocytic anaemia, thrombocytosis, unilobular megakaryocytes and a low propensity to develop into AML. It is more common in females than males. The WHO definition of 5q– syndrome requires the presence of an isolated del(5q), blasts less than 5% in the bone marrow and less than

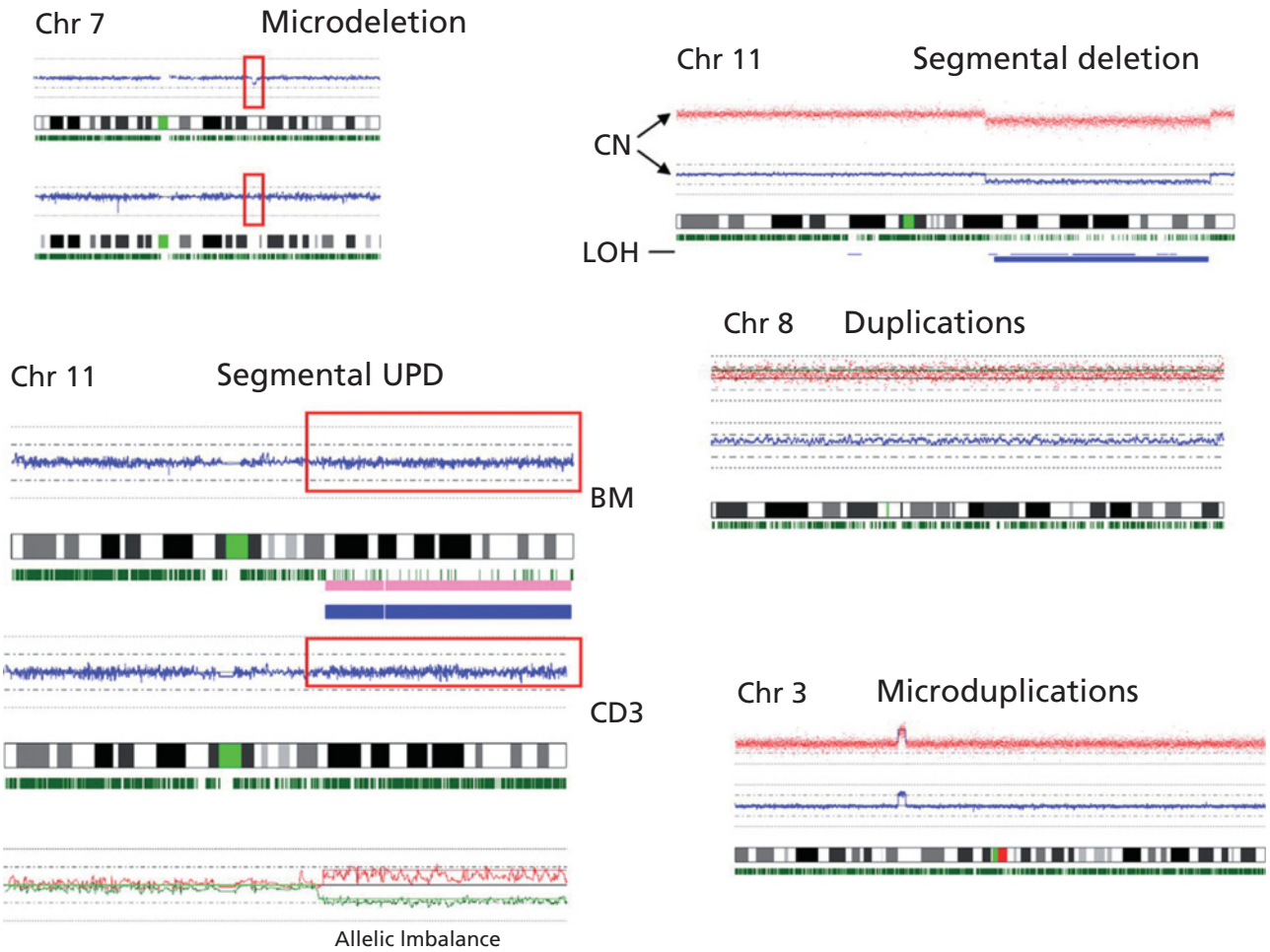


Figure 25.2 Representative examples of genomic mutations identified in MDS patients using single-nucleotide polymorphism analyses with high-density microarrays. BM denotes bone marrow cells and CD3 denotes T lymphocytes. CN, copy number; LOH, loss of heterozygosity; UPD, uniparental disomy. (Source: Maciejewski and Mufti, 2008. Reproduced with permission of the American Society of Hematology.)

1% in the peripheral blood, and absence of Auer rods. However, not all patients with 5q abnormalities have this syndrome, since del(5q), and indeed monosomy 5, are also found in more typical cases of MDS, frequently as part of a complex karyotype with other cytogenetic abnormalities.

The 5q– syndrome, with del(5q) as the sole cytogenetic abnormality, has provided an opportunity to define precisely the molecular defect(s) underlying the pathogenesis of this disease. The search for a minimal common deleted region on the long arm of chromosome 5 has spanned three decades and relied on physical mapping methods. Using FISH analysis, this region has been narrowed down to a 1.5-Mb interval located on 5q32 (the distal CDR) in low-risk diseases, whereas in higher-risk MDS and AML, a proximal CDR at 5q31.2 and 5q31.3 is identified. SNP-A karyotyping in 142 patients with del (5q) AML and MDS identified del (5q) in an additional six patients and UPD at del(5q) in four patients. The deleted regions CDR1 (8.5 Mb located between 5q32 and 5q33.2) was seen in all patients with del(5q) syndrome and the proximal and distal commonly retained regions (CRR1 and 2, respectively) are retained. CRR1 spans 81.7 Mb and ends at band 5q14.2, whereas the smaller CRR2 (1.7Mb) begins at 5q34. In other del(5q) MDS and 5q AML, a 1.92-Mb CDR is centred at 5q31.2 and 5q31.3. Patients with the typical SNP profile of del(5q) with distal CDR deletion and retention of CRR1 and 2 (i.e. a smaller deleted segment) have higher platelet counts and lower-risk disease, with a median survival of 32 months, even if they do not meet the diagnostic criteria of 5q– syndrome. Larger deletions that extend into the CRR are associated with a shorter survival of 14 months and a higher risk of transformation to AML. Cases that transform

to AML without deletions in CRR have other coexisting genetic mutations involving *NPM1* at 5q35.1 or *MAML1*. Chromosome 7 abnormalities and LOH at 17p spanning *TP53* are common in del(5q) MDS that has progressed to AML. Epigenetic inactivation of *CTNNA1* at 5q31 is a potential marker of disease progression.

The search for a tumour-suppressor gene located within this region has involved many different approaches over recent years and resulted in numerous candidate genes. However, it was not known whether 5q– syndrome required the loss of heterozygosity for a particular gene locus or could arise through haploinsufficiency of a particular gene alone.

The best candidate gene for 5q– syndrome was recently identified by way of a functional screening study that used RNA interference (RNAi) to systematically 'knock down' genes in the 5q32 region in normal CD34-positive cells. These cells were cultured in conditions promoting either erythroid or megakaryocytic differentiation. Partial RNAi knockdown of the ribosomal protein S14 gene (RPS14) caused a block in erythroid differentiation, but preservation of megakaryocytic differentiation that recapitulates the phenotype of 5q– syndrome. Moreover, RPS14 was able to rescue the 5q– phenotype by re-expression in CD34-positive cells from 5q– syndrome patients. If this result is confirmed, then there would appear to be a remarkable similarity between 5q– syndrome and the congenital bone marrow disorders of Diamond–Blackfan anaemia and Schwachman–Diamond syndrome that arise due to inherited abnormalities of related ribosomal protein genes, namely RPS19 and SBDS, respectively. While haploinsufficiency of RPS14 explains the clinical features of the 5q– phenotype, mechanisms that lead to growth advantage of the 5q– clone remain to be fully elucidated.

Chromosome 7 abnormalities

Complete loss of chromosome 7 (monosomy 7) or partial deletion of the long arm of chromosome 7 (del7q) is the second most common chromosomal abnormality, occurring in up to 9% of MDS cases overall. MC detects isolated monosomy 7/del(7q) in 27% of patients, with the rest having additional lesions, whereas SNP-A karyotype detects cryptic lesions in additional patients due to which only 15% are isolated monosomy 7/del(7q) and 85% with more complex lesions (including 12% with additional lesions only detected by SNP-A). Even in cases with balanced translocations involving chromosome 7, small deletions (1–5 Mb) are detected at the boundaries of the translocation by SNP-A. SNP-A karyotyping identifies cryptic lesions in 45% of isolated 7q lesions identified by MC. The SNP-A-defined LOH lesions at chromosome 7 are UPD7q, del(7q) and monosomy 7. The presence of SNP-A monosomy 7 correlates highly with hypoplastic MDS (77%), is also found in half the patients with hypoplastic MDS in whom MC fails, and predicts a poorer prognosis due to higher rates of transformation to AML (40% versus 7% without monosomy 7). All cases of chromosome 7 abnormalities in children with JMML and Fanconi's anaemia are monosomy 7. SNP-A monosomy 7 tends to be isolated, with few additional genomic lesions and the absence of LOH at 17p. In contrast, patients with del(7q) present with higher-risk disease (88% AML or Int-2/high-risk MDS) and have a higher number of associated genomic lesions per patient, the commonest of which is del(5q), found in 50% and often involving the centromeric or telomeric extremities of the long arm of 5q. LOH at 17p, which always spans *TP53*, occurs in 20%. Patients with del(7q)MDS and MDS/MPN overlap have a shorter OS and time to leukaemia progression than patients with UPD(7q) or monosomy 7. UPD(7q) occurs more frequently in CMML and has the lowest number of additional genomic lesions when compared to monosomy 7 or del(7q) and possibly predicts for a shorter median overall survival (460 versus 730 days) and a higher rate of transformation to AML (26% versus 13%). The AML cases with UPD(7q) often have antecedent CMML. Clinically, monosomy 7 MDS is characterized by a lower median age of affected patients compared with deletions of 5q, severe refractory cytopenias and tendency to life-threatening infections. It can present as a total or partial monosomy and may be an isolated abnormality or part of a complex karyotype. Monosomy 7 confers a poor prognosis that ranges from 14 months as an isolated abnormality down to 7 months if the karyotype is complex and involves other abnormalities. At least two distinct regions of common deletions have been identified by metaphase cytogenetics: the band 7q22 and the more telomeric regions of 7q31–32 and 7q36. SNP-A karyotyping delineates three CDRs at 7q22, 7q34 and between 7q35 and 7q36.1. The CDR between bands 7q35 and 7q36.1 contains six genes, including a frame shift mutation in exon 19 position Ile 715 of *EZH2*. Sequencing of genes within the CDRs identifies *EZH2* mutations as the only recurring mutation. The relatively better prognosis of patients with large deletions and their occurrence in children suggests these may be founding lesions, whereas the smaller partial deletions of 7q represent a secondary event due to genomic instability.

Another recurrent cytogenetic abnormality in MDS that is associated with characteristic morphological features is isolated del(20q), which typically involves erythroid and megakaryocytic lineages. Deletions of 20q generally carry a favourable prognosis and the pathogenetic gene(s), within the commonly deleted region 20q12, are unknown.

Molecular basis of MDS

Recent discoveries have provided important insights into the pathogenesis of the disease. MDS is a preleukaemic disorder characterized by impaired cellular differentiation that has the potential to transform to AML. Conversely, the myeloproliferative syndromes, including the myelodysplastic–myeloproliferative overlap, are characterized by enhanced

survival and proliferation that has the potential to transform to AML if coupled to a block in differentiation. Thus, conceptually at least, both MDS and myeloproliferative disorders represent different routes to acquiring the multiple mutations necessary for development of AML.

Genetic abnormalities in MDS: candidate genes and whole-genome sequencing

Application of whole-genome sequencing technologies to various morphological subgroups of MDS has radically changed our understanding of the molecular abnormalities that underlie dysplastic haemopoiesis, as well as the progression to acute myeloid leukaemia (Table 25.5). Identification of somatic mutations of *TET2* in patients with copy-neutral loss of heterozygosity of chromosome 4q24 was followed by a large number of studies of the mutational landscape in MDS that have consistently confirmed the presence of specific gene mutations in up to 90% of patients, paving the way for a future molecular classification. The most important of the mutations are discussed below.

Spliceosome mutations

The spliceosome is a complex protein structure responsible for excising introns from primary mRNA, and aligning exons to form the mature mRNA that is transcribed. The spliceosome complex is made up of the U1, U2 and U4/U5/U6 small nuclear ribonucleoproteins (snRNP) or the U11/12 snRNPs. Splicing factor 3B (SF3B) along with splicing factor 3A and a 12s RNA unit form the U2 snRNP. Serine/arginine-rich splicing factor 2 (SRSF2) is a pre-mRNA splicing factor that is important for binding to RNA. *ZRSR2* (zinc finger RNA binding motif and serine/arginine rich) is the gene that encodes U2 small nuclear ribonucleoprotein auxiliary factor 35 kDa subunit-related protein, which is also important in interactions in spliceosome assembly. As splicing proceeds, the formation of the active spliceosome involves an ordered, step-wise assembly of discrete particles on the pre-mRNA substrate. The initiating step is the formation of the early complex (E), which is formed by the binding of the U1 snRNP to the GU sequence at the 5′ splice site. Simultaneously, the proteins involved at the 3′ splice site recognition also begin to bind: SF1 binds to the branch point (BP) and the serine/arginine-rich proteins SRSF2 and ZRSR2 bind to the exon splicing enhancer (ESE) site of the next exon to aid the binding and stability of U2 proteins U2AF1 and U2AF2 (Figure 25.3).

Somatic mutations of the genes encoding the spliceosome proteins that recognise the 3′ splice site and U2snRNP function occur in over half the patients with MDS, and appear to be acquired in the elderly as they are rare in childhood myeloid neoplasms. The mutations target proteins involved in the initiation of splicing and are restricted to the components of the E/A splicing complex, including SF3B1, SRSF2, ZRSR2 and U2AF1. Most mutations are missense and affect conserved positions in *SF3B1*, *SRSF2*, *U2AF1* and other spliceosomal proteins (Table 25.5 and Figure 25.4).

The *SF3B1* gene encodes the 155 kDa subunit of the SF3B complex which targets U2snRNP complex to the branch point during complex A formation. Mutations of SF3B1 are found in over 90% of patients with ringed sideroblasts, but in only 6% of patients with MDS without ringed sideroblasts. The mutations are heterozygous missense mutations due to A to G transversion, leading to substitution of lysine at residue 700 by glutamic acid (K700E) or C to A or C to G substitution at position 662, where histidine is replaced by glutamine (H662Q). The downstream effects of these mutations are that exons in up to 350 genes are differentially used in mutants when compared to wild-type *SF3B1*. Amongst these are *RUNX1*, *CBL1, EZH* and *ASXL1*. Structural differences in iron distribution also occur as a consequence of *SF3B1* mutation in RARS, due to downregulation of pathways affecting mitochondrial ribosomes and electron transport pathways. *SF3B1* also interacts with the polycomb group of proteins and represses *HOX* genes.

In patients with RARS *SF3B1* is mutated in 90% and is always associated with a good prognosis with reduced rate of transformation to AML. Phenotypically patients with mutated *SF3B1* have higher neutrophil and platelet counts, lower blast percentage in the marrow and have ringed sideroblasts.

Mutations of *SF3B1* with *JAK2*V617F mutation has been demonstrated in up to 50% of patients with the overlap condition of RARS with thrombocytosis (denoted RARS-T).This mutation, which is more commonly associated with the myeloproliferative disorders, causes constitutive activation of the JAK2 protein and downstream signalling. Similarly, mutations of the MPL gene, which are described in essential thrombocythaemia, have also been identified in patients with RARS-T.

SRSF2 mutations are the second most frequently mutated splicing gene. These are heterozygous, missense mutations specifically at P95H (proline at position 95 which may be replaced by histidine, leucine, arginine, alanine or threonine). Mutant *SRSF2* is most prevalent in CMML (28%) with a considerably lower frequency in RARS (1.5%), MDS without ringed sideroblasts (11.6%) and AML (7%). In MDS and CMML mutant *SRSF2* correlates with older age and is associated with shorter time to progression to AML and lower overall survival than wild-type *SRSF2*. Mutations of *RUNX1*, *IDH1* and *TET2* occur concomitantly with *SRSF2* mutations. The occurrence of *SRSF2* and *RUNX*1 mutation are associated with an inferior survival. *EZH2* and *TP53* mutations are, on the other hand, exclusive to SRSF2 mutations. The SRSF2 mutations are stable during the disease course of MDS and are not acquired during disease progression, which is due to acquisition of mutations in other genes.

ZRSR2 is a member of the U2AF1-related family that interacts with the U2AF 65-kDa subunit in spliceosome assembly. *ZRSR2* mutations do not appear at 'hot spots', but are dispersed nonsense or frame-shift mutations leading to a loss of function. They

Table 25.5 Heterozygous spliceosomal mutations in MDS.

Gene/Chromosomal location	Mutation frequency/timing	Phenotype	Mutational hotspots	Function	Downstream genes	Prognosis
SF3B1/2q33.1	15–28% Founding mutation	RARS, RCMD-RS Strictly associated with ringed sideroblasts	K700, K666, K662, K622	3′ splice site recognition	*DNMT3A, ASXL1 CBL, EZH2, RUNX3*	Good OS and low risk of leukaemia evolution
SRSF2 (*SC35*) 17q25.1	10–15% Founding mutation	RCMD or RAEB associated with TET2 in CMML	P95 (missense or deletion)	3′ splice site recognition	*TET2, RUNX1*	Poor OS and high risk of leukaemia evolution
U2AF35 (*U2AF1*) 21q22.3	5–16% Founding mutation	RCMD or RAEB MDS CMML	S34, Q157 (missense, N and C terminal domains respectively)	3′ splice site recognition	*ASXL1* or *TET2*	High risk of leukaemia evolution
ZRSR2 Xp22.1	3–11% Founding mutation	MDS CMML	Nonsense, frame shift	3′ splice site recognition	*TET2*	
SF3A1 22q12.2	Rare	Uncertain	None	3′ splice site recognition	–	
U2AF65 (*U2AF2*) 19q13.42	Rare	Uncertain	None	3′ splice site recognition	None	
SF1 11q13.1	Rare	Uncertain	None	3′ splice site recognition	None	
PRPF-40B 12q13.12	Rare	Uncertain	None	3′ splice site recognition (speculative)	None	
PRPF8 17p13.3	Rare	Uncertain	None	Aligning 5′ and 3′ exons		
LUC7L2 7q34	Rare	Uncertain	None	Recognition of non-consensus splice sites		

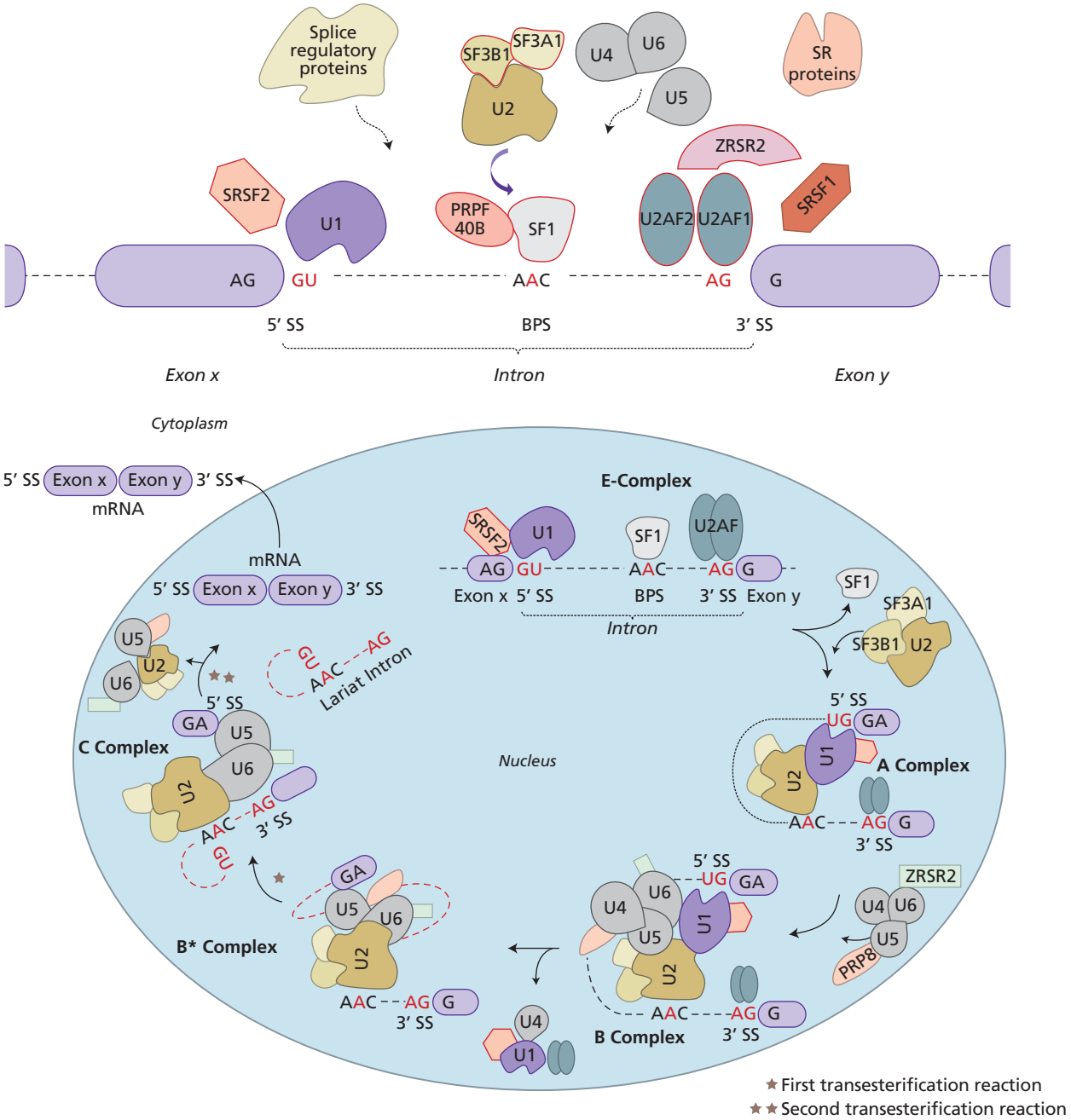


Figure 25.3 mRNA splicing pathway. Schema showing splicing of mRNA leading to the ligation of flanking exons by the spliceosome proteins. The GU dinucleotide at the 5′ splice-site, branch point site and AG dinucleotide at the 3′ splice site are specific recognition sites for splicing (B) The spliceosome components assemble sequentially to form five different complexes E, A, B, B* and C. The components include U1, U2, U4/U5/U6 snRNPs. The rectangles and broken lines represent exonic and intronic regions, respectively. Gene names with red boundaries indicate mutated proteins seen in MDS. (Source: Austin *et al.*, 2013 [*BJH* 2013 **162**(5): 587–605.] Reproduced with permission of John Wiley & Sons.)

occur with similar frequency in MDS and AML, cluster with RAEB-1 and RAEB-2 subtypes of MDS, show severe thrombocytopenia and are associated with higher transformation to AML and poor overall survival.

The *U2AF1* gene encodes the 35-kDa subunit of U2AF, which recognizes the AG dinucleotide that marks the end of the intron. U2AF1 is critical for splice site recognition in introns, and U2AF1-related proteins ZRSR1, ZRSR2 and U2AF1L4 can

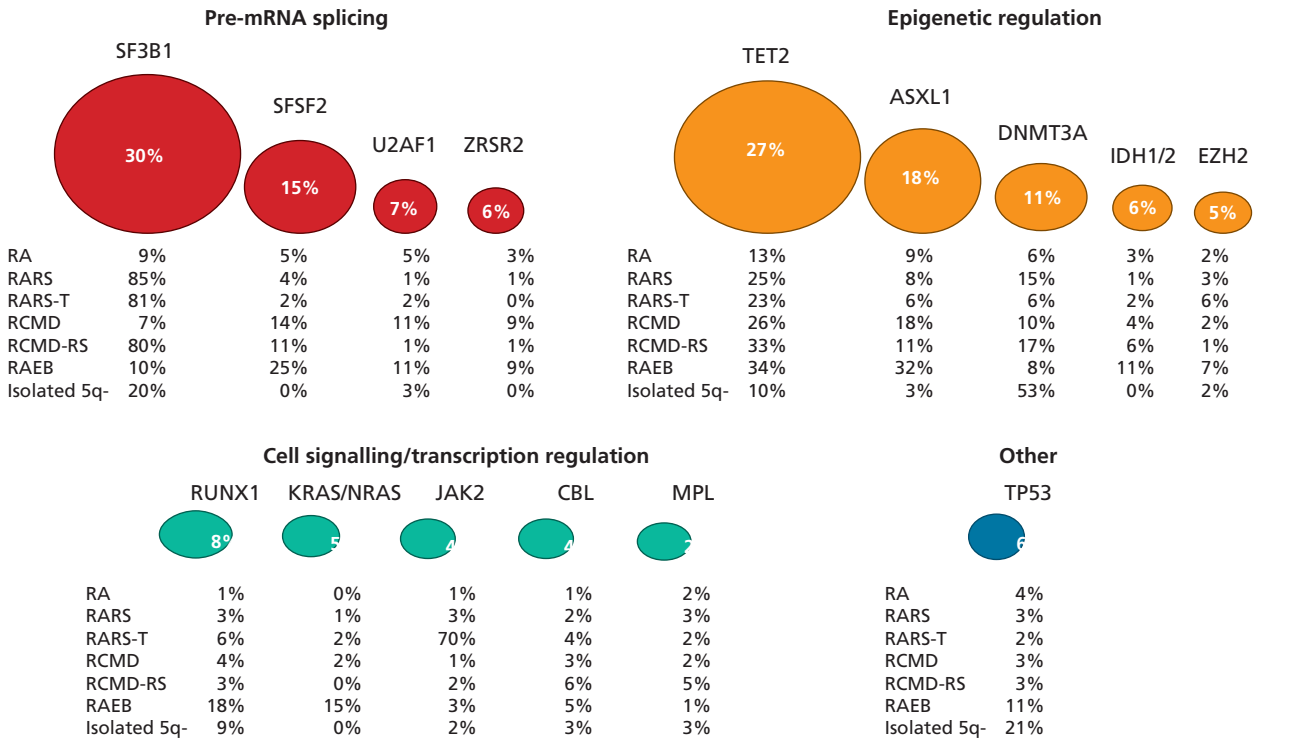


Figure 25.4 Mutational landscape of MDS. Frequency of the commonly mutated genes identified in the sequencing screening in a cohort of 1677 MDS cases. The size of the circle represents proportional frequency of the specific gene mutations in MDS. Each column under the circle represents proportion of gene mutations in each MDS subtype. (Source: Dr Syed Mian, King's College London. Reproduced with permission. The data is obtained from three published studies: Haferlach *et al.*, 2014 *Leukaemia*; Papaemmanuil *et al.*, 2014 *Blood*; Mian *et al.*, 2013 *Haematologica*.)

substitute for U2AF1 in some gene subsets. Mutations of *U2AF1* occur in codons 34 and Q157 with equal frequency in MDS without ringed sideroblasts (11.6%), CMML (23%) and AML (11%), but are not seen in RARS. These mutations associate with shorter time to progression to AML but the impact on overall survival is unclear. IU2AF1 mutations result in abnormal splicing of *RUNX1* and *TET2* and mutations of *ASXL1* and *TET2*.

Deep sequencing of RNA (RNA-seq) from MDS patients harbouring spliceosome mutations show that the downstream effects occur in genes, some of which have already been implicated in the pathogenesis of MDS. For example, splicing abnormalities of *RUNX1* occur in cases with *SRSF2* mutations and splicing abnormalities of *ASXL1*, *CBL*, *EZH1* and *RUNX3* occur in two cases with *SF3B1* mutations. Similarly, gene expression profiles of patients with AML and *U2AF1* mutations show downregulation of splicing and RNA recognition motif (RRM) genes. Spliceosome mutations also show a tendency to aggregate, with genes implicated in aberrant DNA methylation. Crosstalk between spliceosome mutations and specific epigenetic modifiers has also been suggested. Thus *SRSF2* and *ZRSR2* mutations are common in patients with mutations of DNA methylation modifier *TET2*, *SF3B1* mutations coexist with mutations of the methyltransferase *DNMT3A* and *U2AF1* mutations with *ASXL1* or *TET2* mutations. Additional roles for these genes may also be important, for example *SRSF2* is important in the regulation of DNA stability and its depletion results in genomic instability which may account for the coexistence of *SRSF2* with additional mutations and the worse prognosis in patients with this mutation.

Epigenetic abnormalities

Epigenetics refers to various molecular modifications of chromatin that, without altering the DNA sequence, play a critical role in genomic regulation and control of gene expression. There are two important epigenetic modifications relevant to MDS, namely DNA methylation and histone modification. DNA methylation refers to the addition of a methyl group to cytosine, which can occur wherever this is followed by a guanine within a CpG dinucleotide pair. Such CpG pairs are underrepresented in the human genome, but cluster together within so-called CpG islands that tend to be located in the proximity of gene promoter regions. In normal cells, these CpG islands are typically unmethylated, allowing genes to be transcribed. However, if CpG islands are methylated, then transcriptional

activity at the promoter is impeded and the gene is silenced. Thus, aberrant promoter methylation leads to inactivation of the gene, thereby providing a mechanism where genes can be functionally inactive.

The other major epigenetic control mechanism that works in conjunction with DNA methylation is histone modification. Histones form the chromatin scaffold and closely regulate whether the DNA exists locally in a repressed or permissive state. Biochemical alterations to the tails of the histone molecules influence the degree of compaction of the nucleosomes and hence the level of transcriptional activity of nearby genes. Unlike DNA methylation, which is largely irreversible, histone modifications are more dynamic, particularly acetylation/deacetylation of lysine residues of histones H3 and H4. Moreover, there is a close and cooperative interplay between these two epigenetic control mechanisms that together can render a gene permanently silenced. The significance of this is that combination epigenetic therapies that comprise a hypomethylating agent with a histone deacetylase inhibitor may be more effective than single agents in re-expressing silenced tumour-suppressor genes. Genes that control DNA methylation (*TET2*, *DNMT3a* and *IDH1/2*) and chromatin modification (*ASXL1* and *EZH2*) are frequently mutated in MDS and may be responsible for the disordered differentiation manifested.

The TET2 (ten-eleven-translocation) gene is one of the three-member TET family of genes, has two highly conserved regions implicated in the conversion of 5-methylcytosine (5mC) to 5-hydroxymethylcytosine (5hMC). It is located within a region of UPD and microdeletion at chromosomal position 4q24 and is mutated in patients with MDS (15–27%), in CMML (44%), secondary AML (24%) and myeloproliferative disorders (12%). Recurrent somatic mutations in *TET2* can occur in the elderly with clonal haemopoiesis, but not manifest clinically, suggesting that this is an early mutation that can cause haemopoietic stem cell expansion. The mutations are often nucleotide substitutions resulting in missense, frame shift, stop codons, in-frame deletions or amino acid substitutions in extremely conserved residues. A single mutation or two different mutations may occur in the same patient, resulting in heterozygous or homozygous loss of copies of the *TET2* gene. TET2 shares sequence homology with TET1, a fusion partner of the MLL oncogene in AML that catalyses the conversion of 5-methylcytosine to 5-hydoxymethylcytosine. TET2 lacks DNA binding domain and maintains hypomethylation by utilizing Fe^{2+} and 2-oxyglutarate, a product of the Krebs cycle, to add a hydroxyl group to 5-methylcytosine, converting it to 5hMC, which is then converted to cytosine by base excision repair. Levels of 5hMC are consistently reduced in patients with *TET2* mutations. 5hMC is preferentially distributed at transcriptional start sites and along highly expressed exons, and correlates with enhanced gene expression. In MDS, *TET2* mutations are often found in patients with a normal karyotype and older age, and may be a marker of clonality. *TET2* mutations commonly coexist with *SRSF2* mutations in CMML. These mutations occur early in the pathogenesis of MDS, appear not to be predictors of transformation to AML, but may predict positive responses to therapy with hypomethylating agents.

IDH1/IDH2 (isocitrate dehydrogenases 1 and 2) are isoforms of the critical NADP-dependent enzyme in the Krebs cycle catalysing the decarboxylation of isocitrate to α-ketoglutarate (see Chapter 18). Mutations of these enzymes lead to depletion of α-ketoglutarate (α-KG), which is necessary for TET2 and the jumonji family of histone demethylases, and TET inhibition, resulting in widespread promoter hypermethylation. Additionally, the ensuing accumulation of 2-hydroxyglutarate (2-HG) inhibits other α-ketoglutarate-dependent enzymes. Mutations occur most frequently at arginine 132 in *IDH1* and arginine 140 and 172 in *IDH2*, disrupting the isocitrate binding site and hindering the synthesis of α-KG, which is reduced instead to 2-HG. *IDH1*, *IDH2* and *TET2* mutations are mutually exclusive; however, the epigenetic profiles consequent to these mutations are similar to reduced 5hMC, which may lead to impaired HSC differentiation. *IDH1* mutations occur in 4–12% of MDS and predict a poor outcome.

DNMT3A (DNA methyltransferase 3A) is the enzyme mediating *de novo* methylation of CpGs. It is somatically mutated in *de novo* AML and MDS with mutations occurring throughout the coding regions as missense or nonsense heterozygous mutations, but over 50% occur at arginine 882 (R882), which disrupts the methyltransferase domain, DNA binding and reduces enzyme activity (methylation). The frequency of mutation in MDS is lower (10–15%) and does not segregate with any particular cytogenetic subgroup. These mutations occur in older individuals and in RCMD or RAEB. *DNMT3A* mutations are associated with rapid progression to AML, although in RARS the effects of *SF3B1* mutation are dominant and patients are less likely to progress.

Post-translational modifications of histones play an important part in epigenetic regulation of gene expression. Fundamental to mediating these modifications are the polycomb repressor complex proteins 1 and 2 (PRC1 and PRC2) that are known to silence HOX genes through modulation of chromatin. They mediate transcriptional repression by trimethylation of histone 3 at lysine 27 (H3K27me3) and ubiquitination of histone 2A at lysine 119, both of which are markers of repression. PRC2 includes *EZH2* (enhancer of zeste homolog 2), *EED* (embryonic ectoderm development), *SUZ12* (suppressor of zeste 12) and *RBBP4* (retinoblastoma binding protein 4). PRC2 has histone methyltransferase activity and is responsible for initial targeting of genomic regions to be silenced. PRC1 has an E3 ubiquitin ligase activity that monoubiquitinylates H2A on lysine 119 (H2AK119ub1) and is responsible for stabilizing PRC2-mediated silencing and long-term cellular memory.

ASXL1 (additional sex combs-like 1) is part of a family of genes that are enhancers of trithorax and polycomb proteins,

which are involved in the maintenance of activation and the silencing of development-related genes by chromatin remodelling. The *ASXL1* gene is located at 20q11.1. Mutations occur frequently in exon 12, with a common somatic frame shift mutation (Gly646TryfsX12) occurring in 50%, leading to truncation of the protein and loss of the C-terminal PHD domain. Gene-expression profiling reveals deregulation of the retinoic acid receptor activation pathways in patients with mutant *ASXL1*. With loss of the PHD domain, H3K27me3 levels are reduced at the HOXA locus, whereas H2AK119 ubiquitylation is not changed. Additionally, the direct interaction between ASXL1 and PRC2 is disrupted, compromising PRC2 function. Mutations of ASXL1 have been shown to promote myeloid transformation by inhibiting the PRC2-mediated gene repression. *ASXL1* mutations are common in both MDS (16.2%), *de novo* AML (6.5%), secondary AML (30%), CMML (45%) and primary myelofibrosis (34%), and are associated with a poor clinical outcome in all myeloid neoplasms.

EZH2 is a key component of the PRC2 complex and induces H3K27 di- or trimethylation, via a conserved SET domain, with resultant compaction of chromatin, and inhibits transcription. It can also recruit DNA methyltransferase, induce DNA methylation and gene-promoter silencing, thus further repressing the gene. Somatic *EZH2* mutations in MDS can be truncating mutations, which occur throughout the gene, whereas missense, frame shift and loss of function mutations cluster at the carboxy terminal of the SET (Su(var) 3-9 and enhancer of zeste) cysteine-rich domain, CXC and D2 domains and occur in about 6–8% of patients. Disruption of the SET domain leads to global reduction in H3K27 me2/me3 (H3K27 di- and trimethylation). Inactivating *EZH2* mutations, particularly homozygous mutations in low-risk MDS, result in absent H3K27me3 and are an independent indicator of poorer outcomes. Patients with MDS or MDS/MPN with del(7) or monosomy 7, or UPD of 7q have an increased frequency of *EZH2* mutations (up to 71%). Downstream loss of function of *EZH2* de-represses genes such as *HOXA* and *HOXB,* leading to proliferation and stem cell expansion. Mutations of other components of PRC2 complex SUZ12 and EED occur at a low frequency, with reduction of methyltransferase activity.

Integration of these different epigenetic mutations in MDS suggest that they act to modulate the expression of clusters of genes that regulate haemopoietic stem cell renewal, proliferation and differentiation. The epigenetic and spliceosome mutations appear to cooperate. It would appear that spliceosome mutations that are mutually exclusive of each other affect the expression of particular sets of downstream genes and hence impair differentiation of cells generating morphological dysplasia.

Somatic mutations in other pathways

Somatic mutations of the RUNX1 (AML1) gene have been recognized to occur in 7–8% of all patients with MDS, particularly where it is treatment-related or radiation-induced. In contrast to the mutations of *RUNX1* found in M0 AML that are often biallelic, those found in MDS are generally monoallelic and result in haploinsufficiency rather than a dominant-negative protein. Moreover, while the mutations in AML tend to occur in the Runt domain and affect DNA binding, mutations in MDS occur more frequently in the C-terminus of the protein, causing truncation and loss of its transactivating domain. Interestingly, individuals who inherit one abnormal copy of the RUNX1 gene commonly exhibit congenital thrombocytopenia, with a propensity to develop AML, suggesting the gene might be acting as a tumour suppressor in this setting. *RUNX1* mutations are associated with advanced disease, severe thrombocytopenia and poor clinical outcome.

The *RUNX1* gene is also involved in one of the few examples of a recurrent MDS-associated chromosomal lesion, namely the t(3;21) translocation found in some cases of MDS (and in blast crisis in chronic myeloid leukaemia). This translocation rearranges the RUNX11 and EVI1 (also known as MDS1) genes, fusing the N-terminus of AML1 with a small portion of MDS1 and nearly all of EVI-1 and results in a fusion protein that appears to de-regulate normal haemopoiesis. EVI-1 is a zinc-finger protein that when over-expressed can block erythroid differentiation and promote transformation. Translocations that involve only chromosome 3, namely t(3;3) and inv(3), both cause overexpression of EVI1 and are associated with MDS and AML with abnormal megakaryocytes.

TP53 gene is located at chromosome 17p13.1 and encodes for p53, a coordinator of transcription programs and a key cell-cycle control protein at the G1/S checkpoint, which can trigger apoptosis, and is a well-known tumour-suppressor gene. *TP53* mutations occur in 5–10% of patients with MDS, particularly in patients with advanced disease, therapy-related MDS in association with deletions of chromosome 5 and 7, and complex karyotypes that include abnormalities of chromosome 17. However, 30–50% of patients with complex cytogenetics have *TP53* mutations. Patients with *TP53* mutations have an unfavourable clinical outcome and a high risk of leukaemic evolution. *TP53* mutations are particularly associated with del(5q) MDS and may occur at an early stage of the disease and in copy-number alterations of chromosome 5, as 90% of *TP53*-mutated patients have either isolated 5q– or –5/5q–. Their presence predicts for poorer response rates to lenalidomide and an increased likelihood of progression to AML. *TP53* mutations often have fewer mutations in other genes implicated in myeloid disease pathogenesis. Phenotypically, del(17p) is associated with the presence of pseudo-Pelger–Huët cells containing small vacuoles, deletion of *TP53* and a high risk of leukemic transformation. In del(5q), haploinsufficency of *RPS14* disrupts ribosome assembly, accelerates MDM2 degradation and stabilizes *TP53*.

Activating mutations of the Ras superfamily that encode small GTP-binding proteins are involved in intracellular

signal transduction. Mutations in MDS affect *N*RAS (neuroblastoma RAS), *KRAS* (Kirstin RAS), *NF1* (neurofibromatosis 1), *PTPN11* (protein tyrosine phosphatase, non-receptor type 11) and *CBL* (Casitas B-lineage lymphoma). In adults these usually involve NRAS, are found in up to 20% of cases of MDS, especially CMML, and are often associated with RUNX1 point mutations. In juvenile CMML somatic or germline mutations of RAS pathway genes occur in 90% of patients and secondary mutations of *SETBP1* and *JAK3* are implicated in disease progression.

SETBP1 (SET-binding protein 1) encodes a protein that binds the SET nuclear oncogene that is involved in DNA replication. The gene on 18q12.3 is subject to germline mutations causing the Schinzel–Gideon midface retraction syndrome, with skull abnormalities, hydronephrosis and developmental delay. In atypical CML, somatic mutations occur in 25–30% of patients, whereas the frequency is lower, approximating 2% in MDS and 2.5% in secondary AML. The mutations are consistently in the region of codons 858 to 871, with hot spots at D868 and G870, and result in a gain of function as *SETBP1* over-expression promotes self-renewal of myeloid progenitors. Patients with these mutations tend to have a higher white cell count, may have a complex karyotype or monosomy 7 and a worse prognosis compared to wild-type. Other genes occasionally found to harbour mutations in MDS include those encoding various cell-surface receptors with kinase activity, such as FMS (now called CSF1R), KIT, FLT3, PDGFRB and GCSFR, and those encoding transcription factors, such as GATA1, PU.1 (SPI1), CEBPA, MLL and TP53. Mutations of the ATRX gene located on the X chromosome are associated with the rare phenomenon of acquired HbH disease, also termed acquired α-thalassaemia in MDS, in which the dysplastic red cells are markedly microcytic and hypochromic, with HbH inclusions.

The CBL gene at chromosomal position 11q23.3 is another candidate tumour-suppressor gene recently implicated in MDS. It was identified through demonstration of a recurrent UPD at 11q in 12 of 301 patients (4%) with various myeloid malignancies, including MDS/MPN, CMML, CML blast crisis and JMML. Of these, seven patients were found to carry mutations of the c-CBL gene, which functions as an E3 ubiquitin ligase involved with ubiquitination and degradation of active protein tyrosine kinases. The oncogenic mutations occur in a highly conserved α-helix linker joining SH2 tyrosine kinase receptor binding domain with a RING domain. Three unique mutations within or directly adjacent to the RING finger domain result in substitution of cysteine at codon 384 with tyrosine (C384Y) or C404Y or the substitution of arginine at codon 420 with either glutamine or proline (R420Q/P). These mutations are not associated with del(11q). Patients have monoblasts or monocyte proliferation, splenomegaly, surface expression of CKIT on malignant cells, increased reticulin fibrosis and transformation to AML. Megakaryocytes in CMML show hyperchromatic and abnormally lobated raisinoid nuclei, and express aberrant pSTAT5. Mutations in this family of genes are predictive of an adverse prognosis, with a median survival of 5 months.

Nonsense mutations of *CSF3R*, which encodes the receptor for colony-stimulating factor 3, result in a truncated CSF3R protein that is responsible for the development of MDS/AML in severe congenital neutropenia. Mutations of the cohesion complex (a highly conserved four-subunit ring structure that encircles sister chromatids during metaphase) that is critical in enabling the cohesion of sister chromatids and transcriptional regulation and post-replicative DNA repair are found in MDS. The genes *STAG2*, *RAD21*, *SMC1A* and *SMC3* are involved in the cohesion complex and mutations of STAG2 are found in 6% of MDS patients, whereas mutations of all four genes are found in 8% of MDS, 10% of CMML and 12% of patients with AML.

BCOR (BCL 6 corepressor) is located on Xp11.4 and encodes a corepressor of BCL6. Inactivating mutations of BCOR have recently been described in AML with a normal karyotype and in 4.2% of MDS patients. Germline mutations of *BCOR* cause the oculo-facio-cardio dental syndrome with microphthalmia, dysmorphic appearance and dental abnormalities. The somatic mutations occur throughout the gene and result in absent or low expression of truncated BCOR protein. The mutations are mutually exclusive to *NPM1,* but segregate with *DNMT3A* and *RUNX1* mutations in 40–50% of cases. In patients with CMML, the frequency of mutation is 7.4% and concomitant mutations occur most frequently with *U2AF1*.

GATA2 (GATA binding protein 2) is a crucial transcription factor for haemopoietic differentiation and lymphatic formation. It is important in maintaining the proliferation and survival of early HSC and preferential differentiation to erythroid or megakaryocyte lineage. The gene has an N-terminal transactivating domain, with two highly conserved zinc-finger domains. Two classes of mutations, namely N-terminal frame-shift mutations with premature termination, leading to a non-functional protein lacking most of the carboxy terminal, and missense mutations in the C-terminal zinc-finger domain, which disrupts DNA and other transcription factor binding. The most common of these is Thr354met, but Thr358Lys and Leu359Val have also been described. Large frame-shift mutations are associated with lymphoedema and may occur at R398W, R398Q or T354M (Emberger syndrome; autosomal dominant). Germline mutations are responsible for familial MDS/AML, which presents at a younger age and often more than one first-degree relative has AML-MDS with an autosomal dominant inheritance. They can also cause the MonoMAC syndrome with congenital infectious disease, persistent and profound peripheral monocytopenia, B- and NK-cell lymphocytopenia and the near absence of dendritic cells, with an increased susceptibility to mycobacterial and papilloma viral infections. Serum FLT 3 levels may be elevated due to stress haemopoiesis and may help in defining disease progression. Clinically, problems due to cytopenia may predate MDS.

Diagnosis

Clinical features

Approximately 20% of cases of MDS are detected incidentally in patients who have a routine blood count taken for unrelated reasons that reveals unexpected cytopenia or dysplasia. The majority of the remainder present with symptoms and signs of bone marrow failure, notably fatigue due to anaemia in up to 80% and infections or bleeding in up to 20%. Among the most common infections are bacterial pneumonias and skin abscesses, occurring particularly in patients with a neutrophil count of less than 1×10^9/L. Features of lymphadenopathy, splenomegaly and hepatomegaly are rarely found. There is a recognized association between MDS and several rare disorders that seem to have an immunological basis, including neutrophilic dermatosis (Sweet syndrome), pyoderma gangrenosum and cutaneous vasculitis.

Blood count

Macrocytic anaemia is the predominant finding in most patients at presentation, occurring as pancytopenia in 30–50% or in combination with neutropenia or thrombocytopenia in 20–30%. Isolated neutropenia or thrombocytopenia is rarer, accounting for 5–10% of presentations. Occasionally, the blood count is normal and the diagnosis is suggested by abnormal parameters generated by automated cell counters that reflect aberrant morphology, such as the red cell distribution width, which can itself become a useful marker of MDS.

Peripheral blood morphology

Morphological abnormalities in the peripheral blood, while generally non-specific for MDS, can nonetheless be very helpful in arriving at a diagnosis. At least 200 cells in the peripheral blood require evaluation and 10% of any lineage must be found to be dysplastic to count as evidence of dysplasia. There is commonly marked anisocytosis/poikilocytosis and the red cells tend to be macrocytic and oval-shaped. In sideroblastic anaemia, the blood film is classically dimorphic, containing a minority population of hypochromic microcytic cells; Pappenheimer bodies, which can be confirmed with an iron stain, and basophilic stippling may also be seen. Microcytosis is present in the rare variant of acquired HbH disease. Some patients have occasional circulating erythroblasts in the peripheral blood that are often dysplastic or megaloblastic.

Neutropenia is common and neutrophils often exhibit reduced granulation and the acquired Pelger–Huët anomaly (Chapter 19). Hypogranular neutrophils arise due to defective formation of secondary granules, with the agranular ones being highly specific for MDS. The pseudo-Pelger–Huët neutrophil is one that exhibits dense clumping of the chromatin and hypolobulation of the nucleus that is classically bilobed (resembling a pair of spectacles) or even non-lobed (resembling a dumbbell). This acquired abnormality, which resembles the inherited Pelger–Huët anomaly, is sufficiently characteristic of MDS to be almost pathognomonic.

Monocytosis is present, by definition, in CMML and monocytes can often be morphologically abnormal. While often reduced in MDS, basophils and eosinophils might also be raised in the proliferative overlap syndromes. Circulating blasts may be found in all categories of MDS, but if present in significant numbers are more usually indicative of RAEB. Finally, the platelet count is often reduced and platelets may show dysplasia, such as hypogranulated and giant forms. In CMML, circulating monoblasts (large oval nuclei, open chromatin, prominent nucleoli, variable basophilic and some azurophilic granules) and promonocytes (similar to monoblasts, but with indented or convoluted nuclei) may be present.

Bone marrow morphology (see also Chapter 19)

The ability to diagnosis MDS according to the presence of dysplastic morphology is critically dependent on optimal staining of the marrow slides with a Romanowsky stain (Figure 25.5). Ideally, only a small volume of marrow should be aspirated to avoid excess dilution with peripheral blood. There is often considerable interobserver variation, particularly where the dyserythropoiesis or granulocytic hypogranularity is subtle, but better consistency among observers in identifying cells, such as ring sideroblasts (five or more siderotic granules that encircle over a third of the nuclear circumference), pseudo-Pelger–Huët cells, micromegakaryocytes (megakaryocytes that are approximately the size of promyelocytes), and when myeloblasts are clearly increased. Blasts may be agranular or granular and are distinguished from dysplastic promyelocytes, both of which may be granulated and have nucleoli, by the presence of a clearly identifiable Golgi zone in the latter.

The bone marrow is hypercellular in the majority of patients, due to erythroid and/or granulocytic hyperplasia, but can be normocellular or, in 10–20% of cases, hypocellular. Dysplastic features can be recognized in any number of lineages and a minimum of 500 cells (at least 100 of these should be non-erythroid cells, particularly when the erythroid cells are >50% of the bone marrow) should be examined to gain an accurate and representative differential of cells in the myelogram.

Erythropoiesis is usually normoblastic, but may exhibit megaloblastic features. In patients with sideroblastic anaemia, the erythroblasts are often poorly haemoglobinized or show cytoplasmic vacuolation. The list of dysplastic features is considerable and may include binuclearity and multinuclearity, internuclear bridging, nuclear budding and fragmentation, increased pyknosis and basophilic stippling. Cytochemical iron staining should be performed in all cases of suspected MDS in order to quantify iron stores and to detect and enumerate ring sideroblasts

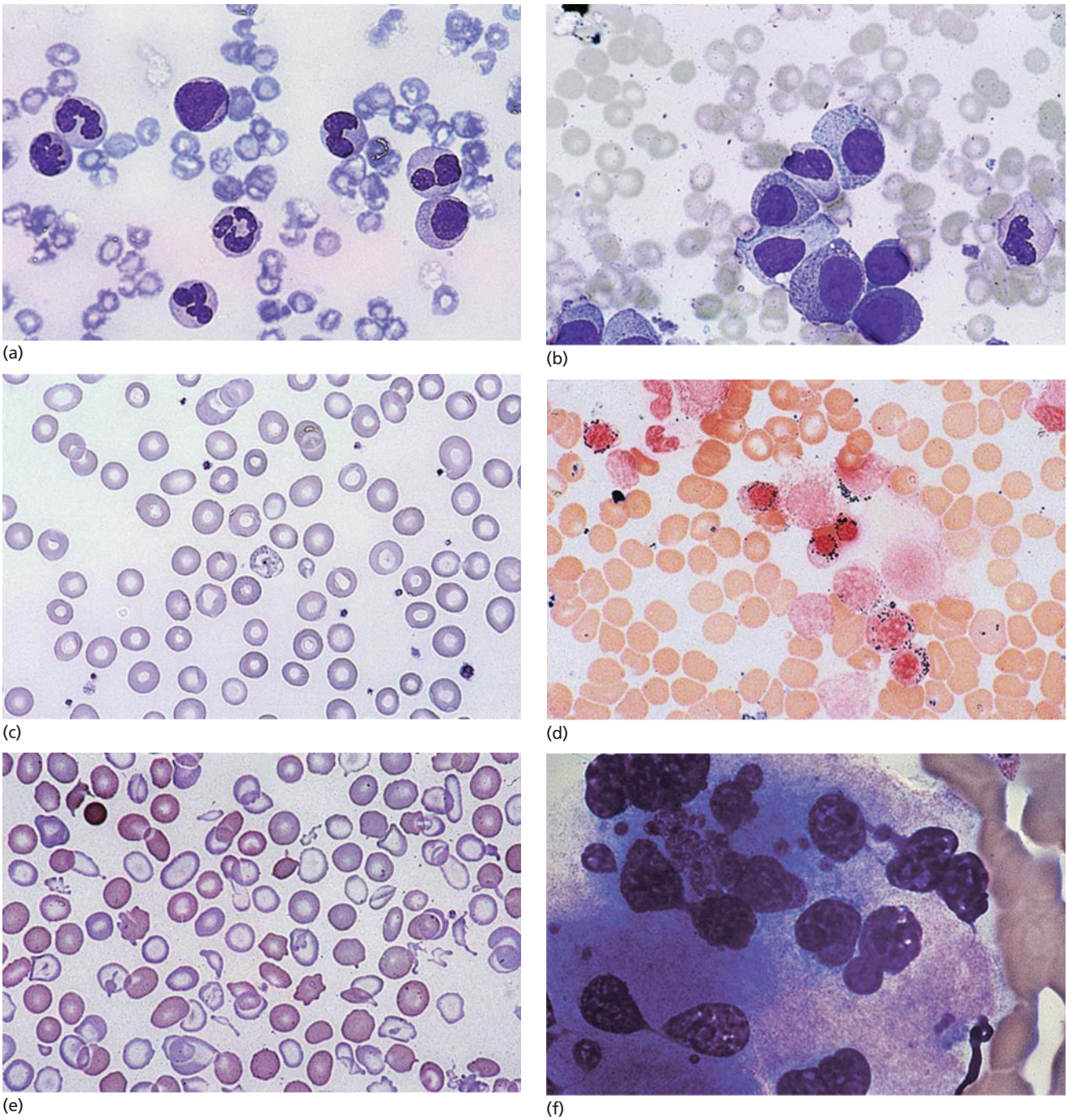
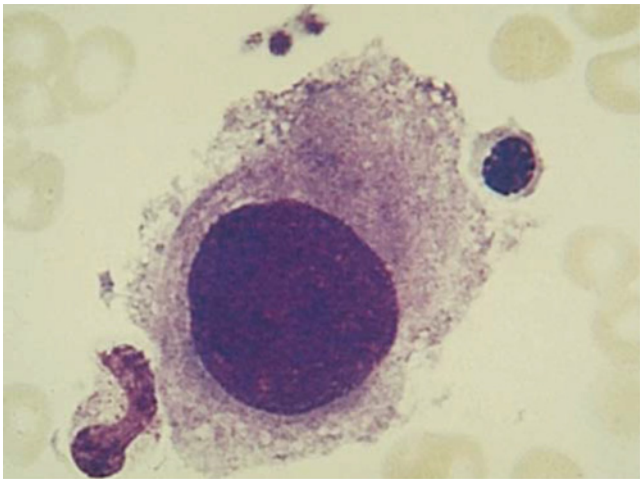
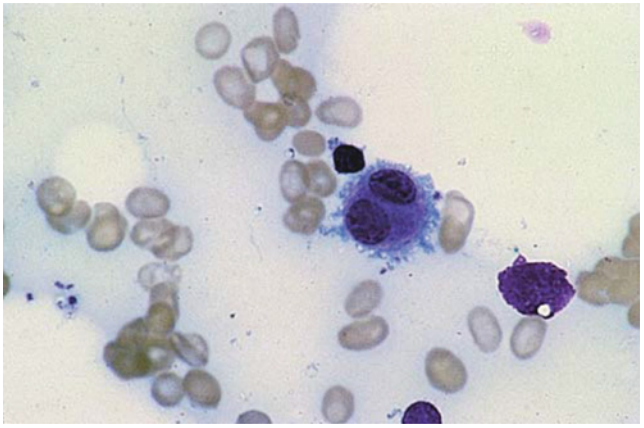


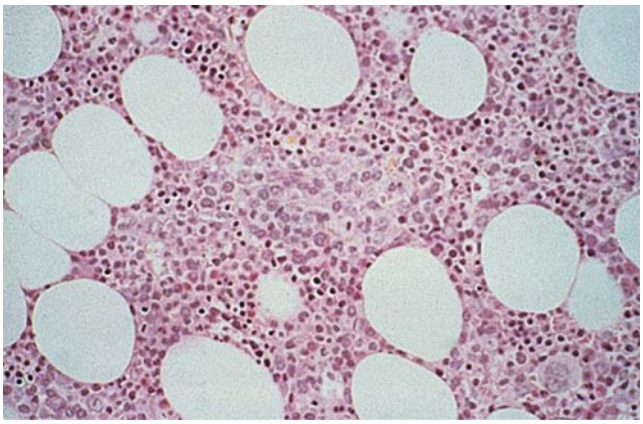
Figure 25.5 (a) Marrow aspirate showing erythroid dysplasia (nuclear irregularity, binucleate late normoblast, basophilic stippling). (b) Marrow aspirate stained with Perls stain from a patient with RARS showing ring sideroblasts. (c) Marrow aspirate showing myeloid dysplasia with neutrophil hypogranularity, pseudo-Pelger–Huët cell, promyelocyte with a Golgi zone. (d) Marrow aspirate showing RAEB with promonocytes, blasts, basophil, vacuolated neutrophil. (e) Marrow aspirate with megakaryocytic dysplasia with a bilobed nucleus. (f) Marrow aspirate with megakaryocytic dysplasia with a hypolobated nucleus.



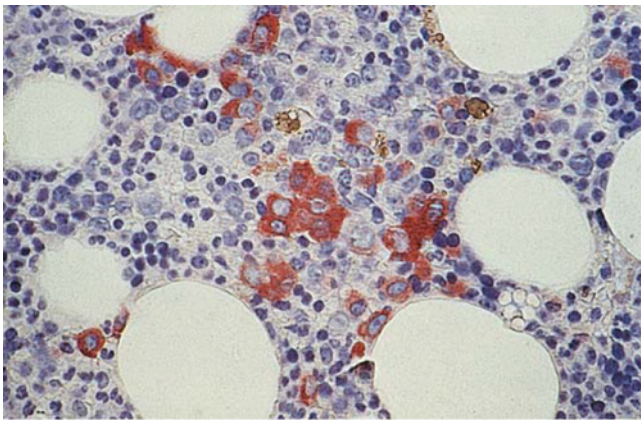
(g)



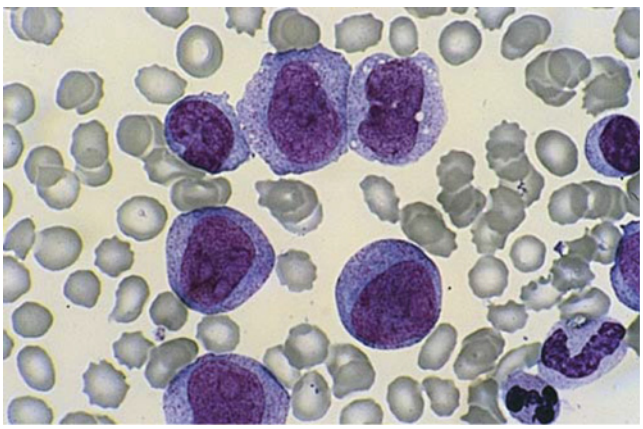
(h)



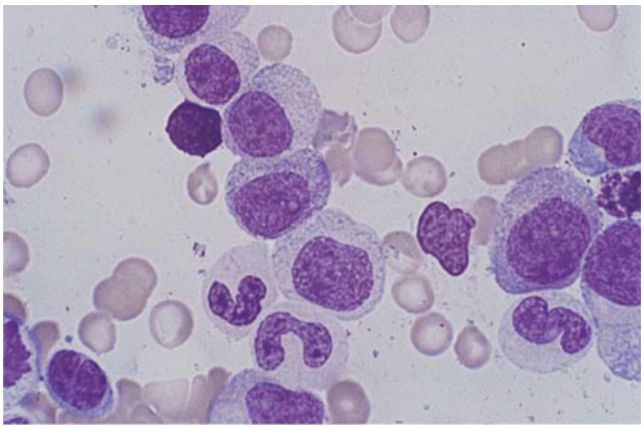
(i)



(j)

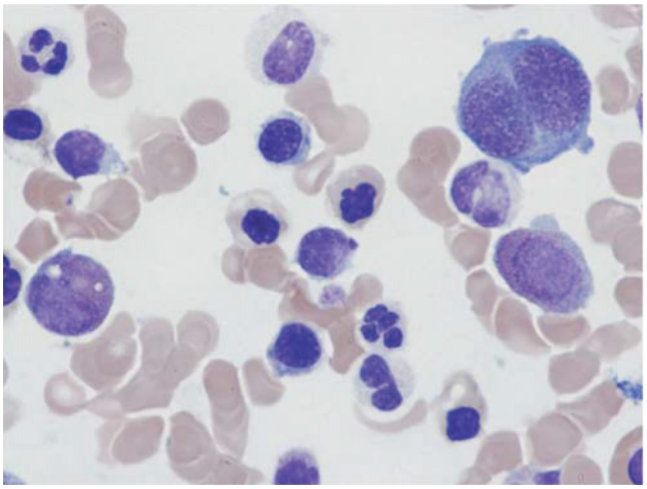


(k)



(l)

Figure 25.5 (*Continued*) (g) Dysplastic megakaryocyte with multiple widely separated nuclei. (h) Marrow aspirate from a patient with 5q– MDS showing hypolobated megakaryocyte. (i) Marrow aspirate from a patient with CMML showing monocytes, promonocytes and blast cells. (j) Marrow aspirate showing a patient with CMML with monoblasts. (k) Peripheral blood smear from a patient with CMML showing monocytes and promonocytes. (l) Bone marrow aspirate from a patient with AML. (Source: Dr Robin Ireland, Department of Haematology, Kings College Hospital, London. Reproduced with permission.)



(m)

Figure 25.5 (*Continued*)

and pathological non-ring forms. Rarely erythropoiesis may be absent due to pure red-cell aplasia that can accompany MDS.

Granulopoiesis is usually hyperplastic. Dysplasia of the granulocytic series is often quite difficult to appreciate and includes defective granulation, agranular or hypogranular cells or the presence of pseudo-Chediak–Higashi granules, and nuclear hypo- or hyperlobulation. Granulocytic precursors may show cytoplasmic basophilia with aberrant granulation. Prominent basophilic and eosinophilic differentiation and increased numbers of blasts may be present. In the context of the WHO 2008 classification, blasts that display Auer rods signify RAEB-2 or CMML-2, regardless of blast percentage.

Megakaryopoiesis is commonly dysplastic, of which the most specific feature is the micromegakaryocyte. This is typically the size of a myeloblast with an ill-defined or blebbed outline and a single monolobed nucleus. Other megakaryocytes may exhibit hypolobulation or contain multiple disparate nuclei due to aberrant maturation. In the acquired 5q– syndrome, there are often increased numbers of megakaryocytes that are variably sized and contain a large non-lobulated nucleus.

Bone marrow histology

Histological analysis of the bone marrow can yield diagnostic information not apparent by aspirate morphology, particularly regarding bone marrow cellularity, architectural changes and stromal reactions. Bone marrow trephine biopsy is recommended at diagnosis in suspected cases of MDS. The majority of patients with MDS have hypercellular marrows, but a significant minority (10%) will have a hypocellular marrow. Cytological evidence of dysplasia can be found in all lineages, but is most easily detected in the erythroid precursors. Abnormal clustering of megakaryocytes is often seen, as are micromegakaryocytes that can be more easily detected by immunohistochemical staining with CD61. Abnormal localization of immature precursors (ALIP) is a much-debated finding that refers to the clustering of promyelocytes in intertrabecular regions of the biopsy. Their presence tends to indicate propensity to leukaemic transformation. The presence of aggregates (3–5 cells) or clusters (>5 cells) of blasts, usually localized in the central portion of the bone marrow away from vascular and endosteal surfaces of trabeculae, are frequently found in RAEB and imply an aggressive course. Reticulin staining may demonstrate increased reticulin that may be mild, but can also be WHO grade 2 or 3. The presence of reticulin is often helpful in differentiating hypoplastic MDS from aplastic anaemia. Additionally, in fibrotic MDS, splenomegaly is notably absent or mild. In some cases an increase in mast cells may be present.

Cytogenetic abnormalities in MDS

Clonal cytogenetic abnormalities occur in about 50% of patients with MDS. Detection of some of these abnormalities are presumptive of MDS if accompanied by refractory cytopenias, even in the absence of morphological dysplasia. These include unbalanced aberrations such as monosomy 7 (–7), del 7q), –5, del(5q), isochromosome 17q i(17q), t(17p), 13, or del(13q), del(11q), del(12p), t(12p), del(9q), idic(X)(q13) and the rarer balanced translocations t(11;16)(q23;p13.3), t(3;21)(q26.2;q22.1), t(1;3)(p36.3;q21.2), t(2;11)(p21;q23), inv(3)(q21q26.2) and t(6;9)(p23;q34). The presence of trisomy 8, –Y or del(20q) are not defining for MDS and require morphological evidence of dysplasia to make the diagnosis. Some of these abnormalities define clinical syndromes such as the 5q– syndrome, which is seen in older women presenting with a refractory macrocytic anaemia, normal platelet count or thrombocytosis and an isolated del(5q) with non-lobated megakaryocytes in the bone marrow. Similarly loss of 17p is associated with pseudo-Pelger-Huët anomaly, small vacuolated neutrophils and an unfavourable clinical course due to *TP53* mutation. The abnormalities of chromosome 3 (inv(3)(q21q26.2) or t(3;3)(q21;q26.2)) are associated with increased abnormal megakaryocytes and del(20q) with erythroid and megakaryocyte abnormalities. Cytogenetic abnormalities form an important prognostic indicator in MDS, as discussed below with prognostic scores. SNP karyotyping in normal cytogenetic MDS can detect deletions or uniparental disomies that are undetected by normal karyotyping, as discussed with pathogenesis.

Molecular abnormalities in MDS

Recurrent acquired somatic mutations have been identified in MDS, but are not specific to MDS. However, detection of these provide evidence of clonality and are useful in patients with normal karyotypes. Thus, although not as yet recommended as a

routine test, the combination of cytogenetics and sequencing may identify an abnormality in over 90% of patients with MDS and CMML.

Other investigations

Immunophenotyping does not play a major role in the diagnosis of MDS and need not be routinely performed. However, various abnormalities are discernible, notably low side-scatter, reduced expression of normal myeloid markers and aberrant patterns of expression of other markers. CD34 expression, and to a lesser degree CD117, often correlates with the blast percentage, while coexpression of CD7 is significant for conferring a worse prognosis.

There are a number of further laboratory tests that are commonly abnormal in MDS, but which are rarely indicated for purposes of diagnosis. These include the following:

1 Granulocyte function tests to demonstrate defective phagocytosis, cell killing and motility.
2 Platelet function tests to demonstrate reduced aggregation and prolonged bleeding time.
3 Haemoglobin electrophoresis, or HPLC, to detect HbH (raised in acquired HbH disease) and HbF (raised in juvenile myelomonocytic leukaemia).
4 Ferrokinetics to assess erythropoiesis.
5 Autoantibodies (found in up to 50% of CMML patients).
6 Serum protein electrophoresis to assess immunoglobulins and detect the presence of a paraprotein.
7 Lymphocyte populations to detect altered numbers of T-cell subsets and natural killer cells.
8 Exclusion of HIV infection and parvovirus B19 infection in hypoplastic MDS.

Natural history and prognostic factors

Natural history

The natural history of MDS is highly variable due to the considerable biological heterogeneity of the disease. The median survival of adult patients with primary MDS is approximately 20 months, but this varies from only a few months for patients with high-risk disease up to nearly 12 years for patients with low-risk disease. At one end of the MDS spectrum are patients whose disease rapidly transforms into AML, while at the other end are patients with modest levels of anaemia not requiring transfusion support. Many factors influence the clinical course of the disease, including the subtype of MDS, the levels of dysplasia and blast percentage, and the number of cytopenias present. The biology of the disease, which dictates the rates of clonal expansion and leukaemic evolution, clearly involves genetic and epigenetic abnormalities.

Cytogenetic analysis represents the most important investigation at diagnosis, both for understanding the biology of the disease and for making prognostic recommendations for the patient. Indeed it is possible to construct a prognostic scheme in MDS based solely on the cytogenetic abnormalities. This was recently carried out by the German–Austrian MDS Study Group using data generated from 1202 patients treated with supportive care only. Their analysis was based on a total of 24 different karyotypic abnormalities, ranked according to the median survival of patients with each abnormality, which were then grouped into four risk subgroups. The most common cytogenetic abnormality in each of the four karyotypic subgroups, along with their respective median survival, were as follows: 5q– in good-risk group (55 months), trisomy 8 in intermediate-I risk group (29 months), chromosome 7 anomalies in intermediate-II risk group (15 months) and complex (more than three anomalies) in poor-risk group (8 months). However, normal karyotype accounted for just over 50% of all the cases studied and these could be categorized into the good-risk group. For these patients, alternative techniques will be required to further delineate the genetic abnormalities, for instance high-resolution SNP microarrays. One such study of low-risk MDS patients with normal karyotypes identified UPD in 46% of cases analysed, compared with deletions in 10%, both of which result in cryptic loss of heterozygosity and potential predisposition to development of MDS. In the future, new analytical tools, including matrix CGH, microarray gene expression analyses, proteomics and methylation profiling, will no doubt add further to our understanding of the pathogenesis of MDS, identification of therapeutic targets, and individualization of prognosis and therapy. Until such time, we are left with a long list of clinical and laboratory criteria by which to determine prognosis and optimal management for an individual diagnosed with MDS.

International Prognostic Scoring System

The International Prognostic Scoring System (IPSS) was published in 1997 by an International MDS Risk Analysis Workshop (Tables 25.6–25.8). It was based on the retrospective evaluation of a clinical, morphological and cytogenetic dataset obtained from 816 patients with MDS that was compiled from seven large, risk-based studies. These patients were either untreated or had received only short courses of low-dose oral chemotherapy or haemopoietic growth factors. The study included those patients diagnosed with MDS according to the FAB criteria, except for those patients with proliferative CMML, defined as having a white cell count greater than 12×10^9/L. A global analysis was performed and critical prognostic variables were evaluated to generate a consensus prognostic scheme, particularly using more refined bone marrow cytogenetic information. In addition to patient age, univariate analysis

Table 25.6 International Prognostic Scoring System (IPSS).

Variable	Score value				
	0	0.5	1.0	1.5	2.0
Bone marrow blasts (%)	<5	5–10	–	11–20	21–30
Karyotype*	Good	Intermediate	Poor	–	–
Cytopenias†	0/1	2/3	–	–	–

*Karyotype status: good category includes normal, –Y, del(5q), del(20q); poor category includes complex (three or more) abnormalities, chromosome 7 anomalies; intermediate category includes other abnormalities.
†Cytopenias defined as haemoglobin <10 g/L, neutrophils $<1.8 \times 10^9$/L and platelets $<100 \times 10^9$/L.

indicated that the major variables having an impact on disease outcome for evolution to AML were cytogenetic abnormalities, percentage of myeloblasts in the marrow and number of cytopenias.

The IPSS generates an overall score using three criteria, namely blast percentage, number of cytopenias and karyotype risk. Based on this score, patients are allocated into one of four risk groups: low, intermediate-1, intermediate-2 or high. The median survival for each of these groups is stratified according to age at diagnosis, but broadly ranges from 5–7 years for low risk, 3–5 years for intermediate-1, 1–2 years for intermediate-2, to 4–8 months for high risk. The value of such a prognostic scoring system is clearly dependent on the type of treatment that a patient receives. The IPSS has been validated in patients undergoing allogeneic transplantation, but the outcome of patients receiving intensive chemotherapy is more dependent on cytogenetic data than the percentage of marrow blasts and the IPSS appears less helpful in this setting. More recently, an alternative prognostic scoring system based on the WHO classification scheme has been promoted to address some of the shortcomings of the IPSS.

WHO Classification-based Prognostic Scoring System

One of the major shortcomings of the IPSS is the fact that the system was designed to be used only at diagnosis and may not be suitable for serial assessment of patients whose disease can evolve over time. In addition, the IPSS predated the revised classification of MDS according to WHO criteria and therefore incorporated some entities, notably CMML, that are now categorized as myeloproliferative overlap disorders. Finally, the IPSS fails to reflect whether or not the patient is dependent on transfusion support, which along with cytogenetics is one of the

Table 25.7 Prognostic outcomes of MDS patients according to IPSS risk score.

Outcome measure	IPSS risk category			
	Low	Intermediate-1	Intermediate-2	High
Combined score	0	0.5–1.0	1.5–2.0	2.5
Leukaemic death	19%	30%	33%	45%
Median time to AML (years)	9.4	3.3	1.1	0.2
Median survival (years)	5.7	3.5	1.2	0.4

The median survival for each of these groups is stratified according to age at diagnosis but broadly ranges from 4 to 12 years for low risk, 2 to 4 years for intermediate-1, 1 to 2 years for intermediate-2, and to <1 year for high risk across all age groups.

Table 25.8 Median survival of MDS patients according to IPSS risk score and age.

Risk group	IPSS score	Median survival (years)			
		<60 years old	>60 years old	<70 years old	>70 years old
Low	0	11.8	4.8	9	3.9
Intermediate-1	0.5–1.0	5.2	2.7	4.4	2.4
Intermediate-2	1.5–2.0	1.8	1.1	1.3	1.2
High	2.5	0.3	0.5	0.4	0.4

Table 25.9 WHO Classification-based Prognostic Scoring System (WPSS).

	Score value			
Variable	**0**	**1**	**2**	**3**
WHO category	RA, RARS, 5q–	RCMD, RCMD-RS	RAEB-1	RAEB-2
Karyotype*	Good	Intermediate	Poor	–

*Karyotype status: good category includes normal, –Y, del(5q), del(20q); poor category includes complex (three or more) abnormalities, chromosome 7 anomalies; intermediate category includes other abnormalities.

main prognostic factors affecting survival. For these reasons, the WHO Classification-based Prognostic Scoring System (WPSS) was developed based on the retrospective analysis of 467 Italian patients and validated against 620 German patients (Tables 25.9 and 25.10). All patients had a diagnosis of MDS according to one of the WHO subgroups, but CMML and other overlap disorders were excluded. Analysis of the data showed that the most significant variables for the model were WHO subgroup, the cytogenetic abnormalities and the blood transfusion requirements of the patient (Figure 25.4). These were incorporated into a scoring system that allowed stratification into five distinct risk groups showing significantly different overall survival and probability of leukaemic progression (Figure 25.5). The median survival (in months) of these five groups was 136 (very low), 63 (low), 44 (intermediate), 19 (high) and 8 (very high). The likelihood of developing leukaemia ranged from a 10-year probability of 7% in the very low-risk group to a 50% probability at 8 months in the very high-risk group.

The major difference from the IPSS is that the WPSS provides prognostic information from initial evaluation through treatment to follow-up. Thus, the WPSS permits a dynamic estimation of survival and risk of AML transformation at multiple time points during the natural course of MDS and is therefore more versatile than the IPSS in clinical decision-making and selection of appropriate treatment options. Notably, patients with normal blast counts and those with only erythroid dysplasia have a better prognosis according to the WPSS model, but this requires validation. No doubt further revisions will emerge in the future that incorporate additional discriminatory risk factors.

Revised International Prognostic Scoring System

The Revised IPSS was developed to address some of the shortfalls in the IPSS system, namely that it included patients with 20–30% blasts now considered to have AML, was only valid at diagnosis and did not include severity of cytopenias and to examine the newer cytogenetic groupings. Data from 7012 primary MDS patients older than 16 years who had not received disease-modifying therapy (hypomethylating agents, intensive chemotherapy or haemopoietic stem cell transplantation) from several international registries was collated by the International Working Group for Prognosis in MDS (IWG-PM). Seven thousand patients were classified morphologically by FAB and five thousand of them by WHO criteria. Defining blood and marrow parameters included marrow blasts ≤30%, peripheral blood blasts ≤19%, white blood cell count $\leq 12 \times 10^9$/L and absolute neutrophil count $\leq 8 \times 10^9$/L and were to be stable for 2 months. The karyotypic abnormalities and risk categories required more than 10 patients to be included as a specific category. Data was censored if and when patients received disease-modifying treatment.

The score utilizes the same variables as the IPSS, but with significantly expanded cytogenetic categories, in recognition of the increased prognostic significance of cytogenetics in comparison with bone marrow blast percentage. Thus there are five cytogenetic categories with very good (del(11q), –Y); good (normal, del(20q), del(5q) alone or with another anomaly, del (12p)); intermediate (+8,del(7q), i(17q), +19, +21, any single or double abnormality not listed, two or more independent clones); poor

Table 25.10 Median survival of MDS patients according to WPSS risk score.

	WPSS risk category				
Criteria	**Very low**	**Low**	**Intermediate**	**High**	**Very high**
Combined score	0	1	2	3–4	5–6
Median survival (months)	136	63	44	19	8

Table 25.11 Revised International Prognostic Scoring System (IPSS).

	0	0.5	1	1.5	2	3	4
CGN	–	–	Good	–	Int	Poor	Very poor
BM Blast %	≤2	–	2.1–4.9%	–	5–10	>10	–
Hb	≥100	–	80–99	<80	–	–	–
Plts	≥100	50–99	<50	–	–	–	–
ANC	≥0.8	<0.8	–	–	–	–	–

(der(3q), –7, double with del(7q), complex with three abnormalities) and very poor prognosis (complex with more than three abnormalities) and these are more heavily weighted in defining the overall score. Blast percentages have been refined in recognition of the similar risk in terms of both survival and AML evolution for those with >10 to ≤20% versus 20 to ≤30% blasts, with elimination of the 20–30% blast category and inclusion of 3% blasts as a risk factor. Cytopenias are assessed individually with increasing risk for depth of individual cytopenias (Tables 25.11, 25.12 and 25.13), notably the ANC cut point has been lowered from 1.8×10^9/L to 0.8×10^9/L. Age-adjusted cut-offs are applied to put patients into five discrete prognostic groups: very low (median overall survival 8.8 years, median time to 25% AML evolution not reached), low (median overall survival 5.3 years, median time to 25% AML evolution 10.8 years), intermediate (median overall survival 3.0 years, median time to 25% AML evolution 3.2 years), high (median overall survival 1.6 years, median time to 25% AML evolution 1.4 years) and very high (median overall survival 0.8 years, median time to 25% AML evolution 0.73 years).

Several factors such as serum ferritin, LDH, performance status and possibly β2-microglobulin also impact survival, but not evolution to AML; however, their impact on prognosis was lower than the five main features and age.

The revised IPSS score is more discriminatory than the IPSS, with 27% of low, Int-1 IPSS being shifted into higher rIPSS categories and 18% of IPSS Int-2, high-risk patients were downstaged into lower rIPSS categories (mainly intermediate). The IPSS R intermediate group has a survival of 3.0 years, which is similar to the 3.5 years for IPSS Int-1 although this group has a distinctly higher risk of AML evolution than the lower rIPSS categories. The score has been validated in an independent cohort and prospectively in several countries, is of relevance in predicting outcomes in response to disease-modifying therapies, such as treatment with azacytidine and allogeneic stem cell transplantation.

Molecular mutations and prognostic scoring systems

The presence and number of acquired somatic mutations is also being studied for their prognostic value in MDS. Bejar *et al.* demonstrated that the presence of somatic mutations in any of the five genes ASXL1, RUNX1, TP53, EZH2 and ETV6 was associated with a worse survival in each of the IPSS groups. This work is being refined further in larger data sets to derive a molecular prognostic scoring system. Furthermore the number of acquired somatic mutations are also prognostic, with a poorer prognosis attached to increasing numbers of mutations.

Management and treatment

The mainstay of managing MDS is good supportive care, as the patients are usually elderly and may have accompanying comorbidities. Increasingly, patients may be carers for partners or other relatives, and these along with their functionality need

Table 25.12 RIPSS cytogenetic subgroups.

Very Good	Good	Intermediate	Poor	Very Poor
Single	**Normal**	**Single**	**Single**	**Complex**
–Y	**Single**	Del(7q)	der(3q)	>3 abnormalities
Del 11q	Del(5q)	+8	–7	
	Del(12p)	I(17q)	**Double**	
	Del(20q)	+19	Incl –7/7q–	
	Double	Any other indep clone	**Complex**	
	Incl del (5q)	**Double**	3 abnormalities	
		Any other		

Table 25.13 IPSS-R prognostic risk categories and scores.

Risk category	Risk score	Survival (median years)	AML 25% evolution
Very Low	≤1.5	8.8	–
Low	>1.5–3	5.3	10.8
Intermediate	>3–4.5	3.0	3.2
High	>4.5–6	1.6	1.4
Very High	>6	0.8	0.73

to be taken into consideration whilst developing a management strategy. Alongside the disease-specific risk scores, assessments of comorbidities such as the MDS-Specific Co-morbidity Index and Comprehensive Geriatric Score are helpful in treating patients.

Three important considerations should be borne in mind in arriving at an appropriate management plan for a patient with MDS: firstly, in asymptomatic patients, a period of observation can be extremely helpful in determining the rate of disease progression, secondly, it can be unwise to base the treatment plan on blood and marrow results during an acute infective episode, since neutrophil and blast counts may be transiently increased and thirdly, the patient's own wishes and expectations, as well as his or her health and social circumstances, should all be part of the overall consideration before arriving at a management plan.

The British Committee for Standards in Haematology (BCSH) and European LeukaemiaNet have developed consensus guidelines based on an evidence-based review of the published literature to guide optimal treatment of patients with MDS. The management recommendations in this chapter are drawn partly from these guidelines, but have also sought to include newer treatment options that have emerged since the guidelines were published.

Supportive care

Anaemia is the commonest symptom in MDS and has a significant impact on quality of life. Moreover, transfusion dependency or the surrogate Hb values of <90 g/L in males and <80 g/L in females predict an increased risk of cardiac comorbidity and reduced overall survival. Thus, patients who exhibit symptoms or signs of clinical anaemia should receive red cell transfusions in order to improve quality of life and ideally maintain Hb above these thresholds. For patients with coexisting cardiac dysfunction, anaemia may precipitate cardiac failure and individualized transfusion goals may help alleviate this. However, the potential risks of blood transfusions should always be considered, notably iron overload in multitransfused patients. Therefore, transfusions should only be used to alleviate symptoms of anaemia and not simply to maintain the haemoglobin above an arbitrary level. Other factors that might accentuate anaemia, such as nutritional haematinic deficiencies, haemorrhage, haemolysis or infection, should be sought and treated as appropriate. Where appropriate consideration should be given to treatment with erythropoietin to alleviate anaemia, as discussed below.

Bleeding is a common and potentially serious complication of MDS, resulting from both thrombocytopenia and the effect of functionally defective platelets. Platelet transfusions are indicated in MDS patients in the treatment of acute bleeding episodes, or as prophylaxis prior to surgery or following chemotherapy. However, their long-term use, for instance in the prevention of recurrent epistaxis or oral bleeding in elderly patients with persistent thrombocytopenia, presents significant logistical issues. Moreover, as for red cell transfusions, platelet transfusions are not without potential complications, including allosensitization that can lead to refractoriness. Antifibrinolytic agents such as tranexamic acid can be useful on occasion, but are not routinely recommended. Where patients are likely to undergo intensive chemotherapy followed by transplantation, the use of leucodepleted products and irradiated products should be considered.

Haemopoietic growth factors

The use of recombinant erythropoietin in the treatment of anaemia in low- or Int-1-risk MDS can normalize haemoglobin levels and enhance quality of life without the risks posed by blood transfusions. Early institution of erythropoietin therapy in patients with a low erythropoietin level (less than 200 IU/L) predicts for response. Early studies indicated that patients with low serum erythropoietin concentration of less than 200 IU/L (usually raised in MDS) and low transfusion requirements (two or less units per month) have a greater than 70% chance of responding, whereas those who have higher erythropoietin levels and transfusion requirements respond in less than 7% of cases. For patients with refractory anaemia with ringed sideroblasts, response rates to erythropoietin are enhanced by the addition of GCSF. In such patients, there is evidence that cytokines can act on haemopoietic progenitor cells to reduce apoptosis and improve erythropoiesis. Additionally, the achievement of a CR (Hb >115 g/L) suggests a longer duration of response of 29 months than the median of 5.5 months, when patients achieve a partial erythroid response (an increase of 15 g/L with elimination of transfusion dependence without achieving an Hb >115).

Factors to be taken into consideration in treating patients with erythropoietin are their baseline renal function (as those with impaired renal function may need lower doses of erythropoietin), iron stores and recognition that patients with MDS often need higher doses of erythropoietin to elicit a response. Various treatment schedules exist, but, in general, patients are treated with the equivalent of 30,000 IU of erythropoietin-α or -β, weekly for 8 weeks, increased to 60,000 IU if an adequate response is not achieved, followed by the addition of GCSF 300 μg/week in 2–3 divided doses in those who do not respond.

Care must be taken not to exceed 120 g/L, due to the increased risk of thrombosis, which is approximately 2% particularly if there are coexisting vascular risk factors such as previous stroke, diabetes or hypertension.

Similarly, darbepoietin is a recombinant long-acting erythropoietin that differs from the native form in having two additional N-linked oligosaccharide chains. Treatment with darbepoietin may be commenced at 300 μg every 14 days, which may be increased after 8 weeks to 300 μg per week for a further 8 weeks.

In patients achieving a complete erythroid response, the dose of erythropoietin can be slowly reduced to the lowest level needed to maintain the response. If the response is lost then functional iron deficiency or other haematinic deficiency should be excluded.

The administration of GCSF to patients with MDS typically results in a dose-dependent improvement in neutrophil number and function, and is usually reserved for patients with severe sepsis or recovering from intensive chemotherapy. There is insufficient evidence to support its prophylactic use for preventing neutropenic infection, although some patients whose quality of life is compromised by recurrent infective exacerbations may respond to such an approach.

Immunosuppression

In MDS, the occurrence of clinical features of autoimmunity, as well as increased numbers of cytotoxic T lymphocytes that inhibit haemopoiesis, and increased numbers of Th17-secreting CD4+ cells implicate dysregulation of the immune system and form the basis for employing immunosuppressive therapies (IST). Retrospective studies of horse ATG in unselected low-risk MDS patients yield response rates of 30–40%. Retrospective evaluation of responders suggest that those less than 60 years of age, with blasts <5%, normal karyotype or trisomy 8 and transfusion dependence are likely to respond to IST. Patients with hypocellular, but also normo- or hypercellular marrows who meet these criteria would also be candidates for such treatment. Treatment with horse ATG over 5 days followed by ciclosporine maintained at 100–200 mg/dL for 6 months is suggested followed by a slow taper to the lowest dose that maintains the response in these patients. Data on the presence of HLADR15 and PNH clones as positive predictors of response have been variable.

Prospective randomized controlled trial data (with a cross-over design), with just over 40 patients in each arm treated with horse ATG/ciclosporine or best supportive care, supports the superiority of immunosuppression. However, the accompanying increased side-effects, such as severe haemorrhage, serum sickness, cardiac events, thrombosis and infections warrant careful selection prior to institution of this treatment.

Due to the increased toxicity of ATG, a Phase I/II study of alemtuzumab, an anti-CD52 antibody, was trialled by the NIH in 31 patients with Intermediate-I risk, Intermediate-II risk patients with transfusion dependence or severe thrombocytopenia (<50) or severe neutropenia (ANC $<0.5 \times 10^9$/L) on a schedule of 10 mg intravenously daily for 10 days without ciclosporin. Eligibility was based on the sum of the patient age in years and months of transfusion with a score of 52 or less if they were HLA DR15 negative and a score of 72 or less if they were HLA DR 15 positive. Responses (CR+PR) were seen in 77% of Intermediate-1 and 57% of Intermediate-2 patients, some of these responses occurring after three months of therapy and in four of seven evaluable responders, cytogenetic complete remissions for clones including a del(13q), –7 and +21, del (5q) were observed at 6–12 months post treatment. No significant viral complications were noted. Patients with a PNH clone experienced a decrease in clone size in 75%, whereas 25% demonstrated expansion of the clone. There was also a differential response in haemolysis due to PNH, with half the patients improving and half experiencing more haemolysis (these had a PNH red cell clone, 30% pretreatment) following treatment. Overall the toxicity profile appeared to be milder than ATG and at a year post treatment, 56% had normal counts and 78% were transfusion independent. Thus alemtuzumab is a viable alternative in this group of patients.

Chelation therapy

The decision to implement iron chelation in a transfusion-dependent MDS patient is not straightforward. There is a paucity of published literature to support the use of routine iron chelation therapy. Furthermore the treatments are expensive and inconvenient. Current BCSH recommendations state that iron chelation should be considered once a patient has received 5 g iron (approximately 25 units of red cells) and has a serum ferritin level of >1000 ng/mL but only in younger patients for whom long-term transfusion therapy is likely, such as those with sideroblastic anaemia, or in those with good-prognosis disease, including isolated del(5q) without significant comorbidities. For those patients who are candidates for allo-SCT there is evidence of increased non-relapse mortality if the ferritin exceeds 2000 ng/mL and these patients may be considered for iron chelation therapy. Desferrioxamine 20–40 mg/kg should be administered by subcutaneous infusion over 10–12 hours for 5–7 days per week with annual audiometry and ophthalmology review (Chapter 4). The target ferritin concentration should be below 1000 μg/L and vitamin C 100–200 mg daily can be added after 1 month to enhance chelation. Deferasirox, an oral iron chelator, is licensed for use in MDS when desferrioxamine is contraindicated or inadequate. It must, however, be used cautiously due to side-effects that may cause renal or hepatic impairment and gastrointestinal haemorrhage. Early experience with deferasirox shows that it is well tolerated and controls or reduces iron levels in chronically transfused MDS patients when used at a daily dose of 20–40 mg/kg. Deferiprone is not recommended in neutropenic patients.

Intensive chemotherapy

The efficacy of remission induction chemotherapy with standard doses of cytarabine 100 mg/m^2 IV every 12 to 24 hours, in combination with an anthracycline (daunorubicin, idarubicin) in MDS is dependent on karyotype, with those with normal karyotype achieving remissions and a median survival of 18 months, whereas patient with complex karyotypes have a median survival of only 4 months. This is accompanied by a risk of chemotherapy-induced aplasia and an early death rate of 10%. Thus, although chemotherapy may offer a remission and accompanying improved QOL, it is often used when the remission achieved will be consolidated with an allogeneic stem cell transplant. The results in MDS are inferior to those from *de novo* AML attributable to increased chemo-resistance as MDS cells over-express p-glycoprotein, which mediates resistance. Alternatively the MDS HSC may be more quiescent, thus reducing its susceptibility to cytotoxic agents. Outcomes for allogeneic stem cell transplantation are dependent on disease characteristics at the time of transplant, in particular those with less than 5% blasts at the time of transplant have improved outcomes when compared with those who have increased blasts. Extrapolation of this to suggest that reduction of blasts to less than 5% pretransplant is the rationale for induction chemotherapy pretransplant. Given the risk of chemotherapy induced aplasia, particularly if the marrow is hypocellular or fibrotic, it would be prudent to tissue type the patient and identify a donor prior to commencing chemotherapy. Outcomes following induction chemotherapy include CR rates of 24–79% with relapse rates of 33–91% and treatment-related mortality of 2–42%, all of which are better in younger patients with a good performance status and favourable cytogenetics. In some patients with less than 10% blasts and smouldering disease, transplant may be undertaken without prior chemotherapy.

On an intention-to-treat basis, relatively few patients intended for an allograft actually receive the transplant, either due to treatment toxicity or refractory disease. Thus an alternative view would be to allograft patients up front or following treatment with 5-azacytidine. In a randomized controlled open label Phase III study, comparisons of azacytidine and induction chemotherapy showed that there was no significant difference in outcomes, although the numbers treated with induction chemotherapy were relatively small. Retrospective analyses of 128 patients who were pretreated with azacytidine (n = 40) or best supportive care prior to an RIC allograft demonstrated equivalent results. Thus the role of pretransplant intensive chemotherapy in MDS is controversial.

Consolidation with high-dose chemotherapy in the form of an autograft is fraught with the difficulty of achieving adequate harvests. Furthermore, the data from the EBMT study on 114 patients who underwent an autograft suggested limited success with 2-year disease-free survival of 34% and high relapse rates of 64%. Thus, autografting is not recommended as standard treatment.

Allogeneic stem cell transplantation

Allogeneic SCT is the only therapeutic modality that leads to long-term disease control/cure in MDS. As per the CIBMTR data, MDS is the second most common indication for allografts. This has been made possible due to the introduction of reduced intensity transplants that use lower doses of chemotherapy to enable engraftment and rely on the graft-versus-leukaemia effect to provide long-term disease control. Numerous conditioning regimens are currently in use and these may be T replete or T deplete. T-cell depletion reduces the risk of acute and chronic GVHD and is very well tolerated at the expense of increased relapse. Cellular therapy in the form of donor lymphocyte infusions, either to achieve a full donor chimerism (pre-emptive) or to treat relapse, may help to reduce relapse rates.

Irrespective of conditioning regime used, full intensity or reduced intensity, T-replete or T-deplete transplants, the outcome at 5 years are similar at approximately 40% disease-free survival (data from EBMT is shown in Table 25.14). The main predictors of outcome remain patient- or disease-related, with high-risk disease, particularly where this is due to complex monosomal karyotypes, showing little benefit from current allografting strategies.

Given that for an individual patient, outcomes from an allograft are uncertain and associated with severe side-effects, including acute and chronic GVHD, life-threatening infections and a significant transplant-related mortality, advice on when to consider an allograft is crucial. Historically Markov modelling examined three different constructs for myeloablative transplant strategies, concluding that patients with low- or Int-1-risk MDS, life expectancy was maximized by delaying transplantation, provided this was performed prior to disease progression to AML, whereas for patients with Int-2- and high risk-MDS immediate transplantation was associated with maximum life expectancy. Recently, similar modelling for RIC transplantation supports early allografting in Int-2/-risk MDS as survival is improved, after adjustment for the presence of GVHD, when compared to treatment with demethylating agents.

For patients with low-risk MDS, transplant outcomes are favourable, with a low risk of relapse and disease-free survival rates near to 60% at 5 years. The most successful outcomes are achieved if patients are younger, disease duration shorter (<12 months) and patients are transfusion independent. Within the low-risk group, the presence of red cell transfusion dependence, severe thrombocytopenia, marrow fibrosis and molecular mutations such as *ASXL1* or *EZH2* or *RUNX1*, may be considered for an allograft.

In those with high-risk MDS, while early transplant is advocated, it is becoming clear that those with complex, particularly monosomal, karyotypes, *TP53* mutations and *DNMT3A*

Table 25.14 Comparative outcomes from standard myeloablative and reduced-intensity conditioning allogeneic stem cell transplantation in MDS patients from the European registry.

Criteria	Standard myeloablative conditioning (N = 621 patients)	Reduced-intensity conditioning (N = 215 patients)	P-value
Patients over 50 years (%)	28	73	<0.001
Acute GVHD (%)	58	43	<0.001
Chronic GVHD (%)	52	45	Not significant
NRM at 3 years (%)	32	22	0.04
Relapse at 3 years (%)	27	45	<0.01
PFS at 3 years (%)	41	33	Not significant
OS at 3 years (%)	45	41	Not significant

Note that the 3-year progression-free survival (PFS) and overall survival (OS) are comparable between the two groups despite significant age differences, but that the risk of graft-versus-host disease (GVHD) and non-relapse mortality (NRM) are higher for standard transplants, while the relapse rate is higher for transplants with reduced-intensity conditioning.

mutations may derive significantly less benefit with current modalities. The reasons for this are currently unclear and need to be elucidated in order to address these results.

In multivariate analyses, age is not an independent risk factor for transplant outcomes and with RIC transplants patients are increasingly being transplanted to age 65, and dependent on functional and comorbidity status, up to the age of 70–75. The evaluation of the HCT-CI is important in attempting to quantify risk and to identify factors, such as cardiac function, that may be optimized pretransplant to minimize risk. For patients with an HCT-CI ≥3, a risk–benefit analysis between RIC and non-transplant options would be justified.

The outcomes based on donor source, particularly between matched sibling or 10/10 matched unrelated donor sources are equivalent. Alternative donors, namely umbilical cord blood stem cells and haploidentical donors have also been used successfully. Recent single-centre data comparing outcomes for patients transplanted with matched siblings, matched unrelated and haploidentical donors suggest that outcomes are comparable and dependent on the disease risk, thus supporting the use of alternative donors when matched donors are not available. Given that most patients are older, umbilical cord blood transplants are often associated with prolonged time to immune-reconstitution and higher morbidity, making them a less attractive option. Thus, in the absence of matched donors, consideration should be given to alternative donor transplants, as CIBMTR outcomes from mismatched donors are less favourable. The stem cell source used most frequently is peripheral blood stem cells due to ease of donation, shorter times to neutrophil and platelet engraftment, similar acute GVHD, but higher chronic GVHD.

Up-front sequential chemotherapy–transplant protocols are particularly attractive in MDS as they circumvent the possibility of chemotherapy-induced aplasia and morbidity, which may preclude transplantation and also the chances of selecting the most resistant clone with cycles of chemotherapy. Data on this approach is encouraging, but limited, and various combinations of fludarabine or cladribine, amsacrine and high-dose cytarabine, followed by a busulfan/ATG-based transplant, are in use.

In the UK campath-based RIC transplants are commonly used, whereas in Europe ATG-based T-cell depletion is common. Long-term follow-up of 192 patients with MDS or secondary AML treated with a fludarabine busulfan (6.4 mg/kg) and campath (100 mg) protocol with a median follow up of 4.5 years was encouraging, with a 5-year overall survival of 44%, EFS of 33% and non-relapse mortality of 26% (Figure 25.6). Whilst chronic GVHD rates were low at 19%, the relapse rate at 5 years was 51% (similar to outcomes reported with fludarabine, melphalan and campath). Relapse rates appeared to increase with age for those over 66, which to some extent could be ameliorated by early pre-emptive DLI with 5-year OS 67%. Thus T-replete

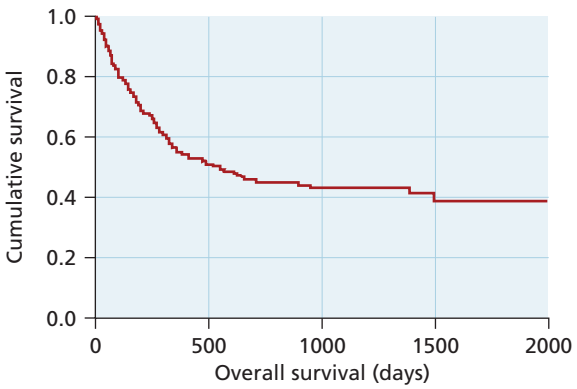


Figure 25.6 Kaplan–Meier survival plot and data for 159 patients with advanced MDS/AML who received a reduced-intensity conditioned allogeneic transplant at King's College Hospital, London. (Source: Dr Ziyi Lim, King's College London. Reproduced with permission.)

RIC transplants may form the basis for early cellular therapy with DLI, with reduced toxicity and improved outcomes.

Hypomethylating drugs

Transcriptional regulation of gene expression is modulated by epigenetic changes such as DNA methylation of CpG islands in gene promoters. These islands have an increased density of CG dinucleotides in CIS on the strand of DNA and incorporation of a methyl group into the cytosine results in silencing of the gene. This incorporation is mediated by the enzymes DNMT1 (DNA methyltransferase 1) and enables methylation patterns to be reproduced in daughter cells during cell division, whereas DNMT3a is responsible for incorporating *de-novo* methylation that appears to be important in haemopoietic stem cell differentiation. Two drugs, azacitidine (a pyrimidine analogue of cytarabine) and decitabine, disrupt methylation by being incorporated into RNA and DNA, respectively, binding DNMT and preventing its function so that methylation is not replicated with cell replication, leading to eventual demethylation. Azacytidine is approved in the USA for use in all subtypes of MDS, but in Europe it is approved for use in Int-2 and high-risk MDS and AML, with blasts 20–30% and non-proliferative CMML. The initial studies in AML used high doses and subsequent unacceptable toxicities, whereas the Cancer and Leukaemia Group B (CALGB) 9221 and AZA 001 studies used the now licensed 75 mg/m^2 SC for 7 days followed by a 21-day break schedule. In the former randomized controlled, cross-over (at 4 months) study, all FAB subtypes of MDS responded, with 47% achieving a complete remission and 36% a haematological improvement. For the first time, the interval to AML progression was prolonged from 12 months in the best supportive care arm to 21 months in azacytidine recipients. The follow-on licensing AZA-001 study enrolled 358 untreated Int-2- and high-risk patients, RAEB-t and non-proliferative CMML. Patients were randomized between azacytidine and physicians' choice, which included low-dose cytarabine (LDAC), best supportive care only or intensive chemotherapy. Physicians preselection prior to randomizations were best supportive care (n = 222), LDAC (n = 94) and intensive chemotherapy (n = 42); on randomization, 175 patients received azacytidine, 102 BSC, 44 LDAC and 19 intensive chemotherapy. Patients were required to receive a minimum of four cycles of azacytidine, based on the median time to respond in the CALGB 9221 study, and continuation of therapy following response or in the absence of progression was at investigator discretion. The study established superiority over BSC and LDAC and there was no difference in outcomes when compared to intensive chemotherapy. Median survival for azacytidine-treated patients was 24.5 months versus 15 months in conventional care arms, with a benefit of over 9 months, and overall survival was almost doubled at 2 years being 51% versus 26%. Haematological improvement, including transfusion independence, occurred in almost 50% of patients and was durable. Patients experienced fewer infections, and appeared to benefit with a 2-year survival of 33%, compared to 8% with conventional care. An important observation was that overall survival was prolonged, even if patients did not achieve a complete remission, conceptually moving the goal posts for therapy in MDS. Post-hoc analyses in AML with 20–30% blasts and patients older than 75 years also support the use of azacitidine in these groups. Azacitidine has been approved in the UK for patients with MDS who are not eligible for allogeneic stem cell transplantation and have IPSS Int-2 or high-risk, MDS or AML with 20–30% blasts and non-proliferative CMML.

Although Azacitidine is now considered the standard of care for treating MDS patients, the low response rates (complete remission rate (17%), partial remission (21%) and stable disease (42%)) have led to development of predictive scores to guide optimal therapy. One such score by the French cooperative group includes assessment of performance status, presence of peripheral blasts, transfusion dependence and cytogenetic risk, and distinguishes likely responders who have a low score and have a median response duration of 30 months from those with a high score and median response duration of less than 6 months (Table 25.15). The score may be of benefit in planning additional intervention such as allogeneic stem cell transplantation.

Mechanisms of resistance or loss of response to azacitidine include the persistence of the MDS stem cell even in remission and re-emergence at relapse. Additionally depletion of enzymes such as uridine cytidine kinase, which phosphorylates azacitidine to its diphosphate form, and ribonucleotide reductase, which converts it to decitabine diphosphate prior to DNA incorporation, may lead to loss of response. This provides a rationale for switching from azacytidine to low-dose cytarabine or decitabine, which do not need this enzyme.

Decitabine, the deoxyribonucleic acid analogue of azacitidine, is also approved in the USA for treatment of MDS. Its dosing schedules have varied and the Phase III randomized controlled trials with best supportive care failed to show superiority in part due to recipients receiving very few cycles of decitabine. The currently licensed dosing schedule is 20 mg/m^2 IV for 5 days in high-risk MDS and yields similar response rates of 17% CR, 15% PR and 18% HI, with overall survival of 19.4 months. In the UK it is not approved for treatment of MDS and is used on a compassionate basis to salvage patients who have failed to respond to azacytidine. Current studies are focused on improving response rates by using lower doses, prolonged schedules and combination therapies with histone deacetylase inhibitors, immunomodulatory agents and growth factors, amongst others.

Lenalidomide

The immunomodulatory drug thalidomide was effective in achieving transfusion independence in some patients with MDS, possibly due to its anticytokine, antiangiogenic and other

Table 25.15 Prognostic score for azacitidine in MDS and its suggested use in planning treatment.

Performance status >2		1	
Peripheral blasts		1	
Transfusion dependence greater than 4 units in 8 weeks		1	
IPSS intermediate risk cytogenetics		1	
IPSS high risk cytogenetics		1	
Azacitidine prognostic score	0	1–3	4–5
Median OS with Aza alone	Not reached	18 months	6 months
2-year OS with Aza alone	85%	50%	35%
Consideration for HSCT	Absence of response to azacitidine or at disease progression	Deliver 6 cycles of azacitidine and proceed to HSCT	HSCT at the earliest possibility

immunomodulatory effects, but its neurologic toxicity limited its use. Its derivative, lenalidomide was initially developed for use in myeloma and has a better side-effect profile. It has been used in four successive trials in low-/Int-1-risk transfusion-dependent or symptomatic anaemic patients with MDS. The initial Phase II study treated 43 such patients with three different dosing regimens (25 mg or 10 mg per day or 10 mg for 21 days of a 28 day cycle). At 16 weeks 77% of patients had achieved transfusion independence; this was higher for patients with del(5q31.1), with 83% responding, compared to 57% of patients with normal karyotype and 12% with other cytogenetic abnormalities. Nine of the 12 del(5q) patients had accompanying cytogenetic remissions. The follow on study (MDS 003) treated 148 patients with del(5q), initially on 10 mg every 21 days, but this was changed to 10 mg daily for 28 days. Transfusion needs were reduced in 112/148 (76%) with a median time to respond of just over 4 weeks. Cytogenetic responses included improvement in 62/85, and 38/62 achieved a complete cytogenetic remission. The median duration of response was not reached at 2 years. The MDS 004 study was a Phase III randomized controlled study, including 205 transfusion-dependent low/Int-1 patients with del(5q) (76% had an isolated del(5q)) randomized to lenalidomide 10 mg daily for 21 days or 5 mg for 21 days of a 28 day schedule or placebo. Responses were superior in those treated with 10 mg daily for 21 days, with 56% achieving TI at 26 weeks versus 42% in the lower dose and 6% in the placebo arms. Cytogenetic responses were observed in 50% versus 25% in the 10 mg and 5 mg arms, respectively, and the duration of response was 82 weeks versus 41 weeks in the two lenalidomide treated arms. In this study, a normal baseline platelet count appeared to predict response (80% response compared to 19% in thrombocytopenic patients). However, the cumulative incidence of AML in both the lenalidomide groups was 16.8% at 2 years and 25.1% at 3 years. Despite this, the study led to approval in the USA of lenalidomide for low/Int-I del(5q) patients. In contrast, approval was delayed in Europe, leading to further examination of this risk in a cohort of treated (95 del(5q) patients treated with 10 mg for 21 days schedule) and untreated (99 patients with del(5q) who were never treated with lenalidomide). The 4-year risk of progression was 9% in the lenalidomide-treated versus 15.8% in the untreated group, which did not achieve statistical significance. Thus the drug has received a positive opinion in Europe for the treatment of transfusion-dependent anaemia due to low or Int-1 MDS with an isolated deletion 5q, when other options are insufficient or inadequate. Lenalidomide has also been approved in the UK for transfusion-dependent patients with an isolated del(5q).

Lenalidomide is also an option in reducing transfusion dependence in low-/Int-1-risk MDS without del(5q). Of 214 patients enrolled, 26% achieved transfusion independence and 19% had a cytogenetic response. The commonest side-effects are neutropenia, which may be supported with GCSF, and thrombocytopenia. Additional side-effects include a skin rash, thrombosis, hypothyroidism and hypogonadism.

Patients with del(5q) have an increased frequency of *TP53* mutations and patients in whom the mutation is detected by strong immunohistochemical staining in >2% of the bone marrow cells tend not to respond to lenalidomide and have a poorer overall prognosis.

Lenalidomide inhibits two phosphatase regulators of the G2/M cell-cycle check point, namely cell division cycle 25 homologue C (cdc25C) and protein phosphatase 2A (PP2A) catalytic subunit α (PP2Acα). The del(5q) cells are already haploinsufficient for these phosphatases, sensitizing them to further inhibition and leading to apoptosis and cell death, thus displaying an example of synthetic lethality. Moreover, resistance to lenalidomide is associated with overexpression of PP2A. Lenalidomide is also known to have potent immunomodulatory and antiangiogenic properties that may clearly be relevant to its therapeutic action. Lenalidomide may also be used at higher doses up to 50 mg per day in higher-risk MDS with a del(5q) abnormality. Responses are short-lived, but may be useful in bridging to a transplant. Its combination with azacitidine also improved CR rates in Phase II studies and is currently in randomized controlled trials. Thus, lenalidomide represents a promising advance in the treatment of selected subtypes of MDS.

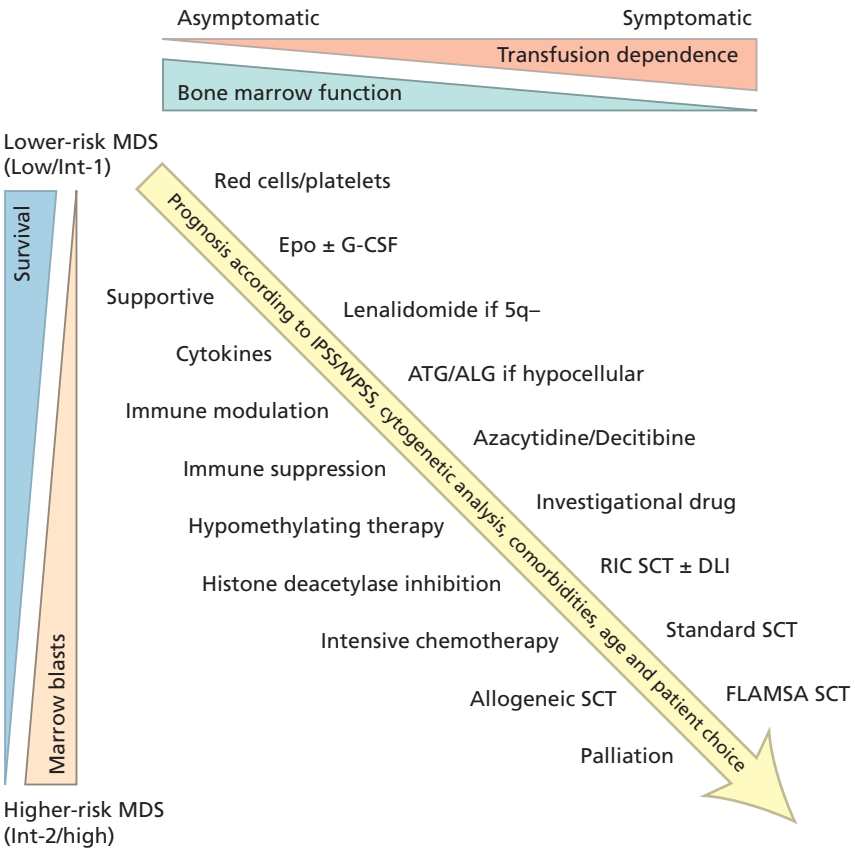


Figure 25.7 Schematic diagram illustrating the treatment options available for patients with MDS.

Therapeutic strategy

Therapeutic goals

Arriving at the best management plan for an individual patient with MDS can be difficult (Figure 25.7). The initial decision to resolve is whether the goal of treatment should be to extend patient survival or to palliate symptoms with supportive care. This decision is helped by considering prognosis according to the IPSS/WPSS, coupled to the patient's age and performance status, always ensuring that any plan is fully in accordance with the patient's own wishes. Treatment of MDS has historically been separated into low-intensity and high-intensity therapies, based on the risk of toxicities. Low-intensity therapies, which include immunosuppressive therapy, biological response modifiers and cytokines, were reserved for lower-risk patients with the goal of improving cytopenias and quality of life without improving survival. Other low-intensity therapies such as low-dose chemotherapy and methyltransferase inhibitors aim to extend survival, whereas high-intensity therapies such as high-dose chemotherapy and allogeneic SCT aim to alter the course of the disease or even achieve cure.

Many patients with MDS receive supportive care alone, particularly lower-risk patients with chronic cytopenias or patients with higher-risk disease who are unable to tolerate high-intensity therapy. Such patients often become dependent on frequent red cell or platelet transfusions and experience repeated infective and haemorrhagic complications. Regardless of disease status, many patients experience significantly impaired quality of life, simply as a consequence of the physical toll caused by frequent laboratory monitoring and transfusions, physician visits and the fatigue that accompanies this. Thus, improvement in quality of life and alleviation of disease-related symptoms are the key goals of therapy.

Assessing response

In 2000, the International Working Group (IWG) of investigators proposed standardization of response criteria by which to assess the results of different therapies in MDS. These criteria have been widely accepted into clinical practice and were updated in 2006. The investigators grouped patients into IPSS risk categories and recommended that the major goal of therapy for patients with lower-risk disease (low-risk and intermediate-1 categories) is to achieve haematological improvement. For higher-risk patients (intermediate-2 and high-risk categories), the focus turns to altering the natural history of the disease and prolonging life. The IWG criteria define four specific aspects of responses based on treatment goals: haematological improvement, cytogenetic improvement, alteration of disease progression and quality of life. Haematological improvement is scored for each lineage according to whether there is a major or minor

response, while cytogenetic improvement is scored according to whether there is a partial or complete response. Alteration of the natural course of the disease is determined according to various measures of disease progression and survival.

- *IPSS low:* The median survival for low-risk MDS is 4.8 years for those aged over 60 years to 11.8 years for those less than 60 years. For this reason, intensive chemotherapy and SCT cannot be justified, given their potential for morbidity and mortality. Individuals should be monitored for disease progression and supported as necessary. Where possible, a trial of erythropoietin with or without GCSF should be undertaken in patients with symptomatic anaemia, while those with del(5q) should ideally receive lenalidomide.
- *IPSS intermediate-1*: For older patients above the age of 65, or less if there are significant comorbidities, supportive care should be offered as for low-risk patients. Immunosuppression with ATG and ciclosporin should be considered for cytopenic patients who are otherwise deemed unfit for intensive chemotherapy, particularly if the marrow is hypocellular with no excess of blasts or adverse cytogenetics. For younger patients below the age of 50 who have either a sibling or unrelated donor available, allogeneic transplantation should be offered as a potentially curative procedure. Whether the conditioning regimen should be myeloablative or non-myeloablative is unclear and should be at the discretion of the local transplant unit. For patients between 50 and 65 years of age, the optimal management is less straightforward, given the significant TRM-associated with allogeneic SCT and the predicted median survival of 5.2 years for patients less than 60. The impressive data emerging for patients treated with RIC transplantation makes this an attractive treatment option, especially when a sibling donor is available, although unrelated donor transplants are also feasible. While trials demonstrate better outcomes if the transplantation is performed prior to disease progression, this is not always a straightforward decision and some patients may reasonably elect to reserve this option for when there is evidence of clonal evolution or early transformation. Iron overload should be avoided in those patients who remain transplant candidates, either by managing anaemia with erythropoietin with or without GCSF or by using iron chelation for those who are transfusion-dependent.
- *IPSS intermediate-2/high*: Patients in these poor prognostic categories are known to have median survivals of less than 2 years for intermediate-2 to only a few months for high-risk individuals. Patients below the age of 65 years and without significant comorbidities, particularly where a sibling donor is available, should be offered intensive chemotherapy and RIC transplantation, with myeloablative conditioning reserved for younger, fitter patients below 50 years at the discretion of the local transplant unit. However, if the patient fails to achieve at least a good partial remission following intensive chemotherapy, then the prospects of successful outcome from allogeneic transplantation are reduced and palliative/experimental treatment with supportive care may be a preferable option, although a sequential chemotherapy-conditioned RIC transplant could be considered in some cases.

With the advent of further novel agents in the future, the most appropriate treatment option is likely to become more complex. However, as with all investigational therapies, patients should be treated within the context of clinical research trials.

Myelodysplastic/myeloproliferative diseases

The WHO classification established a new diagnostic entity for those diseases that share features characteristic of both the myelodysplastic syndromes and the myeloproliferative neoplasms. This overlap category comprises disorders that at the time of initial presentation share clinical, laboratory or morphological findings indicative of underlying dysplastic and proliferative processes. They are usually characterized by hypercellularity of the bone marrow due to proliferation in one or more of the myeloid lineages, with increased numbers of circulating cells that may be morphologically dysplastic. Simultaneously, one or more of the other lineages may exhibit ineffective proliferation so that cytopenias may also be present. The presence of excess blasts is closely correlated with the risk of leukaemic transformation.

The WHO recognizes three distinct entities, namely chronic myelomonocytic leukaemia, atypical chronic myeloid leukaemia and juvenile myelomonocytic leukaemia, and one provisional entity of refractory anaemia with ring sideroblasts and thrombocytosis, with the remainder categorized as myelodysplastic/myeloproliferative neoplasms unclassifiable.

Chronic myelomonocytic leukaemia

CMML is a clonal haemopoietic stem cell disorder with MDS and myeloproliferative overlap features and a propensity to evolve to AML. It is characterized by an absolute monocytosis with an absolute monocyte count of $>1 \times 10^9$/L in peripheral blood. The clonal monocytosis of CMML must be differentiated from reactive monocytosis often caused by viral infections (varicella zoster, arbovirus), chronic infection/inflammatory conditions (tuberculosis, brucellosis, leishmaniasis, subacute bacterial endocarditis, sarcoidosis and connective tissue disorders). CMML has a male predominance (M:F ratio 1.5–3:1) and a median age of presentation of approximately 65–75 years; only 10% of CMML cases occur in individuals less than 60 years. While the aetiology is largely unknown, therapy-related cases of CMML may also occur uncommonly.

Clinical and laboratory features

The WHO 2008 diagnostic criteria for CMML include: (i) persistent peripheral blood monocytosis $>1 \times 10^9$/L, (ii) absence of the Philadelphia chromosome and *BCR–ABL1* fusion

oncogene (as cases resembling CMML may express the p190 isoform, both the p190 and the p210 isoforms should be tested for in the absence of t(9;22)(q34;q11)), (iii) absence of *PDGFRA* or *PDGFRB* gene rearrangements, (iv) less than 20% blasts and promonocytes in the peripheral blood and bone marrow and (v) dysplasia involving one or more lineage. In the absence of dysplasia, persistent monocytosis for greater than 3 months, or the presence of an acquired clonal or molecular abnormality enable the diagnosis to be made. CMML is subclassifed into: CMML-1 (<5% circulating blasts+promonocytes in blood and <10% blasts+promonocytes in the bone marrow) and CMML-2 (5–19% circulating blasts+promonocytes in blood and 10–19% BM blasts or the presence of Auer Rods). Patients with CMML with a peripheral blood eosinophilia $>1.5 \times 10^9$/L may have t(5;12)(q31-q32;p13) leading to the *ETV6 (TEL)-PDGFRB* fusion oncogene which responds to imatinib and is currently classified as a myeloproliferative neoplasm.

Phenotypically CMML may present as an MDS with anaemia, peripheral blood cytopenias, fatigue and bruising. It may also present with myeloid proliferation with leucocytosis, monocytosis, hepatomegaly, splenomegaly and B symptoms such as fatigue, night sweats, weight loss and cachexia. CMML may present with leukaemia cutis. Caution should be made in diagnosing AML when CMML patients have an active infection, as treatment of this may result in reversion to the chronic blood picture. Approximately half of patients have splenomegaly, and often hepatomegaly, at diagnosis. Individuals with high monocyte counts may develop a maculopapular skin infiltration, gum infiltration, and monocytic pleural and pericardial effusions. Lymphadenopathy is uncommon, but when it occurs it may signal a more acute phase with infiltration of lymph nodes by myeloblasts.

Cytogenetic analysis is important for confirming clonality, although abnormalities are only found in 30–40% of cases, notably +8, −7/del(7q) and del(12p), and none is specific for CMML. In a Spanish CMML study, 73% had a normal karyotype, +8,del(5q),+10, del(11q), del(12p), add 17(p), +19, +21, abnormalities of chromosome 7 and complex karyotypes were observed. These were more frequent in patients with increased blasts and dysplastic features. The Spanish cytogenetic risk stratification categorizes patients into high risk (+8, chromosome 7 abnormalities or complex karyotype), intermediate risk (all others not classified as high or low risk) and low risk (normal karyotype, or –Y) with separation of 5-year OS rates of 4%, 26% and 35%, respectively.

Additionally patients with CMML also harbour somatic mutations in similar genes to those discussed in pathogenesis. Of these *EZH2* mutations are less frequent in CMML, whereas spliceosome component mutations *SRSF2* and *U2AF1* are particularly frequent in CMML. SRSF2 is mutated in 28–47% of patients and associates with a normal karyotype, less anaemia and older age. However, it does not have independent prognostic value. SF3B1 mutations are also seen in CMML and ringed sideroblasts, but again do not affect OS or disease progression to AML. *U2AF1* mutations occur in 10% of CMML and are also not independently prognostic. Up to 40% of patients have point mutations of RAS genes, which is higher than for other forms of MDS. Hypermethylation of the CDKN2B gene (which encodes the tumour suppressor p15INK4b), resulting in reduced expression, can be demonstrated in approximately 50% of patients. Recently, the CBL gene, which encodes the E3 ubiquitin ligase, has been implicated in progressive CMML by its presence within a region of UPD on chromosome 11q. GM-CSF-dependent pSTAT5 sensitivity has also been demonstrated in CMML. The hierarchy of mutational architecture is hypothesized to involve epigenetic genes such as *TET2* or *IDH1* or *IDH2* or *ASXL1* at initiation, followed by mutations of the spliceosome component such as SRSF2 leading to mutations in signal transduction genes such as RAS. Mutations of the JAK2 gene should also be excluded.

Bone marrows in CMML are usually hypercellular, with granulocytic hyperplasia and dysplasia. Immunophenotyping may be helpful for identifying myelomonocytic populations with CD13, CD33 positivity, whereas expression of CD14, CD68, CD64 and CD163 may be variable. Immunophenotyping can also give prognostic information, such as reduced CD14 expression indicating monocytic immaturity, aberrant expression of CD2 and CD56, and the CD34-positive cell percentage. Cytochemical stains with non-specific esterase, which are positive in monocytic cells, are the most reliable way of identifying monocytes in the bone marrow, whereas lysozyme and chloracetate esterase are positive in normal granulocyte precursors. Reticulin may be increased in 30% of patients and nodules of mature plasmacytoid dendritic cells (CD123 positive) may occur in 20%. Rarely, bone marrow monocytes without peripheral blood monocytosis may be seen in MDS that later progresses to CMML.

Natural history and prognosis

The clinical course of CMML is highly variable, with a median survival of approximately 2 years, though occasional patients survive beyond 8 years. There is an age-dependent prognosis that can be calculated using the IPSS according to the blast percentage, number of cytopenias and karyotypic abnormalities. The distinction between dysplastic and proliferative forms of the disease is unreliable and unhelpful, but the separation between CMML-1 and CMML-2, which reflects the marrow blast percentage (<10% and 10–19%, respectively), is important and prognostically relevant. Similarly, anaemia and thrombocytopenia are adverse prognostic factors, especially if the patient is transfusion-dependent. A rising lactate dehydrogenase level often reflects a rising blast count, indicating progression to AML with leukaemic transformation occurring in approximately 15–30% of cases. Over the last decade several prognostic scores that include FAB subtype, WHO subtype, haemoglobin, LDH, absolute lymphocyte count, presence of

Table 25.16 CMML prognostic scoring system.

Variable	0	1	2
WHO Subtype	CMML-1	CMML-2	–
FAB Subtype	CMML-MD (WBC < 13)	CMML-MP (WBC > 13)	–
CGN*	Low	Int	High
RBC dependent	No	Yes	–

*Low = normal, −Y; Int = other abnormalities; High = +8, complex (≥3 anomalies), chr 7 anomalies.

Risk group	Overall score
Low	0
Intermediate-1	1
Intermediate-2	2–3
High	4–5

circulating blasts, gender, karyotype and presence of ASXL1 mutations (excluding missense mutation) have been developed (Tables 25.16 and 25.17). These are discriminatory in predicting outcomes, but consensus on which particular system should be used is outstanding.

Management and treatment

Therapy of CMML is unsatisfactory, with the vast majority of patients being elderly (>60 years). Supportive care by way of blood transfusions and antibiotics for infections are the mainstay of management. When the proliferative phase prevails,

Table 25.17 Comparison of scoring systems in CMML.

		Median survival per risk group (months)				
Prognostic score	**Variables included in final model**	**Low**	**Int-1**	**Int-2**	**High**	**Transformation to AML**
MDAPS (Onida *et al.*, 2002)	Haemoglobin <120 g/L Circulating immature cells Absolute lymphocyte count >2.5 × 10^9/L Bone marrow blasts >10%	24	15	8	5	19% developed AML after a median of 7 months
Dusseldorf Score for CMML (Germing *et al.*, 2004)	Bone marrow blasts ≥5% LDH >200 u/L Haemoglobin ≤90 g/L Platelets <100 × 10^9/L	93	26		11	8%, 23% and 23% at 5 years respectively
CMML Prognostic scoring system CPSS (Such *et al.*, 2013)	CMML FAB type (myelodysplastic or myeloproliferative) CMML WHO subtype (CMML1 or 2) CMML specific cytogenetics RBC transfusion dependence	72	31	13	5	13%, 29%, 60% and 73% at 5 years, respectively
GFM (Itzykson *et al.*, 2013)	Age >65 years WBC >15 × 10^9/L Anaemia Platelets <100 × 10^9/L ASXL1 mutation	Not Reached	38.5		14.4	56, 27.4 and 9.2 months AML-free survival
Mayo prognostic model (Patnaik *et al.*, 2013)	Increased absolute monocyte count >10 × 10^9/L Presence of circulating blasts Haemoglobin <100g/L Platelet count <100 × 10^9/L	32	18.5		10	

causing symptomatic organomegaly, infiltrative disease or escalating monocytosis, then oral chemotherapy with hydroxycarbamide (hydroxyurea) is the treatment of choice. A European randomized study comparing hydroxycarbamide with etoposide gave a response rate and survival, respectively, of 60% and 20 months for hydroxycarbamide and 36% and 9 months for etoposide. Intensive chemotherapy alone is of little benefit in this age group. Erythropoietin may be used in CMML, but GCSF should be used cautiously due to the risk of splenic rupture in proliferative patients. Hypomethylating agents such as azacitidine may be used in CMML with a white cell count $<13 \times 10^9$/L and less than 20% blasts (although the numbers of patients in AZA001 were limited). Phase II studies in CMML are encouraging, with overall response rates varying between 25 and 70% and median OS ranging from 12–37 months. Peripheral blasts <5% and absolute monocyte count $<10 \times 10^9$/L were predictive of a response, whereas no consistent correlation between somatic mutations and responses to azacitidine have been reported.

In younger patients, particularly with adverse features, intensive treatment and allogeneic transplantation represent the only possibility of cure. However, the EBMT data for adults with CMML show relatively low long-term outcomes with this approach, with 5-year overall survival after transplant (both related and unrelated) of only 21%. Another study of 85 patients allografted for CMML (32% RIC HSCT, 62% PBSC) showed 42% overall survival at 5 years with over 30% dying from non-relapse causes and just under 25% dying due to relapse. More recent data from the Societe Francaise de Greffe de Moelle et de Therapie Cellulaire in 72 patients (43 RIC HSCT) demonstrated a 3-year OS of 32%, NRM 36% and CIR of 35% with 28% developing grade II–IV aGVHD and 34% cGVHD. As with other forms of MDS, outcomes are better if the patient is transplanted before significant disease progression, but this must be weighed against the considerable TRM rate. A better understanding of the biological basis of CMML may lead to trials that target these anomalies.

Juvenile myelomonocytic leukaemia

Juvenile myelomonocytic leukaemia (JMML) is a clonal haemopoietic disorder of childhood characterized principally by proliferation of the granulocytic and monocytic lineages. It occurs with an incidence of 1.2 per million, comprising approximately 2–3% of all childhood leukaemias, but 40% of childhood MDS. The majority of cases of JMML occur in children under 3 years of age and twice as commonly in boys than girls. There are associations with neurofibromatosis type 1 and Noonan syndrome due to germline mutations in the NF1 and PTPN11 or KRAS genes, respectively. There is marked *in vitro* hypersensitivity of myeloid progenitors to granulocyte/macrophage colony-stimulating factor that is a hallmark feature of JMML and suggestive of defective RAS–MAP kinase signalling that is often attributable to RAS mutations. Monosomy 7 is the most common chromosomal abnormality, found in 25% of cases. A marked increase in the synthesis of HbF is a recurrent finding that has poor prognostic implications.

Clinically, most patients present with constitutional symptoms or evidence of infection and are found to have marked hepatosplenomegaly. Lymphoid and tonsillar enlargement is also common. Typically, there is a leucocytosis comprising neutrophils, myeloid precursors and monocytes, with blasts constituting less than 5% of cells. The marrow is hypercellular and dysplastic features are minimal. The prognosis of JMML is variable, with a median survival of 1 year. It is usually rapidly fatal without treatment, causing organ failure, especially respiratory failure, due to leukaemic infiltration, while blast transformation occurs infrequently. Although responses are seen to cytarabine-containing regimens, allogeneic transplantation offers a cure in up to 50% of patients in some series.

Refractory anaemia with ring sideroblasts and thrombocytosis

The precise nature of refractory anaemia with ring sideroblasts and thrombocytosis (RARS-T) is a controversial and unresolved issue. These patients meet the criteria for RARS, but also have persistently elevated platelet counts over 450×10^9/L. The majority (50–60% of cases) carry the V617F mutation of the JAK2 gene that is more commonly associated with myeloproliferative disorders. As it is unclear whether RARS-T represents a JAK2-positive myeloproliferative neoplasm with acquired dysplastic features or, conversely, a form of MDS with an acquired proliferative mutation, RARS-T exists as a provisional entity in the current version of the WHO classification.

Myelodysplasia of childhood

Childhood MDS is recognized as an entity in its own right in the current version of the WHO classification, but excludes JMML, which is grouped within the myelodysplastic/myeloproliferative category. MDS associated with Down syndrome, which previously accounted for up to 25% of paediatric MDS, is now grouped within a new entity of Down-syndrome-related myeloid leukaemia. Most remaining cases of childhood MDS fall within one of the subgroups of conventional MDS, namely RA, RARS, RCMD or RAEB-1/RAEB-2. The term 'refractory cytopenia of childhood' is reserved for cases of MDS associated with persistent cytopenias, less than 5% blasts in the marrow and less than 2% blasts in peripheral blood. About 75% of children with RCC show considerable hypocellularity of the bone marrow, making it difficult to differentiate from congenital bone marrow failure syndromes that can lead to secondary myelodysplasia in affected children. These congenital syndromes include disorders such as Fanconi's anaemia, dyskeratosis congenita, Schwachman–Diamond syndrome, amegakaryocytic thrombocytopenia, and pancytopenia with radioulnar synostosis. Understanding the mechanism by which they predispose children to developing MDS has helped shed light on the aetiology of

acquired MDS in adulthood. However, a discussion of the biology of these disorders is beyond the scope of this chapter.

Future directions

The identification of somatic driver mutations in the spliceosome and epigenetic machinery has opened the doors to understanding the roles of these critical genes in both normal haemopoiesis and in the pathogenesis of MDS. A deeper understanding of the function of these genes in normal haemopoiesis will in turn reveal mechanisms by which MDS and clonal evolution to AML occur, leading to therapeutic targets. Therapies such as IDH1 and -2 inhibitors, Toll-like receptor antagonists, s100a9 ligand receptors, HDAC inhibitors, such as pacritinib, that are already in trials will be developed. These are likely to complement or supplant existing therapies, which in themselves continue to be refined in trials.

Selected bibliography

Bejar R, Stevenson K, Abdel-Wahab O *et al*. (2011): Clinical effect of point mutatations in myelodysplastic syndromes. *New England Journal of Medicine* **364**: 2496–506.

Collin M, Dickinson R, Bigley V (2015) Haematopoietic and immune defects associated with GATA2 mutation. *British Journal of Haematology* **169**(2): 173–87.

Delhommeau F, Dupont S, Della Valle V *et al*. (2009) Mutation in TET2 in myeloid cancers. *New England Journal of Medicine* **360**: 2289–301.

Ebert BL, Pretz J, Bosco J *et al*. (2008) Identification of RPS14 as a 5q– syndrome gene by RNA interference screen. *Nature* **451**: 335–9.

Fenaux P, Mufti GJ, Hellstrom-Lindberg E *et al*. (2009) Efficacy of azacitidine compared with that of conventional care regimens in the treatment of higher-risk myelodysplastic syndromes: a randomized, open-label, phase III study. *Lancet Oncology* **10**: 223–32.

Greenberg P, Cox C, LeBeau MM *et al*. (1997) International scoring system for evaluating prognosis in myelodysplastic syndromes. *Blood* **89**: 2079–88.

Greenberg P, Teuchler H, Schanz J *et al*. (2012) Revised international prognostic scoring system for myelodysplastic syndromes. *Blood* **120**: 2456–65.

Haferlach T, Nagata Y, Grossman V *et al*. (2014) Landscape of genetic lesions in 944 patients with myelodysplastic syndromes. *Leukaemia* **28**: 241–7.

Jonas BA, Greenberg PL (2015) MDS prognostic scoring systems – past, present, and future. *Best Practice Research Clinical Haematology* **28**(1): 3–13.

Killick SB, Carter C, Culligan D *et al*. (2003) Guidelines for the diagnosis and therapy of adult myelodysplastic syndromes. *British Journal of Haematology* **164**: 503–25.

Kordasti SY, Ingram W, Hayden J *et al*. (2007) CD4+CD25high Foxp3+ regulatory T cells in myelodysplastic syndrome (MDS). *Blood* **110**: 847–50.

Lim ZY, Ho AY, Ingram W *et al*. (2006) Outcomes of alemtuzumab-based reduced intensity conditioning stem cell transplantation using unrelated donors for myelodysplastic syndromes. *British Journal of Haematology* **135**: 201–9.

List A, Kurtin S, Roe DJ *et al*. (2005) Efficacy of lenalidomide in myelodysplastic syndromes. *New England Journal of Medicine* **352**: 549–57.

List A, Dewald G, Bennett J *et al*. (2006) Lenalidomide in the myelodysplastic syndrome with chromosome 5q deletion. *New England Journal of Medicine* **355**: 1456–65.

Maciejewski JP, Mufti GJ (2008) Whole genome scanning as a cytogenetic tool in hematologic malignancies. *Blood* **112**: 965–74.

Malcovati L, Germing U, Kuendgen A *et al*. (2007) Time-dependent prognostic scoring system for predicting survival and leukaemic evolution in myelodysplastic syndromes. *Journal of Clinical Oncology* **25**: 3503–10.

Malcovati L, Hellstrom-Lindberg E, Bowen D *et al*. (2013) Diagnosis and treatment of primary myelodysplastic syndromes in adults: recommendations from the European LeukaemiaNet. *Blood* **122**: 2943–4564.

Mohamedali A, Mufti GJ (2009) Van-den Berghe's 5q– syndrome in 2008. *British Journal of Haematology* **144**: 157–68.

Mohamedali A, Gäken J, Twine NA *et al*. (2007) Prevalence and prognostic significance of allelic imbalance by single-nucleotide polymorphism analysis in low-risk myelodysplastic syndromes. *Blood* **110**: 3365–73.

Mohamedali AM, Smith A, Gaken J *et al*. (2009) Novel TET2 mutations associated with UPD4q24 in myelodysplastic syndrome. *Journal of Clinical Oncology* **27**: 4002–6.

Mufti GJ, Bennett JM, Goasguen J *et al*. (2008) Diagnosis and classification of myelodysplastic syndrome: International Working Group on Morphology of myelodysplastic syndrome (IWGM-MDS) consensus proposals for the definition and enumeration of myeloblasts and ring sideroblasts. *Haematologica* **93**: 1712–17.

Pappaemanuil E, Gerstung M, Malcovati L *et al.* (2013) Clinical and biological implications of driver mutations in myelodysplastic syndromes. *Blood* **122**: 3616–27.

Pappaemanuil E, Cazzola M, Boultwood J *et al.* (2011) Somatic SF3B1 Mutation in myelodysplasia with Ring Sideroblasts. *New England Journal of Medicine* **365**: 1348–95.

Patnaik MM, Tefferi A (2015) Refractory anemia with ring sideroblasts and RARS with thrombocytosis. *American Journal of Hematology* **90**(6): 549–59.

Schanz J, Tuchler H, Solé F *et al*. (2012) New comprehensive cytogenetic scoring system for primary myelodysplastic syndromes (MDS) and oligoblastic acute Myeloid Leukaemia after MDS derived from an international database merge. *Journal of Clinical Oncology* **30**: 820–9.

Schmid C, Schleuning M, Ledderose G, Tischer J, Kolb HJ (2005) Sequential regimen of chemotherapy, reduced-intensity conditioning for allogeneic stem-cell transplantation, and prophylactic donor lymphocyte transfusion in high-risk acute myeloid leukaemia and myelodysplastic syndrome. *Journal of Clinical Oncology* **23**: 5675–87.

Swerdlow SH, Campo E, Harris NL *et al*. (eds) (2008) *WHO Classification of Tumours of Haematopoietic and Lymphoid Tissues*. IARC Press, Lyon.

Van den Berghe H, Cassiman JJ, David G, Fryns JP, Michaux JL, Sokal G (1974) Distinct haematological disorder with deletion of long arm of no. 5 chromosome. *Nature* **251**: 437–8.

CHAPTER 26

Myeloproliferative neoplasms

26

Peter J Campbell[1], Claire Harrison[2] and Anthony R Green[3]

[1]Cancer Genetics and Genomics at the Institute, Wellcome Trust Sanger Institute, Cambridge, UK
[2]Department of Haematology, Guy's and St Thomas' NHS Foundation Trust, London, UK
[3]Cambridge Institute for Medical Research, Wellcome Trust/MRC Stem Cell Institute, and Department of Haematology, University of Cambridge, Cambridge, UK

Introduction

For the purposes of this chapter, the term myeloproliferative neoplasms (MPNs) will refer to clonal disorders of haemopoiesis that lead to an increase in the numbers of one or more mature blood cell progeny. The chronic myeloid leukaemias would fit this definition and share pathogenetic features with some of the MPNs, but have, historically (since the discovery of the Philadelphia chromosome), been studied separately from the MPNs and are described in Chapter 24. The myelodysplastic syndromes (MDS) can also, in a minority of cases, fit our working definition of MPN, in being associated with increased numbers of mature cell progeny, but dysplasia is a major feature and there are, usually, coexisting cytopenias (Chapter 25). Not surprisingly, a small number of patients do not fit neatly into a single category and exhibit features of both MPN and MDS.

This chapter will focus on the classical MPNs: polycythaemia (rubra) vera (PV), essential thrombocythaemia (ET) and primary myelofibrosis (PMF), three related disorders originally grouped together by William Dameshek in 1951. They share clinical, morphological and molecular features and can transform, in their course, into one another. They are clonal disorders of the pluripotent haemopoietic stem cell and have, to varying degrees, the potential to transform into acute myeloid leukaemia (AML). Secondary (non-clonal) polycythaemias and thrombocytoses will also be discussed in this chapter, as they often enter the differential diagnosis of their clonal counterparts.

In addition, some of the less common MPNs will be described, namely mastocytosis and its variants, and the clonal eosinophilic syndromes. The ontogeny of the target cell for transformation is less well established in these disorders, but there is accumulating evidence implicating the pluripotent haemopoietic stem cell in at least some cases.

The polycythaemias

True polycythaemia refers to an absolute increase in total body red cell volume (or mass), which usually manifests itself as a raised haemoglobin concentration (Hb) and/or haematocrit/packed cell volume. A raised Hb (or haematocrit) can also be secondary to a reduction in plasma volume, without an increase in total red cell volume; this is known as apparent (or relative) polycythaemia.

True polycythaemia is further subdivided into primary polycythaemia (in which haemopoiesis is intrinsically abnormal e.g. PV), and secondary polycythaemia, which results from an increased erythropoietin drive, either in the presence or in the absence of hypoxia (Figure 26.1).

Polycythaemia vera

The central pathological feature of PV is an expansion in the total red cell mass, although elevations in the platelet and/or neutrophil counts are relatively common. The first description of PV was by Vaquez in 1892. Osler, in 1903, published the first series of patients, identifying salient clinical features setting PV apart from other erythrocytoses. Considerable information has been gathered about PV since the work of these pioneers, much of it due to the work of the Polycythaemia Vera Study Group

Postgraduate Haematology, Seventh Edition. Edited by A Victor Hoffbrand, Douglas R Higgs, David M Keeling and Atul B Mehta.

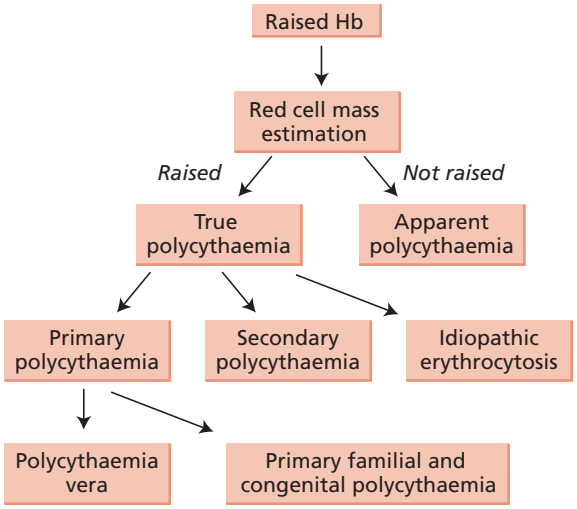


Figure 26.1 An aetiological classification of polycythaemia.

(PVSG), which was set up in 1967 with the aim of optimizing diagnosis and management of PV. The discovery of the V617F mutation in the pseudokinase domain of the tyrosine kinase JAK2, in nearly all cases of PV, was a major advance in our understanding of this disorder.

Pathophysiology

PV is a stem cell disorder characterized by hyperplasia of all three major myeloid cell lineages. The first line of evidence in support of the stem cell origin of PV came in the form of clonality studies. Using X-chromosome inactivation patterns (XCIPs) in the mid-1970s Fialkow and colleagues showed that neutrophils, erythrocytes and platelets originated from the same clone. Large studies have since confirmed these findings.

Erythropoiesis in PV is autonomous and does not rely on erythropoietin (EPO). Plasma levels of this hormone are reduced in PV patients and PV progenitor cells, unlike normal ones, can survive *in vitro* and give rise to erythroid colonies (BFU-E) in the absence of added erythropoietin (endogenous erythroid colonies, EECs). PV erythroid progenitors show an increased sensitivity to EPO, but also to several other growth factors, including insulin-like growth factor-1, thrombopoietin, interleukin-3 and granulocyte/monocyte colony-stimulating factor. Germline mutations in the erythropoietin receptor (EPOR) are known to occur in inherited polycythaemia, but such mutations are absent in patients with PV.

In 2005, several groups identified a unique acquired mutation in the cytoplasmic tyrosine kinase JAK2 in myeloid cells from the great majority of patients with PV. JAK2 lies downstream of several cell-surface receptors including EPOR. Upon EPO binding to EPOR, JAK2 becomes phosphorylated and in turn phosphorylates downstream targets, most important of which are the STATs (signal transducers and activators of transcription), leading to stimulation of erythropoiesis. Valine 617 is located in the JH2 domain of JAK2, which acts to repress its kinase activity (Figure 26.2). The V617F mutation leads to increased kinase activity, confers cytokine independence and results in erythrocytosis in a mouse transplant model. The mutation appears to be fairly specific to the classical MPNs and although it has been reported in small numbers of patients with related myeloid neoplasms, it is not present in lymphoid or non-haemopoietic cancers. Intriguingly, the mutation is homozygous in a large proportion of patients with PV and PMF, but this is rare in ET. Indeed one of the proposed mechanisms for this single mutation occurring in three different entities is that gene dosage influences phenotype with mutation burden being higher in PV and PMF than ET. There is no current role for quantitating mutant allele burden for JAK2 *V617F* or other mutations except perhaps in the post stem cell transplant setting. It appears that the moderate familial predisposition to MPNs can be largely attributed to an inherited haplotype block surrounding the *JAK2* gene (the 46/1 haplotype), although the mechanism underlying this interaction between somatic and germline genetics has not been identified.

When appropriately sensitive methods are used for the detection of the V617F mutation, about 95% of PV patients are positive. Recently, mutations elsewhere in *JAK2* have been described in most of the V617F-negative patients who have PV by strict diagnostic criteria. These mutations cluster in exon 12, and can be of several different types, but all seem to affect the pseudokinase domain, leading to constitutive activation of the JAK2 kinase. Interestingly, compared to PV patients with the V617F mutation, those carrying exon 12 mutations tend to be younger and have a more isolated erythrocytosis, with less frequent thrombocytosis and leucocytosis. The molecular basis for these differences in phenotype has not been identified.

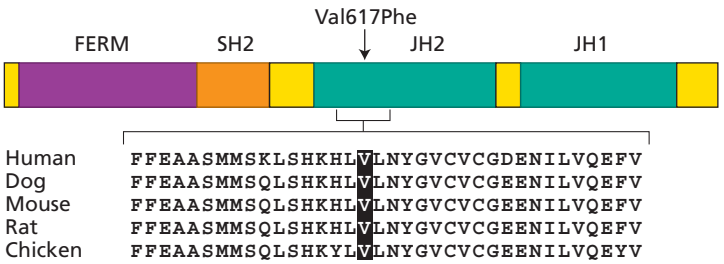


Figure 26.2 Diagrammatic representation of JAK2 indicating the location of valine 617 and the very high degree of cross-species amino-acid homology in its JH2 domain. The JH2 domain normally acts to repress the kinase activity of JAK2, but its ability to do so is impaired in the presence of the JAK2 V617F mutation. (Adapted from Baxter *et al.*, 2005 [*Lancet* 2005; **365**: 1054–61]. Reproduced with permission of Elsevier.)

Truncating mutations and deletions in a tumour-suppressor gene, TET2, have been reported in ~15–20% of patients with MPNs; consequences of these mutations remain unclear but recent data suggests order of acquisition mutations effects phenotype and therapy response. Recurrent somatic mutations in a number of other genes have also been found in a minority of patients with PV. These genes often encode epigenetic regulators, such as *TET2*, and are also mutated in other MPNs, as well as in MDS and AML.

Clinical features

Epidemiology

The annual incidence of PV is reported to be around 2–3 per 100,000 of the population, with a male–female ratio of 1.2:1. The median age at onset is 55–60 years and although incidence increases with age, PV can occur at any age even, rarely, in childhood.

Thrombotic complications

Thrombosis is the most common serious complication of PV. Untreated PV patients run a greatly increased risk of thrombosis, which can be arterial, venous or microvascular. The increased haematocrit leads to an increased blood viscosity, rheological abnormalities and abnormal platelet–endothelial contact. Additionally, procoagulant changes in platelets (e.g. decreased response to prostaglandin D_2), thrombocytosis and pre-existing vascular disease can all conspire to dramatically increase thrombotic risk.

Arterial occlusions can lead to myocardial infarcts, strokes, transient ischaemic attacks, amaurosis, scotomata, and mesenteric and limb ischaemia. Less commonly, microvascular occlusions affecting the extremities and erythromelalgia can occur.

In the venous circulation, unusual sites, such as the splanchnic vessels can be involved. As a result, mesenteric, splenic and hepatoportal thromboses (Figure 26.3) are recognized presenting features of PV. Recent data indicate that this propensity to venous thrombosis in atypical sites is particularly strongly correlated with the presence of the *JAK2* V617F mutation, and indeed patients presenting with otherwise unexplained splanchnic vein thrombosis will often be found to have the mutation, even in the absence of an overt MPN. Superficial thrombophlebitis, conventional deep venous thromboses and pulmonary emboli are also seen.

Neurological features

Over and above the consequences of occlusive vascular lesions, the sluggish cerebral blood flow secondary to the increased haematocrit is thought to underlie features such as headaches, drowsiness, insomnia, amnesia, tinnitus, vertigo, chorea and even depression. Transient visual disturbances also occur.

Pruritus

This symptom occurs in about one-quarter of PV patients and in some it may be severe. It is characteristically aquagenic, precipitated by warm baths and can be associated with erythema, swelling or even pain. Pruritus is often relieved by controlling the haematocrit, but its aetiology remains elusive. Basophilia, hyperhistaminaemia and iron deficiency may have a role and there is an increased incidence in patients with a lower mean corpuscular volume (MCV).

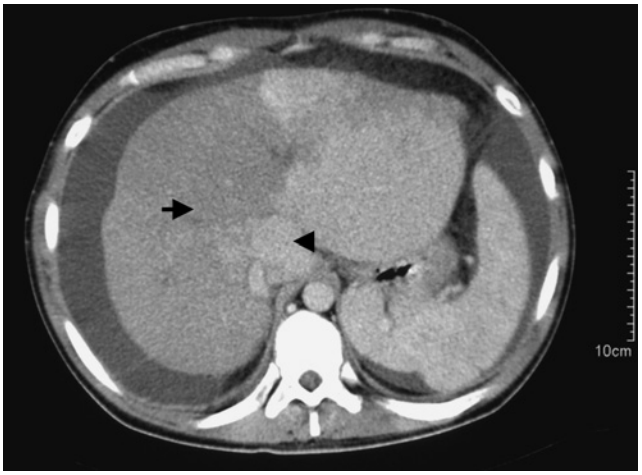


Figure 26.3 Polycythaemia vera presenting with the Budd–Chiari syndrome in a 28-year-old man: contrast computerized tomography showing reduced enhancement of the right lobe of the liver (arrow) with characteristic sparing of the caudate lobe (arrowhead). On this occasion, the left lobe was also relatively spared. Marked ascites and a bulky spleen are also seen.

Skin

Plethora, dilated conjunctival vessels and rosacea-like facial skin changes are not uncommon at presentation. Brown discoloration of the skin, erythromelalgia and, rarely, Sweet's syndrome may be seen.

Splenomegaly

Palpable splenomegaly is seen in 30–50% of cases of PV. It is unclear if its presence affects prognosis, but it may be associated with an increased risk of progression to myelofibrosis.

Hypertension and gout

Hypertension is probably more common in patients with PV, as is hyperuricaemia, with gout seen in about 5% of cases.

Leukaemic transformation

This is perhaps the most feared complication of PV, but the risk of developing acute leukaemia in PV patients treated only with venesection is thought to be small in the short term (1–3%). This risk, however, increases dramatically (more than tenfold) when radioactive phosphorus (^{32}P), chlorambucil or irradiation are used as treatment. The median time interval between first starting such therapy and developing acute leukaemia is 5–8 years. Interestingly, in about half of patients transforming to acute

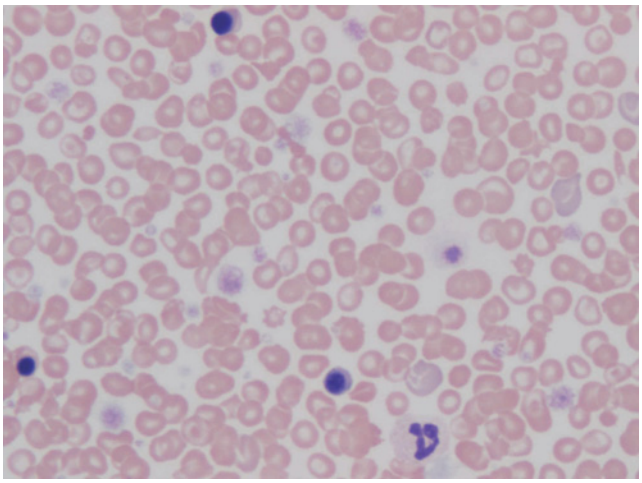


Figure 26.4 Blood film from a case of postpolycythaemic myelofibrosis after splenectomy. Note the presence of nucleated erythrocytes, giant platelets and features of splenectomy including target cells, spherocytes and acanthocytes.

leukaemia from a preceding *JAK2*-positive MPN, the leukaemia is *JAK2*-negative. Recent long-term studies suggest that the risk of myelodysplasia (MDS) or acute myeloid leukaemia (AML) in PV patients treated with hydroxycarbamide, is approximately 20% at 20 years, although it is not clear whether this represents an effect of the drug or part of the natural history of the disease.

Myelofibrosis

Progression to myelofibrosis, so-called post-PV myelofibrosis (PPV-MF) (Figure 26.4) occurs in around 10–20% of PV cases at 15 years after diagnosis. This figure is approximate, not least because different studies have used distinct criteria to define myelofibrotic transformation. Transformation often occurs gradually over many years, is thought to reflect accumulation of additional mutations and is associated with an increased risk of leukaemic conversion. The management of these patients is similar to that of PMF.

Investigations

The diagnosis of PV requires the identification of features in support of this diagnosis, as well as the exclusion of secondary and apparent polycythaemia. The original set of diagnostic criteria was formulated by the PVSG in the 1970s. The BCSH guidelines for diagnosing PV, taking into account the latest advances in the molecular understanding of these disorders, are shown in Table 26.1.

Most patients can now be diagnosed on the basis of a raised haematocrit (>0.52 for men and >0.48 for women) together with the presence of the *JAK2* V617F mutation. It is also important to emphasize that the V617F mutation has been identified in a small proportion of other haematological malignancies (for example, AML, chronic myelomonocytic leukemia and MDS). However, none of these disorders is associated with a raised red cell mass and clinical distinction from PV is rarely an issue. The role of bone marrow biopsy in the evaluation of PV remains controversial, but it may be useful as a baseline investigation for later comparison in younger patients with the disorder, or if there are atypical clinical and laboratory features at presentation.

A minority of patients with clinical PV is negative for the V617F mutation, even when tested using sensitive detection methods. Many reference laboratories now offer highly sensitive molecular testing for *JAK2* exon 12 mutations, and this should be considered in patients who have a typical clinical and laboratory presentation of PV, but who are negative for the V617F mutation. Rarely, patients will be negative for both V617F and exon 12 *JAK2* mutations. In these patients, the diagnosis of PV can be made if the other criteria set out in Table 26.1 are met, but it should be remembered that this group of truly *JAK2*-negative polycythaemia vera is very rare, and secondary erythrocytosis, idiopathic erythrocytosis and relative erythrocytosis are all much more likely diagnoses in this clinical setting.

Table 26.1 Diagnostic criteria for polycythaemia vera, as recommended in the BCSH guidelines.

***JAK2*-positive polycythaemia vera**	
A1	High haematocrit (>0.52 in men, >0.48 in women) OR raised red cell mass (>25% above predicted)*
A2	Mutation in *JAK2*
Diagnosis requires both criteria to be present	
***JAK2*-negative polycythaemia vera**	
A1	Raised red cell mass (>25% above predicted) OR haematocrit >0.60 in men, >0.56 in women.
A2	Absence of mutation in *JAK2*
A3	No cause of secondary erythrocytosis
A4	Palpable splenomegaly
A5	Presence of an acquired genetic abnormality (excluding BCR–ABL) in the haematopoietic cells
B1	Thrombocytosis (platelet count >450 × 10^9/L)
B2	Neutrophil leucocytosis (neutrophil count >10 × 10^9/L in non-smokers; >12.5 × 10^9/L in smokers)
B3	Radiological evidence of splenomegaly
B4	Endogenous erythroid colonies or low serum erythropoietin
Diagnosis requires A1 + A2 + A3 + either another A or two B criteria	

*Dual pathology (coexistent secondary erythrocytosis or relative erythrocytosis) may rarely be present in patients with a JAK2-positive MPN. In this situation, it would be prudent to reduce the haematocrit to the same target as for polycythaemia vera.

Treatment

Identification of the JAK2 V617F mutation has generated a lot of interest in the development of therapeutic JAK2 inhibitors. However, as patients with PV currently have a very good prognosis, new agents will have to display an excellent safety profile. Preliminary data is now becoming available in this setting, as discussed below. For the time being, the treatment of PV should employ existing treatment modalities whose effectiveness has been validated, as described below.

In the absence of extreme leucocytosis or thrombocytosis, progressive splenomegaly or thrombosis, regular venesection remains the mainstay of treatment for PV in patients who can tolerate it. A target haematocrit of ≤0.45 is widely used, following the demonstration that in patients with PV, higher haematocrit targets are associated with a significantly increased risk of thrombosis. This has recently been confirmed in a randomized controlled trial, CYTO-PV. Venesection has little impact on the haematocrit in the short term: the purpose of regular venesection is to induce iron deficiency, such that the haematocrit remains chronically below the target threshold of 0.45. Thus, typical venesection regimens start with phlebotomy every 2–3 weeks until the haematocrit is controlled, and thereafter phlebotomy is generally needed every 1–3 months, depending upon factors such as dietary iron intake and erythropoietic activity.

Cytoreductive therapy is recommended for patients unable to undergo venesection and those with marked thrombocytosis, leucocytosis, and either progressive splenomegaly or prior thrombosis. However, the level of thrombocytosis, leucocytosis or splenomegaly that warrants cytoreduction is not well defined. Hydroxycarbamide (hydroxyurea) is the most commonly used drug, it is orally bioavailable and generally very well tolerated, and it will reduce both the haematocrit, leucocyte count and platelet count. The commonest complications are leucopenia or thrombocytopenia, which are dose dependent and can usually be avoided by close monitoring of the blood count when the drug is first introduced. In susceptible patients it can cause photosensitivity, painful leg ulcers and gastrointestinal side-effects. Other important cutaneous toxicities of hydroxycarbamide include increased risk of sun-induced squamous skin lesions, these may be both cancerous and precancerous, and pigmentation can also occur. Patients should be warned about these effects and counselled about sun exposure. Hydroxycarbamide is also genotoxic and as a radiomimetic it should be stopped during radiotherapy. The usual dose is 0.5–2 g daily.

It has been suggested that hydroxycarbamide may increase the inherent leukaemogenic risk associated with PV. This concern is largely based on studies involving small numbers of patients or patients who have also required other cytotoxic agents (and may therefore represent a subgroup with more aggressive disease). At present, there are no convincing data to show that hydroxycarbamide, when used as a single agent, significantly increases the risk of leukaemia, although a small effect cannot be excluded. The only available comparator studies have shown that the comparator treatments (e.g. pipobroman) are more leukaemogenic. Preliminary data from studies of hydroxycarbamide in sickle cell disease are reassuring, with most recently reported follow-up of 17.5 years; however, sickle cell disease is not an acquired clonal disease with an inherent tendency to leukaemic transformation, so whilst this data is reassuring it is not directly transferable to MPN patient populations.

Interferon-α is effective in controlling both the platelet/leucocyte count and the haematocrit, and there are some data that it may induce impressive reductions in the burden of V617F-positive cells in the blood. It is not widely used because of its cost, route of administration (subcutaneous injection) and its side-effects (including fatigue, flu-like symptoms, depression, autoimmune phenomena). It can be useful, however, in young patients who are reluctant to take other cytotoxic agents, in pregnancy and in patients with intractable pruritus. The usual dose range is 3–5 mU three times per week. There is increasing interest in pegylated formulations of interferon, which may be better tolerated, but comparator studies with any type of interferon are currently lacking.

Anagrelide can be useful in controlling the platelet count and can be combined with hydroxycarbamide allowing lower doses of both agents. Approximately 10% of patients are completely refractory to anagrelide. The usual dose is 1–2 mg daily, but occasional patients may require doses of up to 8 mg daily. Its side-effects are mainly secondary to its inotropic and vasodilatory properties (e.g. headaches, palpitations and fluid retention) although marked anaemia may also occur. Anagrelide is currently only approved for patients with ET and as second-line therapy.

Ruxolitinib continues to be investigated in PV, it has been approved for PV patients resistant or intolerant to hydroxycarbamide but its effects on thrombosis and transformation are unclear.

Busulfan is sometimes used in elderly patients or when all other treatments are not tolerated. It is very convenient as it need only be administered intermittently, but may increase the risk of leukaemia. The usual dosage is 25–75 mg as a single dose every 2–3 months, or 2–4 mg daily for 7–14 days repeated every 4 weeks until the target blood count is reached.

Low-dose aspirin (75–100 mg daily) reduces thrombotic complications in PV and is used in most patients without contraindications to this drug. Its use is supported by a randomized study (ECLAP).

Pruritus often improves with control of the haematocrit, but paroxetine, antihistamines and aspirin (in some cases) can help. There are also reports that psoralen-UV light therapy, IFN or the JAK1/2 inhibitor ruxolitinib can be useful in intractable cases.

In view of the age of most patients and the relatively benign natural history of treated PV, bone marrow transplantation is not advocated for stable disease. The role, if any, of transplants

employing reduced-intensity conditioning regimes is not yet clear.

Prognosis

In the first half of the twentieth century, untreated polycythaemia had a dismal prognosis, with a 50% survival of less than 2 years. However, adequately treated PV now has a relatively benign natural history with a life expectancy of over 11 years, bearing in mind that the average age of onset is 60 years. Factors predictive of poorer prognosis and increased complication rates in PV include *JAK2* mutation burden and white cell count at diagnosis, both factors probably serving as surrogate markers of disease activity.

Other causes of erythrocytosis

All disorders with an increased red cell mass which are not due to a clonal proliferation of haemopoietic progenitors are included under this heading. They are most conveniently subclassified into primary and secondary causes. In primary polycythaemia, the defect is intrinsic to the red cell precursors, which are hypersensitive to erythropoietin. In secondary polycythaemia, the defect is upstream of the red cell precursors. The latter group can be further subdivided into polycythaemias in the presence or in the absence of systemic hypoxia. A small group of patients do not fall into any of these categories and are given the diagnosis of idiopathic erythrocytosis (IE) (Table 26.2). The clinical management of many of these syndromes is not well defined.

The term apparent polycythaemia refers to a raised haematocrit in the absence of a raised red cell mass and is discussed later in this chapter.

Table 26.2 Causes of an erythrocytosis.

Causes of absolute erythrocytosis (i.e. red cell mass 125% of predicted)

Primary (abnormality within RBCs)

- Congenital
  - Truncated erythropoietin receptor
  - Mutations within *LNK*
- Acquired
  - Polycythaemia vera

Secondary (abnormality outside RBCs)

- Congenital
  - Abnormal haemoglobin with increased oxygen affinity
  - Reduced 2,3-bisphosphoglycerate
  - Mutation in von Hippel–Lindau (VHL) gene
  - Mutations in proline dehydroxylase genes
  - Mutations in hypoxia inducible factor (HIF) genes
- Acquired (increased erythropoietin)
  - Conditions causing low oxygen levels – high altitude, chronic lung disease, some congenital heart diseases
  - Renal disease – tumours (hypernephroma), cysts (usually benign), hydronephrosis, following kidney transplantation
  - Liver disease – hepatoma, cirrhosis, hepatitis
  - Tumours – bronchial cancer, fibroids in the uterus, cerebellar haemangiomata
  - Endocrine abnormalities – Cushing's syndrome, phaeochromocytoma
  - Drugs – erythropoietin and androgens

Idiopathic (undefined primary or secondary)

- May resolve or pathology may be masked initially

Causes of apparent erythrocytosis (i.e. normal red cell mass, but elevated haematocrit)

- Normal variant
- Early absolute erythrocytosis
- Obesity, fluid loss, diuretics, smoking, hypertension, alcohol, renal disease, psychological stress

Primary erythrocytosis

Inherited/congenital erythrocytosis

Primary familial and congenital polycythaemia (PFCP) is a rare disorder in which erythropoiesis is intrinsically overactive. The disorder is usually transmitted in an autosomal dominant manner, with some cases appearing sporadically. Clinical features include the presence of isolated erythrocytosis without evolution into leukaemia or other myeloproliferative neoplasms, absence of splenomegaly, normal white blood cell and platelet counts, low or normal plasma erythropoietin (Epo) levels, normal haemoglobin–oxygen dissociation curve/P_{50}, and hypersensitivity of erythroid progenitors to Epo. Mutations in the gene encoding the erythropoietin receptor (EpoR) have been described in several (but not all) families with PFCP. In most cases, the mutations lead to a C-terminal truncation of the EpoR protein, with increased sensitivity to Epo.

Secondary erythrocytosis

Erythrocytosis in the presence of systemic hypoxia

Chronic lung disease and hypopnoea

Lung disease is the predominant cause of chronic systemic hypoxia at sea level. Hypoxaemic chronic obstructive pulmonary disease (COPD) is the commonest syndrome, but any lung–airway disease leading to chronic hypoxia could cause polycythaemia. Syndromes such as obstructive sleep apnoea and hypoventilation due to muscle weakness or paralysis can also occasionally be associated with secondary erythrocytosis.

Where possible, hypoxia should be ameliorated by treating the lung disease or with home oxygen therapy. Erythrocytosis has opposing effects on oxygen delivery as it increases the oxygen-carrying capacity, while also increasing blood viscosity. Unfortunately, there is little evidence from clinical trials to guide management. In practice, many specialists suggest that venesection should be performed for haematocrits above 0.55, with a target of 0.50.

High altitude

Residents at altitudes above 4000 m compensate for the ambient hypoxia by multisystem adaptation, including mild erythrocytosis, increase in capillary perfusion and lung diffusion capacity, as well as biochemical changes in metabolic enzymes and myoglobin. Excessive altitude polycythaemia (haematocrit ≥0.65) is seen in a proportion of cases and is often accompanied by hyperuricaemia and proteinuria. Some of these individuals eventually decompensate and develop chronic mountain sickness (Monge's disease). Such people deteriorate steadily and develop extreme polycythaemia (sometimes haematocrit ≥0.75), arterial desaturation and right heart failure. Resettlement at lower altitudes halts disease progression and can partly reverse it. Treatment with angiotensin-converting enzyme (ACE) inhibitors has been shown in randomized studies to reduce the haematocrit and proteinuria seen in excessive altitude polycythaemia.

Congenital cyanotic heart disease

Congenital heart defects leading to a right-to-left shunt can cause dramatic erythrocytosis (up to a haematocrit of 0.80). Surgery to correct the cardiac defect should be undertaken when possible. A few inoperable patients survive to adulthood and the management of their erythrocytosis is not straightforward. As in patients with chronic lung disease, the increase in oxygen-carrying capacity afforded by erythrocytosis is countered by an increased viscosity and associated haemodynamic changes. Here, however, we are often dealing with young patients with responsive vasculatures, which can usually accommodate such changes. Experts advocate allowing the haematocrit to rise further and venesect for symptoms such as recurrent haemoptysis, marked fatigue or deteriorating exercise tolerance.

High-affinity haemoglobins (see also Chapter 6)

High-affinity haemoglobins release less oxygen for a given oxygen partial pressure, and may thus give rise to tissue hypoxia. This leads to an erythropoietin-driven erythrocytosis, which tends to re-normalize erythropoietin levels. The pathognomonic anomaly is a left shift in the oxygen dissociation curve (Figure 26.5). The precise variant can be identified by mutational screening of DNA or by protein mass spectrometry.

There are over 40 haemoglobin variants with an increased affinity for oxygen, all dominantly inherited. Most are due to mutations in β-globin, with a small number due to mutations in α-globin. Mutations are clustered in regions of the globin chains involved in the regulation of the transition between tense (T) and relaxed (R) states of haemoglobin. Normally, oxy-HbA is in the T state and has low affinity for oxygen, and deoxy-HbA is in the R state and has a high affinity for oxygen. Mutations at the αβ contact site (e.g. Hb San Diego), the C-terminus (e.g. Hb Bethesda) and the 2,3-BPG (2,3-bisphosphoglycerate) binding site (e.g. Hb Helsinki) are the commonest causes of this finding.

Most people with a high-affinity Hb are in good health and are either diagnosed coincidentally or after being noticed to be plethoric. Some experience excessive muscle fatigue after vigorous exercise. Hyperviscosity is rarely a problem and surveys have failed to identify increased cardiovascular morbidity or mortality. Pregnancy is not adversely affected, even in mothers with haemoglobin affinity that exceeds that of HbF. Generally, therefore, no management is required, unless patients are symptomatic.

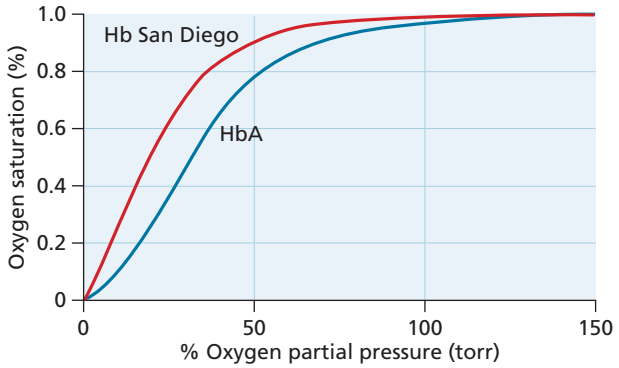


Figure 26.5 Haemoglobin–oxygen dissociation curves from a 27-year-old man with a raised red cell mass and a normal person (HbA) showing the presence of a high-affinity Hb ('left shift'). Mass spectrometric analysis showed the man to be heterozygous for Hb San Diego, a high-affinity β-chain variant.

Red cell metabolic defects

Very rare cases of erythrocytosis are due to abnormalities in red cell metabolism that lead to a reduction in intraerythrocytic 2,3-DPG. The best-characterized defect is a mutant 2,3-DPG mutase. The only well-characterized family with this disorder showed an autosomal recessive inheritance, although heterozygous family members had a decreased P_{50} and, in some cases, a moderate erythrocytosis. This disorder is excluded by the finding of a normal P_{50} in a fresh blood sample.

Methaemoglobinaemia

Oxidation of haem iron converts it from its normal ferrous (Fe^{2+}) to the ferric (Fe^{3+}) form and, correspondingly, haemoglobin (HbA) becomes methaemoglobin (Met-HbA). This constitutes an important antioxidant mechanism for the red cell, and conversion of Met-HbA back to HbA requires the generation of NADH from glycolysis. The rate of HbA autooxidation is about 20 times slower under normal circumstances than the rate of Met-HbA reduction, thus preventing Met-HbA accumulation.

Methaemoglobin has an increased affinity for oxygen and a left-shifted oxygen dissociation curve. Pathological acquired methaemoglobinaemia can result from exposure to strong oxidants (e.g. dapsone, paraquat, benzocaine) and can be life-threatening when severe, but is rarely sufficiently long-lived to give rise to polycythaemia.

Hereditary methaemoglobinaemias can be due to haemoglobin mutations involving amino acids around the haem pocket (haemoglobin M disease) (Chapter 6), or secondary to enzymatic deficiencies that interfere with the generation of NADH, which is required for day-to-day methaemoglobin reduction (namely NADH reductase and cytochrome *b*5 reductase). In view of the chronic nature of the methaemoglobinaemia in these disorders, secondary erythrocytosis can develop and is likely best managed in the same manner as for high-affinity haemoglobin mutants.

Heavy smoking

Heavy smoking can lead to mild polycythaemia in the absence of hypoxic lung disease. The underlying cause of this is a raised carboxyhaemoglobin (CO-Hb) level, resulting from chronically raised carbon monoxide (CO) levels. CO-Hb levels in urban-dwelling non-smokers are rarely higher than 2%, with levels ranging from 3% to 20% in smokers. The short half-life of CO in the body (3–5 h) leads, in smokers, to a rise in CO-Hb during the day and a fall during sleep, making it difficult to compare measurements taken at different times. Binding of CO to Hb, as well as displacing O_2, leads to a conformational change similar to that seen in methaemoglobinaemia, with a similar left shift in the oxygen dissociation curve and a fall in P_{50}. Venesection is usually not used in these patients unless there is a clear thrombotic history or high risk.

Erythrocytosis in the absence of systemic hypoxia

Chuvash polycythaemia (CP) is an autosomal recessive condition that is endemic in the Russian mid-Volga river region of Chuvashia. Patients have increased levels of circulating erythropoietin but do not carry mutations of the Epo receptor. CP was recently shown to be associated with an Arg200Tr*p* substitution in the von Hippel–Lindau (VHL) gene. The *VHL* gene is pivotal for ubiquitination and subsequent degradation of HIF-1 transcription factor, which is central to the oxygen-sensing pathway (Chapter 2). This *VHL* mutant leads to a reduced rate of degradation of HIF-1 and upregulation of downstream targets including Epo, leading to polycythaemia. Recently, some of the rare non-Russian families with inherited polycythaemia were also shown to carry *VHL* mutations, and occasional pedigrees with mutations elsewhere in the oxygen-sensing pathway have been described, including the *PHD2* and *HIF-2A* genes.

Abnormal erythropoietin secretion

Abnormal Epo secretion is a well-recognized cause of secondary erythrocytosis and is most commonly secondary to renal pathologies such as renal tumours (benign and malignant), polycystic kidney disease and diseases associated with local hypoxia, such as renal cysts, hydronephrosis and renal artery stenosis. The polycythaemia usually responds to treatment of the underlying renal pathology.

Erythrocytosis occurs in 20–30% of patients after renal transplantation. The biggest risks to such patients are hypertension, strokes and cardiovascular complications. In many cases, the erythrocytosis and associated hypertension respond to ACE inhibitors. Theophylline may also be effective in some cases. Patients who remain polycythaemic despite such treatments should be treated with repeated venesections to maintain their haematocrit below 0.45; 30–40% of cases resolve spontaneously.

Non-renal tumours can rarely be associated with polycythaemia. The commonest reported ones are hepatocellular carcinoma, cerebellar and other haemangiomata and large uterine fibromyomata. Polycythaemia responds to removal of the tumour in most of these cases.

Endocrine disorders

The mechanism underlying the development of polycythaemia in most endocrine disorders lies in the over-production of androgens, which can produce polycythaemia by increasing Epo levels and also, probably, through a direct action on bone marrow progenitors. Uncontrolled diabetes is also an important cause, which is usually easily identified.

Idiopathic erythrocytosis

In a small proportion of patients with polycythaemia the criteria for the diagnosis of PV are not met and no other aetiology for the raised red cell mass can be identified. This group is heterogeneous and likely to include patients with germline mutations causing polycythaemia as well as some that will go on to develop overt PV. With the advent of increasingly sophisticated diagnostic tests and the identification of the molecular lesions in many inherited forms of polycythaemia, idiopathic erythrocytosis is becoming a rare entity. These patients are usually treated with low-dose aspirin unless this is contraindicated and venesection instituted to a target of 0.45.

Apparent polycythaemia

Apparent polycythaemia refers to a raised haematocrit in the presence of a normal red cell volume (less than 25% above the predicted mean normal value). The raised haematocrit is due to a reduction in the plasma volume. Smoking, hypertension, obesity, excessive alcohol and diuretic therapy have all been associated with apparent polycythaemia. Pathogenesis is uncertain and almost certainly heterogeneous.

It is not clear whether apparent polycythaemia is associated with increased rates of thrombosis, but it seems sensible to encourage affected individuals to avoid known predisposing factors. There are no convincing data that routine venesection is beneficial, but a haematocrit beyond 0.54 is likely to be associated with thrombotic risk and many haematologists would maintain the haematocrit below this level to an individually tailored target.

Essential thrombocythaemia

The fundamental pathological feature of essential thrombocythaemia (ET) is a persistent elevation in the platelet count. However, ET has been poorly understood largely because of a lack of positive diagnostic criteria, together with the fact that cases labelled as ET are likely to be pathogenetically heterogeneous. In 1934 Epstein and Goedel first described a patient with persistent elevation of the platelet count in association with megakaryocyte hyperplasia and tendency for venous thromboses and haemorrhage. Subsequently Ozer and Gunz independently described two series of patients in 1960, thus confirming ET as a specific clinical entity. The discovery that approximately half the cases of ET carry the *JAK2* V617F mutation, as do half the cases of PMF and nearly all cases of PV, has enhanced our understanding of the relationship between the three disorders. Many JAK2-unmutated patients with ET or PMF harbor mutations in calreticulin (*CALR*), an endoplasmic reticulum chaperone, and a small proportion carry mutations of the thrombopoietin receptor *MPL*. A number of different mutations occur in the *CALR;* they all affect a common region of the gene exon 9. The commonest are a 52-bp deletion, so-called type-1 mutation, or a 5-bp insertion, so-called type-2. Thus far the functional consequences of these different mutations are unclear, but type-1 mutations appear to predominate in PMF and type-2 in ET.

Pathophysiology

X-chromosome inactivation patterns (XCIP) provided the first evidence that ET may be a clonal stem cell disorder involving granulocytes, platelets and red cells, but not T cells. For 80–90% of patients, a molecular basis for the clonality can be found, with ~55%, 30% and 5% of ET patients positive for the *JAK2*, CALR and MPL mutations, respectively.

Clinical and pathological features vary significantly between patients with ET, suggesting that the disease is heterogeneous. Prospective data from over 800 patients with ET has demonstrated that the presence or absence of the JAK2 V617F mutation divide ET into two biologically distinct disorders. Mutation-positive ET exists along a continuum with PV, as it displays multiple features of the latter, with significantly increased haemoglobin levels, neutrophil counts and bone marrow erythropoiesis, more venous thromboses and a higher incidence of polycythaemic transformation. Mutation-negative patients do nonetheless exhibit many clinical and laboratory features characteristic of a myeloproliferative neoplasm, including the presence of endogenous erythroid colonies and a risk of transformation to acute leukaemia. Compared to JAK2-mutant ET, patients with an MPL mutation present with a more isolated thrombocytosis and less hypercellular bone marrow, whereas those with a *CALR* mutation have lower haemoglobins, higher platelet counts, less risk of thrombosis and perhaps a higher risk of myelofibrotic transformation. Patients lacking all three mutations (triple-negative) are often young and also have a lower thrombosis risk.

Clinical features

Epidemiology

The annual incidence of ET is similar to that of PV at around 1.5–2.0 cases per 100,000 of the population. The median age at onset is 50–55 years, with a small second peak in women of reproductive age and, although it can occur at any age, it is rare in childhood.

Thrombotic complications

As with PV, thrombotic complications are the main cause of morbidity and mortality in ET. Thromboses are present in around 15–20% of patients at presentation and may be arterial or venous. The range of clinical syndromes is similar to PV, but the frequency of splanchnic thromboses is probably lower and strongly correlated with presence of the *JAK2* V617F mutation.

A number of risk factors are associated with an increase in the risk of thrombosis in patients with ET. The best characterized are aged over 60 years and have a prior history of thrombosis. Other risk factors for thrombosis in ET are likely to include diabetes, hyperlipidaemia, hypertension and cigarette smoking. More recently, it has been reported that white cell count at diagnosis and reticulin levels at diagnosis have predictive value for thrombosis (and indeed other complications) in ET. The exact role these and other predictors have in individualizing treatment regimens remains unclear.

Haemorrhagic complications

Bleeding is less common and less well studied than thrombosis in ET, but can be dramatic when it happens. Efforts to correlate the thrombotic risk to platelet function abnormalities have generally been fruitless and this investigation is also unable to predict haemorrhagic risk. Bleeding is, however, more common in patients with platelet counts above 1000 and, in at least some cases, this is due to an acquired von Willebrand disease, with a decrease in circulating high-molecular-weight multimers caused by adsorption to the surface of the excessive platelets. Routine testing for acquired von Willibrand disease is not generally recommended, however.

Splenomegaly and hyposplenism

Splenomegaly is present in about 5% of ET patients at diagnosis and it is rarely more than mild. Progressive enlargement of the spleen during the course of ET should raise suspicion of evolving myelofibrosis. It has been suggested that over time, some patients with ET develop splenic atrophy secondary to silent microinfarcts in the splenic microcirculation. Frank hyposplenism and its complications are rare, however.

Transformation to myelofibrosis or polycythaemia vera

Transformation to myelofibrosis (termed post-ET myelofibrosis) and, more rarely, to PV, are recognized complications of ET. Some, but not all, cases of apparent polycythaemic transformation may represent resolution of prior iron deficiency, as can happen with iron supplementation and after the menopause. The insidious onset of myelofibrotic transformation and the reluctance to serially study bone marrow trephine biopsies have hampered attempts to define its nature and frequency. Nonetheless, recent studies have shown that progression of reticulin levels over time shows extensive interindividual variability, and can be influenced by choice of therapy, being more marked in patients treated with anagrelide (see below). From the available data, myelofibrotic transformation occurs in less than 10% and polycythaemic transformation in less than 1–2% of ET patients over 10 years.

Leukaemic transformation

ET can evolve into an MDS or AML, even in untreated cases, but only rarely. The presence of cytogenetic abnormalities and treatment with alkylating agents increase this risk. Approximately 3% of patients treated with hydroxycarbamide alone develop MDS or AML if followed for a median time of 8 years. As with PV, there are no data that demonstrate that hydroxycarbamide as a single agent significantly increases the risk of leukaemia inherent to this disease, but a small effect cannot be excluded. Data from the Swedish cancer registry suggests that up to 25% of MPN patients who develop AML have not been treated with a cytoreductive agent.

Investigations

The lack of pathognomonic features and the existence of many other causes of a raised platelet count have posed significant hurdles in the diagnosis of ET. The identification of a *JAK2* V617F, CALR or *MPL* mutation now provides a very useful positive diagnostic criterion for approximately 80–90% of ET patients. However, for so-called triple-negative patients ET remains a diagnosis of exclusion and one which can only be made after other clonal blood disorders and reactive thrombocytosis have been ruled out. A proposed diagnostic schema for ET, as outlined by the BCSH guidelines, is given in Table 26.3.

An alternative set of diagnostic criteria, based largely around salient bone marrow morphological features, have been proposed as part of the 'WHO classification of tumours'. Included in these is the concept of 'prefibrotic myelofibrosis', a putative group of patients previously labeled as ET, who reportedly have increased rates of transformation to myelofibrosis over time. These proposals remain highly controversial in the MPN literature. In particular, studies of interobserver reliability for the histological component of this classification have shown it to be poorly reproducible, and a prospective multicentre study of the prognostic discrimination achieved by such a label found it to be minimal. Nevertheless, bone marrow histological features such as giant, multilobated megakaryocytes and megakaryocyte clustering (Figure 26.6) can be of value in making the diagnosis of ET.

Table 26.3 Diagnostic criteria for essential thrombocythaemia (ET), as recommended in the BCSH guidelines for ET.

Diagnosis requires A1–A3 or A1 + A3–A5	
A1	Sustained platelet count >450 × 10^9/L
A2	Presence of an acquired pathogenetic mutation (e.g. in the *JAK, CALR* or *MPL* genes)
A3	No other myeloid malignancy, especially PV[1], PMF[2], CML[3] or MDS[4]
A4	No reactive cause for thrombocytosis and normal iron stores
A5	Bone marrow aspirate and trephine biopsy showing increased megakaryocyte numbers displaying a spectrum of morphology with predominant large megakaryocytes with hyperlobated nuclei and abundant cytoplasm. Reticulin is generally not increased (grades 0–2/4 or grade 0/3).

[1]Excluded by a normal haematocrit in an iron-replete patient.
[2]Indicated by presence of significant bone marrow fibrosis (greater or equal to 2/3 or 3/4 reticulin) AND palpable splenomegaly, blood film abnormalities (circulating progenitors and tear-drop cells) or unexplained anaemia.
[3]Excluded by absence of *BCR–ABL1* fusion from bone marrow or peripheral blood.
[4]Excluded by absence of dysplasia on examination of blood film and bone marrow aspirate.

Reactive thrombocytosis

Thrombocytosis is most commonly reactive and secondary to increased levels of circulating cytokines that stimulate thrombopoiesis. Inflammatory, vasculitic and allergic disorders, acute and chronic infections, malignancies, haemolysis, iron deficiency and blood loss can all lead to an increased platelet count (Table 26.4). Reactive thrombocytosis can sometimes be marked and, occasionally, the platelet count can be greater than 1000 × 10^9/L. There is usually evidence of on-going inflammation in the form of a raised erythrocyte sedimentation rate (ESR) or C-reactive protein but this is not always the case.

Other clonal thrombocytoses

A number of other haematological malignancies can be associated with thrombocytosis. Chronic myeloid leukaemia (CML) can be excluded by demonstrating the absence of the *bcr–abl* fusion transcript. Iron deficiency can mask the typical raised

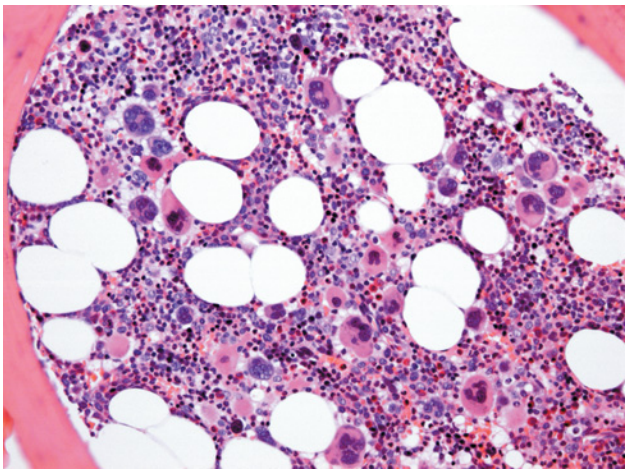


Figure 26.6 Bone marrow trephine section (haematoxylin and eosin, H&E) from a 60-year-old man with essential thrombocythaemia. Note the hypercellularity and marked increase in megakaryocyte numbers, consisting largely of clusters of mature, multilobated forms.

red cell mass of PV, although the clinical utility of distinguishing iron-deficient PV from ET is probably minimal. Established PMF can be excluded by the absence of significant bone marrow fibrosis and other characteristic laboratory features, such as leucoerythroblastic blood film and splenomegaly.

Lastly, myelodysplastic syndromes can also be associated with a thrombocytosis in a minority of cases, but there are usually coexisting cytopenias, dysplastic features or specific cytogenetic abnormalities (e.g. deletion 5q). Sometimes, refractory anaemia with ringed sideroblasts can be associated with thrombocytosis and the presence of the *JAK2* V617F mutation: this illustrates the overlap among these chronic myeloid disorders.

Table 26.4 Causes of a reactive thrombocytosis.

- Iron deficiency
- Blood loss (acute or chronic)
- Hyposplenism/splenectomy
- Surgery
- Acute bacterial infection
  - Pneumonia, septicaemia, meningitis, diverticular abscess etc.
- Chronic inflammation
  - Vasculitides
  - Inflammatory bowel disease
  - Connective tissue disorders
  - Rheumatoid arthritis
  - Chronic infections
- Malignancies
- Rebound thrombocytosis
  - Following treatment of immune thrombocytopenic purpura
  - Recovery from chemotherapy
- Drugs
  - Vincristine

Treatment

When considering the management of patients with ET, it is helpful to stratify patients into risk groups, according to their risk of vascular complications, and take treatment decisions on this basis. Patients may be assigned to a high-, intermediate- or low-risk category.

High-risk patients

High-risk patients are those over 60 years old and those with one or more high-risk features, i.e. a platelet count $\geq 1500 \times 10^9$/L, a prior history of thrombosis or significant thrombotic risk factors such as diabetes or hypertension. In high-risk patients, control of the platelet count with hydroxycarbamide reduces thrombotic events compared to no cytoreductive therapy. The MRC PT-1 trial compared treatment with hydroxycarbamide plus low-dose aspirin to anagrelide plus low-dose aspirin. Patients receiving anagrelide plus aspirin were significantly more likely to reach the composite primary endpoint (arterial thrombosis, venous thrombosis or major haemorrhage) and more likely to discontinue their allocated treatment. Compared to hydroxycarbamide plus aspirin, treatment with anagrelide plus aspirin was associated with a significantly increased rate of arterial thrombosis, major haemorrhage and myelofibrotic transformation, but a decreased rate of venous thromboembolism. These results suggest that hydroxycarbamide plus aspirin should remain first-line therapy for high-risk patients. Anagrelide is a useful second-line agent, but the decision to use concurrent aspirin should depend on the relative risks of arterial thrombosis and haemorrhage in the individual patient. A subsequent smaller study (the ANAHYDRET trial) reported non-inferiority between hydroxycarbamide and anagrelide in patients with ET, as strictly defined by the WHO-diagnostic criteria, but was underpowered to see the differences observed in the PT1 study.

Analysis of patients from the PT-1 trial according to *JAK2* status has shown that compared with V617F negative patients, mutation-positive patients share many features with PV, including a higher risk of venous thrombosis. Moreover V617F-positive patients were more sensitive to hydroxycarbamide than anagrelide, raising the possibility that hydroxycarbamide is particularly effective in these patients.

IFN-α can give good control of the platelet count in ET, but as discussed under PV, its significant side-effect profile, subcutaneous administration and cost prevent its widespread use. It has a clearer role in the management of ET in pregnancy (*vide infra*) and some favour it in young patients. Comparator studies between interferon and hydroxycarbamide will be very

important and are underway. Busulfan can achieve good control of the platelet count, but is only rarely used because of concerns over its long-term leukaemogenic potential. This is true for other alkylating agents and for radioactive phosphorus (^{32}P).

Intermediate-risk patients

Intermediate-risk patients are those between 40 and 60 years old who lack any of the high-risk features listed above. It is not clear whether it is beneficial to lower the platelet count in this group. Most receive either aspirin alone, or hydroxycarbamide and aspirin. For those electing to receive aspirin, the threshold for initiating therapy is 1500×10^9/L.

Low-risk patients

Low-risk patients are those younger than 40 years old who lack any high-risk features. Low-risk patients are usually given low-dose aspirin alone, unless there is a contraindication, such as previous peptic ulceration or allergy to salicylates. Antiplatelet agents such as dipyridamole or clopidogrel should be considered in these cases. In the setting of ET-associated haemorrhage, it is probably best to avoid antiplatelet agents.

Prognosis

Few studies have directly addressed survival in ET and these have reached different conclusions. Some suggest that mortality at 10 years is that of age-matched controls, whereas others found it to be worse. In high-risk patients, hydroxycarbamide reduces vaso-occlusive events from 10.7 to 1.6 per 100 patient-years.

ET and pregnancy

ET is the MPN encountered most frequently in women of childbearing age, hence the most data exists for this compared to other MPN. In pregnancy, the commonest complication of ET is first-trimester miscarriage, which occurs in up to 30% of pregnancies, and is thought to reflect placental microinfarcts and insufficiency. Other less frequent complications include intrauterine death, growth retardation, premature delivery and pre-eclampsia. The risk of maternal thrombosis and haemorrhage is higher than in normal pregnancy; nonetheless a successful outcome (live birth) is achieved in around 60% of cases and no maternal deaths were seen in a recent review of 220 pregnancies. Data for PV and PMF in pregnancy is much more sparse; however, pregnancy outcomes are similar, the main difference for PV is the need to ensure the haematocrit remains in the middle of the gestation-appropriate range (i.e. 0.45 is too high for women beyond the first trimester of pregnancy).

The optimal management of ET in pregnancy has not yet been fully defined. There is conflicting evidence about the effectiveness of aspirin but, given the good documentation of its safety in a large unrelated study of pre-eclampsia, it should probably be given to most patients who are pregnant or planning a pregnancy. Pregnancies deemed at 'high risk' (e.g. patients with complications in a previous pregnancy or prior thrombosis; see Table 26.5) should be considered for therapy with a combination of IFN, aspirin and low-molecular-weight heparin. Hydroxycarbamide and anagrelide should not be used on account of their teratogenic potential. Fetal growth and successful placentation should be closely monitored and therapies adjusted accordingly. The blood count may rise dramatically in the postpartum period. Hydroxycarbamide and anagrelide are excreted in breast milk so that breast-feeding is contraindicated while a patient is receiving either of these agents. Although IFN-α is also excreted in breast milk, it is unlikely to be absorbed intact by the baby and there are anecdotal reports of successful breast-feeding while the mother was receiving IFN.

Table 26.5 Criteria for high-risk pregnancy in myeloproliferative neoplasms (MPN).

*Sustained rise in platelet count rising to above 1500×10^9/L**
Previous venous or arterial thrombosis
*Previous haemorrhage attributed to MPN**
Previous pregnancy complication
- ≥1 unexplained deaths of a morphologically normal fetus ≥10 weeks of gestation.
- ≥1 premature delivery of a morphologically normal fetus <34 weeks gestation because of:
  - **i** Severe pre-eclampsia or eclampsia defined according to standard definitions.
  - **ii** Recognized features of placental insufficiency.
- ≥3 unexplained consecutive miscarriages <10 weeks gestation, with maternal and paternal factors (anatomic, hormonal or chromosomal abnormalities) excluded.
- Otherwise unexplained intrauterine growth retardation.
- Significant antepartum or postpartum hemorrhage requiring transfusion.

Abnormal uterine artery Dopplers at 20 weeks (mean pulsatility index >1.4)

*These would represent indications for interferon only rather than interferon plus low molecular weight heparin.

Primary myelofibrosis

Also known as agnogenic myeloid metaplasia and idiopathic myelofibrosis, primary myelofibrosis (PMF) has the poorest prognosis of the MPNs. PV and ET can develop into a condition that resembles PMF, usually after a latency of many years. The first reported case of PMF is probably that reported by Hueck in 1879 as a 'peculiar leukaemia'. It was not until Dameshek's seminal work in 1951 that PMF was recognized as a myeloproliferative neoplasm. The identification of the *JAK2* V617F mutation in approximately half the cases of PMF and ET and

nearly all cases of PV, revolutionized our understanding of the relationship between the three disorders. We now also know that many patients with JAK2-unmutated ET or PMF carry CALR mutations, and a smaller number have MPL mutations. Analogous to the concept of chronic and accelerated phases in CML, it seems likely that PMF represents the presentation in accelerated phase of a previously undiagnosed MPN. Consistent with this concept, patients with PMF harbour more mutations, have more cytogenetic abnormalities and display an increased risk of leukaemic transformation.

Pathophysiology

Primary myelofibrosis is a clonal myeloproliferative neoplasm of the pluripotent haemopoietic stem cell, in which the proliferation of multiple cell lineages is accompanied by progressive bone marrow fibrosis. Marrow fibrosis is thought to be secondary to the release of proinflammatory cytokines from abnormal clonal cells (primarily megakaryocytes), which act to stimulate fibroblast proliferation and fibrosis. In support of this premise, transgenic mice expressing high levels of TPO rapidly develop myelofibrosis in association with increased megakaryocyte numbers. Additionally, mice expressing reduced levels of the transcription factor GATA-1, which impairs the ability of their megakaryocytes to differentiate into platelets, also develop myelofibrosis in association with increased expression of cytokines such as transforming growth factor-β1, platelet-derived growth factor and vascular endothelial growth factor in the bone marrow.

In the peripheral circulation there is an increase in the number of CD34-positive cells, together with increased numbers of progenitors capable of giving rise to a variety of haemopoietic colonies. As with PV and ET, erythroid and megakaryocytic colonies can also be derived in the absence of exogenous growth factors.

The same molecular abnormalities seen in chronic-phase MPNs, such as *JAK2*, *MPL, CALR* and *TET2* mutations, are found in patients with PMF, underscoring the inter-related nature of these disorders. Nonetheless, other mutations and epigenetic abnormalities are more frequently found in PMF, in keeping with the more aggressive phenotype, poorer prognosis and later stage of the disease. For example, cytogenetic abnormalities are found in up to 60% of cases. The commonest are deletions of 20q and 13q, trisomy 8, and abnormalities of chromosomes 1, 5, 7 and 9. Oncogene mutations are not infrequently found, and include point mutations in the *RAS* genes, *KIT* and *TP53*.

Clinical features

Epidemiology

The estimated annual incidence of PMF is around 0.5–1.5 per 100,000 of the population, with most patients diagnosed in the sixth decade and roughly equal involvement of the two sexes. Up to a third of patients are asymptomatic at diagnosis and many of these are discovered after unrelated blood tests show modest abnormalities, such as anaemia and thrombocytopenia.

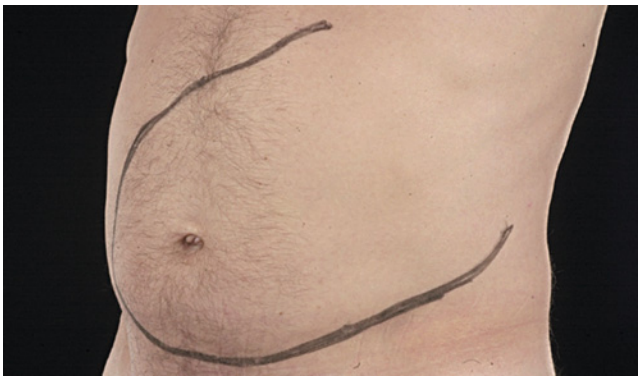


Figure 26.7 Massive splenomegaly in a 53-year-old man with an 8-year history of primary myelofibrosis (PMF).

Splenomegaly

An enlarged spleen is found in almost all patients at presentation and splenic pain/discomfort is a common presenting symptom of PMF. Most cases develop moderate to marked splenomegaly during the course of the disease and about 10% of cases develop massive splenomegaly, with the spleen extending to the right iliac fossa (Figure 26.7). This dramatic increase in splenic mass (up to 20–30 times normal) can lead to a substantial increase in splenic blood flow which, in the most severe cases, can lead to portal hypertension with oesophageal varices and ascites. Painful and painless splenic infarcts are common sequelae of splenomegaly in PMF.

Extramedullary haemopoiesis

The spleen is the commonest site of extramedullary haemopoiesis in PMF. The liver is also usually involved and this can lead to significant hepatomegaly. Unusual sites can sometimes be affected, leading to haemopoietic tumours surrounded by a capsule of connective tissue. Such sites include lymph nodes, central nervous system, skin, pericardium, peritoneum, pleura, ovaries, kidneys, adrenals, gastrointestinal tract and lungs. Many such cases remain asymptomatic, but involvement of the central nervous system can be a cause of serious morbidity. Treatment with radiotherapy or surgery, when required, almost always leads to resolution of these masses.

Systemic symptoms

A hypermetabolic state presenting with fevers, anorexia, weight loss and night sweats develops in many cases of PMF, sometimes early on in the disease. The presence of such symptoms is associated with a poor prognosis.

Anaemia

Mild to moderate anaemia is found in most patients at presentation and worsens as myelofibrosis progresses. The anaemia is in large part due to reduced erythropoiesis, but may be compounded by hypersplenism, bleeding and iron or folate deficiency. Acquired HbH disease is a rare complication.

Platelet abnormalities

Platelet counts are raised in up to one-half of the cases at presentation and can be associated with thrombotic complications. However, progressive thrombocytopenia is a frequent occurrence and becomes increasingly troublesome as the disease progresses. Dysmegakaryopoiesis and abnormal platelet function further add to the risk of haemorrhagic complications.

White cells and leukaemic transformation

The presence of immature myeloid as well as erythroid progenitors is a characteristic feature of PMF (Figure 26.8). Neutrophilia is common, as are modest elevations in basophil and eosinophil counts. As the disease progresses, leucopenia increases in frequency and is believed to be secondary to progressive hypersplenism, dysmyelopoiesis and progressive replacement of the bone marrow by fibrotic tissue. In end-stage PMF, myeloid precursors become increasingly common relative to mature cells, as do circulating blasts. Transformation to AML, as defined by the persistent presence of 20% blasts in blood or bone marrow occurs in 20–30% of cases of PMF and is usually rapidly fatal.

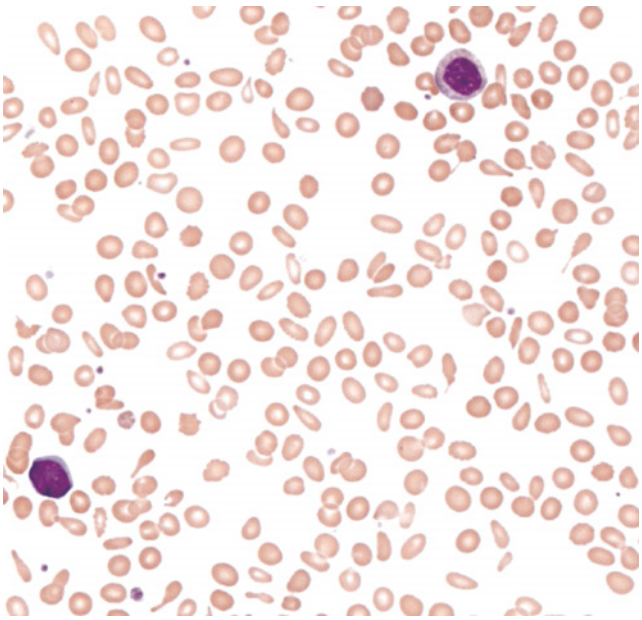


Figure 26.8 Peripheral blood film in PMF, showing a blast, an abnormal myelocyte, teardrop red cells and marked anisopoikilocytosis.

Table 26.6 Diagnostic criteria for primary myelofibrosis.

***JAK2*-positive primary myelofibrosis**	
A1	Reticulin ≥grade 3 (on a 0–4 scale)
A2	Mutation in *JAK2, MPL* or *CALR*
B1	Palpable splenomegaly
B2	Otherwise unexplained anaemia (Hb <115 g/L for men; <100 g/L for women)
B3	Tear-drop red cells on peripheral blood film
B4	Leucoerythroblastic blood film (presence of at least two nucleated red cells or immature myeloid cells in peripheral blood film)
B5	Systemic symptoms (drenching night sweats, weight loss >10% over 6 months OR diffuse bone pain)
B6	Histological evidence of extramedullary haemopoiesis
Diagnosis requires A1 + A2 and any two B criteria	
***JAK2*-negative primary myelofibrosis**	
A1	Reticulin ≥grade 3 (on a 0–4 scale)
A2	Absence of mutation in *JAK2*
A3	Absence of *BCR–ABL* fusion gene
B1	Palpable splenomegaly
B2	Otherwise unexplained anaemia (Hb <115 g/L for men; <100 g/L for women)
B3	Tear-drop red cells on peripheral blood film
B4	Leucoerythroblastic blood film (presence of at least two nucleated red cells or immature myeloid cells in peripheral blood film)
B5	Systemic symptoms (drenching night sweats, weight loss >10% over 6 months OR diffuse bone pain)
B6	Histological evidence of extramedullary haemopoiesis
Diagnosis requires A1 + A2 + A3 and any two B criteria	

Investigations

Diagnostic criteria for PMF are shown in Table 26.6. Other causes of bone marrow fibrosis are listed in Table 26.7.

Peripheral blood

The presence of myeloid and erythroid precursors in the peripheral blood (leucoerythroblastic blood picture) is common in PMF (see Figure 26.8). Other causes of a leucoerythroblastic blood film include bone marrow infiltration, severe sepsis, severe haemolysis and a sick neonate. Teardrop poikilocytes (dacryocytes), basophilic stippling, macrocytosis (which may or may not be secondary to folate deficiency), giant platelets and megakaryocyte fragments may also be present.

Bone marrow

Attempts at bone marrow aspiration often yield a dry-tap or a haemodilute sample, making aspirate morphology of limited diagnostic value. Sufficient material can often be obtained from either bone marrow or peripheral blood to assess the karyotype, which can help exclude diagnoses such as CML. Other

Table 26.7 Differential diagnosis of marrow fibrosis.

Haematological malignancies	**Drugs/toxins**
Primary myelofibrosis	Benzene
Chronic myeloid leukaemia	Thorotrast
Acute myelofibrosis (AML M7)	Irradiation
Myelodysplasia	**Bone disease**
Myeloma	Paget's disease
Hairy-cell leukaemia	Osteopetrosis
Non-Hodgkin lymphoma	**Inflammatory diseases**
Hyperparathyroidism	Systemic sclerosis
Hodgkin disease	Systemic lupus
Hypoparathyroidism	**Other**
Systemic mastocytosis	Grey plt syndrome
Metastatic carcinoma	
Infections	
Tuberculosis	
Leishmaniasis	

chromosomal abnormalities may be found in up to 60% of cases, as detailed above. Abnormalities of chromosomes 5 and 7 are usually found in patients with prior exposure to genotoxic agents and are associated with a poor prognosis.

Bone marrow trephine biopsy is essential to make a diagnosis of PMF. Initial stages are characterized by an increase in bone marrow cellularity in association with a disorganization of marrow architecture and the presence of abnormal large megakaryocytes often occurring in clusters (Figure 26.9). Bone marrow fibrosis becomes increasingly dominant and progressively replaces haemopoiesis. Intrasinusoidal haemopoiesis can sometimes be seen at this stage (Figure 26.10). The degree of fibrosis is best demonstrated using silver impregnation, which stains reticulin fibres (see Figure 26.9). Collagen fibres are best demonstrated using a trichrome stain. The degree of fibrosis can be graded from 0 to 4 (Bauermeister) or MF-0 to 3 (EU consensus) according to severity. In a minority of cases of advanced PMF, osteosclerosis ensues, with thickening of the trabecula and extensive deposition of osteoid. Such changes may be evident on plain radiography and MRI imaging.

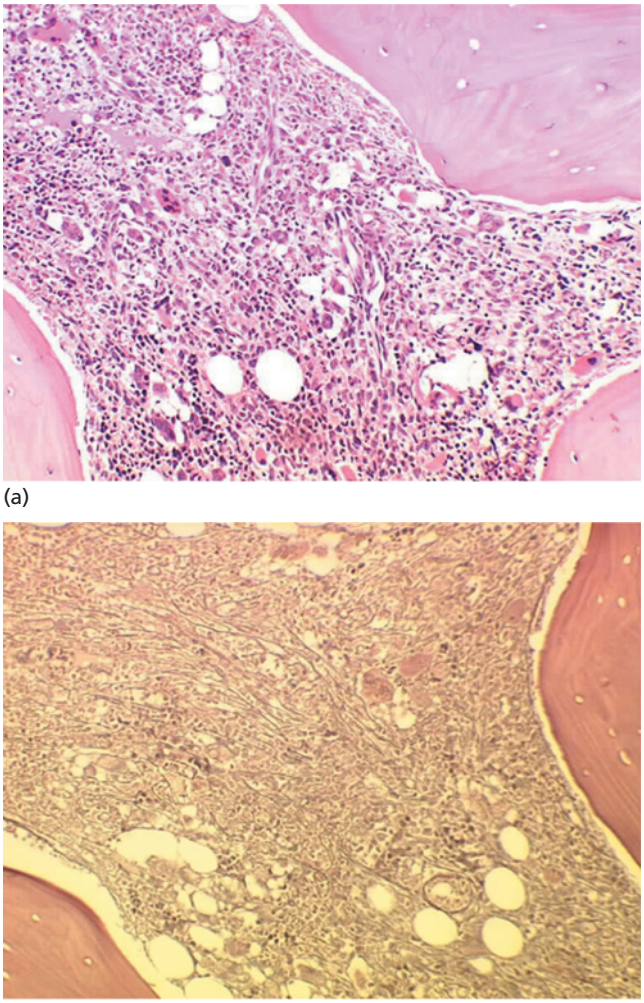


Figure 26.9 Bone marrow trephine sections from a patient with early-stage PMF. The H&E stain (a) shows hypercellularity, disorganized architecture, increase in megakaryocyte numbers and prominent sinusoids. The silver stain (b) also shows a marked increase in reticulin fibres.

Treatment

The only curative treatment for PMF is allogeneic stem cell transplantation, but this is only appropriate for a small proportion of patients. In the remaining cases therapy remains supportive and aimed at alleviating symptoms, but has little impact on the relatively poor survival of PMF patients. For this reason the development of therapeutic *JAK2* inhibitors and other novel agents was most eagerly awaited for these patients, rather than those with PV or ET, who have a much better prognosis. In recent years, JAK inhibitors, specifically ruxolitinib, have radically altered the therapeutic algorithm for the management of PMF.

Conventional allogeneic stem cell transplantation, usually a reduced-intensity conditioning is only a realistic option in young patients who represent perhaps 10% of all cases. The decision to proceed to transplantation should always be made in light of the patient's specific prognosis, age and general fitness. Only small series have been reported and the long-term survival of patients ≤45 years is approximately 50%, with a 30% transplant-related mortality. For patients ≥45 years, outcomes are much worse, with long-term survival of 10–20%. The utility of JAK1/2 inhibition specifically with ruxolitinib has radically altered the treatment landscape for PMF; as evaluated in the Phase III COMFORT trials, this agent delivers benefit in control of symptoms and splenomegaly, as well as improving survival for patients

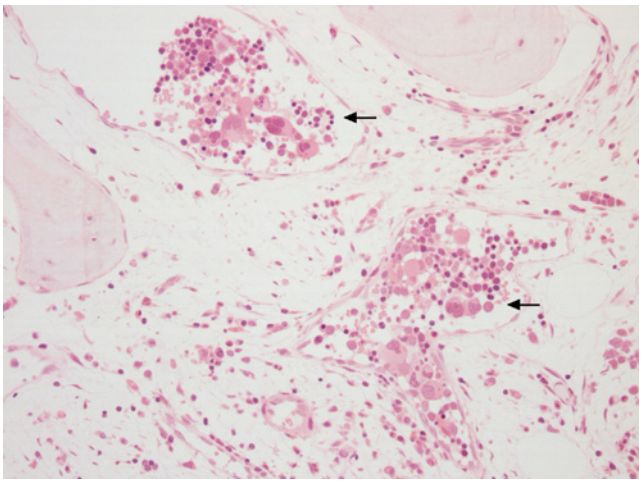


Figure 26.10 Bone marrow trephine section in PMF (H&E). Note the presence of two dilated sinusoids (arrows) containing immature haemopoietic cells, including megakaryocytes.

with intermediate-risk-2 and high-risk PMF, regardless of their *JAK* mutation status. These studies led to its approval. As anticipated given its mode of action, ruxolitinib causes anaemia (a predictable fall in haemoglobin levels by 10%, peaking at 12–16 weeks and gradually resolving thereafter) and thrombocytopenia (fall in platelet levels by 40%). Both anaemia and thrombocytopenia can be attenuated by dose-modification and may be moderated by addition of agents such as Epo and androgens. Two other important adverse effects of ruxolitinib merit attention: rebound of symptoms and splenomegaly after withdrawal, which therefore should be managed carefully, and the increased risk of both herpes virus re-activation and opportunistic infections. Given the beneficial effect of ruxolitinib upon survival and generally excellent tolerability, attention has focused subsequently upon its potential use in several other settings: earlier stages of disease, data is available for intermediate-1-risk patients, but for low-risk patients data is limited; to improve the outcome of allogeneic transplant and, in combination with other agents such as histone deacetylase inhibitors, to improve disease response. Other JAK inhibitors are under development, but several have been associated with neurological toxicity, raising the possibility that JAK-responsive targets cause this toxicity and urging careful evaluation of this class of drug.

Anaemia responds to treatment with androgens in up to one-third of cases, with the best responses seen in patients without massive splenomegaly and with a normal karyotype. The drugs most commonly used are oxymethalone (50–150 mg daily) and danazol (400–600 mg daily). Both of these can have virilizing effects and can lead to abnormal liver function. Patients with a reduced red cell survival may respond to treatment with corticosteroids. Human recombinant erythropoietin has recently shown promise in small clinical studies of anaemia in PMF, usually where endogenous Epo levels are below 125 IU. Splenectomy may also have a residual role for transfusion-dependent anaemia (see below). Despite these treatments, most patients become transfusion dependent eventually.

The antiangiogenic drug thalidomide has been reported to improve anaemia, thrombocytopenia or splenomegaly in approximately one-third of the patients, but these changes are clinically significant only in a small proportion (20%). Thalidomide is poorly tolerated at conventional doses (>100 mg per day) with more than one-half of the patients being unable to tolerate it beyond 3 months. Furthermore, thalidomide increases the risk of extreme thrombocytosis and probably that of venous thrombosis, and needs special precautions for use in view of its extreme teratogenicity. Trials of lower doses of the drug alone or in combination with prednisolone suggest that such doses are better tolerated and may be similarly efficacious. Pomalidomide has also been reported to be successful in the management of anaemia in PMF, but unfortunately a Phase III study (RESUME) reported negative findings; further studies are underway. Lenolidomide has been evaluated in MF, but is generally recommended for the rare patients who have MF and a 5q– cytogenetic abnormality.

Cytoreductive therapy can be useful in the management of some aspects of PMF, such as hepatosplenomegaly, constitutional symptoms and troublesome thrombocytosis. Hydroxycarbamide is the most widely used agent, but anagrelide has also been used for thrombocytosis. The utility of these agents, as evaluated in the COMFORT studies, is extremely limited in the majority of cases and efficacy was comparable to that of placebo for symptom and spleen control, as well as impact upon survival. Interferon may be of utility in early stages of this disease and has occasionally been reported to be associated with fibrosis reversal.

The indications for splenectomy include splenic pain, constitutional symptoms, portal hypertension and transfusion-dependent anaemia. In contrast, there is no good evidence that thrombocytopenia responds to splenectomy. The procedure has significant mortality (around 10%) and morbidity, particularly in elderly patients. Problems include perioperative bleeding, infection and thrombosis, as well as rebound thrombocytosis and progressive hepatomegaly. It is particularly important to correct any coagulation abnormalities prior to surgery.

Splenic irradiation is an alternative to splenectomy in some cases and it can significantly reduce splenic size, albeit transiently. This procedure is not without complications, as it can lead to life-threatening cytopenias. Radiotherapy can also be useful for treating pockets of extramedullary haemopoiesis involving vital organs or bodily cavities. With the increasing use of JAK inhibitors, indications for splenectomy and splenic irradiation are likely to reduce in the future, except as a last resort. Ruxolitinib has also been shown to be effective for treating extramedullary haemopoiesis outside the spleen and liver.

Table 26.8 Risk stratification for PMF patients.

IPSS prognostic score				
Prognostic variable	***0***	***1***		
Age in years	≤65	>65		
WBC count × 10^9/L	≤25	>25		
Haemoglobin g/L		<100		
Peripheral blast %	<1	>1		
Constitutional symptoms	No	Yes		
Risk assignment: Low = 0; Intermediate 1 = 1; Intermediate 2 = 2; High = 3+ Median survivals are 135, 95, 48 and 27 months, respectively				
DIPSS prognostic score				
	0	***1***	***2***	
Age in years	≤65	>65		
WBC count × 10^9/L	≤25	>25		
Haemoglobin g/L			<100	
Peripheral blast %	<1	>1		
Constitutional symptoms	No	Yes		
Risk assignment: Low = 0; Intermediate 1 = 1 or 2; Intermediate 2 = 3 or 4; High = 5 or 6 Median survival: not reached, 14.2, 4, and 1.5 years, respectively				
DIPSS plus prognostic score				
Points from DIPSS	***0***	***1***	***2***	***3***
DIPSS prognostic group points	Low risk 0	Intermediate 1 risk 1	Intermediate 2 risk 2	High risk 3
To the DIPSS prognostic group add one point each for:				
Platelet count × 10^9/L		<100		
Red cell transfusion required	No	Yes		
Unfavourable karyotype*	No	Yes		
Risk assignment number of points: Low = 0; Intermediate 1 = 1; Intermediate 2 = 2 or 3; High = 4 to 6 Corresponding median survival estimates: 185, 78, 35 and 16 months				

*Unfavourable karyotype includes +8, −7/7q−, i(17q), inv(3), −5/5q−, 12p−. 11q23 rearrangements and complex karyotypes.

Prognosis

The median survival is 3–5 years, but the range is very wide; ruxolitinib and improvements in general care have improved this in recent years. The most widely used algorithms to determine prognosis are shown in Table 26.8 and are the International Prognostic Score System (IPSS) and Dynamic (D)IPSS or DIPSS-plus. Increasingly molecular markers are having an impact upon prognosis, but these are not yet included in standard prognostic scores; current evidence suggests that patients with *CALR* mutations may have fewer thrombotic events, but more myelofibrotic transformation, while those patients having any one of the following mutations appear to have an overall worse prognosis – *ASXL-1, EZH2, IDH1/2*, and *SRSF2*.

Mastocytosis

Mastocytosis comprises a rare group of disorders characterized by a pathological increase in mast cells in tissues, including the skin, bone marrow, liver, spleen, lymph nodes and gastrointestinal tract. Mastocytosis can be an isolated finding or can form part of other haematological disorders, including myelodysplastic syndromes, myeloproliferative neoplasms or AML. Some

Table 26.9 Classification of mastocytosis.

Cutaneous Mastocytosis (CM)
- Maculopapular cutaneous mastocytosis (formerly: urticaria pigmentosa)
- Diffuse cutaneous mastocytosis
- Mastocytoma

Systemic Mastocytosis (SM)
- Indolent systemic mastocytosis
  - Isolated bone marrow mastocytosis
  - Smouldering systemic mastocytosis
- Systemic mastocytosis with an associated clonal hematological non-mast-cell-lineage disease
- Aggressive systemic mastocytosis
- Mast-cell leukaemia
- Mast-cell sarcoma
- Extracutaneous mastocytoma

(Source: Valent *et al*., 2001 [*Leukemia Research* 2001; **25**: 603–25]. Adapted with permission of Elsevier.)

cases involve just the skin (cutaneous mastocytosis), whereas others involve multiple tissues and are associated with systemic symptoms (systemic mastocytosis). Paediatric mast-cell disease is generally a reactive condition rather than a clonal myeloproliferative neoplasm, and will not be further discussed here. A proposal for the classification of mast cell diseases was recently put forward (Table 26.9).

The first case of urticaria pigmentosa (a form of cutaneous mastocytosis) was described in 1869 by Nettleship, and systemic disease due to increased mast cells was first documented by Ellis in 1949. The observation that stem cell factor (SCF) is an essential growth factor for mast-cell development has led to significant advances in our understanding of this group of diseases.

Pathophysiology

After a search for abnormalities of SCF failed to identify any pathological changes, researchers turned to c-Kit, the tyrosine kinase receptor for SCF. A *c-kit* point mutation leading to a single amino-acid substitution (Asp816Val) was thus identified. This mutation leads to ligand-independent phosphorylation of c-Kit and a consequent clonal expansion of mast cells.

Asp816Val was originally identified in cases of mastocytosis with associated haematological disorders, but is now known to be present in up to 90% of adults presenting with maculopapular cutaneous mastocytosis (formerly urticaria pigmentosa) or indolent systemic mastocytosis. More recently, other activating mutations affecting the same codon have been identified in a minority of adult cases of cutaneous mastocytosis (Asp816Tyr, Asp816Phe). A different mutation has been reported in a small number of cases of paediatric mastocytosis, namely Lys839Glu,

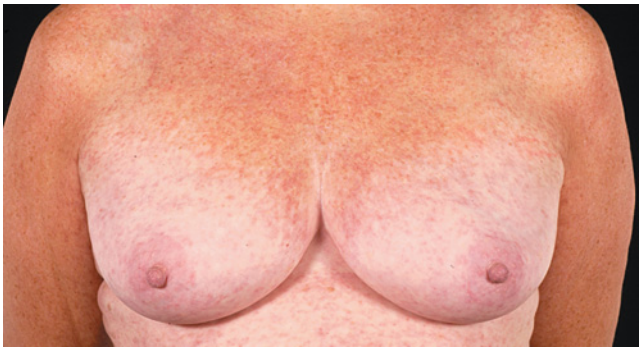


Figure 26.11 Urticaria pigmentosa in a 55-year-old woman. Note the widespread pink or brownish macules that become confluent in areas.

which surprisingly gives rise to a dominant-negative (inactivating) form of c-Kit; the significance of this observation remains unresolved.

Recent reports have suggested that in some cases of mastocytosis the *c-kit* mutation is found in other haemopoietic cells such as B cells, myeloid cells and T cells. These results suggest that mastocytosis, as with the classical MPNs, may be a clonal disorder of the haemopoietic stem cell.

Clinical features

Cutaneous manifestations

Maculopapular cutaneous mastocytosis is the usual presenting feature in children and adults with isolated mastocytosis. Yellowish-brown lesions, usually macular and sometimes papular, appear in a patchy distribution. Less commonly, there is diffuse involvement of the skin, which becomes thickened and darker brown (Figure 26.11). Pruritus is common, as is flushing, and some cases develop haemorrhagic bullous disease. Wealing of lesions upon rubbing is known as Darier's sign.

Systemic disease

Systemic manifestations are very heterogeneous and are thought to be largely secondary to mast-cell mediator release. Episodes of flushing, angioedema or even anaphylaxis, with or without any specific trigger, can arise as a result of systemic histamine release. Gastrointestinal symptoms include abdominal pain, diarrhoea, nausea and vomiting. Gastritis and peptic ulceration may occur secondary to hyperhistaminaemia and severe cases may develop malabsorption. Osteoporosis is well recognized and can sometimes lead to pathological fractures. Peripheral blood cytopenias may arise secondary to mast-cell infiltration of the bone marrow. Hepatosplenomegaly is more common in cases associated with another clonal haematological disorder. Fever, fatigue and weight loss can sometimes ensue and may result from the release of cytokines such as tumour necrosis factor-α (TNF-α) and IL-1.

Table 26.10 Criteria for the diagnosis of systemic mast-cell disease.

Major
- Multifocal dense infiltrates of mast cells in bone marrow and/or other extracutaneous tissues

Minor
- More than 25% of mast cells on bone marrow smears or tissue biopsies are atypical or spindle-shaped
- Identification of a codon 816 c-kit point mutation in blood, bone marrow or lesional skin
- Mast cells in bone marrow, blood or other lesional tissues expressing CD25 or CD2
- Baseline total serum tryptase greater than 20 ng/mL

Major and one minor, OR three or more minor criteria needed for diagnosis

(Source: Valent *et al*., 2001 [*Leukemia Research* 2001; **25**: 603–25]. Adapted with permission of Elsevier.)

Symptoms of organ failure due to infiltration are characteristic of aggressive systemic mastocytosis. Depending on the organs involved, cytopenias, pathological fractures, impaired liver function, ascites and malabsorption can all be seen. Mastocytosis can present in association with a clonal haematological disease (e.g. MDS or other MPN); it is important to be aware of dual pathology in these circumstances.

Investigations

Clinical features of mastocytosis can be highly suggestive of the disease, but diagnosis usually requires histological and biochemical confirmation. An algorithm has been proposed for the diagnosis of systemic mast-cell disease and is shown in Table 26.10.

Routine investigations should include a full blood count, liver function tests, bone profile and a random serum tryptase. Tests for histamine metabolites in 24-h urine specimens are probably no more useful than measurements of serum tryptase. Plasma levels of soluble CD25 and CD117 (kit) have shown promise as novel markers of mast-cell disease, but the most widely used marker remains serum tryptase, though it is important to understand this will be falsely elevated in the time immediately surrounding a major allergic episode.

Bone marrow aspiration and trephine biopsy allow assessment of bone marrow involvement. Mast-cell aggregates can be visualized on conventional haematoxylin and eosin-stained sections (Figure 26.12), but stand out much more clearly with stains such as toluidine blue (Figure 26.13). Immunochemistry using antitryptase antibodies can also be very useful, being highly specific for mast cells, and CD2, CD25, CD30 are important tools. Flow cytometry to look for expression of CD2 and CD25 in bone marrow mast cells may be useful, as this phenotype is not seen in normal mast cells.

Abdominal ultrasound or computerized tomography should be performed to look for hepatosplenomegaly and lymphadenopathy. Plain radiography and bone densitometry can be used to look for bone involvement and osteoporosis. Endoscopy and biopsy can be useful if gut involvement is suspected.

Treatment

Despite significant advances in the understanding of its pathophysiology, no curative treatment exists for mastocytosis, the management of which remains symptomatic.

There are four main components to the management of mastocytosis:

1 Avoidance of factors that can trigger mediator release from mast cells
2 Treatment of acute mediator release
3 Treatment of chronic mediator release
4 Reduction of the mast cell burden/organ infiltration.

Avoidance of triggers of mast-cell mediator release is primarily an exercise in patient education. Severe reactions due to systemic mast-cell mediator release are difficult to predict in patients with mastocytosis and do not correlate well with disease category, mast-cell burden or severity of other symptoms. All patients and relevant healthcare workers should be warned of particular triggers, including general anaesthesia, contrast radiography and insect stings. Known mast-cell activators such as morphine and dextran should only be introduced with great caution. Patients, regardless of whether they have had previous anaphylaxis or not, should carry injectable adrenaline and they, their family and friends should be instructed in its intramuscular administration. Local mediator release in cutaneous mastocytosis can be moderated by avoidance of triggering factors such as friction and heat.

Acute systemic mast-cell mediator release should be treated in much the same way as other forms of anaphylaxis. Treatment with adrenaline and intravenous fluids should be started as soon as possible, with early involvement of intensive care specialists in severe cases. Antihistamines (H_1 and H_2 blockers) should be introduced and continued long term if the episode was particularly severe or recurrent.

Symptoms of chronic mediator release are the commonest clinical problem in mastocytosis. Symptomatic cutaneous disease should be managed with the help of a dermatologist. Treatments include H_1 and H_2 blockers, topical corticosteroids and PUVA for severe disease. Non-life-threatening systemic symptoms such as flushing, abdominal pain and diarrhoea should be treated with H_1 and H_2 blockers, sodium cromoglycate and corticosteroids. Inhibitors of prostaglandin synthesis, such as aspirin and non-steroidal anti-inflammatory drugs, can also be useful. Aspirin should always be started with caution as it can initially lead to acute mediator release. Such drugs can be used

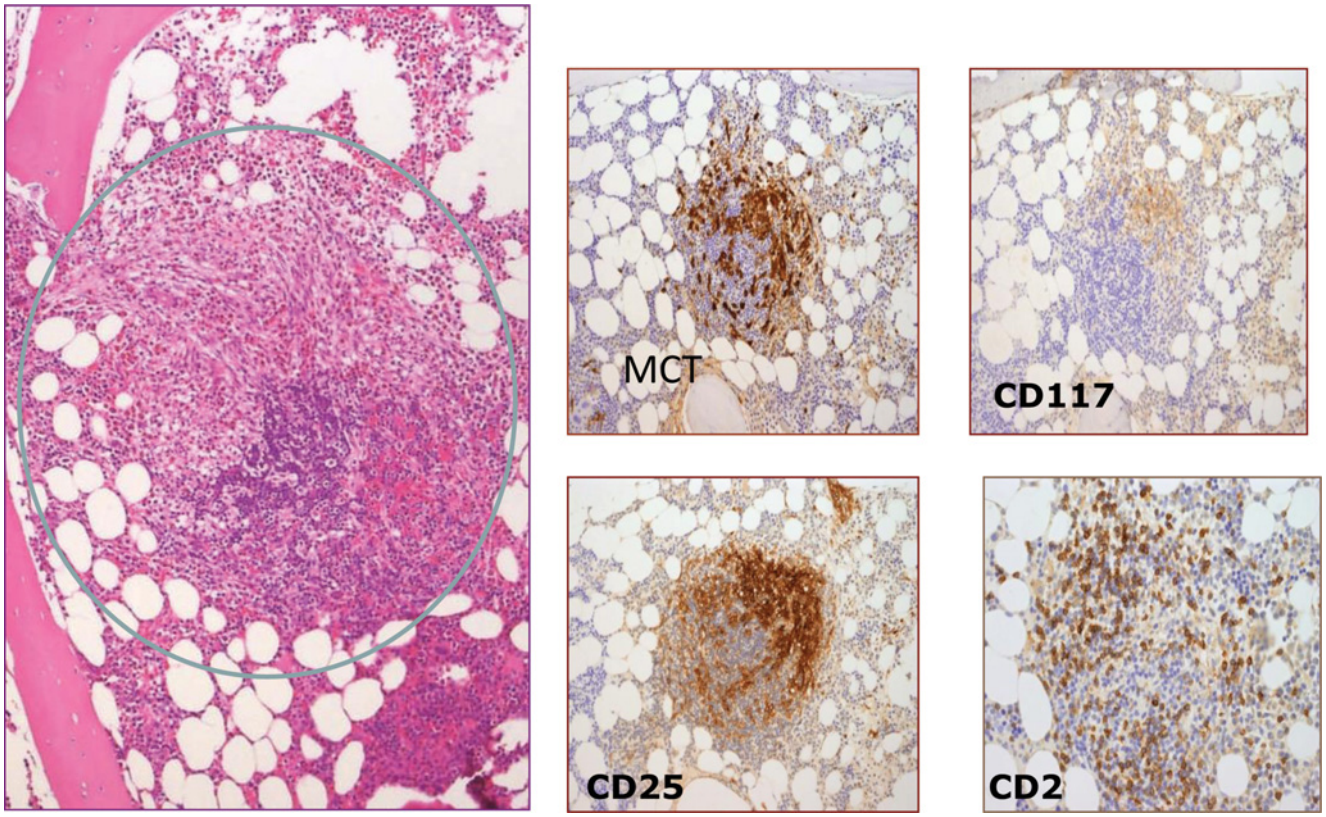


Figure 26.12 Systemic mastocytosis involving the bone marrow (H&E). Malignant whorls of rounded and spindle-shaped mast cells are seen infiltrating the bone marrow in a paratrabecular distribution (circle). Expression of CD2, CD25 and mast cell tryptase is seen (Source: Dr Deepti Radia. Reproduced with permission.).

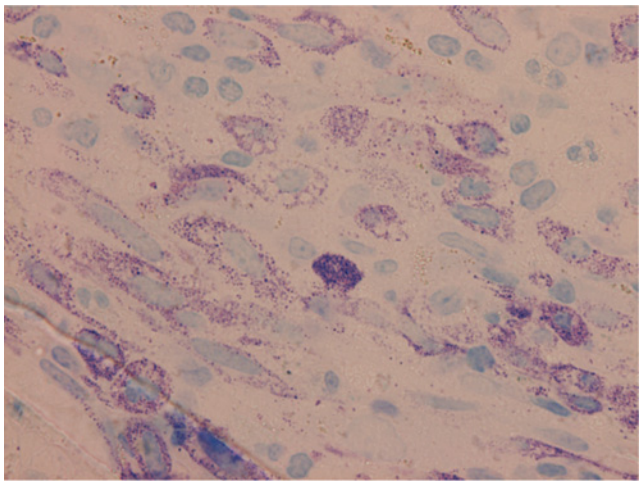
prophylactically if symptoms recur frequently. Gastrointestinal disease usually responds to the drugs used to treat chronic systemic symptoms. Mast-cell stabilizers such as sodium chromoglycate anecdotally help with abdominal cramps and diarrhoea. Leucotriene antagonists may also be useful. Peptic ulcer and reflux disease should be treated with proton pump inhibitors. Osteoporosis should be treated with bisphosphonates and may be prevented with bisphosphonates or calcium and vitamin D supplementation. Bone density should be recorded and monitored according to the severity of osteoporosis. Radiotherapy can help with severe localized pain.

For patients in whom adequate symptomatic control cannot be achieved, and for those with aggressive mastocytosis, IFN-α, usually given in combination with oral corticosteroids, should be considered. Splenectomy may help reduce the mast-cell burden and associated systemic symptoms. Cladribine has been found effective in isolated cases of aggressive systemic mastocytosis.

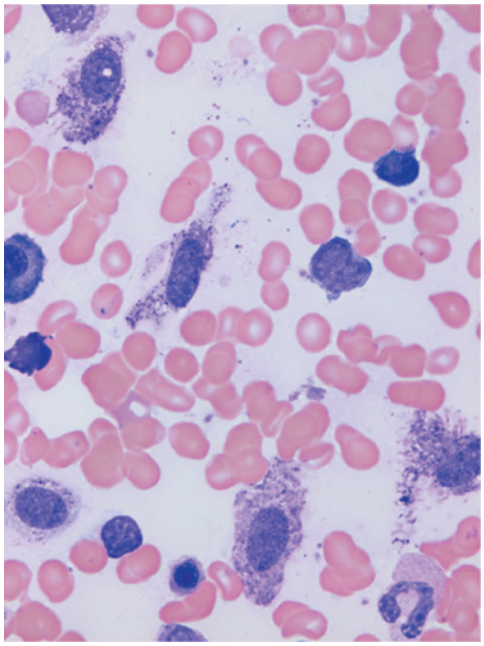
Treatment with chemotherapy is usually reserved for cases of progressive aggressive mastocytosis, mast-cell leukaemia and mast-cell sarcoma, and those with refractory mastocytosis and excessive symptoms, but published data are not encouraging. Mast cell sarcoma may also respond to local radiotherapy when appropriate. Allogeneic bone marrow transplantation should also be considered. Treatment of any associated haematological disorder should be undertaken as appropriate for that disorder and the overall prognosis is usually that of the latter.

Future treatments

Treatments that target the mutant c-kit tyrosine kinase have attracted a lot of interest recently. Midostaurin is one such agent and has shown promising interim results in aggressive systemic mastocytosis or mast-cell leukaema; final results of a Phase II international study are pending. Imatinib was known to inhibit wild-type c-kit in *in vitro* studies and be active against juxtamembrane mutants of c-kit found in gastrointestinal stromal tumours. By contrast, the drug does not have the same effect on malignant mast cells carrying codon 816 mutations, probably because the mutant c-kit does not allow access of imatinib to the site, hence conferring resistance to this drug in a similar way to acquired imatinib resistance in CML. In keeping with this, there have been early reports of its lack of efficacy in the presence of codon 816 mutations. A recent report showing that imatinib was effective in patients with mastocytosis with



(a)



(b)

Figure 26.13 Toluidine blue stain of bone marrow aspirate (a) and trephine biopsy section (b) in systemic mastocytosis. Abnormal spindle-shaped mast cells containing metachromatic granules are seen.

associated eosinophilia, but without demonstrable *KIT* mutations this awaits confirmation, particularly as these cases may represent variants of chronic eosinophilic leukaemia (CEL). It is particularly important to exclude the PDGFRA, B and FGF rearrangements in this clinical scenario, as the diseases associated with these lesions are often sensitive to imatinib. Novel tyrosine kinase inhibitors and other agents, which can inhibit the Asp816Val mutant *KIT in vitro*, are under investigation at present.

Prognosis

Age and disease category are the most important determinants of outcome. The most benign syndrome is paediatric mastocytoma, which disappears with time in over 50% of cases. Paediatric urticaria pigmentosa also has a good prognosis and resolves in about one-half of the cases.

In adult mastocytosis, urticaria pigmentosa is usually associated with mast-cell deposits in the marrow or other tissues, making this a systemic syndrome. Indolent systemic mastocytosis carries a favourable prognosis and usually persists as a chronic low-grade disorder, although it rarely progresses to aggressive mastocytosis or mastocytosis associated with another haematological malignancy. Aggressive systemic mastocytosis can show a slowly progressive or a rapid clinical course, but its overall prognosis has not been well defined in clinical studies. Mast-cell leukaemia is rare, but has a grave prognosis with a median survival of less than 6 months.

Clonal hypereosinophilic syndromes

The term hypereosinophilic syndrome (HES) was coined by Hardy and Anderson in 1968, who gave it the definition that is still in use to date. Recently, major progress has been made in elucidating the molecular pathogenesis of clonal eosinophilia as described below and see also Chapter 14.

Pathophysiology

Eosinophilia can be divided into three categories: reactive, idiopathic and clonal (Table 26.11). Reactive eosinophilia, which is by far the most common is discussed in Chapter 14. Idiopathic eosinophilias are those in which the cause is obscure. Within this category, HES describes patients with an unexplained elevation of peripheral blood eosinophils ($>1.5 \times 10^9$/L) for more than 6 months associated with end-organ damage (see Chapter 14). Many of these are probably cases of chronic eosinophilic leukaemia for which the molecular defect has not been identified.

Clonal eosinophilias are those in which the eosinophilia is part of a clonal haematological malignancy. CEL is defined as an eosinophil count $>1.5 \times 10^9$/L, with evidence of eosinophil clonality or an increased blast account in blood or bone marrow. The distinction between this entity and HES is blurred, as it relies on the availability of a clonal marker. Indeed it has been shown recently that 25–50% of cases labelled as HES in fact have a microdeletion on chromosome 4, which results in the fusion of the *FIP1L1* and *PDGFRA* genes and the generation of a constitutively active tyrosine kinase. Importantly, patients carrying this fusion respond well to the tyrosine kinase inhibitor imatinib. Some patients with HES that do not carry this fusion gene also

Table 26.11 Causes of eosinophilia.

Reactive eosinophilia	
Infections	Parasitic
	Others (rarely)
Vasculitides	Polyarteritis nodosa
	Churg–Strauss syndrome
Connective tissue disorders	Rheumatoid arthritis
	Systemic sclerosis
	Systemic lupus erythematosus
Allergic and inflammatory disorders	Asthma
	Eczema
	Bullous skin diseases
	Inflammatory bowel disease
Drug reactions	Hypersensitivity
	L-Tryptophan
Immunodeficiencies	Wiskott–Aldrich syndrome
	Job's syndrome (hyper-IgE)
Neoplasia	Hodgkin disease
	Non-Hodgkin lymphoma
	Peripheral blood T-cell clones
	Some ALL
	Non-haematological cancers (rare)
Clonal eosinophilia	
Chronic eosinophilic leukaemia	Atypical chronic myeloid leukaemia (PDGFRA fusions)
	8p11 Myeloproliferative syndrome (FGFR1 fusions)
	Chronic myeloid leukaemia (BCR–ABL fusion)
	Acute myeloid leukaemia, e.g. carrying inv(16)
	Acute lymphoblastic leukaemia (occasionally)
Idiopathic eosinophilia	
Hypereosinophilic syndrome	

respond to imatinib, suggesting that in such cases other tyrosine kinases may be dysregulated.

In addition to CEL, a number of other haematological malignancies may be associated with increased numbers of clonal eosinophils, and in many cases, this reflects tyrosine kinase dysregulation. *PDGFRA* rearrangements (e.g. *TEL–PDGFRB*) may present as chronic myelomonocytic leukaemia (CMML) or atypical CML. The 8p11 myeloproliferative syndrome (EMS) is associated with rearrangements in the *FGFR1* gene (e.g. *ZNF198–FGFR1*) and leads to a chronic myeloproliferative neoplasm that frequently presents with eosinophilia and associated T-cell lymphoblastic lymphoma. CML, a consequence of the *BCR–ABL* tyrosine kinase fusion protein, may also be associated with clonal eosinophilia.

Eosinophilia as part of the malignant clone also occurs in patients with AML associated with inversion of chromosome 16 and the *SMMHC–CBFB* rearrangement. It has been reported that rare cases of acute lymphoblastic leukaemia (ALL) may be associated with clonal eosinophilia, but in this disease the eosinophilia is more usually secondary to growth-factor release. Growth-factor release is also believed to underlie the reactive eosinophilia seen in Hodgkin disease and in cases with clonal T cells in the peripheral blood.

Sustained hypereosinophilia can lead to symptomatology and end-organ damage, regardless of its aetiology, but does not always do so. The reasons for this are unclear, but may lie in the heterogeneity of eosinophilia and genetic differences between individuals that affect the propensity of eosinophils and other granulocytes to inflict tissue damage.

Clinical features (see also Chapter 14)

Much of the tissue damage in eosinophilia is believed to be secondary to eosinophil degranulation and release of mediators such as eosinophil cationic protein and major basic protein. Eosinophil mediators act mainly locally in tissues infiltrated by eosinophils to cause tissue damage. The recent finding that a raised serum tryptase in a subset of cases with clonal eosinophilia hints at a role for other cells (mast cells) in some cases.

Patients can present with constitutional symptoms such as fatigue, muscle aches or fevers. Pruritus, angioedema, diarrhoea and cough may also be present. Many tissues can be involved, but cardiac disease is the major cause of mortality. The heart can be affected by endomyocardial fibrosis, pericarditis, myocarditis and intramural thrombus formation. Death is usually due to dilated cardiomyopathy.

Involvement of the central and peripheral nervous systems can result in mononeuritis multiplex, paraparesis, encephalopathy and even dementia. Pulmonary involvement can take the form of pulmonary infiltrates, fibrosis or pleural disease with effusions. Gastrointestinal involvement can manifest as diarrhoea, gastritis, colitis, hepatitis or the Budd–Chiari syndrome. The skin can be affected by pruritus, angioedema, papules or plaques. Rarely, other tissues such as the kidneys and bones can be involved.

Investigations

There are two aims in the investigation of eosinophilia: one is to establish its aetiology and the other to look for evidence of end-organ damage. As regards the former, given the diverse nature of the aetiologies of eosinophilia, a full history, including family history, drug history and travel history, can provide valuable clues. Investigations will usually aim to exclude reactive causes and will be guided by the clinical picture.

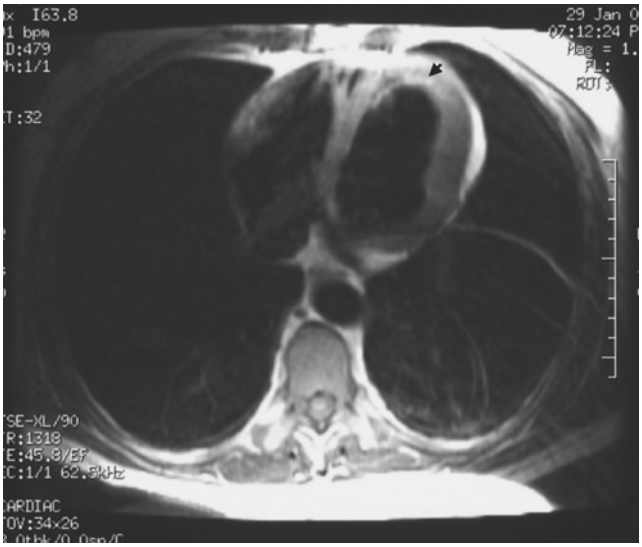


Figure 26.14 Cardiac magnetic resonance imaging scan in a 65-year-old man with HES, showing a rim of subendocardial fibrosis (arrow).

Bone marrow aspiration will reveal morphological abnormalities associated with haematological malignancies and allows cytogenetic analysis. It is also important to look for clonal T-cell receptor (TCR) gene rearrangements, the *FIPIL1–PDGFRA* and *BCR–ABL* fusion genes, as well as rearrangements of the *PDGFRB* and *FGFR1* genes. It has been reported that serum tryptase is raised in patients with the *FIPIL1–PDGFRA* fusion.

Investigations to assess end-organ damage will depend on the clinical presentation. However, echocardiography should be performed and repeated annually in patients with sustained eosinophilia, particularly as cardiac disease correlates poorly with the eosinophil count. If there is strong clinical suspicion of cardiac damage then cardiac magnetic resonance imaging (MRI) can be useful, as this is more sensitive in detecting early disease (Figure 26.14). If there is doubt as to the aetiology of cardiac disease, endomyocardial biopsy may demonstrate eosinophil infiltration. Serial monitoring of pulmonary function may be required if there is evidence of lung involvement.

Treatment

Treatment should be used to halt or reverse organ damage. Eosinophilia without evidence of end-organ damage does not usually require treatment. When underlying clonal or non-clonal disorders are identified they should be treated appropriately.

Patients with rearrangements of the *PDGFRA* or *PDGFRB* genes respond well to imatinib, with normalization of eosinophil counts within weeks. A trial of imatinib is also reasonable in patients with HES who lack a clonal marker, as a proportion of these patients also respond.

For patients who do not respond to imatinib, prednisolone is the initial treatment of choice. Steroids reduce blood eosinophilia and the inflammation resulting from tissue infiltration. Cardiac disease may respond, even in the absence of a significant reduction in the eosinophil count. Hydroxycarbamide and IFN-α may benefit patients resistant to steroids. Cladribine and ciclosporin were also found to be of use in some cases.

Prognosis

The reported prognoses of CEL and HES are highly variable, with estimates of 3-year survival ranging from 23% to 96%. This is likely to reflect heterogeneity within these two categories of patients. In patients with HES, indicators of a poor prognosis include lack of response to steroids, a markedly elevated eosinophil count, normal IgE levels, splenomegaly, dysplastic features and male sex. Many of these adverse prognostic indicators may simply be markers of clonal (versus reactive) eosinophilia.

Chronic neutrophilic leukaemia

Chronic neutrophilia is a very common entity, and is usually secondary to chronic infections, chronic inflammation or malignancy. A very small subgroup of patients with chronic neutrophilia have chronic neutrophilic leukaemia (CNL), a clonal haematological disorder (Figure 26.15). Given the absence of a specific marker for this disease, CNL, like ET, is a diagnosis of exclusion (see also Chapter 24).

The first description of CNL was probably that by Tuohy in 1920. In the ensuing 75 years, a total of less than 200 cases have been reported in the literature, mostly as isolated case reports. This rarity has hampered progress in the understanding of its pathogenesis.

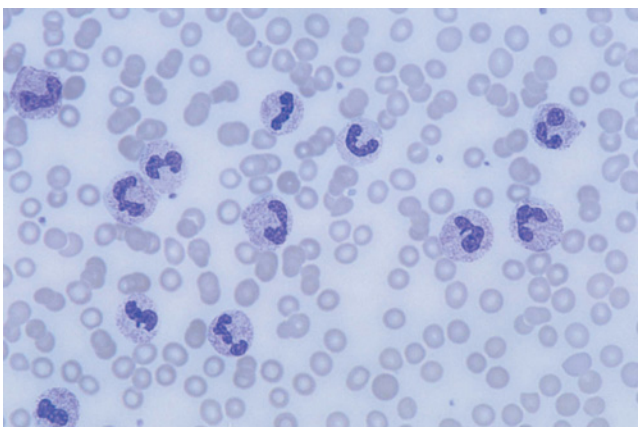


Figure 26.15 Abundance of mature neutrophils and band forms in a blood film from a patient with chronic neutrophilic leukaemia.

Pathophysiology

Isolated mild to moderate neutrophilia is commonly seen in many clinical contexts associated with inflammation, ranging from infections to tissue trauma/infarction, haemorrhage, arthritis, inflammatory bowel disease and many other ailments. Additionally, it can be seen in smokers, after vigorous exercise and in patients taking corticosteroids. In these contexts, the aetiology is usually apparent and such cases are rarely referred to a haematologist.

Marked chronic neutrophilia (neutrophils >20 × 10^9/L) is usually secondary to a chronic infection or an underlying malignancy. Neutrophilia is particularly common in metastatic cancer but can predate overt malignancy by months or years. Such leukaemoid reactions are thought to reflect the release of cytokines such as granulocyte colony-stimulating factor (G-CSF) and granulocyte–macrophage colony-stimulating factor (GM-CSF) from tumour cells.

Neutrophilia is clonal in a small number of cases, which can be subdivided into three main groups: 'true' CNL, atypical chronic myeloid leukaemia and neutrophilic chronic myeloid leukaemia (N-CML). The two groups are only distinguishable by the presence in the latter of a rare type of *BCR–ABL1* rearrangement that produces a 230-kDa fusion protein (p230). In addition, rare patients with a myelodysplastic syndrome can closely mimic CNL, but exhibit dysplastic features and there are anecdotal reports of PV evolving into a disorder indistinguishable from CNL. The literature also includes reports of what was thought to be clonal chronic neutrophilia in association with plasma cell dyscrasias. However, data are accumulating that this type of neutrophilia is non-clonal and probably a result of cytokine release from clonal plasma cells.

Diagnostic features of CNL are shown in Table 26.12.

Table 26.12 Diagnostic features of chronic neutrophilic leukaemia.

- Peripheral blood leucocytosis >25 × 10^9/L
- Segmented neutrophils and bands >80% of white blood cells
- Immature granulocytes <10% of white blood cells
- Myeloblasts <1% of white blood cells
- Hypercellular bone marrow biopsy
- Neutrophilic granulocytes increased in percentage and number
- Myeloblasts <5% of nucleated marrow cells
- Normal neutrophil maturation pattern
- Hepatosplenomegaly
- No identifiable cause of reactive neutrophilia
- No evidence of another haematological malignancy
- No Philadelphia chromosome or *BCR–ABL* fusion*
- No evidence of another myeloproliferative disorder (i.e. normal PCV, platelets <600, no bone marrow fibrosis or other features of PMF)
- No evidence of a myelodysplastic syndrome (i.e. no dysplasia, monocytes <1 × 10^9/L)

*Presence of *CSF3R* or *SETBP1* not yet incorporated into diagnostic criteria.
(Source: Valent *et al*., 2001 [*Leukemia Research* 2001; **25**: 603–25]. Adapted with permission of Elsevier.)

Clinical features and treatment

A review of the literature identified only 33 cases that fulfilled criteria for CNL. These cases exhibited a 2:1 male:female ratio and a median age at diagnosis of 62.5 years (range 15–86 years). The median survival was 30 months, with only 28% of cases surviving to 5 years. Transformation to AML ensued in 21% (7 out of 33) and this was invariably lethal. Other causes of death included sepsis and haemorrhage.

Haemoglobin was normal and platelet counts above 100 × 10^9/L in most cases. The mean leucocyte count at diagnosis was 54.3 × 10^9/L with mature and band forms in the peripheral blood. Vitamin B_{12} levels were raised and neutrophil alkaline phosphatase (NAP) levels were not low in most cases. Hyperuricaemia and gout were common. The spleen may be moderately enlarged. Bone marrow biopsies were markedly hypercellular and showed a marked granulocytic proliferation, as did the bone marrow aspirates. Cytogenetic abnormalities were seen in about a third, with deletion of del(20q), del(11q), del(12p), +8, +9, +21 being mentioned in WHO (2008) and being the only recurrent abnormality identified thus far. In a small number of cases, XCIP studies were used to demonstrate the clonal nature of CNL in patients lacking a cytogenetic marker. More series of different mutations in *CSF3R* and *SETBP1* were reported in patients with CNL, they are also seen in atypical chronic myeloid leukaemia.

The optimal treatment of CNL remains unclear. Oral cytoreductive agents such as hydroxyurea and busulfan can control the neutrophil count, as can IFN. The only potentially curative modality is allogeneic bone marrow transplantation and this option should be considered in younger patients. Some patients with CNL who have the *CSF3R* and *SETBP1* mutations have been reported to respond to ruxolitinib.

Neutrophilic chronic myeloid leukaemia

This entity is probably even more rare than CNL, with only a handful of documented cases in the literature. The reported cases followed a more benign course than conventional CML, with a lower white cell count, lower proportion of immature granulocytes, milder anaemia, less marked splenomegaly and a lower propensity to acute transformation. Given recent advances in the treatment of BCR–ABL1-related diseases, it is important to consider and exclude neutrophilic chronic myeloid leukaemia (N-CML) during the investigation of chronic neutrophilia.

Transient abnormal myelopoiesis of Down syndrome

Children born with Down syndrome show a 10–20-fold increased risk of leukaemia, despite not showing an increased incidence of other cancer types. One haematological disorder, characteristically found only in neonates with Down syndrome, is transient abnormal myelopoiesis (TAM), also known as transient myeloproliferative disorder or transient leukaemia.

Incidence, clinical features and treatment

About 10% of neonates with Down syndrome are thought to develop TAM, although this may be underestimated since many otherwise healthy babies with trisomy 21 do not routinely have blood tests performed. Additional problems with recognition arise from the fact that many normal neonates with Down syndrome show mild abnormalities of blood counts, including polycythaemia and thrombocytopenia. About 25% of neonates with TAM are asymptomatic, with the blood film showing sometimes significantly elevated numbers of circulating immature myeloid cells, including basophilic blasts, nucleated red cells, megakaryocyte fragments and thrombocytosis or thrombocytopenia. In symptomatic babies, the clinical features are variable, but can include neonatal jaundice, bleeding problems, respiratory distress and, rarely, liver failure.

The natural history of TAM is intriguing. In the majority of babies, the disorder resolves spontaneously by 3 months of age, without the need for treatment. Severely symptomatic infants, especially those with respiratory or hepatic dysfunction, can be treated very effectively with low-dose cytarabine chemotherapy. About 20% of neonates with TAM will subsequently develop acute megakaryoblastic leukaemia (AMKL) before the age of 4 years, although sometimes the leukaemia develops without antecedent TAM.

Pathophysiology

The fact that TAM only occurs in neonates points to it being a disorder of fetal haemopoiesis, and this presumably also explains why it is self-limiting in most cases. Importantly, a prospective analysis of clinical findings, blood counts and smears, and GATA1 mutation status in 200 neonates with Down's syndrome showed that all subjects had multiple blood count and smear abnormalities. Surprisingly, 195 of 200 (97.5%) had circulating blasts. GATA1 mutations were detected in 8.5% of cases (all with blasts >10%). Furthermore, low-abundance GATA1 mutant clones were detected by targeted next-generation resequencing (NGS) in a further 20% of cases that were clinically and haematologically unremarkable. The term 'silent TAM' for neonates with Down syndrome and with GATA1 mutations detectable only by NGS was suggested.

The characteristic association between TAM and AMKL indicates that a multistep model of mutation is operative in this clinical progression. TAM appears to be caused by a single GATA1 mutation in the context of constitutive trisomy 21. Subsequent AMKL evolves from a pre-existing TAM clone through the acquisition of additional mutations, with major mutational targets including multiple cohesin components (53%), CTCF (20%), and EZH2, KANSL1 and other epigenetic regulators (45%), as well as common signaling pathways, such as JAK family kinases, MPL, SH2B3 (LNK) and multiple RAS pathway genes (47%).

Selected bibliography

Campbell PJ, Green AR (2006) The myeloproliferative disorders. *New England Journal of Medicine* **355**(23): 2452–66.

Campbell PJ, Scott LM, Buck G *et al*. (2005) Definition of subtypes of essential thrombocythaemia and relation to polycythaemia vera based on JAK2 V617F mutation status: a prospective study. *Lancet* **366**(9501): 1945–53.

Cervantes F, Vannucchi AM, Kiladjian JJ *et al*. (2013) Three-year efficacy, safety, and survival findings from COMFORT-II, a phase 3 study comparing ruxolitinib with best available therapy for myelofibrosis. *Blood* **122**(25): 4047–53.

Gotlib J, Cross NC, Gilliland DG (2006) Eosinophilic disorders: molecular pathogenesis, new classification, and modern therapy. *Best Practice and Research in Clinical Haematology* **19**(3): 535–69.

Gotlib J, Maxson JE, George TI, Tyner JW (2013) The new genetics of chronic neutrophilic leukemia and atypical CML: implications for diagnosis and treatment. *Blood* **122**(10): 1707–11.

Harrison CN, Robinson SE. Myeloproliferative disorders in pregnancy. (2011) *Hematology and Oncology Clinics of North America* **25**(2): 261–75.

Harrison CN, Bareford D, Butt N *et al*. (2010) Guideline for investigation and management of adults and children presenting with a thrombocytosis. *British Journal of Haematology* **149**(3): 352–75.

James C, Ugo V, Le Couedic JP *et al*. (2005) A unique clonal JAK2 mutation leading to constitutive signalling causes polycythaemia vera. *Nature* **434** (7037): 1144–8.

Klampfl T, Gisslinger H, Harutyunyan AS *et al*. (2013) Somatic mutations of calreticulin in myeloproliferative neoplasms. *New England Journal of Medicine* **369**(25): 2379–90.

Landolfi R, Marchioli R, Kutti J *et al*. (2004) Efficacy and safety of low-dose aspirin in polycythemia vera. *New England Journal of Medicine* **350**(2): 114–24.

McMullin MF, Bareford D, Campbell P *et al*. (2005) Guidelines for the diagnosis, investigation and management of polycythaemia/erythrocytosis. *British Journal of Haematology* **130**(2): 174–95.

Metcalfe DD (2008) Mast cells and mastocytosis. *Blood* **112**(4): 946–56.

Nangalia J, Massie CE, Baxter EJ *et al*. (2013) Somatic CALR mutations in myeloproliferative neoplasms with nonmutated JAK2. *New England Journal of Medicine* **369**(25): 2391–405.

Ortmann CA, Kent DG, Nangalia J *et al*. (2015) Effect of mutation order on myeloproliferative neoplasms. *New England Journal of Medicine* **372**(7): 601–12.

Pikman Y, Lee BH, Mercher T *et al*. (2006) MPLW515L is a novel somatic activating mutation in myelofibrosis with myeloid metaplasia. *PLoS Medicine* **3**(7): e270.

Reilly JT, McMullin MF, Beer PA *et al*. (2012) Guideline for the diagnosis and management of myelofibrosis. *British Journal of Haematology* **158**(4): 153–71.

Roberts I, Alford K, Hall G *et al*.; (2013) Oxford-Imperial Down Syndrome Cohort Study Group. GATA1-mutant clones are frequent and often unsuspected in babies with Down syndrome: identification of a population at risk of leukemia. *Blood* **122**(24): 3908–17.

Scott LM, Tong W, Levine RL *et al*. (2007) JAK2 exon 12 mutations in polycythemia vera and idiopathic erythrocytosis. *New England Journal of Medicine* **356**(5): 459–68.

Tefferi A, Thiele J, Orazi A *et al*. (2007) Proposals and rationale for revision of the World Health Organization diagnostic criteria for polycythemia vera, essential thrombocythemia, and primary myelofibrosis: recommendations from an ad hoc international expert panel. *Blood* **110**(4): 1092–7.

Vannucchi AM, Kiladjian JJ, Griesshammer M *et al*. (2015) Ruxolitinib versus standard therapy for the treatment of polycythemia vera. *New England Journal of Medicine* **372**(5): 426–35.

Verstovsek S, Mesa RA, Gotlib J *et al*. (2013) Efficacy, safety and survival with ruxolitinib in patients with myelofibrosis: results of a median 2-year follow-up of COMFORT-I. *Haematologica* **98**(12): 1865–71.

Wilkins BS, Erber WN, Bareford D *et al*. (2008) Bone marrow pathology in essential thrombocythemia: interobserver reliability and utility for identifying disease subtypes. *Blood* **111**(1): 60–70.

Yoshida K, Toki T, Okuno Y *et al*. (2013) The landscape of somatic mutations in Down syndrome-related myeloid disorders. *Nature Genetics* **45**(11): 1293–9. Erratum in: *Nature Genetics* **45**(12): 1516.

CHAPTER 27

Chronic lymphocytic leukaemia and other chronic B-cell disorders

27

Emili Montserrat[1] and Peter Hillmen[2]

[1]Institute of Haematology and Oncology, Hospital Clinic, University of Barcelona, Barcelona, Spain
[2]Department of Haematology, Leeds Teaching Hospitals NHS Trust, Leeds, UK

Introduction

Chronic B-cell lymphoproliferative disorders include a number of entities arising from mature B lymphocytes and which involve primarily the blood, bone marrow and lymphoid organs such as the lymph nodes and spleen. A constant finding in all these entities is the presence in peripheral blood of leukaemic cells in various degrees. Some of these conditions could be considered as primary leukaemias. Others represent the leukaemic phase of indolent non-Hodgkin lymphomas (NHL) (Table 27.1). Important progress has been made in the management of these disorders.

General aspects of diagnostic methodology (see also Chapter 19)

Examination of the morphology of leukaemic cells in well-prepared peripheral blood and bone marrow films is still the first diagnostic procedure and cannot be replaced by any of the newer, more modern diagnostic techniques. The second element for the diagnosis of any type of lymphocytosis (defined as a permanent, non-transient absolute blood lymphocyte count $>5 \times 10^9$L) is to determine the immunophenotype (B or T) and whether the process is clonal or polyclonal by light-chain restriction analysis. Once a monoclonal B-cell proliferation is established, the diagnosis can be further clarified by using a panel of monoclonal antibodies and other markers (Table 27.2).

When the absolute lymphocyte count is high ($>5 \times 10^9$/L) and the blood film and immunophenotype shows unequivocal features of CLL, PLL or HCL, lymph node histology is not required. If the diagnosis of NHL is a possibility it is essential to obtain tissue for histology to classify the disease. Bone marrow biopsies can confirm a diagnosis of CLL or NHL and provide information about the mechanism of anaemia or thrombocytopenia. For disorders with an enlarged spleen, such as B-PLL and HCL, SMZL and some forms of MCL, spleen histology may be of diagnostic value, in selected cases. Fluorescence *in-situ* hybridisation (FISH) analysis allows the assessment of cytogenetic abnormalities on interphase cells. However, there is no specific genetic abnormality in CLL. FISH, however, is important for excluding the two NHLs that have characteristic abnormalities, MCL with t(11;14)(q13;q32) and follicular lymphoma with t(14;18)(q32;q21). Next-generation sequencing techniques are unfolding the molecular heterogeneity of B-cell chronic lymphoproliferative disorders, but are not employed in routine diagnosis. Finally, whenever possible, biological material not used for diagnostic purposes should be preserved, which can be useful for both research purposes and further studies of clinical interest for the patient.

Chronic lymphocytic leukaemia

Definition

Chronic lymphocytic leukaemia (CLL) is due to the relentless accumulation in bone marrow, peripheral blood, and

Postgraduate Haematology, Seventh Edition. Edited by A Victor Hoffbrand, Douglas R Higgs, David M Keeling and Atul B Mehta.

Table 27.1 B-cell chronic lymphoproliferative disorders with leukaemic expression.

Leukaemias
Chronic lymphocytic leukaemia
B-prolymphocytic leukaemia*
Hairy-cell leukaemia
Lymphomas with leukaemic expression
Follicular lymphoma
Splenic marginal zone lymphoma
Mantle-cell lymphoma
Splenic lymphomas unidentified
Hairy-cell leukaemia variant**
Diffuse red pulp splenic lymphoma**

*Corresponds in most cases to MCL
**Overlapping with SMZL

lymphoid tissues of monoclonal B lymphocytes with a distinct immunophenotype (i.e., $SmIg^{weak}$, CD5+, CD19+, CD23+, CD200+). CLL and small lymphocytic lymphoma (SLL) are considered to be the same mature B-cell neoplasm that mainly differ in the extent to which the tumor involves blood (CLL) or lymphoid tissue (SLL). Within the SLL/CLL spectrum, about 10% of cases present as SLL and 90% as CLL. In many instances SLL evolves to CLL over time. CLL is a heterogeneous disorder from both the biological and clinical points of view. The median survival of patients with CLL is around 10 years, but the individual prognosis is highly variable, ranging from a few months to a normal lifespan. In spite of important progress in its therapy, CLL remains incurable.

Epidemiology

CLL is the most frequent form of leukaemia in Western countries. In the United States the age-adjusted incidence rate is 4.2 cases per 100,000/year. The disease predominates in males (1.5–2:1) and the median age of patients at diagnosis is around 70 years. The incidence of CLL dramatically increases with age, reaching >20 cases per 100,000/year in individuals older than 70. In Europe, the incidence rate for all subtypes of leukaemia is 7.2 per 100,000, of which 34% of cases are CLL, translating to approximately 12,500 new CLL cases per year. Each year, CLL is responsible for approximately 6000 deaths across Europe. A remarkable epidemiologic characteristic of CLL is its racial diversity. While in Western countries CLL accounts for around 30% of all cases of adult leukaemia, in Asian countries (and also in Asian-American people) and in the Far East (Japan) CLL is an infrequent disease.

Genetic predisposition

The relative risk of developing CLL in patients' relatives is three to eight times higher than in the normal population The molecular basis of familial CLL is currently being unravelled by means of genome-wide association studies, including both families and sporadic cases. Familial CLL often cosegregates with other B-cell disorders, mainly NHL, suggesting that the predisposition to CLL is mediated through pleiotropic genes.

Pathogenesis

The key feature in CLL pathogenesis is the progressive accumulation of CD5-positive B cells because of a failure in programmed cell death or apoptosis. Neoplastic CLL B cells are antigen-experienced, expressing CD23, CD25, CD27, CD69

Table 27.2 Inmunophenotype and other markers in B-cell chronic lymphoproliferative disorders.

	SmIg	CD20	CD5	CD10	CD23	CD11C	CD103	CD200	OTHERS
CLL	−/+	−/+	+	−	+	−/+	−	+	FMC7(−)
B-PLL	+	+	−/+	−/+	−/+	−	−	−	*TP53, c-MYC* mutations
HCL	+	+	−	−	−	+	+	+	ANNEXIN A1(+), DBA44(+) (∗) *BRAF* V600E mutation
LPL	+	+	−	−	−		−	−	Igc. (+), trisomy 4, *MYD88* mutation, *CXCR4* mutation
SMZL	+	+	−/+	−	−/+	−/+	−	+	BCL-2(+); del(7q32)
MCL	+	+	+	−	−	−	−	−	t(11;14)(q13;q32); CYCLIN D1(+); Trisomy 3; SOX11(+), CD200(-)
FL	+	+	−	+/−	−/+	−	−	−	t(14;18); BCL-2(+), BCL6(+), CD43(-)

B-PLL, B- prolymphocytic leukaemia; CLL, chronic lymphocytic leukaemia(∗) immunohistochemistry; FL, follicular lymphoma; HCL, hairy-cell leukaemia; LPL, lymphoplasmacytic lymphoma; MCL, mantle-cell lymphoma; SMZL, splenic marginal zone lymphoma.

CD71, ROR1 and CD200. In half of cases the immunoglobulin variable region heavy chain (*IGHV*) gene is mutated, but in the rest is unmutated. The immunoglobulin gene repertoire in CLL shows biased (non-random) V_H usage compared to normal B cells. This restricted V_H usage is linked to the degree of somatic hypermutation.

Approximately 20% of immunoglobulin subsets show similar (homologous) patterns of the complementarity-determining region (CDR3) in both heavy and light chains, designated 'stereotypes'. These are more common in unmutated IGHV cases and provide evidence for antigen selection in the pathogenesis of CLL. Unmutated IGHV cases are associated with autoreactivity and polyreactivity to certain molecules and have a more proliferative pattern of disease, with a more aggressive clinical course. There is no evidence that IGHV-mutated and IGHV-unmutated CLL represent two distinct diseases, but rather two clinicobiological forms of the same disorder.

Although most cells are in $G_{0/1}$ phase, there is a fraction of cells (0.1–1%) rapidly duplicating and renovating every day. The resistance of these cells to apoptosis related to the upregulation of the antiapoptotic proteins such as BCL2 and MCL1 and downregulation of proapoptotic proteins such as BCLX. Microenvironment (a complex mixture of stromal cells, 'nurse-like' cells, T-cells and macrophages) plays an essential role in protecting CLL cells from death through a number of molecules. This is demonstrated by the fact that *in vitro* CLL cells rapidly enter into apoptosis unless they are cocultivated with stromal cells or in a stromal-mimicking milieu. The pathogenesis of CLL implies therefore an intimate relationship of B cells, through the B-cell antigen receptor (BCR), microenvironment and T cells. Normal B cells have a phenomenal ability to proliferate in response to antigens in that single B cells will proliferate to many millions in a very short period of time to create the primary and secondary immune response. CLL cells utilize BCR signalling to rapidly proliferate. Furthermore, normal B cells that are self-reacting are deleted or anergized to prevent autoimmunity. In CLL there is evidence that many of the antigens to which the CLL cell is responding are autoantigens. Finally, normal B cells have to be deleted (apoptosed) when the antigen they are responding to has been cleared with a very small proportion becoming long-lived memory cells. All of these features are part of the biology of a CLL cell, which is proliferative when given the correct signals, is anergised, but may lead to autoimmunity and has dysfunctional apoptotic mechanisms (ubiquitous expression of BCL-2). This pattern of abnormalities leads to the accumulation of mature CLL B cells.

Finally, all CLL cases are preceded by a subclinical phase consisting of a high number of polyclonal B cells that would be subsequently selected and expanded upon the influence of some genetic events. This would give origin to a monoclonal B-cell lymphocytosis (MBL)-like picture (see Diagnosis). Next, accumulation of genetic alterations might cause transformation of a small proportion of MBL clones to CLL. Genetic drivers of this effect are the deletion of miR15 and miR16 in chromosome 13q14, trisomy 12 and *MYD88* mutations. Over time, other mutations appear, including those involving *TP53, ATM, NOTCH1* and *SF3B1*. Finally, in the most advanced cases *C-MYC and CDKN2A* mutations can be also present. Through a selection process, a minute clone harbouring these genetic lesions can overcome other clones and become predominant. Importantly these events may occur either spontaneously or, and more frequently, are triggered by treatment.

Clinical features

Approximately, 80% of patients are diagnosed while asymptomatic. In the rest of the patients, painless generalized peripheral lymphadenopathy (i.e. neck, axillae, inguinal) is frequent. In contrast, mediastinal and retroperitoneal lymphadenopathy are uncommon. On some occasions, the investigation of bacterial infections, particularly pneumonia, or herpes virus infections may lead to the diagnosis. Rarely, the disease is discovered during the diagnostic workup of an autoimmune haemolytic anemia (AIHA) or, even less frequently, an immune thrombocytopenia (ITP). A peculiar feature in some patients is a high sensitivity to insect, mainly mosquito, bites. Rarely, CLL may involve extrahaematological tissues, such as skin, liver, kidney or central nervous system; in such cases, however, disease transformation needs to be ruled out. In addition, vasculitis, hypercalcaemia and nephrotic syndrome are occasionally observed. Spontaneous regression of the disease can be observed in 1–2% of patients per year. The remission is rarely complete, most patients having persistent disease by flow cytometry and the remission is usually transient. The mechanisms behind spontaneous remissions are poorly understood, but could involve immune-mediated reactions.

Laboratory features (see also Chapter 31)

The hallmark of the disease is an increased white blood cell (WBC) count with a high percentage of lymphocytes and absolute lymphocytosis (Figure 27.1). A proportion of prolymphocytes (1–10%) are nearly always seen when WBCs are higher than 30×10^9/L. If the proportion of prolymphocytes is greater than 10%, it represents a variant designated chronic lymphocytic/prolymphocytoid leukaemia (CLL/PL) (Figure 27.2). Some patients have a mixed pattern of small and large cells and others have lymphoplasmacytoid features or even cells with nuclear clefts. These are often associated with other atypical features, such as trisomy 12 or FMC7 positivity.

In asymptomatic patients, anaemia is found in less than 10%. Importantly, anaemia is not always due to the infiltration of the bone marrow by the disease; other causes such as autoimmunity, iron, folate or vitamin B12 deficiency may cause anaemia, particularly in the elderly or poorly nourished

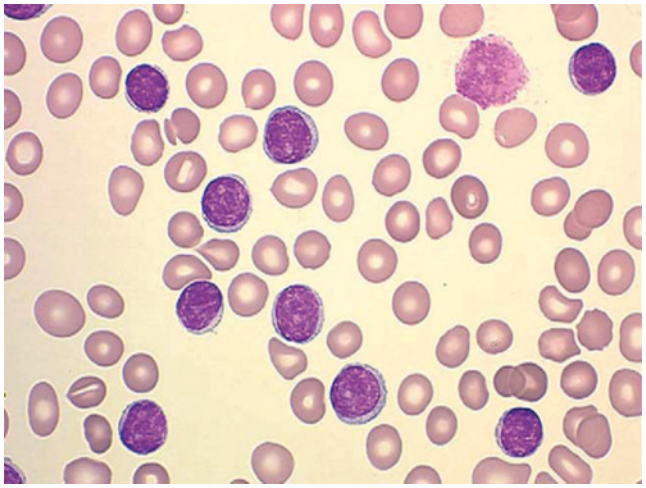


Figure 27.1 Blood film of a typical case of CLL. Small lymphocytes with scanty cytoplasm and tenuous nucleoli. In the upper right corner a Gumprecht cell is observed.

people. Likewise, thrombocytopenia is infrequent at diagnosis. The appearance of a marked thrombocytopenia (e.g. $<20 \times 10^9$/L) in a patient with previous normal counts should raise the possibility of an immune origin, particularly in the absence of anaemia. Hypogammaglobulinaemia is frequent (30% of patients) and tends to worsen over the course of the disease. Serum immunofixation can demonstrate an M component (usually of the IgM type) in around 10–15% of patients. A positive direct antiglobulin test (DAT) is observed in around 5% of the patients at the time of diagnosis, with clinically apparent AIHA being less frequent.

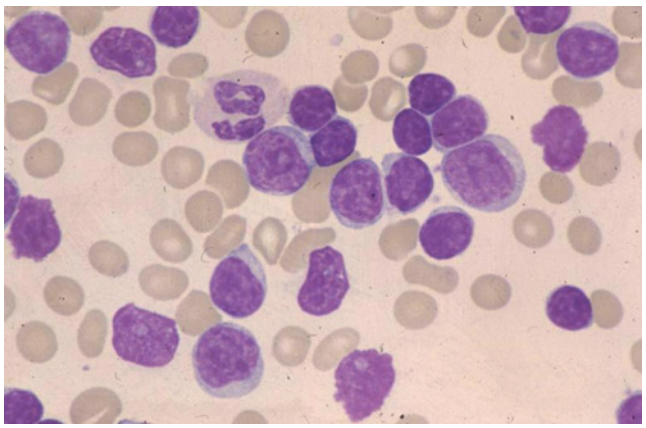


Figure 27.2 Blood film from a case of CLL/PL. Note the dual population of small lymphocytes and larger nucleolated prolymphocytes.

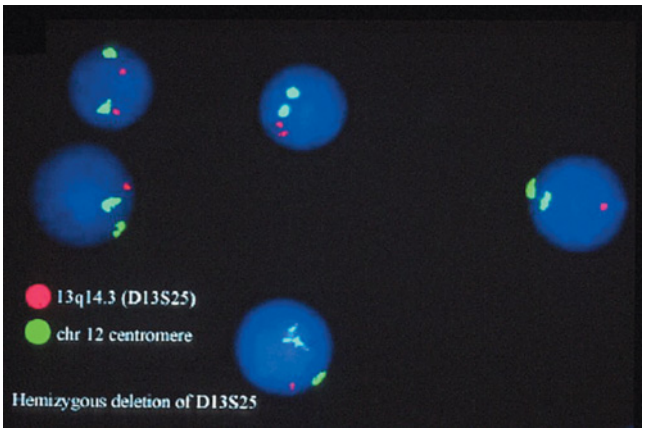


Figure 27.3 FISH analysis with a 13q14 probe showing missing red dots in several cells denoting 13q deletion; all the cells are disomic for chromosome 12 (normal pattern).

Genetic and molecular features

The analysis of the karyotype of CLL lymphocytes by conventional techniques does not yield informative results in many cases. FISH analysis uses specific probes to detect the most common genetic changes on interphase cells (Figures 27.3 and 27.4). However, none of the abnormalities found in CLL are unique or specific to this disease and can be found in other B-cell disorders. There are no unique translocations in CLL either. The most frequent cytogenetic abnormalities are del(13q), del(11q), trisomy 12, del(6q) and del(17p), which are associated with some peculiar clinical features and outcomes (Table 27.2). Del(17p) is observed in 5–50% of cases, depending on the characteristics of

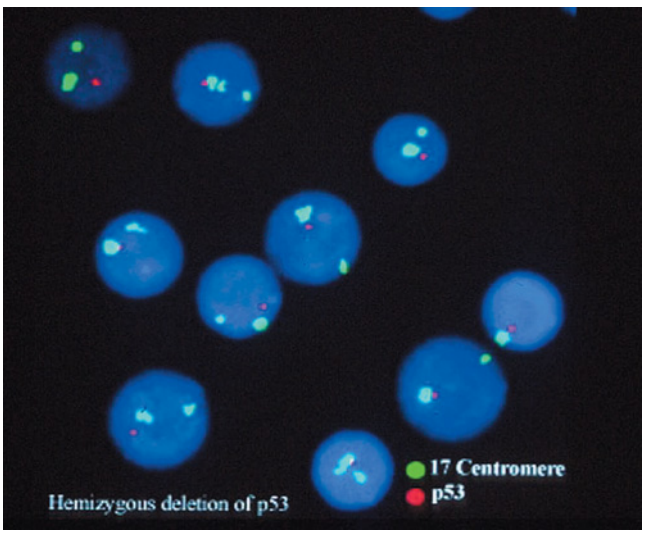


Figure 27.4 FISH analysis with a *TP53* probe (red) and a control 17 probe. Note many cells with only one red dot (instead of two), indicating *TP53* deletion.

the studied population (e.g. untreated, treated). Translocations (e.g. t(14;19) or t(14;18)) have been reported in a few cases, but whether these represent typical CLL cases is equivocal. Complex karyotypes (more than three abnormalities) can be observed. In around 30% of cases, cytogenetic evolution occurs over the course of the disease. *TP53* mutations are usually observed in del(17p), but in 5–15% of cases *TP53* mutations can occur in the absence of del(17p). Mutations of other genes such as *MYD88, NOTCH-1,* SF3B1 and *BIRC3* can be observed in 5–15% of cases. Epigenetic lesions are present in CLL and are the subject of many studies to better understand the landscape of genetic lesions in this disease.

Pathology features

In typical cases, the histology of lymph nodes shows effacement of the architecture by small lymphocytes with a pseudofollicular pattern of pale areas that correspond to proliferation centre. Proliferation centre contain numerous T cells, a fine network of dendritic cells and B cells with an increased expression of proliferation-associated markers and antiapoptotic molecules compared to the non-proliferation component. There are patients whose lymph nodes have expanded proliferation centre (pseudofollicles) with a high proliferation rate ('accelerated CLL'). Differences observed in lymph node histopathology are consistent with the concept of CLL as an evolving disorder, from typical CLL to 'accelerated CLL' and eventually disease transformation (see Disease Transformation). The spleen shows a diffuse infiltration, mainly affecting the white pulp.

The bone marrow displays a variable degree of infiltration by the disease and, in contrast to follicular lymphoma, there is no paratrabecular infiltration. Four infiltration patterns have been described: nodular, interstitial, mixed (nodular+interstitial) and diffuse. While in early clinical phase nodular and interstitial patterns predominate, in advanced phase a diffuse infiltration is the norm.

Diagnostic criteria

According to the International Workshop on CLL (IWCLL), the diagnosis of CLL is sustained by the following parameters:

- Presence in peripheral blood of $>5 \times 10^9$/L monoclonal B lymphocytes persisting for at least 3 months
- Demonstration of the clonality of the population (κ/λ analysis)
- Characteristic immunophenotype: $SmIg^{weak}$, $CD5^+$, $CD19^+$, $CD20^{weak}$, $CD23^+$.

Regarding the immunophenotype, FMC7 (a CD20 epitope) and CD79b are usually absent or weakly expressed. Based on immunophenotypic characteristics and giving one point to each one of the following: $CD5^+$, $CD23^+$, $FMC7^{weak}$, SmIg (κ/λ staining)^{weak} and $CD79b^{weak}$ Matutes and Catovsky showed that in patients with a score of 4–5, the diagnosis is virtually always CLL, while in those cases with a score <3, the diagnosis of CLL is extremely unlikely. The typical immunophenotype for CLL as accepted in current guidelines should be revisited because of the availability of markers highly characteristic of CLL, such as CD200 and ROR1.

Bone marrow aspirate and biopsy are not necessary for diagnosis, but can be useful in cases where diagnosis is complex, and to provide invaluable information about the origin of cytopenias (i.e. bone marrow failure because of disease infiltration versus 'peripheral mechanisms' such as hypersplenism or autoimmunity).

Differential diagnosis

Usually, the diagnosis of CLL is straightforward and does not present difficulties. However, NHL in leukaemic phase, particularly mantle-cell lymphoma, lymphoplasmacytic lymphoma and marginal zone lymphoma, can mimic CLL. In CLL a small percentage (i.e. 1–10%) of atypical lymphocytes (e.g. prolymphocytes, lymphoplasmacytoid cells, cleaved cells or centrocytes) can be present in blood. If the percentage of atypical cells is >10% but <55%, CLL with prolymphocytoid cells (CLL/PL) is accepted as an atypical variant of CLL. However, the diagnosis of atypical CLL should not be accepted without having excluded a lymphoma in leukaemic phase.

Lymphoplasmacytic lymphoma (Waldenström's macroglobulinaemia) is characterized by a proliferation of an admixture of small lymphocytes, plasmacytoid lymphocytes and plasma cells; a monoclonal IgM peak is a hallmark of the disease. The diagnosis requires the exclusion of other B-cell disorders with plasmatic differentiation. *MYD*88(L265P) mutations are found in nearly all patients, and mutations of *CXCR4* are present in around 30% of cases. In some cases a moderate increase of lymphocytes in blood is observed. The presence of more than 55% prolymphocytes is a feature that arbitrarily defines B-cell prolymphocytic leukaemia (B-PLL), an extremely infrequent disorder that in most cases corresponds to mantle-cell lymphoma (MCL) in leukaemic phase. In fact B-PLL is a morphological diagnosis, which may represent cases that are either poor-risk CLL or MCL, and is probably not a discrete entity in its own right. CLL does not evolve to B-PLL.

Immunophenotyping of the leukaemic cells is useful in the differential diagnosis. In particularly difficult cases, biopsy of involved lymph nodes and bone marrow, as well as genetic and molecular studies, can be of help (Table 27.2). Regarding cytogenetic and molecular studies, it is debatable whether translocations of the *IGH* locus are a feature of CLL. Nevertheless, systematic interphase FISH with an *IGH* probe can detect up to 5% of cases with such translocations. Excluding t(11;14)(q13,q32), a feature of MCL, some apparently genuine cases of CLL with t(14;18)(q32;q21) or *IGH/BCL2*, t(14;19)(q32;q13) or *IGH/BCL3*, and very rarely t(8;14)(q24.1;q32) involving the *MYC* gene have been reported. Of note, the latter cases were

Table 27.3 FISH cytogenetics: biological and clinical correlates.

Abnormality	(%)	Correlates		Median survival (years)
		Biological	Clinical	
Del(13q14) isolated	25–40	Mutated *IGHV* genes *MYD88* mutations	Good prognosis	>15
Trisomy 12	15–20	Atypical morphology (prolymphocytes) Atypical immunophenotype (CD20+FMC7+, CD11c+) *NOTCH1* mutations	Short treatment-free survival	~ 7.5
Del(11q22-23)	20–25	*ATM* mutations (30%) *BIRC 3* mutations	Tumoral forms Shorter time to progression. Longer PFS with FCR vs. FC or F	~ 6
Del(17p13.1)	5–3	Unmutated *IGHV* genes *TP53* mutations Clonal evolution Complex karyotype	Resistance to therapy	~ 4

described as CLL/PL and some as typical B-PLL. *IGH* translocations have often been described as atypical and associated with poor prognosis. Cases with t(14;19), although rare, have been more extensively characterized. This translocation has been reported in a heterogeneous group of B-cell malignancies, including marginal zone lymphomas and, in about half of cases, with a diagnosis of CLL frequently associated with trisomy 12, unmutated *IGHV* genes and low frequency of del(13q).

Chronic lymphocytic/prolymphocytoid leukaemia (CLL/PL) is a proliferative form of CLL displaying 10–55% atypical lymphocytes in blood, not infrequently associated with trisomy 12, FMC7 expression, *TP53* deletion, high rate of cell divisions (as shown with the antibody Ki-67) and histologically by many proliferating centres in lymph nodes and bone marrow sections ('accelerated CLL').

When atypical features are present, including those of CLL/PL, it is important to consider the differential diagnosis of NHL evolving with lymphocytosis, notably MCL. The best method to exclude (or confirm) MCL is to test for t(11;14) by FISH and/or to demonstrate expression of cyclin D1 by immunohistochemistry. Other NHLs can be excluded by histological criteria (see Chapter 33) and the combination of peripheral blood morphology, immunophenotype and genetic features (Table 27.2).

A rare abnormality is polyclonal B-cell lymphocytosis. In affected individuals, the absolute number of B cells is increased, but, as shown by κ/ λ staining, they are polyclonal. Disease features are median age 40 years, predominance in women, lymphocyte count 3–15 × 10^9/L (median 6.5), splenomegaly (15%) and mild bone marrow infiltration, sometimes showing intrasinusoidal distribution; 95% of patients are smokers and 90% have HLA-DR7 in their genotype. The cell markers are non-clonal cell proliferation, $CD19^+$, $FMC7^+$, $CD5^-$, $CD23^-$, $CD27^+$ and the patients have elevated levels of polyclonal IgM in serum. Despite the benign and non-clonal nature, distinct abnormalities have been shown in the B lymphocytes such as an extra copy of the long arm of chromosome 3 in the form of iso(3q) and also the presence of *IGH/BCL2* rearrangements. A minority of subjects (around 3%) may develop NHL. The blood picture is pleomorphic and frequently shows binucleated lymphocytes (Figure 27.5), which are characteristic of this condition. These persons should not receive therapy, but be closely followed to detect a potential disease transformation.

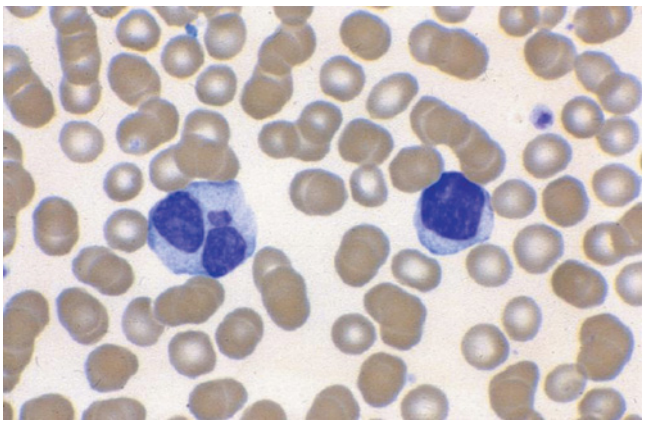


Figure 27.5 Blood film from a case of polyclonal B-cell lymphocytosis showing a binucleated cell characteristic of this condition.

Table 27.4 Differences between CLL, SLL, and MBL.

	CLL	SLL	MBL
Clonal B cells in peripheral blood	≥ 5000/μL	<5000 /μL	<5,000/μL
Lymphadenopathy	Possible	Yes	No*
Organomegaly	Possible	Possible	No
Anaemia, thrombocytopenia	Possible	Possible	No
Bone marrow involvement	Yes	Possible	Yes
Disease-related symptoms	Possible	Possible	No
Lymph node biopsy	Not required	Required	Not required

*Negative imaging studies required

Diagnosis of CLL-related conditions

Most cases of MBL and all cases of CLL and SLL are part of the same spectrum of CD5+ B-cell chronic lymphoproliferative disorders, whose main differences are shown in Table 27.4 and are discussed here. Also described is disease transformation that frequently represents the latest stage in the history of both CLL and SLL.

Monoclonal B-cell lymphocytosis (MBL)

In the absence of lymphadenopathy, organomegalies (detected clinically or radiologically), cytopenias and clinical symptoms, the presence of less than 5×10^9/L monoclonal B lymphocytes in blood with a CLL phenotype is defined as MBL. The incidence of MBL in the normal population is about 3–10% and is higher in first-degree relatives of patients with CLL and older subjects. Its detection highly depends on the sensitivity of the flow-cytometry technique. The cut-off value of 5×10^9/L monoclonal cells for diagnosis of MBL was selected arbitrarily for consistency with CLL diagnostic guidelines, and is not based on objective clinical data. Many studies have compared MBL versus Rai stage 0 CLL, finding that a higher cut-off lymphocyte level (i.e. $>10 \times 10^9$/L) is better for identifying those patients more likely to progress to symptomatic CLL. On the other hand, MBL represents a miscellaneous category ranging from patients with minute clonal B-cell populations detected through general population studies, to patients with absolute lymphocytosis that are referred for consultation, currently defined as 'clinical' MBL. There are cases in which there is more than one B-cell clone, and even others in which a mixture of monoclonal B cells and monoclonal T cells can be detected. Although the immunophenotype is in most cases identical to that of CLL, cases with atypical immunophenotype (e.g. CD5-negative) or resembling MZL or lymphoplasmacytic lymphoma have been described. The clonal population can spontaneously regress over time or, conversely, increase and evolve to CLL, which occurs in around 1–2% of cases per year.

Small lymphocytic lymphoma (SLL)

The diagnosis of SLL requires the presence of lymphadenopathy or organomegaly (detected either clinically or by imaging studies), cytopenias or other disease-related features with $<5 \times 10^9$/L B lymphocytes in blood. By definition, SLL cells display the same immunophenotype as CLL cells, although the diagnosis should be confirmed by histologic evaluation of a lymph node or other tissues. The prognosis and treatment is the same as in CLL. Patients with SLL requiring treatment should be included in CLL trials.

Richter's syndrome (RS)

Richter's syndrome (RS) represents the development of an aggressive lymphoma, most commonly a diffuse large B-cell lymphoma (DLBCL) in a patient previously (or simultaneously) diagnosed with CLL. The diagnosis needs to be substantiated by the biopsy of a lymph node or another lymphoid tissue. Rarely, RS presents in peripheral blood or bone marrow smears (Figure 27.6) (see also Disease complications).

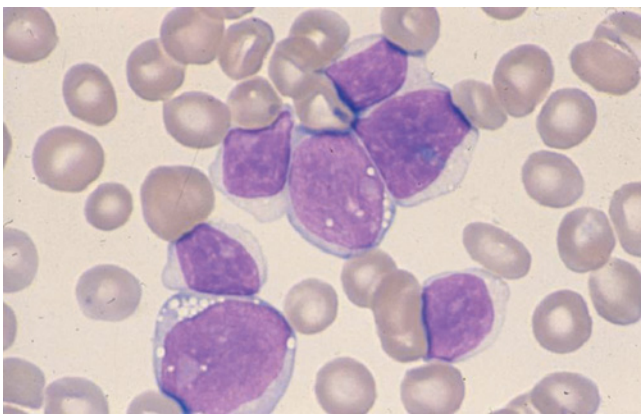


Figure 27.6 Blood film of a patient with Richter's syndrome with circulating large blast cells that were positive for surface membrane immunoglobulin.

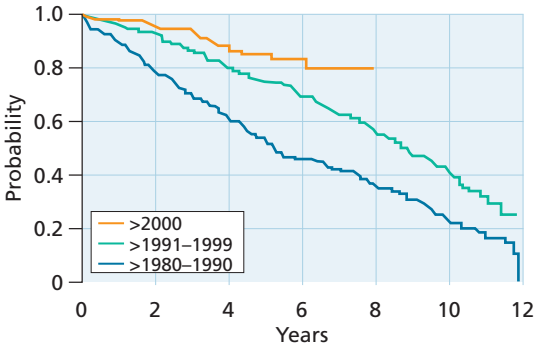


Figure 27.7 CLL patients survival over the years (Hospital Clínic Barcelona).

Prognosis

The median survival of patients with CLL has improved over the last years and is now around 10 years (Figure 27.7). This is due at least in part to better therapies, particularly in younger patients. Other factors include a more accurate diagnosis of B-cell chronic disorders (allowing separation of CLL from other, more aggressive conditions) and an early diagnosis, inherently accompanied by a longer survival.

The individual prognosis, however, is heterogeneous and ranges from a few months to a normal lifespan. Prognostication is therefore an essential component in the management of patients with CLL. Although somewhat overlapping, it is useful to distinguish, under the umbrella of outcome predictors, those parameters that predict disease progression, and hence need of therapy (prognostic factors), and those informing about the probability that an individual patient will respond to a given therapy (predictive factors) (Table 27.5; see also Treatment).

Table 27.5 Most important Outcome Predictors in CLL.

Prognostic factors	
Consolidated	**Complementary value**
Clinical stages	Blood lymphocyte count
IGHV mutational status	ZAP70
Cytogenetics	CD38
LDT	Beta-2 microglobulin
Predictive factors	
Consolidated	**Requiring more studies**
Del(17p)/*TP53* mutation Del(11q)	*NOTCH1* mutation (*)

LDT, lymphocyte doubling time; (*), no benefit from anti-CD20 MoAbs treatment.

It is important to underscore that the mere correlation of a parameter with a given outcome does not necessarily qualify it as outcome predictor. Outcome predictors need to fulfill a number of requisites to be considered as such. For example: standardization, inter- and intrareproducibility, independent prognostic value from other predictors of the same outcome and direct consequences in the management of a substantial proportion of patients. Importantly, outcome predictors can complement, but not replace, clinical expertise and sound clinical judgment.

Age

Not surprisingly, older patients have a poorer prognosis than younger ones. Age by itself should not be a criterion to decide treatment modality; as discussed in the Treatment section, 'biological' age (as reflected by performance status, frailty and comorbidity) rather than 'chronological' age should be considered.

Sex

The incidence of CLL is higher in men than in women. The reason for this imbalance is unclear. In addition, females appear to have a longer survival than males. Whether this is because of the gender itself or the characteristics of the disease in women is unclear. This difference might be due at least in part to biological disparities, including a predominance of *IGHV*-mutated forms among women with CLL, and also to variations in pharmacokinetics between men and women, as demonstrated in lymphoma patients.

Comorbidity

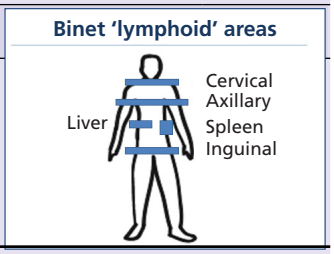
The median age of patients at diagnosis is over 70 years. As a result, the presence of concurrent diseases is frequent. Although a specific index to evaluate comorbidity and frailty in patients with this form of leukaemia does not exist, a modified Cumulative Illness Rate Score (CIRS) of greater than 6 has been used by the German CLL Study Group to define patients considered unfit for fludarabine-based therapy. Particularly important is an impaired renal function (creatinine clearance <70 mL/min), which is associated with poor tolerance to chemoimmunotherapy and higher toxicity.

Clinical stages

Although developed more than 30 years ago, the clinical stages independently formulated by Rai and Binet continue to be the backbone of estimating prognosis and indicating therapy. Clinical stages are based on the concept that CLL cells first accumulate in blood and bone marrow, subsequently in lymph nodes and spleen, and eventually in bone marrow, leading to its functional failure (Table 27.6). Rai and Binet clinical staging systems give reliable information on survival probability. Patients with low-risk disease (Rai 0, Binet A) have a median survival

Table 27.6 The clinical stage so of Rai and Binet.

Rai	**0**	**I**	**II**	**III**	**IV**
Blood lymphocytosis	Yes	Yes	Yes	Yes	Yes
Lymphadenopathy	No	Yes	Yes/no	Yes/no	Yes/no
Spleen/liver enlargement	No	No	Yes	Yes/no	Yes/no
Hb <110 g/L	No	No	No	Yes	Yes/no
Platelets $<100\times10^9$/L	No	No	No	No	Yes
Median survival (years)	**>15**	**7–9**	**6–8**	**3–5**	**3–5**
Binet	**A**	**B**	**C**		
Blood lymphocytosis	Yes	Yes	Yes		
< 3 lymphoid areas* enlarged	No	Yes	Yes/no		
Hb <110g/L and/or Platelets $<100\times10^9$ /L	No	No	Yes		
Median survival (years)	**>15**	**7–9**	**3–5**		



*cervical, axillary, inguinal lymph nodes (either uni- or bilateral), spleen, and liver.

greater than 15 years, those with intermediate-risk disease (Rai I, II; Binet B) have a median survival of 5–7 years, and patients with high-risk disease (Rai III, IV; Binet C) have a median life expectancy of less than 3–5 years. Importantly, assigning a clinical stage to a given patient only requires a physical examination and a complete blood cell count; such simplicity is a tremendous advantage, enabling them to be applied globally.

Other validated prognostic factors

Blood lymphocyte doubling time

A rapid blood lymphocyte doubling time (LDT) (e.g. <12 months) indicates progressive disease. LDT can be easily anticipated by plotting in linear regression absolute lymphocyte counts obtained at different time points over 2–3 months since diagnosis, or since the first time in which lymphocytosis was observed, but not further studied. LDT behaves as a continuous prognostic variable, the shorter it is, the poorer the outcome. Not infrequently, a short LDT is associated with other signs of progressive disease. In addition, if a patient has a short LDT prior to initial therapy, an equally rapid or even more rapid doubling time will be observed at disease relapse, indicating that the disease is inherently proliferative.

Cytogenetics

Most patients with CLL (~90%) present with cytogenetic abnormalities that can be detected by FISH. The most common abnormalities are del(13q), trisomy 12, del(11q), del(6q) and del(17p). All these abnormalities have important clinical correlates (Table 27.3). Döhner *et al*., in a seminal publication, proposed a hierarchical model for prognostic purposes. This model implied that patients are classified according to the worst cytogenetic abnormality, as shown in Table 27.2, independently of the number of lesions. It should be noted that deletions of 13q14, present in up to 50% of cases, confer a good prognosis, provided they are found as the only change. Del(11q)(22-23) frequently involves the *ATM* gene and its prognostic significance mainly depends on whether or not the gene in the other allele is mutated or not. Del(17p) is found in 5–30% of cases, its incidence being higher in patients with advanced and heavily pretreated CLL. In the case of complex karyotype (more than three abnormalities) the lesions usually include *TP53* mutations. Del (11q) and particularly *TP53* mutations are a strong predictor of response to therapy (see Response predictors).

Genetic mutations

IGHV gene mutational status: In landmark studies independently conducted by Damle *et al*. and Hamblin *et al*. in 1999, it was demonstrated that in CLL, the *IGHV* gene can be either mutated (<98% homology to the nearest germline sequence) or unmutated in a proportion of 50–60% and 30–50%, respectively, and that *IGHV* mutational status correlates with biological and clinical features. Importantly, *IGHV* mutational status has prognostic value in itself, independently of other parameters. While patients with mutated *IGHV* genes usually have indolent disease not needing therapy and good prognosis, those with unmutated *IGHV* genes tend to have rapidly aggressive disease, respond poorly to therapy and have short survival. Notably, many of the poor prognostic features, such as advanced clinical stage, male gender, atypical morphology and unmutated *IGHV*

Table 27.7 *IGHV* gene mutational status: biological and clinical correlates.

Mutated *IGHV*	**Unmutated *IGHV***
• Early stage (A, O) • WBC count $<30\times10^9$/L • LDT >12 months • Low ZAP70 • Low CD38 • Low B2M • Low-risk genetics – Normal, de(13q) (isolated) – *MYD88* mutations	• Advanced stage (C, III,IV) • WBC count $\geq 30\times10^9$/L • LDT ≤ 12 months • High ZAP70 • High CD38 • High B2M • High-risk genetics – Del(17p) – Complex karyotype – *TP53, NOTCH1, SF3B1, BIRC3* mutations
Median survival >15 years	**Median survival: 4–5 years**

genes are found together. There are some exceptions, notably *IGHV3-21*, which is more often mutated and is associated with poor outcome (comparable to unmutated cases) and has predominantly light chain expression. Trisomy 12 and de(11q), del(17p), *TP53, NOTCH1* and *SF3B1* mutations predominate in unmutated cases, whereas the opposite is true for the mutated ones. *IGHV* mutational status is thus not only a major marker of two different forms of disease, but also a central prognostic feature around which revolve many other prognostic factors (Table 27.7).

Other genetic mutations: The del(17p) at the *TP53* locus is observed in 5–50% of cases, depending on the clinical situation (i.e. early, untreated disease versus late, heavily treated and resistant disease) and in most cases correlates with a point mutation in the other allele, resulting in total inactivation of the *TP53* suppressor gene. In 2–5% of patients, however, *TP53* mutation can be present in the absence of del(17p), meaning that the absence of del(17p) does not rule out mutations in the *TP53* gene. Prognostically, *TP53* mutations are the most important predictor of response. Importantly, next-generation sequencing studies have shown that *TP53* mutations can be observed in small CLL subclones, bearing the same poor prognosis (see Response predictors). In turn, del(11q) deletions are usually monoallelic, frequently large and include a minimally deleted region (MDR) which encompasses the *ATM* gene in 30–40% of cases. *BIRC3*, a negative regulator of non-canonical NFKB signalling located at 11q22, has been found to be deleted or mutated in a mutually exclusive manner with *TP53*.

The European Research Initiative on CLL (ERIC) undertook a study to determine the correlations and clinical impact of genetic mutations in CLL. To that purpose 3490 patients with CLL from different institutions were screened. *BIRC3* mutations (2.5%) were associated with unmutated *IGHV* genes, del(11q) and trisomy 12, whereas *MYD88* mutations (2.2%) were exclusively found among mutated CLL. In multivariate analysis, *SF3B1* mutations and *TP53* along with del(11q) and unmutated CLL, but not *NOTCH1* mutations, showed independent significance as predictors of time-to-first treatment.

Biologic prognostic markers

CD38: A low CD38 expression correlates with stable disease, while a higher expression predicts a short time to disease progression and worse overall survival. The best threshold for CD38 positivity is unclear. Early reports suggested 30% as a cut-off, but more recent studies found lower figures (7%, 20%) to be more clinically relevant. Of note, CD38 expression may vary over time, thus undermining its prognostic value.

ZAP70: Several studies have shown a good correlation between ZAP70 expression and *IGHV* mutational status, both adding discriminating power to clinical stages. This is particularly evident in patients with stage A disease, in whom early assessment of the need for future therapy may be desirable. Although ZAP70 expression may vary over time, changes are rarely observed in cases with clear-cut low or high values.

CD49d: The prognostic significance of CD49d has been outlined in many studies, including a large analysis in which CD49 was found to be the most reliable immmunophenotypic marker regarding prognosis and independently of other markers such as *IGHV* mutational status, but this finding requires further validation.

Serum markers: β-2 microglobulin (β2M) has been extensively studied and validated from the prognostic standpoint. β2M is considered to reflect tumour burden and disease kinetics, but can also be increased as a result of an abnormal renal function. In many, but not all, studies it has been found to predict poor response to therapy and short survival. Likewise, serum levels of CD23 (and its doubling time) as well as those of thymidine kinase, both reflecting cellular turnover, have been linked to prognosis. Increased LDH levels are infrequent in uncomplicated CLL and also correlate with outcome. However, extremely high LDH levels, particularly if sharply increasing, should lead to the suspicion of disease transformation (Richter's

syndrome). Other less well-studied markers include vascular-growth-endothelial-factor (VGEF), soluble interstitial adhesion molecule 1 (SICAM), CD44, CD20, and CD27.

microRNAs: microRNAs (MIRs) are a large family of short, non-coding, single-stranded molecules that regulate almost one-third of human genes. Two microRNAs, *MIR15*a and *MIR16-1*, are located at 13q14, a region frequently deleted in CLL and which is an initial event in CLL pathogenesis. Prognostically, these microRNAs distinguish CLL cases according to *IGHV* mutation status and ZAP70 expression, and have been found to correlate with poor outcome. Besides, many other MIRs have been associated with aggressiveness and outcome of CLL, but no MIR has been developed as a routine prognostic tool.

Prognostic systems and scores

Prognosis is related to many factors, and thus a single parameter rarely predicts outcome. Because of this, prognostic assessment is based on the combination of different parameters in the form of stages, prognostic groups or scores. The best, and oldest, examples are Binet and Rai staging systems in which lymphadenopathy, organomegalies, anaemia and thrombocytopenia are combined to identify different risk groups (Table 27.6). Other systems to evaluate prognosis in either the whole patient population or selected groups have been proposed. For example, the German CLL Study Group (GCLLSG) has identified a series of prognostic parameters such as age (>60 years), sex (male), ECOG (>0) *IGHV* mutational status (unmutated), serum β2M (>3.5 mg/L), serum TK (>10.0 U/L), del(17p) and del(11q), whose combination allows delineation of four prognostic groups with survival ranging from 95% to 19% at 5 years from diagnosis.

Among elderly patients (>70 years old), the Barcelona group has shown that advanced stage (Binet B or C), increased serum β2M microglobulin (>2.3 mg/L), high ZAP 70 (or unmutated *IGHV* gene) and a CIRS score of >6 identify favourable (less than two factors; median survival: 11.4 years; OS at 10 years: 60%) and unfavourable (at least two factors; median survival 6.8 years; OS at 10 years: 10%) risk groups.

In daily practice, patients with early clinical stage (Binet A, Rai 0), mutated *IGHV* genes, no poor cytogenetics, normal serum β2M, a low blood lymphocyte count ($<30 \times 10^9$/L) and a long LDT (>12 months) have a very good prognosis, with a survival close to that of the general population. In contrast, subjects with advanced clinical stage (Binet B/C, Rai I–IV), unmutated *IGHV* genes, poor-cytogenetics (11q–, 17p–, complex karyotype), increased β2M, a high blood lymphocyte count ($\geq 30 \times 10^9$/L) and short LDT (<12 months) have a very poor prognosis.

Disease complications

Autoimmunity

AIHA occurs in around 10% and ITP in 5–7% of patients, and can appear before or concurrently with the diagnosis, spontaneously over the course of the disease or be triggered by treatment. The fear that purine analogues could trigger AIHA has been over-rated, largely due to biased data from retrospective studies in patients that were heavily pretreated. Actually, AIHA can be observed with any kind of therapy, including alkylators. Recent studies have shown that the incidence of AIHA in patients treated with fludarabine-based regimens such as fludarabine, cyclophosphamide and rituximab is lower than with fludarabine alone. However, in one of such studies (UK CLL4 trial) all fatal cases of AIHA followed fludarabine monotherapy. Therefore, patients who experience AIHA secondary to fludarabine alone should only be re-exposed to fludarabine with great caution, and a fludarabine-based combination regimen rather than fludarabine alone should be given.

Most patients with CLL and AIHA have anaemia with positive IgG DAT in the context of reticulocytosis, raised bilirubin and low haptoglobin serum levels; serum LDH is less discriminating, since it may be elevated due to active CLL. In some cases, conventional DAT techniques can be negative, mainly in cases with milder AIHA forms with a low amount of red blood cell (RBC)-bound immunoglobulin. Likewise, reticulocytosis may not be striking in the context of a bone marrow heavily infiltrated by leukaemic cells or in patients receiving treatment. There is the risk, therefore, of underdiagnosing AIHA.

ITP is most commonly an incidental finding on a routine blood count. Diagnosing ITP may present difficulties, particularly because there is not a sensitive and specific diagnostic test. Nevertheless, thrombocytopenia can be considered as immune mediated when platelet counts suddenly decrease in the absence of splenomegaly, infection or chemotherapy, and with abundant megakaryocytes in the bone marrow. Staining with Factor VIII may help to reveal the presence of megakaryocytes in a massively infiltrated marrow. In advanced disease, anaemia usually occurs before thrombocytopenia. Therefore, isolated thrombocytopenia, particularly if extremely low, is more likely to be immune in origin. Response to corticosteroids may be the final, *post-hoc* diagnostic test. On some occasions AIHA and ITP are found together (i.e. Evans' syndrome). In those areas in which this is a prevalent problem, *Helicobacter pylori* infection should be excluded.

Erythroblastopenia or pure red cell aplasia (PRCA) is defined by the lack, or maturation arrest, of red blood cell precursors in bone marrow, anaemia and low absolute reticulocyte count. In the presence of anaemia, the reticulocyte percentage can be misleadingly normal; therefore, the reticulocyte percentage corrected according to the haematocrit and the absolute reticulocyte count are more informative. The bone marrow shows characteristic defects of erythroblast maturation. In the presence of a bone marrow heavily infiltrated by lymphocytes, the identification of red cell precursors can be difficult. In this setting, antiglycophorin immunohistochemistry may facilitate the identification of red cell precursors. Importantly, PRCA can be seen in association with other immune cytopenias, particularly AIHA.

In addition, any patient with CLL and anaemia with a low reticulocyte count should be evaluated for viral infections that can be associated with PRCA, namely *cytomegalovirus, Epstein–Barr virus*, and *parvovirus*.

Autoimmune granulocytopenia is exceedingly rare (<1%). Likely, most early reports of this complication did not occur in the context of *bona fide* CLL cases, but of misdiagnosed T-cell large granular lymphocytosis.

The prognostic significance of autoimmune cytopenias is controversial, but there is some proof that when patients with advanced clinical stage (Binet C, Rai III, IV) are stratified based on the mechanism of the cytopenia, those with advanced stage 'immune' do better than those with advanced stage 'infiltrative', in which the cytopenia is due to a packed bone marrow. The reason for this difference, of important consequences in prognosis and clinical management, is that patients with autoimmune cytopenia tend to respond well to corticosteroids, while those with a high tumour burden do not always respond satisfactorily to CLL therapy.

Treatment of patients with CLL and autoimmune cytopenia is largely based on expert opinion and analysis of retrospective series rather than on prospective clinical studies. It is accepted that in those patients with immune cytopenia in the context of quiescent CLL, the treatment should be the same as in idiopathic AIHA: initially with corticosteroids, and then, in patients who fail to respond or relapse quickly, anti-CD20 monoclonal antibodies. If there is no response to these two approaches, alternative immunosuppression (e.g. ciclosporine, mycophenylate or azathioprine) or splenectomy should be considered. Good results have also been achieved with rituximab plus corticosteroids, rituximab, cyclophosphamide and dexamethasone or rituximab, dexamethasone and ciclosporine. Intravenous immunoglobulin can be useful where a rapid response is needed (e.g. a patient with thrombocytopenia and significant bleeding), though as a single agent it will not give lasting effects. The thrombopoietin receptor agonist eltrombopag may be effective in ITP associated with CLL, as in primary ITP. Splenectomy can be useful in individual cases impossible to control by other measures, but is less and less indicated. Supportive care should include blood product transfusion as clinically indicated, folic acid in AIHA and local measures to control bleeding in ITP. Importantly, failure of autoimmune cytopenia to respond to conventional treatment is an indication for anti-CLL therapy, particularly if the disease shows signs of progression.

Non-haematological autoimmunity is very rare, the only disorders that have been reliably linked to CLL being cold-aglutinin disease, paraneoplastic pemphigus and acquired angioedema.

Disease transformation

CLL can transform into B-cell diffuse large cell lymphoma (DLCL), an event historically known as Richter's syndrome (RS). RS occurs in around 5% of patients and the risk is estimated to be around 10% at 10 years from diagnosis. Risk factors predicting for the development of RS are high CD38 expression, unmutated and stereotyped IGHV genes, and presence of *NOTCH1*, *TP53* and *CDKN2A/B* mutations; whether these are predictors or hallmarks of the disease is debatable. Current evidence suggests that intense immunosuppression with agents such as fludarabine and alemtuzumab may trigger transformation. However, in these cases, RS is not clonally related to the underlying CLL and is actually EBV-driven, analogous to post-transplant lymphoproliferative disorder (Chapter 35).

RS is extremely heterogeneous in its biology and prognosis. In 80% of cases, the transformation occurs in the CLL clone, which usually acquires mutations of *TP53*, *NOTCH1* and *CDKN2A/B*, *MYC* abnormalities and/or clonal evolution. The prognosis of these patients is poor, with a median survival inferior to 2–3 years. In the remaining 20% of patients, the DLCL arises in a CLL-unrelated B-cell clone, and these patients may have a better prognosis.

In practice, disease transformation should be suspected whenever the patient suffers an abrupt worsening of the general status, fever, enlarging lymph nodes or extranodal involvement; also, a sharply increasing serum LDH level is an important clue to suspect transformation. However, the diagnosis may be difficult because RS can be localized in an isolated organ (e.g. spleen, lymph nodes of a given territory, extralymphatic tissue). Rarely, the transformation takes place in the bone marrow. In such cases the large blast-like cells spill over to the peripheral blood and the morphology may resemble that of an acute leukaemia (Figure 27.6). RS is often associated with monoclonal proteins in the serum or free light chains in the urine. The diagnosis can be reliably suspected on clinical grounds, but needs to be confirmed by biopsy. In this regard, and in contrast to uncomplicated CLL, PET/CT can show areas of hyperactivity corresponding to the transformed tissue and be useful in identifying the site to be biopsied. In some cases, EBV infection can be demonstrated in the involved tissue.

Besides DLCL, cases of transformation into Hodgkin lymphoma can be observed, mainly in older men presenting with mixed-cellularity histological subtype, and EBV driven. In such cases the prognosis is controversial, but seems to be as poor as in DLCL transformation. Patients respond badly to conventional lymphoma regimens and allogeneic stem cell transplantation should be considered in eligible patients.

Second neoplasias

Around 10–15% of patients with CLL present with other cancers; the relative risk has been calculated to be two times higher than in the general population. Melanoma, lung carcinoma, lymphoma, Kaposi sarcoma, CNS and gastrointestinal tumouros are the cancer types most frequently observed. Skin cancers other than melanoma can be also observed, but at a frequency no different from that of the general population. Patients with CLL have a substantially increased risk for a rare skin tumour known

as Merkel cell carcinoma, and vice versa. It is usually considered that there is no relationship between treatment and the incidence and type of secondary solid tumours. The possibility of a second tumour should be entertained whenever a given patient presents with unexpected, unusual symptoms. Finally, secondary MDS/AML may be observed in around 5% of heavily pretreated patients.

Infections

Infections are frequent and a common cause of death. Their pathogenesis is multifactorial, including hypogammaglobulinaemia, immunosuppression and treatment-related myelotoxicity. In some studies it has been found that infections are more related to prior therapy and diminished bone marrow reserve than to hypogammaglobulinemia. With chlorambucil, most infections are bacterial and frequently involve the respiratory tract. The pathogenesis of infections with purine analogues is related to T-cell abnormalities induced by these agents, with herpes virus infections being very frequent. Infections by *Pneumocystis, Listeria, Mycobacteria, Aspergillus* and *Candida* can be also observed. The use of alemtuzumab is frequently complicated (10–25% of patients) by CMV reactivation, which deserves close monitoring and pre-emptive treatment. Likewise, re-activation of HBV infection may occur under therapy with anti-CD20 monoclonal antibodies. Importantly, anti-CD20 monoclonal antibodies (e.g. rituximab, obinutuzumab, ofatumumab) carry a risk of hepatitis B reactivation among patients positive for hepatitis B surface antigen (HBsAg) or antibodies against hepatitis B core antigen (anti-HBc). All patients should be screened for hepatitis B prior to starting treatment. Patients with evidence of prior hepatitis B infection should be monitored for clinical and laboratory signs of reactivation during therapy and for several months after its completion. Anti-CD20 monoclonal antibodies should be discontinued in patients with hepatitis re-activation.

Treatment

Criteria to start therapy

The diagnosis of CLL is not equivalent to the necessity of starting therapy, which is only indicated when any of the following features is present:

- Fever without evidence of infection, extreme fatigue, night sweats or weight loss
- Increasing anaemia or thrombocytopenia due to bone marrow infiltration by leukaemic cells
- Bulky or progressive lymphadenopathy or splenomegaly
- Autoimmune cytopenias that do not respond to standard treatment
- Rapid blood lymphocyte doubling time (e.g. <6 months); baseline lymphocyte count should be $>30 \times 10^9$/L.

Hypogammaglobulinaemia or increased WBC counts, in the absence of any of the above criteria, do not justify treatment. Likewise, younger age, poor-prognosis biomarkers such as *IGHV* unmutated status, high ZAP 70, *TP53* or other mutations are not per se criteria to initiate treatment. Since the diagnosis of CLL rarely constitutes a medical emergency, the need for therapy and its modality should be ideally decided after 1–3 months of observation, a period of time useful to assess the pace of the disease and comprehend the patient's needs and desires.

In practice, most patients in Binet stage A or Rai stage O disease present with no symptoms or signs of active disease and should not be treated, whereas the majority (but not necessarily all) of patients in intermediate or advanced stage (Binet B, C; Rai I to IV) present active disease and need therapy. The notion that only patients with active disease should be treated is supported by trials and meta-analysis showing that in patients with asymptomatic low-risk disease, early treatment with chlorambucil was, if anything, deleterious. Thus, the old wisdom of watch and wait should continue to prevail for patients with asymptomatic disease.

Pretreatment evaluation

This should include a complete medical history, physical examination to detect enlarged lymphoid areas, complete blood cells counts, DAT, renal function, general biochemistry, and immunoglobulin levels. Imaging studies should be performed if retroperitoneal, lung or mediastinal involvement is suspected, but they are not part of the routine assessment. A bone marrow biopsy can provide important information about tumour burden and origin of cytopenias, and is also useful as baseline information. A normal creatinine clearance is required if chemoimmunotherapy is to be given, and hepatitis B serology should be performed in patients in whom treatment will contain an anti-CD20 monoclonal antibody.

Predictive factors

Response to therapy is the most important survival predictor in patients requiring therapy. As previously discussed (see Prognosis), predictive or response factors anticipate the probability of response to a specific therapy. Importantly, genetic abnormalities, which are the most important predictive factor, can change over time in approximately 30–40% of patients; because of this, their study should be performed before initiating therapy, and should be repeated before changing treatment.

Unfortunately there are very few robust predictive factors for CLL. The most important is del(17p)/*TP53* mutations that anticipate a low probability of achieving a good and sustained response with fludarabine and other purine-analogue based therapies. Patients with del(17p)/TP53 mutations constitute a treatment challenge. Alemtuzumab, which is no longer licensed for CLL therapy, but can still be obtained through a named programme, is transiently useful in del(17p)/*TP53* mutated cases without a high tumour burden. There is also some evidence that lenalidomide can be effective across all genetic

groups. The better results, however are obtained with ibrutinib, idelasilib and BCL2 antagonists that are effective across all genetic subgroups, although the results are somewhat inferior in del(17p)/*TP53*-mutated cases. For patients with del(11q) there is proof of shorter response duration with fludarabine alone, while this effect is circumvented by adding cyclophosphamide (FC) to the treatment and particularly the combination of fludarabine, cyclophosphamide and rituximab (FCR). The relationship between *NOTCH1*, *SF3B1* and *BIRC3* mutations, and response to different treatment modalities should be prospectively investigated. In the UK LRF CLL4, CLL harbouring *NOTCH1* and *SF3B1* mutations had a significantly shorter OS survival than patients without these mutations, but still better than that of patients carrying *TP53* abnormalities. Results from the GCLLSG CLL8 trial indicate that both *SF3B1* and *NOTCH1* mutations are independent predictors of short PFS after treatment with FCR. Also, it appears that patients carrying *NOTCH1* mutations do not benefit from adding anti-CD20 monoclonal antibodies (i.e. rituximab, ofatumumab) to fludarabine and cyclophosphamide, but this observation needs further investigation. Besides abnormalities detected by FISH analysis, a complex karyotype (more than three abnormalities) is being increasingly recognized as a predictor of poor response and outcome. It is important therefore not to forget to perform conventional cytogenetic analysis prior to therapy.

Minimal residual disease

There is increasing evidence that achieving minimal residual disease (MRD) negativity is clinically important. Patients in CR with no detectable MRD have a longer PFS and OS than those with persistent MRD. However, MRD negativity is an arbitrary concept that depends on the technique used to assess MRD and could be a surrogate of a more benign, sensitive to treatment, disease. Under the leadership of the Leeds group, ERIC has undertaken a number of studies aimed at determining the best combination of monoclonal antibodies to detect MRD at a 0.1% cell level, finding that the sensitivity of flow cytometry could be refined by using it in combination with high-throughput sequencing. An important caveat is that increasing or prolonging treatment to reach MRD-negative status may convey unnecessary risks, such as myelotoxicity and infections. In the GCLLSG CLL8 trial, patients who achieved MRD-negative CR had the same favourable outcome regardless of the given treatment (FC versus FCR). In line with this, the MD Anderson group has shown that the outcome of patients achieving MRD-negative CR upon treatment with FCR is the same, regardless of the number of cycles of therapy received. Currently, MRD eradication as a treatment endpoint is only justified within clinical trials.

Treatment approaches

Whenever possible, patients requiring treatment should be included in clinical trials. For patients not included in trials, plausible treatments are shown in Table 27.8 and are discussed below. Treatment possibilities discussed in this section are mainly based on agents approved by the FDA and the EMA.

Younger and fit patients

The standard intervention is chemoimmunotherapy, namely the combination of fludarabine, cyclophosphamide and rituximab (FCR) (Tables 27.8 and 27.9). The response rate with FCR is over 90%, including 30–40% of patients achieving MRD-negative CR. In a long-term follow-up of 300 patients initially treated with FCR, the 6-year OS was 77%, the median time to progression 80 months and the treatment-related deaths <1%. In fact, some patients remain in remission over 10 years after having received

Table 27.8 Main treatment options for patients with CLL not included in trials.

Clinical situation	Younger, fit patients	Older, unfit patients
Front-line treatment	FCR BR (patients 60–70 years)	CLB CLB + anti-CD20
Late relapse(>2–3 years)	Switch to BR if prior FCR	Switch to CLB + anti-CD20 if prior CLB alone BR (patients 60–70 years)
High-risk disease – Del(17p)/TP53 mutations – Fludarabine-refractoriness – PFS <2–3 years	Allogeneic should be kept as to differentiate this treatment option from autologous HSCT Ibrutinib, Idelasilib Alemtuzumab (+/− corticosteroids)	Ibrutinib, Idelasilib Alemtuzumab (+/− corticosteroids)

BR, bendamustine+rituximab; CLB, chlorambucil; FCR, fludarabine, cyclophosphamide, rituximab; HSCT, allogeneic stem cell transplantation.

Table 27.9 Treatments for younger, fit patients.

Treatment	Phase	No. of patients	ORR (%)	CR (%)	PFS (median), months
Fludarabine →rituximab vs.	I	51	77	28	42
FR		53	90	47	42
FCR	II	300	85	62	80
Bendamustine + rituximab (patients 60–70 years old)	II	171	88	23	33
PCR	II	64	91	41	33
FC vs.	III	371	88	27	33
FCR		390	95	52	43
FCR vs.	III	284	98	48	92% at 2 years
BR		280	98	38	96% at 2 years

B, bendamustine; C, cyclophosphamide; F, fludarabine; P, pentostatin; R, rituximab.

FCR and with no detectable disease, most presenting mutated *IGHV*.

The downside of fludarabine or other purine analogue-based chemoimmunotherapy is the myelotoxicity and immunosuppressive effects. Frequent toxic events include fever and infections, which are particularly frequent in the first two years following treatment. Therefore, care should be taken to use prophylactic measures, such as cotrimoxazole to prevent *Pneumocystisi* pneumonia and acyclovir to prevent herpes virus re-activations. There are other complications, such as HBV re-activation and, rarely, progressive multifocal leucoencephalopathy (PML), which are of concern. The first dose of rituximab is often associated with infusional symptoms (fever, chills, hypotension, dyspnea, nausea, vomiting), which are manageable with acetaminophen or corticosteroid therapy. Subsequent infusions of rituximab are usually uneventful. Late, usually transient and asymptomatic neutropenia can be observed in patients treated with rituximab. Importantly, FCR (or similar regimens) can be only safely given to younger and fit patients, who are a minority within the population of patients with CLL.

Patients between 60 and 70 years of age poorly tolerate FCR. Treatment alternatives include the combination of fludarabine and rituximab (FR), FCR 'lite', pentostatin + cyclophosphamide + rituximab (PCR) or bendamustine and rituximab (BR). The combination most widely recommended is BR. In a randomized trial comparing FCR versus BR, the ORR was identical in both arms (98%), but the CR rate was higher with FCR (47% versus 38%; $p = 0.031$). There was no difference in OS rate for the FCR versus BR arms (94% versus 96% at 2 years $p = 0.593$). Importantly, in patients ≥ 65 years, there were no differences in median PFS between FCR versus BR (not reached versus 46 months; $p = 0.757$). FCR-treated patients had significantly more frequent severe, grade 3 to 5, adverse events during the whole observation period (91% versus 79%; $p < 0.001$).

Elderly and unfit patients

Although defining who is an old person is difficult and subject to permanent revision, 65–70 years of age is the cut-off used to separate 'younger' from 'older' patients. Of note, CLL is a disease that mainly affects subjects older than 70 years (>50% of patients). Importantly, 'biological' rather than 'chronological' age should be taken into consideration when advising therapy. However, only a small proportion of subjects older than 65–70 years can safely receive intensive cytotoxic therapy because of frailty or comorbidity. Although a specific system for evaluation of fitness in patients with CLL does not exist, a Cumulative Index Rating Scale (CIRS) score >6 is used in some studies to identify patients for whom intensive treatments are inappropriate.

Historically, chlorambucil has been the drug given to patients who cannot safely receive more intensive treatments. However, chlorambucil yields a small proportion of CRs (5–10%), and although it improves symptoms, survival is not substantially improved. The dose usually given ranges from 40 to 70 mg/m^2 per month and is given intermittently (i.e. 10 mg/m^2 per day for 7 days or 0.5 mg per kg every two weeks, in monthly courses); the duration of the treatment depends on the degree of response, but usually ranges from 6 to 12 months.

The addition of rituximab to chlorambucil results in better responses and longer PFS than with chlorambucil alone. However, it is unclear whether survival is improved. Addition of a novel anti-CD20 monoclonal antibodies (obinutuzumab or ofatumumab) to chlorambucil results in an even greater improvement in response rates and PFS in elderly patients, and initial results suggest that OS might be improved with obinutuzumab (Tables 27.9 and 27.10).

In a Phase III trial, 447 patients with previously untreated CLL not eligible for fludarabine-based therapy were randomly assigned to ofatumumab + chlorambucil versus chlorambucil alone, each for up to 12 cycles. The addition of ofatumumab

Table 27.10 Treatments for older, unfit patients.

Treatment	Phase	No. of patients	ORR (%)	CR (%)	PFS (median), months
CLB + rituximab	II	100	82	9	24
Bendamustine + rituximab	II	117	88	23	33
Ibrutinib	IB/II	31	71	13	96% at 2 months
Idelasilib + rituximab	IB/II	64	97	19	93% at 2 months
CLB vs.	III	221	69	1	13
CLB + ofatumumab		226	82	12	22
CLB vs.	III	118	31	0	11
CLB + rituximab vs.		330	65	7	16
CLB + obinutuzumab		333	77	22	27

CLB, chlorambucil.

resulted in a higher rates of OR (82% versus 69%) and CR (12% versus 1%) rates, longer median duration of response (22 versus 13 months), and longer median PFS (22 versus 13 months). Based on this study, ofatumumab was approved for use in combination with chlorambucil for the treatment of patients with previously untreated CLL who are not candidates for fludarabine-based treatment.

In another trial, 781 elderly patients with previously untreated CLL were randomized to receive single-agent chlorambucil versus chlorambucil+rituximab versus chlorambucil + obinutuzumab. The addition of rituximab to chlorambucil resulted in superior OR (67% versus 30%) and CR (8% versus 0%), although with a higher incidence of severe (grade 3–4) neutropenia (25% versus 15%), but similar rates of infection (11% versus 14%). At a median follow-up of 19 months, rituximab improved progression-free survival (median 16 versus 11 months), but not OS. The combinations of rituximab + chlorambucil (or bendamustine) are accepted treatments for untreated patients with CLL who are unfit for receiving fludarabine-based regimens. Obinutuzumab and ofatumumab, each combined with chlorambucil, have not been compared directly in randomized trials. A choice between these agents should take into consideration availability, cost, patient comorbidity and clinician experience (Table 27.10).

The immunomodulatory agent lenalidomide has proved to be effective, particularly if combined with rituximab. In an analysis from the MD Anderson group, based on 137 patients with relapsed/refractory CLL treated on three sequential Phase II studies of lenalidomide monotherapy, lenalidomide + rituximab (LR) and lenalidomide + ofatumumab (LO), respectively, the addition of CD20 monoclonal antibody was associated with higher ORR (32% versus 66% versus 70% for lenalidomide monotherapy versus LR versus LO, respectively) and longer time-to-treatment failure (TTF) (median: 17 versus 16 months, respectively) and a median survival of 21, 61 and 47 months, respectively ($p = 0.039$). There was no significant difference in toxicity comparing monotherapy to combination therapy. Patients with complex metaphase karyotypes had inferior TTF (6 versus 15 months) and survival (16 versus 61 months). Lenalidomide is also investigated as an adjunct to chemoimmunotherapy. Adverse events of lenalidomide include bone marrow toxicity resulting in anaemia, neutropenia and thrombocytopenia. A typical reaction to lenalidomide is a 'flare-up' of the lymph nodes (sudden and painful increase of their size along with fever and other general symptoms), although this is well managed by anti-inflamatory agents and corticosteroids. Also combinations of anti-CD20 monoclonal antibodies with lenalidomide drastically reduce this complication. Lenalidomide, however, is not an approved agent for CLL treatment.

Alemtuzumab is an anti-CD52 monoclonal antibody effective in previously untreated CLL. Alemtuzumab, either alone or in combination with high-dose methylprednisolone, has activity in del(17p) cases. Toxicities of alemtuzumab include rigours, chills, fever, immunosuppression and lymphocytopenia. Opportunistic infections can be observed. Cytomegalovirus (CMV) reactivation is a problem that deserves close monitoring and preemptive therapy. Responses to alemtuzumab vary for different disease sites, being higher in peripheral blood and bone marrow than in spleen or lymph nodes. If used before allogeneic stem cell transplantation, a time window of 2–3 months should be kept to reduce the risk of graft failure because of T-cell depletion. Alemtuzumab is no longer licensed for CLL therapy, although it can be obtained through a personalized CLL treatment programme. Furthermore, its role in CLL therapy is decreasing, and eventually will probably disappear, due to the availability of new and more effective agents such as ibrutinib and idelasilib.

Patients with high-risk disease

High-risk disease is defined by refractoriness to fludarabine-based therapy, del(17p)/*TP53* mutations or short time to progression (<2–3 years) after treatment with a fludarabine-based regimen and, particularly, chemoimmunotherapy. The

Table 27.11 BTK-targeting agents for CLL therapy.

Treatment	No. of patients	Risk	ORR (%)	CR (%)	PFS	OS
Ibrutinib (PCYC-1102)	85	R/R	68–88	2	57–93 (2 yr)	70–93 (2 yr)
Ibrutinib (PCYC-1103/1102)	101	R/R	56–81	NA	46–68 (30 m)	66–80 (30 m)
Ibrutinib (RESONATE)	195	R/R	43–66	NA	83–88 (6 m)	90 (12 m)
Ibrutinib + rituximab	40	HR	90–100	10	78 (18 m)	84 (18 m)
Ibrutinib + rituximab + bendamustine	30	R/R	93	17	78 (15 m)	NA
CC-292	66	R/R	25–40	NA	NA	NA

HH, high-risk; m, months; NA, not available; R/R, relaps/refractory; yr, years.

outcome of these patients is extremely poor (median survival <2 years) and their management a challenge. Alemtuzumab (discussed above) is also effective, albeit transiently, in patients with del(17p)/*TP53* mutations. Ofatumubab is an approved treatment for patients refractory to fludarabine and alemtuzumab. The role of BCR-signal inhibitors, BCL2 antagonists and haemopoietic stem cell transplantation in the treatment of high-risk disease is discussed below.

Patients with relapsing disease

Disease relapse is not by itself a criterion to re-initiate treatment, and only patients with relapsed and progressive disease should be treated (see Criteria to start therapy). The type of intervention will depend on several factors, of which the most important are the PFS length and the modality of prior therapy. Not surprisingly, the shorter the PFS, the poorer the outcome. A PFS inferior to 2–3 years indicates poor-risk disease in patients treated with fludarabine-based regimens (see High-risk disease). Conversely, in patients with a longer PFS, the prognosis is not considered poor, and most guidelines consider that such cases can be re-treated with the initial therapy. However, the case can be made as to whether changing the type of therapy, even when the PFS has been long, is not a preferable approach. Furthermore, repeating treatment with FCR is hardly feasible because of undue bone marrow toxicity. A regimen containing an anti-CD20 monoclonal antibody should be considered in those cases not having received it previously. Accepted treatments for relapsed/progressing patients include RB, ofatumumab, ofatumumab + chlorambucil, idelasilib + rituximab and ibrutinib (Table 27.10). There is not enough experience to know the most appropriate treatment for patients progressing after treatment with ibrutinib, idelasilib or BCL2 antagonists (see further on).

BCR-signal inhibitors and BCL2 antagonists

A major breakthrough in the treatment of CLL has been the arrival of agents specifically targeting the B-cell antigen receptor (BCR inhibitors, BCRi) or antagonizing the antiapoptotic protein BCL2 (BCL2 antagonists, BCL2a). BCR signalling is a key component of B-cell and CLL-cell proliferation and survival. This is mediated through a complex series of tyrosine kinases, which modulate signals from the CLL cell surface to the nucleus. Two specific targets within this pathway are particularly relevant – the delta isoform of phosphatidylinositol-3-kinase (PI3Kδ), which is inhibited by idelalisib (formerly known as CAL-101 or GS-1101), and Bruton's tyrosine kinase (Btk), which is irreversibly inhibited by ibrutinib (Figure 16.13; Chapter 16).

The ORRs for ibrutinib and idelalisib in relapsed/refractory patients are respectively 48–71% and 39% if used as monotherapy. If 'partial remissions with persistent lymphocytosis' are considered, the response rate to ibrutinib is over 90%. The ORR is further increased by combining these agents with chemotherapy and/or anti-CD20 antibodies (Tables 27.10, 27.11 and 27.12). A feature of inhibiting BCR signalling is that there is a redistribution of CLL from the tissue to the blood in the early weeks of therapy. This is manifest by a very rapid reduction in nodal disease and increase in peripheral blood lymphocytosis, which can take several months to resolve. Despite this lymphocytosis, most patients feel very well in themselves. Both idelalisib given with rituximab and ibrutinib monotherapy are orally available.

BCRi have shown to be the most effective agents to treat CLL with del(17p)/*TP53* mutations. In a randomized trial involving 220 patients with relapsed CLL, including 44% with del(17p)/*TP53* mutations, comparing idelasilib + rituximab to placebo + rituximab, the ORR rates (81% versus 13%) and PFS at

Table 27.12 PI3Kδ-targeting agents for CLL therapy.

Treatment	No. of patients	Risk group	ORR (%)	CR (%)	PFS (%)	OS (%)
Idelasilib	54	R/R	24–72	0	5–41	75 (36 m)
Idelasilib + rituximab (ofatumumab)	40	R/R	73–83 NA	8	19–20 (20 m)	NA
Idelasilib + rituximab/BF o C	45	R/R	70–83	2	NA	NA
Idelasilib (pooled analysis)	168	R/R	67–83	5	NA	NA
IPI-145	52	R/R	47–50 NA	5	NA	NA

R/R, relapse/refractory; m, months; NA, not available.

6 months (93% versus 46%) strongly favored the idelalisib + rituximab arm. In addition, investigators from the Ohio State University reported a retrospective study on 174 patients with CLL harbouring del(17p). After a median follow-up of 12 months, 1-year PFS and OS in the 27 patients treated with ibrutinib were 77% and 81%, respectively, and thus significantly superior to that of the 58% and 89% of patients treated with cyclin-dependent kinase inhibitors or conventional regimens. Altogether, the effectiveness of BCRi seems to result in a longer disease control of relapsed/refractory CLL, even in the presence of high-risk criteria, than with any other currently available treatment with the possible exception of allogeneic transplantation. Similarly, ibrutinib monotherapy resulted in significantly better ORR, PFS and OS when randomized against ofatumumab in a prospective trial enrolling 391 patients with refractory/relapsed CLL. Importantly, these agents have a relatively favourable toxicity profile. The most severe and potentially fatal adverse effects of ibrutinib are thrombocytopenia, diarrhoea, upper respiratory tract infections, neutropenia, and anaemia. Ibrutinib may cause atrial fibrillation, particularly in patients with high-risk cardiovascular factors. In addition, ibrutinib should not be given to patients receiving anticoagulation, because of the risk of bleeding. Mutations of *C481S-BTK* are linked to the mechanisms of refractoriness to and progression upon therapy with ibrutinib. 'Black box' warnings for idelasilib include hepatic failure, severe diarrhea, colitis and pneumonitis, with some fatal cases of these complications having been reported. CC-292 toxic effects are mild or moderate diarrhoea, nausea, hypertension and fatigue.

The BCL2 inhibitor, ABT-263, demonstrated very high response rates in relapsed, refractory CLL, but its tolerability was limited by an inevitable thrombocytopenia due to it inhibiting BCLXL as well as BCL2, since BCLXL is expressed in platelets. The next-generation BH3 mimetic, ABT-199 (or GDC-0199), is much more specific for BCL2, does not inhibit Bcl-XL and therefore does not result in thrombocytopenia. ABT-199 monotherapy has an overall response rate in relapsed, refractory CLL of approximately 85% without the lymphocytosis. An unexpected complication of ABT-199 was tumour lysis syndrome (TLS), which has meant a much more gradual escalation of the initial dosing to make this manageable. In addition, mild to moderate gastrointestinal complications and respiratory infections appear to be frequent. The few cases of higher-grade toxicity are largely due to neutropenia.

Role of haemopoietic cell transplantation

Autologous transplants are not useful in CLL. Allogeneic stem cell transplantation exerts its therapeutic effect through the T-cell antileukaemia-mediated effect (graft-versus-leukaemia) and overcomes the negative impact of poor prognostic factors such as del(17p). In most series, the PFS and OS are around 60% and around 40%, respectively. These results are independent of whether the donor is a family sibling or an unrelated donor. However, the procedure-related mortality, even if using reduced-intensity conditioning regimens in experienced institutions, is around 10% at 1 year and 30% at 3 years from the transplant. Most deaths are due to graft-versus-host-disease and its complications. Disease relapses can be observed, but they are rare beyond 4 years from transplant. Transplant results are better in younger patients, with no comorbidity, transplanted in remission and receiving a transplant from a fully matched HLA donor.

The effectiveness of BCRi and BCL2a makes it necessary to reconsider the place of allogeneic stem cell transplantation in the treatment algorithm of patients with high-risk CLL. BCRi have been shown to be highly effective in patients with high-risk disease. In a position paper from ERIC and the European Bone Marrow Transplantation Group (EBMT) it is recommended that all patients with high-risk disease are initially treated with BCRi or BCL2a. If no response is obtained, allogeneic transplantation should be considered. If a response is obtained, there are two possibilities: either to continue with the drug with which the response has been achieved or to perform a transplant. Treatment recommendations should be made on an individual

basis. Factors potentially favouring immediate transplantation are: (1) coincidence of relapsed/refractory high-risk CLL with del(17p)/*TP53* mutations and/or del(11q); (2) younger age, no significant comorbidity (comorbidity), (3) bone marrow dysplastic features and cytogenetic abnormalities consistent with secondary MDS/AML; (4) availability of a well-matched donor. Conversely, factors supporting HSCT deferral are: (1) absence of a relapsed/refractory situation, (2) relapse/refractory disease in the absence of del(17p)/*TP53* mutations and del(11q); (3) age >70 years, significant comorbidity or (4) only a partially matched or mismatched donor available. Treating physicians need to keep in mind that allotransplantation should not be considered as a last, desperate attempt in refractory and heavily pretreated patients, when the possibilities of a successful transplant are very few.

CAR-T cells

Chimeric antigen receptors (CARs) usually combine the antigen binding site of a monoclonal antibody with the signal activating machinery of a T-cell, freeing antigen recognition from major HLA complex restriction and thus breaking one of the barriers to more widespread application of cellular therapy (Chapter 16). T cells expressing CARs are highly targeted, move to tumour sites, undergo *in vivo* expansion and persist for long periods of time. The most frequent antigen target for CAR-T cells is CD19, although ROR-1, which is specific for CLL, is an appealing target, which is being actively investigated, along with other alternatives (e.g. double antigen targeting). In a non-systematic review of the literature, 44 patients treated with ibrutinib were identified. The overall response rates were: CR 18% and PR 25%. CAR-T cells can trigger cytokine-release syndrome (hypotension, rigors, muscle aching, fever, anorexia). To prevent cytokine-release syndrome, chemotherapy is given before the infusion of CAR-T cells. Other important toxic events are profound and long-lasting hypogammaglobulinaemia and B-cell depletion. Further studies refining CAR-T cell therapy for CLL are needed. Also, clinical expertise and a broad range of supporting measures are needed to embark on this form of therapy.

Criteria for response

The following response categories are recognized:

- Complete response (CR)
  - CR MRD-negative
  - CR MRD-positive
- Complete response incomplete (CRi)
- Partial response (PR)
- Nodular partial response (PRn)
- Partial response with lymphocytosis (PRL)
- Stable disease (SD)
- Progressive disease (PD).

The achievement of CR requires normalization of blood counts and normal bone marrow as assessed by biopsy. Since cleaning up bone marrow from the disease can take time, the bone marrow biopsy should be performed 3 months after completing therapy. In clinical trials, it is recommended to further assess CR by measuring MRD. Thus there are two types of CR: MRD-negative and MRD-positive, the outcome of patients with the latter not being as good as in those with MRD-negative CR. Immunohistochemistry should always be performed if residual nodules are present in bone marrow biopsy. CT should be repeated only if it was abnormal at the start of therapy. Incomplete CR denotes cases with persistent anaemia, thrombocytopenia or neutropenia, due to myelotoxicity, who otherwise fulfil the criteria for CR. If the bone marrow is hypocellular, it is recommended that the biopsy be repeated once the blood counts have recovered. Not infrequently, the degree of response improves over time.

Nodular PR defines cases that fulfil CR criteria, but with discrete or moderately large nodules in bone marrow. In this context, it is important to carry out immunohistochemistry, as in some patients, the nodules may contain mainly T cells rather than B cells. PR requires the regression of at least 50% organomegalies and abnormal blood cell counts. Treatment with BCRi results in a transient blood lymphocytosis that should not be confused with treatment failure and disease progression. Because of this treatment-related effect, a response category named PR with lymphocytosis is accepted in patients treated with BCRi. SD refers to no changes in clinical status, and PD to progression in signs and symptoms with respect to those present at treatment initiation.

Other B-cell chronic disorders

B-cell prolymphocytic leukaemia

B-cell prolymphocytic leukaemia (B-PLL) is an extremely rare disorder the individuality, and even existence, of which is being reconsidered. In fact, most cases of B-PLL correspond to MCL in leukaemic phase. The diagnosis is mainly based on morphology and cell markers (Table 27.2). Studies addressing differences in gene expression profiling between B-PLL and CLL have yield discordant results. The mean age of patients is 70 years and the chief presenting features are splenomegaly without lymphadenopathy and a high WBC count, usually >100 $\times$ 10^9/L. Anaemia and thrombocytopenia are seen in at least 50% of cases. Other laboratory findings are no different from high-WBC-count CLL, but the incidence of a monoclonal band appears to be higher than in CLL.

The main diagnostic criterion is the identification of prolymphocytes as the predominant population in blood films. A percentage of prolymphocytes >55% is the arbitrary cut-off to diagnose B-PLL. The prolymphocyte is twice the size of a small CLL lymphocyte, has moderately condensed nuclear chromatin, a prominent central nucleolus and a lower nucleus-to-cytoplasm ratio than CLL cells (Figure 27.8).

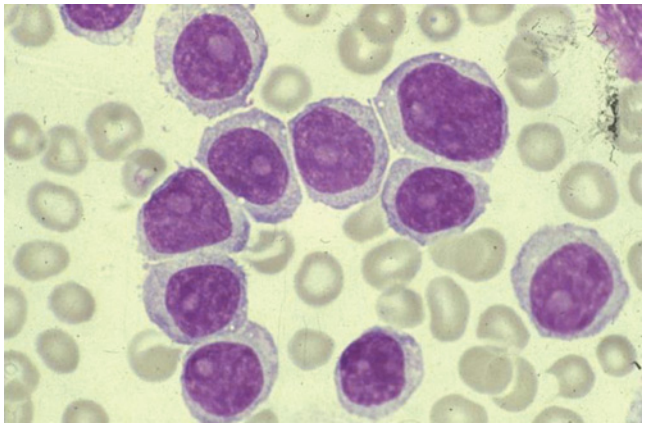


Figure 27.8 Typical blood film from a case of B-prolymphocytic leukaemia with the characteristic large size, abundant cytoplasm and prominent nucleolus.

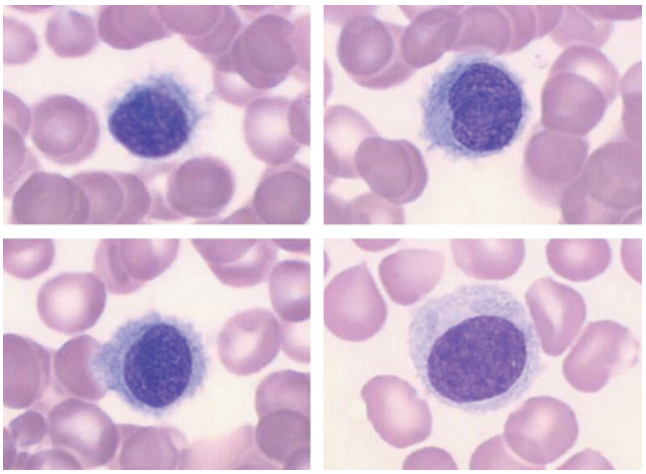


Figure 27.9 Individual hairy cells from blood films from two patients with HCL.

Alkylating agents are of little value. Splenectomy can remove a major proliferative focus and tumour bulk in order to relieve hypersplenism and facilitate further therapy, but is not a practical proposition in elderly and unfit patients. There are insufficient data in B-PLL on the use of fludarabine combinations, including those using monoclonal antibodies, e.g. FCR, FCMR. R-CHOP, BR is a treatment alternative, and in younger patients allogeneic stem cell transplantation should be considered.

Hairy-cell leukaemia

Hairy-cell leukaemia (HCL) accounts for 3% of all leukaemia and is a well-recognized clinicopathological entity that affects males more frequently than females (ratio 4:1), usually over the age of 60 years. The disease features result from the pancytopenia that most patients present. Anaemia is due to the reduced bone marrow production and splenic pooling; haemolysis is exceptional. The WBC count may be low, normal or high (rarely above 20×10^9/L), but neutropenia and monocytopenia are constant, as is an increased MCV. Patients with low WBC count and no circulating hairy cells present diagnostic problems, a bone marrow biopsy being necessary for diagnosis. Platelet counts are below 100×10^9/L in most cases. HCL should be considered in the differential diagnosis of any single patient with pancytopenia.

The main physical sign is splenomegaly. Lymphadenopathy is rare at diagnosis, but can be observed in advanced HCL. Routine CT may show a higher incidence than hitherto suspected of large abdominal nodes. These are more common in relapsed patients and those with long-standing HCL and tend to correlate with bulky disease at presentation.

The recognition of typical hairy cells in peripheral blood films is useful for suggesting the diagnosis. However, since the number of hairy cells in blood is usually low, a high degree of suspicion is required to make the diagnosis, and ultimately a bone marrow biopsy is required. Hairy cells are large, twice the size of a normal lymphocyte, and have abundant cytoplasm (low nucleus-to-cytoplasm ratio) that is characteristically villous in its outline (Figure 27.9). The nucleus is round, oval or slightly indented, and occasionally bilobed. A smooth nuclear chromatin, absence of a visible nucleolus and low nucleus-to-cytoplasm ratio are hallmarks of typical hairy cells. Cells from the rare HCL variant have similar cytoplasmic features, but have a round nucleus with more condensed chromatin and a distinct nucleolus (Figure 27.10). A typical HCL immunophenotype includes pan-B-cell markers (CD19, CD20, CD22) with coexpreesion of CD11c, CD25, CD103, CD123 and CD200. *BRAF* mutations are a typical feature (Table 27.2).

In a characteristic manner, bone marrow aspirates are unsuccessful, as no fragments and few cells are obtained. Therefore, a trephine biopsy is essential. It shows diffuse interstitial infiltration of variable degree; occasionally the infiltration is focal.

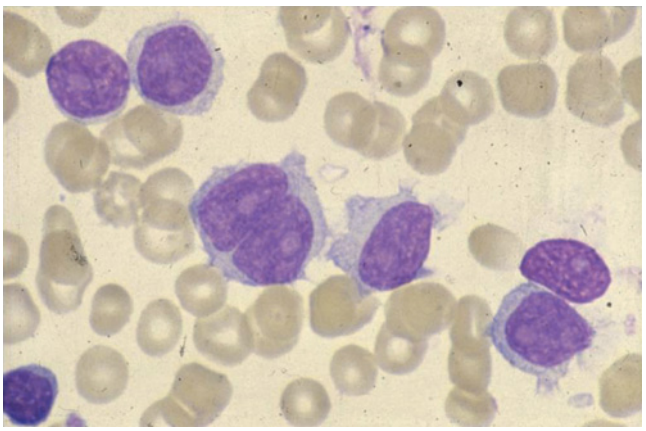


Figure 27.10 Blood film from a case of HCL variant showing large nucleolated villous cells, one of them binucleated.

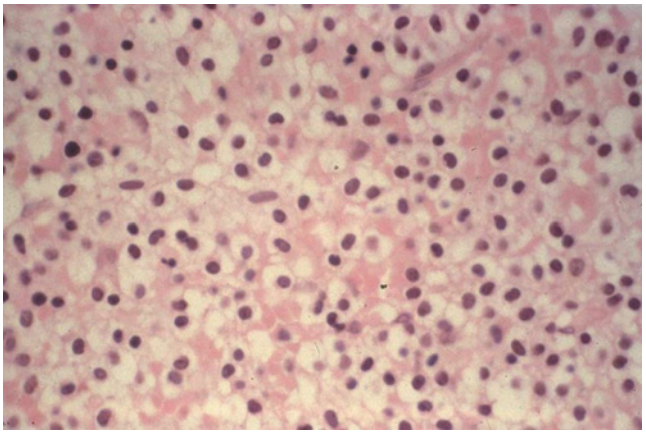


Figure 27.11 Bone marrow trephine section of a case of HCL showing the typical clear zone around the hairy cells in a paraffin-embedded section.

A typical feature is the arrangement of the cellular infiltrate, which is loose, leaving plenty of space between cells, often with a clear zone around each cell, which is unique to this condition (Figures 27.11 and 27.12). Reticulin is always increased and the cells stain strongly with anti-CD20. Hypocellular bone marrows can be observed, which along with the pancytopenia may raise the possibility of aplastic anaemia as a diagnosis. In cases with hypoplastic marrows, immunohistochemistry (e.g. CD20, annexin1, DBA44, CD11c) is important for demonstrating clusters of hairy cells, as it is in assessing response after therapy. Residual marrow disease is virtually observed in all cases, including those seemingly in CR.

Spleen histology shows distinct diagnostic features: infiltration by mononuclear cells with a blunt nucleus in the red pulp with little residual white pulp, and formation of pseudo-sinuses filled with erythrocytes. Tartrate-resistant acid phosphatase (TRAP) demonstrated on hairy cells corresponds to a unique isoenzyme 5 and is specific for HCL. This enzyme is now tested by means of a monoclonal antibody, which can be used in bone marrow biopsies.

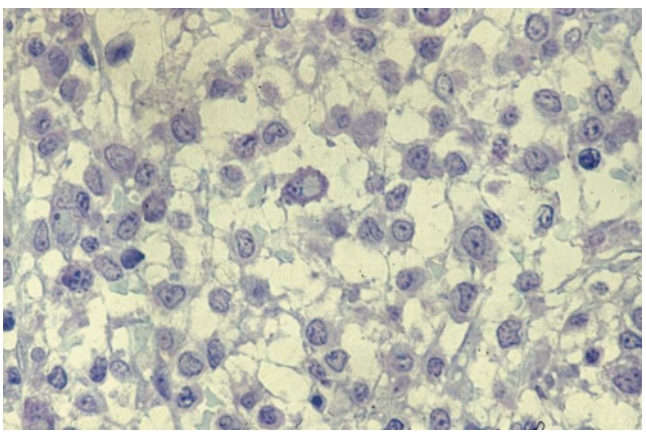


Figure 27.12 Bone marrow trephine from a case of HCL: section embedded in methacrylate giving good morphological detail of the hairy cells.

The HCL variant (HCLv) is very rare and is characterized by a high WBC count (50–80 $\times$ 10^9/L) and splenomegaly. The cells are irregular, large nucleolated villous cells (Figure 27.10). Immunophenotypically, these cells do not express CD25, CD123 or CD200 and the MCV is normal. The main diagnostic difficulty in HCL variant is with SMZL, except for spleen histology, which in the former mainly involves the red pulp and in the latter the white pulp. HCLv is better regarded as a form of splenic small cell lymphoma/leukaemia, which along with splenic diffuse red cell lymphoma, is pending further characterization.

Treatment is indicated in presence of active disease. The main objectives of treatment are the normalization of blood counts (pancytopenia is the main source of complications) and to induce prolonged CR. The two agents currently used in HCL, pentostatin and cladribine, achieve CR in 80–85% of patients. Long-term results are excellent, with very few patients actually dying of HCL. Splenectomy and interferon are treatment approaches whose interest is mainly historical, although they can be useful in selected cases in which cytotoxic therapy is not feasible.

Monoclonal antibodies against CD20 (rituximab) and CD22 (coupled with an immunotoxin) are active in HCL. Hairy cells strongly express CD20 and CD22, and these antibodies seem to achieve good responses in the group of relapsing patients. Its main role may be, as in other B-cell disorders, when used in combination with purine analogues. Elderly patients may also benefit from antibodies like rituximab as they do not cause myelotoxicity. However, since almost all cases have an activating *BRAF* mutation, a *BRAF* inhibitor (i.e. vemurafenib) has been shown to be effective.

Patients with the controversial diagnosis of HCL variant respond less well to cladribine and pentostatin and not at all to interferon-α. Palliation of symptoms and improvement in blood counts can be achieved with splenectomy, which has been the most successful modality in this relatively resistant disease.

Supportive care for HCL is confined to patients undergoing therapy with nucleoside analogues, which can cause transient neutropenia in the early phases and prolonged lymphocytopenia later on. Long-term cotrimoxazole until lymphocyte counts rise above 1 $\times$ 10^9/L and also long periods on acyclovir are recommended. Major infections are only seen in treated patients or those responding poorly to therapy.

The leukaemic phase of indolent NHL

There are three types of NHL which not infrequently evolve with lymphocytosis that mimics, and can be confused with,

CLL: follicular lymphoma, SMZL (formerly described as splenic lymphoma with circulating villous lymphocytes) and MCL. In most cases the clinical picture is dominated by an enlarged spleen, hence the term 'primary splenic lymphomas' is used by some authors to refer collectively to these entities (see Chapters 31, 33).

Follicular lymphoma

Around 15% of patients with follicular lymphoma may have circulating lymphoma cells (5–20 × 10^9/L), and this correlates with bone marrow involvement. The proportion of patients with inconspicuous blood involvement is much higher, as demonstrated by flow cytometry or molecular techniques (BCL-2 detection). A minority of patients may present with a WBC count in excess of 40 × 10^9/L (up to 200 × 10^9/L) and extensive disease, generalized lymphadenopathy, hepatomegaly and a prominent splenomegaly. Usually the circulating cells are very small, with almost no visible cytoplasm, the nuclear chromatin is smooth without clumps of heterochromatin and no visible nucleolus, and the nuclear shape is angular and has a small cleft (Figure 27.13). Lymph node histology is essential for confirming a diagnosis of follicular lymphoma. Patients with very high WBC counts may have a poor prognosis; cases with a minor degree of spill-over behave better, as do cases without circulating cells. Late-stage follicular lymphoma may have circulating blasts (centroblasts) characterized by a peripherally located nucleolus, and this corresponds with histological transformation to a large-cell lymphoma or the so-called blastic form of follicular lymphoma. The immunophenotypic characteristics of follicular lymphoma cells are shown in Table 27.2 (see Chapter 33).

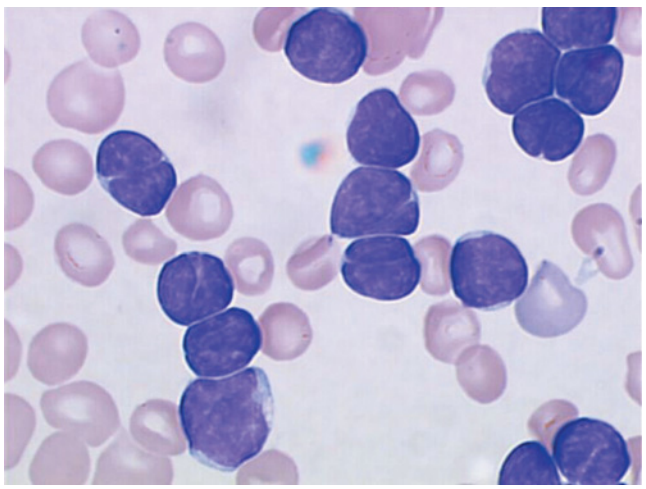


Figure 27.13 Blood film of a patient with follicular lymphoma presenting with significant lymphocytosis. The cells show a cleaved nucleus, angular shape, homogeneous chromatin pattern, high nucleus-to-cytoplasm ratio and small size.

Splenic marginal zone lymphoma

Two-thirds of patients with splenic marginal zone lymphoma (SMZL) (1–2% of all lymphomas) have circulating villous lymphocytes. A minority does not have frankly villous cells despite being clonal, and in some the diagnosis can only be established after splenectomy. The true incidence of this condition without circulating villous cells is unknown. The boundaries of this disorder with the HCL-variant and B-cell splenic lymphoma with red pulp involvement have not been settled.

The median age of patients is around 70 years and the majority present with splenomegaly. Anaemia and/or thrombocytopenia are seen in 40% of cases. One-third have a serum monoclonal band, usually IgM below 30 g/L. Diagnosis is suspected on a typical blood film (Figure 27.14) showing small lymphocytes with an irregular membrane outline and short villi, often confined to one pole of the cells. The immunophenotype (see Table 27.2) is different from CLL. SMZL cells can be distinguished from HCL and the HCL variant (now separated as a form of splenic lymphoma) because they are negative for CD103 and CD123 but are always positive for CD11c.

Bone marrow aspirates may show the same cells as in the blood, with variable degrees of infiltration. A distinct pattern of intrasinusoidal infiltration is characteristic of SMZL, which can be highlighted with anti-B-cell antibodies (CD20, CD79a). This pattern of infiltration is less common in other types of NHL and is not seen in CLL. In addition, the pattern of bone marrow infiltration can be nodular, interstitial and/or paratrabecular.

The spleen histology shows a characteristic bizoned pattern in the white pulp, with a central zone of small lymphocytes with scanty cytoplasm and a peripheral zone of larger cells with more dispersed chromatin and more abundant cytoplasm. The red pulp is always infiltrated by both the smaller and the larger cells.

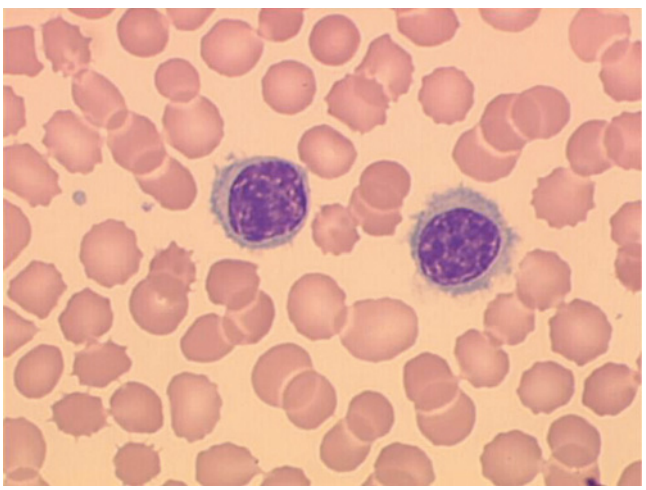


Figure 27.14 Lymphocytes from a case of SMZL. Villous lymphocytes are observed.

Plasmacytoid differentiation may be seen and transformation into large-cell lymphoma occurs in 10% of cases.

There is no unique cytogenetic abnormality in SMZL, but unbalanced translocations and deletions of 7q22–32 have been described in about 30%; 50% of cases have a monoallelic deletion at 13q14. Trisomy 3, an abnormality also seen in extranodal marginal zone lymphoma, has been found in 17% of cases. Abnormalities of *TP53* (mainly deletions) are found in around 15% and are associated with worse prognosis. In lymphoplasmacytic lymphoma *MYD88* (L265P) mutation is found in almost 100% of cases, whereas such mutation is infrequent in SMZL.

An association with hepatitis C has been reported, particularly in the form of mixed cryoglobulinaemia, but its incidence is unknown. Interferon-α has been reported to induce CR in some hepatitis-C-positive SMZL patients. An association with hyperreactive malaria and tropical splenomegaly has been described in African cases.

Splenectomy is the treatment of choice when the spleen is a prominent feature. Splenectomized patients fare significantly better than those treated with chemotherapy only, perhaps reflecting a higher tumour mass in those patients requiring chemotherapy. Rituximab-based regimens have been reported to induce remissions. In most cases, the overall survival is superior to 10–15 years (see Chapter 33).

Mantle cell lymphoma

Blood and bone marrow involvement are common in MCL and can be demonstrated in almost 100% of cases if sensitive enough flow cytometry and immunohistochemical techniques are used. Close to 40% of cases evolving with splenomegaly have significant lymphocytosis, sometimes mimicking atypical CLL (Figure 27.15). The presence of blastic cells and high labelling by Ki-67 are the most important prognostic factors. Marker studies show that MCL cells are always positive for CD5, CD19 and CD20, but negative for CD23 and also CD200 (Table 27.2). The bone marrow biopsy shows a rather monotonous infiltration by slightly irregular cells. In contrast with CLL, no proliferating centres are seen in MCL. The pattern of bone marrow infiltration is often diffuse in advanced cases, but in other instances is nodular and/or paratrabecular. The characteristic translocation t(11;14)(q13;q32) is demonstrated in all cases by conventional cytogenetics or by FISH. The rearrangement of the *CCND1* gene at 11q13 results in the over-expression of cyclin D1, one of the proteins that controls the cell cycle. Cyclin D1 can be demonstrated in histological sections. Presence of SOX11 is also a typical feature of MCL (see Chapter 33). The numbers of proliferating (Ki-67 positive) as well as blastic cells are strong outcome predictors.

Not all cases of MCL require therapy at diagnosis. In some instances, particularly females in which the clinical picture is dominated by splenomegaly, the clinical course can be relatively indolent. These patients can be followed with no therapy for months, and their outcome is superimposable on those treated immediately after diagnosis. Once treatment is required, rituximab-based therapy, including high doses of cytarabine and cyclophosphamide is the most widely used treatment, and stem cell transplantation should be considered in fit patients. No differences in outcome have been observed between autologous and allogeneic transplants. Bortezomib has shown effectiveness in trials. Ibrutinib is highly active and has been approved for MCL treatment, and is likely to change the treatment landscape for MCL.

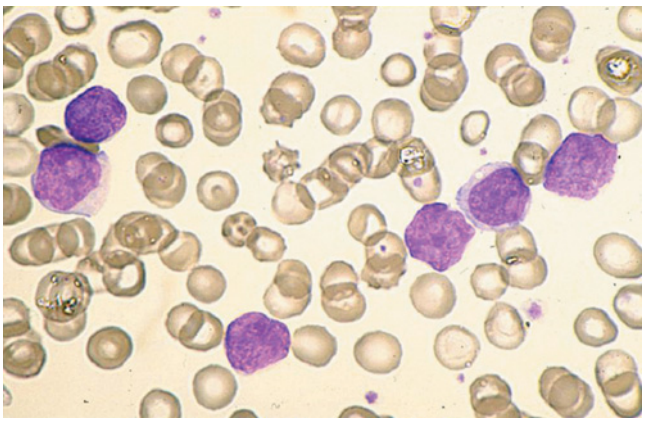


Figure 27.15 Blood film from a patient with mantle-cell lymphoma and lymphocytosis. The cells are medium size, slightly irregular in shape and have a speckled nuclear chromatin pattern.

Acknowledgements

The authors are thankful to Professor Daniel Catovsky who contributed to this chapter in past editions. The authors wish to apologize to those colleagues whose contributions are referred to in this chapter, but could not be quoted because of the limitation in the number of references.

Selected bibliography

Chronic lymphocytic leukaemia

Byrd JC. (2015) Introduction to a series of reviews on chronic lymphocytic leukemia. *Blood* **126**: 427–85.

Gruber M, Wu CJ (2014) Evolving understanding of the CLL genome. *Seminars in Hematology* **51**: 177–87.

Hallek M, Cheson B, Catovsky D *et al.* (2008) Guidelines for the diagnosis and treatment of chronic lymphocytic leukemia: a report from the International Workshop on Chronic Lymphocytic Leukemia updating the National Cancer Institute-Working Group 1996 guidelines. *Blood* **111**: 5446.

Hodgson K, Ferrer G, Pereira A *et al.* (2011) Autoimmune cytopenia in chronic lymphocytic leukaemia: diagnosis and treatment. *British Journal of Haematology* **154**: 14–22.

Rossi D, Gaidano G. (2013) Richter syndrome. *Advances in Experimental and Medical Biology* **792**: 173–91.

Zang S, Kipps TJ (2014) The pathogenesis of chronic lymphocytic leukemia. *Annual Review of Pathogenic Mechanisms of Disease* **9**: 103–18.

Hairy-cell leukaemia

Grever MR, Blachly JS, Andritsos LA (2014) Hairy-cell-leukemia: update on molecular pathology and therapeutic advances. *Blood Reviews* **28**: 197–203.

Topp ZZ, Saven A (2014) Hairy-cell leukemia: a 'hair-raising' update. *Expert Review of Hematology* **7** (5): 659–69.

B-cell prolymphocytic leukaemia

Dearden C (2012) B- and T-cell prolymphocytic leukemia: antibody approaches. *Hematology American Society of Hematology Education Program* **2012**: 645–51.

The leukaemic phase of NHL

Gordon LI, Bernstein JH, Jares P *et al.* (2014) Recent advances in mantle-cell lymphoma: report of the 2013 Mantle-cell lymphoma consortium workshop. *Leukemia and Lymphoma* **55**(10): 2262–70.

Sakata S, Tsuyama N, Tekenchi K (2014) Pathology of indolent B-cell neoplasias other than follicular lymphoma. *Journal of Clinical Experimental Hematopathology* **54**(1): 11–22.

Sohani AR, Zukerberg LR (2014) Small B-cell lymphoma of the spleen: how to tell them apart. *Journal of Hematopathology* **7**: 109–121.

CHAPTER 28

T-cell lymphoproliferative disorders

28

Pier Luigi Zinzani and Alessandro Broccoli

Institute of Haematology "L. e A. Seràgnoli", University of Bologna; Bologna, Italy

Introduction

Chronic T-cell lymphoproliferative disorders are a group of neoplastic diseases that result from the clonal proliferation of mature post-thymic T-lymphocytes, being therefore clearly separated from diseases arising from T-cell (prethymic or thymus-derived) precursors, such as T-cell acute lymphoblastic leukaemia and lymphoblastic lymphoma.

The World Health Organisation (WHO) classification provides clear definitions for the diagnosis of the major subtypes of T-cell lymphoproliferative disorders, which are classified on the basis of their clinical behaviour (indolent versus aggressive) and presentation (nodal, extranodal or leukaemic). Mycosis fungoides and primary cutaneous $CD30^+$ T-cell lymphoproliferative disorders represent extranodal (cutaneous) diseases with a prolonged natural history; conversely, anaplastic large-cell lymphoma, angioimmunoblastic T-cell lymphoma and peripheral T-cell non-Hodgkin lymphoma, not otherwise specified, to cite the most relevant, represent the aggressive nodal counterpart. Enteropathy-associated T-cell lymphoma and hepatosplenic T-cell lymphoma have an extranodal presentation and usually behave aggressively, as well as some rarer cutaneous T-cell lymphoma entities, such as primary cutaneous aggressive epidermotropic $CD8^+$ T-cell lymphoma and cutaneous γ/δ T-cell lymphoma. T-prolymphocytic leukaemia and large granular lymphocyte leukaemia are indeed diseases with a primary leukaemic presentation, as they arise in the bone marrow and then evolve with a peripheral blood leukaemic presentation. Finally, Sézary syndrome may be considered as the leukaemic phase of a T-cell cutaneous lymphoma, since the tumour arises in an extranodal organ – i.e. the skin – but also displays a characteristic leukaemic picture.

This chapter focuses on the most significant clinical entities, by grouping together chronic T-cell leukaemias, peripheral T-cell non-Hodgkin lymphomas and cutaneous T-cell non-Hodgkin lymphomas, and provides the description of disease features and morphology, the basis for their diagnosis and the therapeutic approaches, with a closer look at new molecules and innovative treatment strategies.

Chronic T-cell leukaemias

Two distinct chronic T-cell lymphoproliferative disorders present with a primary leukaemic phase: T-prolymphocytic leukaemia (T-PLL) and large granular lymphocyte (LGL) leukaemia.

T-prolymphocytic leukaemia

T-PLL is a rare, post-thymic, mature T-cell neoplasm, accounting for 2% of mature lymphoid leukaemias. It generally displays an aggressive clinical course, with only a minority of patients being asymptomatic at presentation. Disease progression may occur rapidly, even in asymptomatic patients. Older adults are generally affected (median age of 61 years), with a male predominance.

Clinical features

Splenomegaly, seen in almost two-thirds of patients, and marked peripheral blood lymphocytosis, with circulating prolymphocytes, are distinctive clinical aspects. Hepatomegaly and lymphadenopathies may be seen in up to half of patients, as well as skin nodules, maculopapular rash or erythroderma. Peripheral

Postgraduate Haematology, Seventh Edition. Edited by A Victor Hoffbrand, Douglas R Higgs, David M Keeling and Atul B Mehta.

oedema (periorbital or conjunctival), pleuroperitoneal effusions or central nervous system involvement may also be present, especially during disease progression. Anaemia and thrombocytopenia are usually documented, along with a lactate-dehydrogenase elevation.

Diagnosis

It is established on both morphological and immunophenotyping tests (see also Chapter 19). Prolymphocytes are medium-sized cells, displaying a high nuclear-to-cytoplasm ratio, with round, oval, irregular or convoluted nuclei, and a single prominent nucleolus; the cytoplasm is agranular, with blebs (Figure 28.1). In up to 20% of cases, prolymphocytes may resemble small lymphocytes or show a cerebriform nuclear appearance. Bone marrow is constantly infiltrated, with a diffuse and interstitial pattern, as well as the spleen, which shows an expansion of both red and white pulp. Tissue histology, however, is not mandatory for diagnosis. Immunophenotyping by flow cytometry demonstrates a positivity for CD7 (strong), CD5 and CD2, with lack of expression of TdT and CD1a. CD52 is expressed at high density. T-cell receptor β/γ genes are rearranged in all cases.

Therapy

Since the disease is, most of the time, symptomatic, clinically aggressive and rapidly progressive, a prompt initiation of a systemic treatment is required. Intravenous alemtuzumab, a humanized anti-CD52 antibody, is the first-line treatment of choice, yielding up to an 80% of complete response in treatment-naïve patients. Although prolonged responses to alemtuzumab have been documented, all patients do eventually relapse: therefore, a consolidation with an allogeneic stem cell transplantation in first remission may offer a chance for a better long-term survival. Autologous stem-cell transplantation may also be beneficial for patients, but it does not result in a cure. Re-treatment with alemtuzumab at disease relapse is feasible if remission duration is longer than 6 months. Purine analogues, such as pentostatin, nelarabine or fludarabine may be a good treatment alternative, with some evidence of activity in T-PLL.

Prognosis

Response to alemtuzumab is the main outcome predictor: non-responders have a median survival of only 4 months. Five-year overall survival is nearly 21%, with high white cell counts, older age and short lymphocyte doubling time being adverse prognostic factors.

Large granular lymphocyte leukaemia

LGL leukaemia is characterized by the clonal expansion of large granular lymphocytes, both of T (T-LGL) or natural killer (NK-LGL) origin. The 2008 WHO classification includes T-LGL in the subgroup of mature peripheral T-cell and NK-cell neoplasms, along with a provisional entity termed chronic lymphoproliferative disorder of NK cells, and a clinically distinct disease, which is the aggressive NK-cell leukaemia.

The frequency of the disease has never been accurately established; however, it ranges from 2 to 6% of the chronic lymphoproliferative disorders. Men and women are equally affected, at a median age of 60 (range 12–87) years, with less than 25% of patients being younger than 50 years.

The aetiology remains unclear, although a prolonged response to an antigenic (possibly viral) stimulus may be hypothesized. A dysregulated apoptosis has been documented as the most relevant pathogenetic mechanism, with leukaemic LGLs being able to escape a Fas-FasL-mediated cell death.

Clinical features

The disease displays a chronic and indolent course, although at least two-thirds of patients are symptomatic at diagnosis. Clinically relevant aspects are similar for both T- and NK-LGL, and are represented by peripheral cytopenias, namely neutropenia and anaemia; splenomegaly (in up to 50% of cases), without hepatomegaly or lymphadenopathy; rheumatoid arthritis in 11 to 36% of patients, with a positive rheumatoid factor and the detection of antinuclear antibodies (along with polyclonal

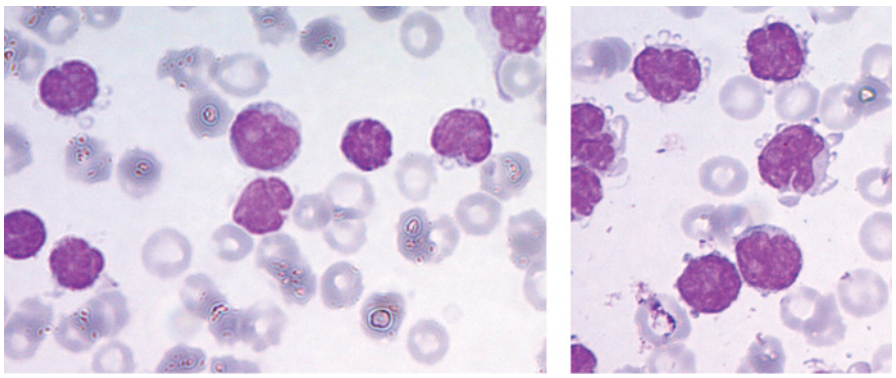

(a) (b)

Figure 28.1 Peripheral blood smear showing T-prolymphocytes. (a) Nuclei are round to markedly irregular (indented, cerebriform), the chromatin is clumped, with inconspicuous nucleoli. (b) Blebs and cytoplasm protrusions are easily seen.

hypergammaglobulinemia) in 40 to 60% of patients (less frequently observed in NK-LGL) and recurrent infections associated with neutropenia, which account for a mucocutaneous or respiratory involvement. Anaemia may be severe, due to multiple mechanisms, including autoimmune haemolysis and pure red cell aplasia, and may require transfusion support. Lymphocytosis ($>4 \times 10^9$/L) is seen in more than half of patients, with mean circulating LGLs varying from 0.5 to 10×10^9/L. Bone marrow involvement is documented in more than 70% of cases, without any linear correlation between the extent of marrow infiltration and the severity of symptoms or cytopenias.

Diagnosis

Diagnosis is established once an increased number of clonal LGLs is documented in a compatible clinical context, both morphologically and by immunophenotyping. LGLs are larger than quiescent lymphocytes, and show an abundant cytoplasm with azurophilic granules and a reniform or round nucleus. A peripheral count of $>2 \times 10^9$/L LGLs is a diagnostic criterion; in case of lower LGL counts, a bone marrow biopsy (showing interstitial infiltration with linear arrays of CD8, granzyme B, perforin, TIA-1 positivity) is recommended to confirm the diagnosis. T-LGLs are CD3+, CD8+, CD57+, $CD5^{dim}$, CD45RA+, and show evidence of clonal T-cell receptor γ gene rearrangement (which makes the distinction from reactive LGL proliferation). On the other hand, NK-LGLs are CD3–, CD2+, CD4–, CD8+, CD16+, CD56+ and CD57±.

Therapy

Therapy should be instituted in case of severe neutropenia (neutrophils <500/μL), recurrent infections (regardless of the absolute neutrophil count), symptomatic or transfusion-dependent anaemia or associated immune diseases (mainly rheumatoid arthritis). No standard treatment exists; however, immunosuppressive treatments represent the mainstay of therapy. Single-agent methotrexate (10 mg/m^2/week, orally) is adequate for patients with neutropenia and arthritis, to be administered for at least 4 months before response assessment. For those with anaemia or pure red cell aplasia, oral cyclophosphamide at a daily dose of 50–100 mg for at least 4 months (up to 12 months in responding patients) is the treatment of choice, comparing favourably with methotrexate in terms of response rates (66 and 55%, respectively). Ciclosporine A is an alternative first-line therapy in patients with anaemia and a suitable second-line, to be maintained indefinitely, as long as it is tolerated.

Prognosis

LGL leukaemia generally behaves indolently; it is difficult to evaluate the amount of patients who will require a systemic therapy due to the worsening of cytopenias or the appearance of symptoms (it varies from 33 to 80% in reported series). Aggressive NK disease, with high LGL peripheral counts and marrow infiltrate, massive organomegaly and lymphoma-related B symptoms displays a poor prognosis, and patients appear refractory to any treatment.

Peripheral T-cell non-Hodgkin lymphomas

Peripheral T-cell non-Hodgkin lymphomas (PTCL) are a distinct form of rare chronic T-cell lymphoproliferative disease, clearly separate from cutaneous T-cell lymphomas (CTCL), discussed below. The incidence of PTCL in the United States is less than 1 case per 100,000 people; worldwide, the prevalence of each specific disease entity profoundly depends on the geographic region, although the biologic reason for the observed discrepancies is still largely unknown. In Western countries, in fact, PTCL accounts for 15 to 20% of aggressive lymphomas, and 5 to 10% of all non-Hodgkin lymphomas. In Asia, however, incidence is higher, and 15 to 20% of all lymphomas are classified as PTCL or NK/T-cell lymphoma. As a consequence, PTCL-not otherwise specified (NOS) appears to be the most common subtype in both North America and Europe, whereas NK/T-cell lymphoma and adult T-cell leukaemia/lymphoma (both entities discussed elsewhere) are very common in Asia, the latter being specifically represented in Japan. Anaplastic large-cell lymphoma (ALCL) is more represented in the United States, whereas angioimmunoblastic T-cell lymphoma (AITL) is more frequent in Europe than in Asia or in North America.

In the context of the 2008 WHO classification, PTCL represent a group of mature T-cell aggressive diseases, primarily with a nodal origin (PTCL-NOS, ALCL, AITL being the most represented and more clinically relevant entities), although extranodal subtypes do exist – being far less common – described by their tissue tropism (enteropathy-associated T-cell lymphoma and hepatosplenic T-cell lymphoma are the most relevant).

Expert haematopathologists can apply the 2008 WHO classification to diagnose a PTCL, however with a heterogeneous agreement on diagnosis depending on the specific subtype of disease they are looking at. For example, if diagnostic accuracy is very good (>97% agreement among experts) for anaplastic lymphoma kinase-positive ALCL, agreement is definitely poor (<85%) for the other lymphoma subtypes, being only 75% for the most common subtype, PTCL-NOS. This indicates how better and more specific diagnostic markers and diagnostic criteria are needed to increase the diagnostic accuracy and reproducibility for these entities.

PTCL are staged according to the Ann Arbor staging system, with the same techniques used for B-cell lymphomas. Much more PTCL patients, with the sole exception of ALCL, present with an advanced (III or IV) stage disease rather than patients with diffuse large B-cell lymphoma. Positron emission tomography (PET) scan has proved its usefulness in the staging of

PTCL, although maximum standard uptake values in patients with PTCL seem somewhat lower than in aggressive B-cell lymphomas, and generally less pronounced for extranodal lesion than for nodal localizations of disease. Data regarding the role of post-treatment PET scan are still inconclusive, since negative scans may not be always associated with a really improved clinical outcome.

The International Prognostic Index (IPI) remains the mainstay of the prognostic stratification of patients with PTCL, being inversely related to survival in a manner similar to B-cell lymphomas, although overall survival in each category is lower if compared to diffuse large B-cell lymphoma.

Herein we describe the relevant clinical and morphological aspects of the most represented PTCL entities; the treatment approach is then discussed at the end of the section.

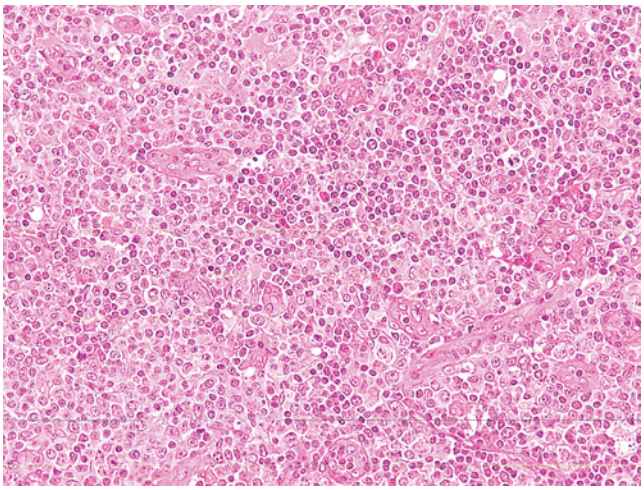


Figure 28.2 Peripheral T-cell lymphoma, not otherwise specified (haematoxylin and eosin stain, magnification 200×).

Peripheral T-cell non-Hodgkin lymphoma, not otherwise specified

PTCL-NOS is the most common subtype, accounting for at least 25% of PTCL. However, it does not represent a single entity, since as many as 20% of cases have shown a gene expression profile characteristic of other PTCL subtypes, such as AITL; moreover, another subgroup shows features of cytotoxic T-cells, thus suggesting a separate disease entity, with a worse prognosis than any other PTCL-NOS. The diagnosis of PTCL-NOS should then be established once other categories of PTCL have been excluded.

Clinical features

The disease tends to affect adult patients (median age at presentation is 60 years), with a male predominance. Children can also be affected. The nodal involvement is the most relevant feature at diagnosis, although any organ can be affected (including bone marrow, liver, spleen and skin), and a combination of nodal + extranodal involvement can be frequently encountered. Advanced stage at presentation is relatively common (nearly 70% of cases); lymphoma-related symptoms, lactate-dehydrogenase elevation and poor performance status are documented in more than half of the cases; almost two-thirds of the patients present an intermediate to high IPI score.

Morphology

Lymph node infiltration is generally diffuse, with neoplastic elements showing a variable morphology, ranging from small cells with irregular nuclei to large cells with prominent nucleoli and mitotic figures (Figure 28.2); cells with clear cytoplasm and occasionally Reed–Sternberg-like cells can also be encountered. Bone marrow is hypercellular and the normal hematopoietic tissue is extensively replaced; cell pleomorphism can be documented, along with reticulin thickening and prominent vascularity. Three rare morphologic variants have been described: the lymphoepithelioid (Lennert's) variant, with the presence of a diffuse or interfollicular infiltrate of epithelioid histiocytes; the T-zone variant, characterized by a preserved lymph node architecture and the presence of interfollicular lymphomatous involvement; the follicular variant, with a pattern of growth intimately related to follicular structures. Neoplastic cells stain positively for CD2, CD3, CD5 and CD7, with a variable expression of CD4 and CD8; aberrant T-cell phenotypes, with lack of expression of some antigens, can also be possible. The CD30 antigen is variably expressed, mostly in large-cell variants.

Prognosis

The IPI is significantly associated with treatment outcomes, which appear better – albeit dismal – for lower scores. Patients with IPI score 0/1, in fact, have a 33% 5-year failure-free survival and a 50% 5-year overall survival, in contrast with those with IPI score 4/5, whose 5-year overall survival is only 11%, with a 5-year progression-free survival of 6%. A newly introduced prognostic index specifically designed for PTCL (PIT) includes age, lactate-dehydrogenase elevation, performance status and bone marrow involvement, and better allows the stratification of patients into more specific prognostic groups, with a 5-year overall survival for the most favourable group (i.e. the one with no adverse factors) of 62%, and of 18% for patients presenting with three or four adverse prognostic factors.

Angioimmunoblastic T-cell lymphoma

AITL makes up nearly 18.5% of cases of PTCL, being the second most common variety. Disease incidence is highest in Europe (29% of all cases), and lower in both North America (16%) and Asia (17.9%), and appears nearly the same in both sexes.

Clinical features

The disease typically affects older adults (median age 65 years), and it is clinically characterized by an aggressive clinical course, an advanced stage at presentation (nearly 90% of patients show stage III/IV) and by the presence of lymphoma-related symptoms (in about 70% of patients since disease onset). Patients complain of a generalized lymphadenopathy, which is usually associated with hepatosplenomegaly and cutaneous rash. Bone marrow is involved in 28% of cases. Lactate-dehydrogenase elevation, polyclonal hypergammaglobulinaemia and autoimmune manifestations (such as autoimmune haemolytic anaemia, cold agglutinin disease, rheumatoid factor and anti-smooth-muscle antibody positivity) are frequently associated findings, as well as the presence of peripheral oedema, pleural or peritoneal effusions and joint pain.

Morphology and biology

From a histological point of view, the architecture of the lymph node appears completely effaced by a T-cell infiltrate, which extends beyond the node capsule, but characteristically sparing the subcapsular sinus, which appears dilated (Figure 28.3). The neoplastic cells are CD3+, CD4+, CD10+, CXCL-13+, PD1+ and sometimes BCL-6+. Reactive lymphocytes, eosinophils, plasma cells, follicular dendritic cells and histiocytes represent accompanying non-neoplastic populations, along with the presence of scattered large CD20+ immunoblastic cells, usually staining positively for EBV-encoded RNA (EBNA). A characteristic feature of AITL is a prominent vascular proliferation, as a consequence of the over-expression of the vascular endothelial growth factor (VEGF)-A gene, both in lymphoma and endothelial cells. It has been shown that a higher expression of this gene, as well as the presence of the VEGF-receptor 1 molecule, correlates with poor survival outcomes, and biologically identifies a disease that is extremely prone to the invasion of nearby tissues.

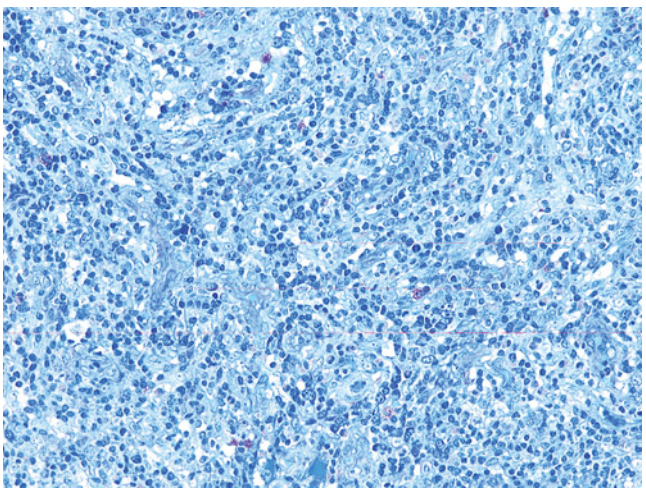


Figure 28.3 Angioimmunoblastic T-cell lymphoma (Giemsa stain, magnification 200×).

Prognosis

Rapid disease progression is seen in most patients, accompanied by frequent infectious complications, yielding to a median survival of less than 3 years.

Anaplastic large-cell lymphoma

First described as a group of tumours displaying large cells with bizarre morphology, prominent sinusoidal invasion and Ki-1 antigen (nowadays recognized as CD30) expression, ALCL has been recognized as a distinct clinical entity since the third edition of the WHO classification, which grouped together a significant proportion of cases characterized by the expression of the anaplastic lymphoma kinase (ALK), and a smaller subcategory, lacking the expression of ALK (ALK-negative ALCL). When it became clear that ALK-positive ALCL represented a relatively homogeneous disease subtype, also displaying a better clinical course in respect of ALK-negative ALCL, the two disease entities were clearly separated: the 2008 WHO classification now recognises ALK-positive ALCL as a distinct clinical and biological entity, whereas ALCL with no ALK expression is regarded as a provisional entity, since it is still controversial whether this entity should be considered as a phenotypic variant of ALCL, or a different disease. Primary cutaneous ALCL (C-ALCL), which is classified within primary cutaneous CD30-positive T-cell lymphoproliferative disorders, is discussed separately.

Clinical features

Nearly 70% of patients with ALCL present with an advanced stage; peripheral or abdominal nodal disease is associated in nearly one-third of patients with an extranodal involvement (especially in patients exhibiting the ALK-positive variant), with bone marrow, skeleton and lung being the most represented sites.

Biology and prognosis

In more than two-thirds of cases, the ALK protein is over-expressed, mainly due to the translocation of the ALK gene on chromosome 2 to the nucleolar phosphoprotein 1 (NPM1) gene on chromosome 5, most frequently as a result of the chromosomal translocation t(2;5)(p23;q35), although at least 11 variant translocations involving the ALK gene at 2p23 have been described. Patients with ALK-positive lymphoma are much younger (median age of 34 years) than those with the ALK-negative variant (median age of 58 years). The expression of the ALK protein makes a substantial prognostic difference in terms of overall survival and failure-free-survival, which both favour patients with ALK-positive AITL in respect of ALK-negative patients (5-year overall survival, 70% versus 49%; 5-year

failure-free survival, 60% versus 36%). However, patients aged over 40 years tend to display a similar outcome in terms of survival, independently from the expression of the ALK protein.

Morphology

ALCL exhibits a wide range of cytological appearances, with no distinction on morphological grounds between the ALK-positive and the ALK-negative variants: five morphologic patterns have been recognized in the 2008 WHO classification (Figure 28.4). The ALCL common pattern is composed predominantly of pleomorphic large cells, with irregular, eccentric, kidney-shaped nuclei; the normal lymph node architecture is frequently obliterated and an intrasinusoidal cell growth is often documented, mimicking a metastatic malignancy. The ALCL lymphohistiocytic pattern is characterized by the presence of tumour cells admixed with a variable amount of histiocytes, whereas the small cell pattern shows a predominance of small- to medium-sized neoplastic cells, with round or irregular nuclei and sometimes with a clear cytoplasm. ALCL Hodgkin-like pattern has morphologic features mimicking nodular sclerosis classic Hodgkin lymphoma, with a significant number of tumour cells resembling Reed–Sternberg cells. Composite patterns may be seen in up to 20% of cases, since features of more than one previously described pattern can be appreciated within a single lymph node specimen.

All ALCL are – by definition – positive for CD30, which is expressed on the cell membrane and in the Golgi region, predominantly in larger tumour cells. The great majority of ALCL express T-cell or NK-cell antigens, although several pan-T-cell antigens may be lost or unexpressed, yielding to an apparently T-null phenotype. The ALK staining may be cytoplasmic, nuclear, nucleolar, or it may be restricted to either the cytoplasm or the cell membrane. ALK expression is virtually specific for ALCL, being absent from all normal postnatal human tissues, except some rare cells in the brain.

Other disease entities

Enteropathy-associated T-cell lymphoma (EATL)

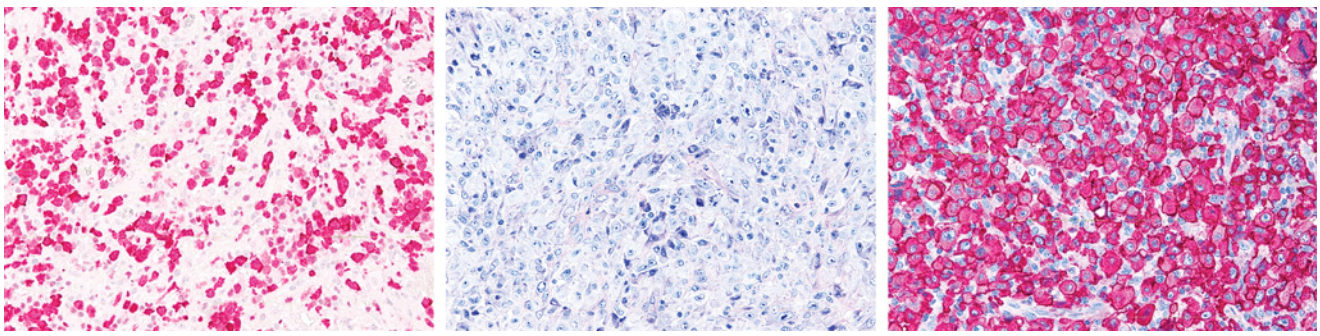
EATL is an aggressive disorder, frequently associated with coeliac disease; its incidence varies geographically, and it overlaps that of gluten-sensitive enteropathy, although not all patients with EATL have an overt coeliac sprue. EATL is frequent in the 6th decade of life, with a similar incidence among men and women. It usually presents with abdominal pain, fatigue and anorexia, as well as with an exacerbation of typical symptoms of coeliac disease, such as diarrhoea and weight loss, together with hypoalbuminaemia and anaemia. Gastrointestinal perforation, obstruction or haemorrhage may be frequent complications, since the tumour infiltration may be transparietal, with frequent ulcerations and induration. The jejunum is mainly affected, although disease extension may be multifocal, also involving other intestinal segments.

Patients with EATL still have a poor prognosis, one of the poorest of any subtype of PTCL, with a median 5-year overall survival of 20% and a progression-free survival of only 4%, which also remains dismal for those patients presenting with a more favourable IPI risk. Many patients, in fact, are unable to tolerate an aggressive chemotherapy approach, which should be comprehensive for an autologous stem cell transplantation.

Hepatosplenic T-cell lymphoma

This subtype accounts for nearly 1.4% of PTCL, being more frequent in young men, who present with systemic symptoms, mainly fever, hepatic and splenic enlargement, along with peripheral cytopenias. It has been reported in the context of immune system disorders or alterations, including Crohn's disease and solid organ transplantation. The tumour spreads to the spleen, the liver and the bone marrow, and infiltrates organs without forming growing and coalescent nodal masses.

The present treatments are largely unsatisfactory, and this disease displays a severe prognosis, with a 5-year overall



(a) (b) (c)

Figure 28.4 Anaplastic large-cell lymphoma (ALCL). (a) ALK-positive ALCL (immunohistochemistry for ALK protein, magnification 200×). (b) ALK-negative ALCL (Giemsa stain, magnification 200×). (c) ALK-negative ALCL (immunohistochemistry for CD30, magnification 200×).

survival of 7% and progression-free survival of 0%. Prolonged survival intervals have been anecdotically reported after induction with alemtuzumab, cladribine or after allogeneic stem cell transplantation.

Treatment of peripheral T-cell non-Hodgkin lymphomas

So far, the treatment strategies most widely applied in patients with PTCL are mainly derived from B-cell lymphomas, with those doing best working well in diffuse large B-cell lymphoma, as the paradigm of an aggressive lymphoma. However, PTCL are biologically and clinically different from the B-cell counterparts, thus explaining the extremely different ranges of response to various therapies, and also the huge prognostic gap that exists between lymphomas of B or T origins.

The treatment approach for PTCL is discussed in this section. A chemotherapy-based first-line treatment is aimed at inducing a first clinical (complete) remission, which may be further consolidated with an autologous stem cell transplantation, depending on the specific PTCL subtype and IPI score at presentation. Novel approaches are under consideration, on the one hand to offer an alternative to those patients proving to be refractory after induction or relapsing after an autologous stem cell transplantation; on the other, to provide innovative treatment strategies to enhance the traditionally applied anthracycline-based induction schemes.

Induction therapy

The cyclophosphamide, doxorubicin, vincristine and prednisone (CHOP) regimen, administered every 21 days, or CHOP-like schemes, still represent the backbone of induction therapy, and are nowadays the most widely used worldwide. No randomized clinical trials are able to prove, however, its superiority over different treatment schemes. ALCL patients are those who best respond to CHOP and CHOP-like regimens, in comparison to other PTCL subtypes, with 5-year overall response rates higher than 60%. AITL patients have a 5-year overall survival of 32% and a failure-free survival of 18% when treated with CHOP, and in patients with PTCL-NOS there is no clear evidence that the use of an anthracycline-containing regimen during induction does any better compared with an anthracycline-free first-line therapy. Of note, nearly 40% of patients with ALK-positive disease and more than 60% of those with an ALK-negative ALCL fail to respond to therapy, and taken as a whole, PTCL display 5-year overall survival rates lower than 40% when treated first-line with CHOP and CHOP-like strategies: therefore, new drug combinations are required to enhance response rates, to prolong survival and to allow higher proportions of patients to be transplanted in first complete remission. There is some evidence that the combination of CHOP plus etoposide (CHOEP) can be more effective than CHOP alone in PTCL-NOS and ALCL; also the addition of alemtuzumab, bortezomib or dinileukin diftitox to CHOP has proven a certain effectiveness in PTCL-NOS patients.

Autologous transplantation

Although the place for bone marrow transplantation in the treatment of PTCL is still uncertain, autologous stem cell transplantation (ASCT) represents an alternative for patients with relapsed disease, as well as a consolidative strategy for patients in first complete remission and with chemosensitive disease, especially for those with intermediate to high IPI scores at disease presentation and with histologies other than ALK-positive ALCL. Unfortunately, many patients tend to progress when they are still on induction therapy, thus ASCT – although feasible and safe – is ineffective in most of cases, if not contraindicated. A dose-dense approach followed by up-front ASCT has demonstrated a 5-year overall and progression-free survival of 51 and 44%, respectively, across a variety of PTCL, with the best results seen in ALCL, even in the ALK-negative subtype. Disease status at transplantation is a major predictor of success.

Allogeneic transplantation

The rationale behind allogeneic transplantation rests on both the fact that allogeneic haematopoietic stem cells are free of tumour contamination and that a graft-versus-lymphoma effect can be hypothesized, since donor-derived immune cells are potentially capable of mediating an antitumour effect. Once again, disease status at transplantation and chemosensitivity are outcome predictors: a high percentage (almost 70%) of patients who are in remission at the time of transplantation can be cured of their disease, while only 25% to 30% of refractory patients may take advantage of this procedure. No differences have been documented in the relapse rates when a myeloablative conditioning is compared to reduced-intensity conditioning regimens, although a higher rate of non-relapse mortality with the myeloablative approach still exists.

New drugs

A variety of new agents have been recently investigated in relapsed or refractory PTCL patients, and some of them are now approved for the treatment of patients in such a disease context. An innovative antifolate drug, pralatrexate, is now approved in the United States for relapsed or refractory PTCL, as a single-agent strategy to re-induce a clinical response. Histone deacetylase inhibitors, such as romidepsin and belinostat, have also shown promising results in poorly responding patients with PTCL, along with other new agents such as dasatinib and aurora kinase inhibitors (alisertib). The prominent vascular proliferation seen in AITL suggests an attractive rationale for the use of antiangiogenic drugs, like lenalidomide or bevacizumab, in patients who do not satisfactorily respond to conventional or high-dose therapies and who may require an allotransplant procedure. Finally, the anti-CD30 antibody conjugate brentuximab vedotin represents an extremely exciting new agent in the

treatment of CD30-positive PTCL at disease reoccurrence: objective response rates higher than 85%, with 57% of complete responses and a median duration of response of 13 months, have been registered in a recent Phase II trial in ALCL patients.

Cutaneous T-cell non-Hodgkin lymphomas

Cutaneous T-cell lymphomas (CTCL) represent the vast majority (more than 75%) of primary cutaneous lymphomas, which are by definition largely confined to the skin at diagnosis, without significant involvement of lymph nodes or other extranodal organs. Mycosis fungoides and Sézary syndrome are the most relevant disease entities, together accounting more than two-thirds of CTCL.

The incidence of CTCL has been increasing to 6.4 cases per million people, and it strongly correlates with age, with a median age at onset in the mid-50s, and a fourfold increase in patients over 70. The highest incidence rates are described in males and African-Americans.

Establishing a correct diagnosis is mandatory to define the best treatment strategy, both at disease onset and in a context of refractory disease. A multidisciplinary approach, with the involvement of clinical haematologists, experienced pathologists and dermatologists, is therefore recommended at every step of patient management. Several new agents have already proved their efficacy within recently conducted clinical trials, but many innovative drugs are now under investigation in order to expand the therapeutic options available for the treatment of these rare entities.

Below is a description of the most frequent CTCL (Table 28.1, highlighted in bold), along with some therapeutic recommendations.

Table 28.1 World Health Organization/European Organization for Research and Treatment of Cancer (WHO/EORTC) classification of primary cutaneous T-cell lymphomas. Highlighted entities are discussed in detail in this chapter.

Mycosis fungoides
Mycosis fungoides, variants and subtypes
Folliculotropic mycosis fungoides
Pagetoid reticulosis
Granulomatous slack skin
Sézary syndrome
Adult T-cell leukaemia/lymphoma
Primary cutaneous CD30+ lymphoproliferative disorders
Primary cutaneous anaplastic large cell lymphoma
Lymphomatoid papulosis
Subcutaneous panniculitis-like T-cell lymphoma
Extranodal NK/T-cell lymphoma, nasal type
Primary cutaneous peripheral T-cell lymphoma, unspecified
Primary cutaneous aggressive epidermotropic $CD8^+$ T-cell lymphoma*
Cutaneous γ/δ T-cell lymphoma*
Primary cutaneous $CD4^+$ small/medium-sized pleomorphic T-cell lymphoma*

*Provisional entity.

Mycosis fungoides

Mycosis fungoides represents the most frequent CTCL, and consists of a mature neoplasm in which T-lymphocytes infiltrate the epidermis (early stages of the disease) or acquire the capacity to deeply proliferate and invade the dermis (plaques and tumour stages of the disease). It is a disease of the middle-aged and older adult, and its incremented diagnostic rate over the last 30 years clearly represents the ability of physicians to detect and diagnose early, a patchy – and mainly asymptomatic – disease.

Clinical features

The clinical history of this disease is characterized by an evolution over time in three different stages: the patch stage, the plaque stage and the tumour stage. The evolution reflects a progressively increased aggressiveness of the disease; it is associated with a different clinical appearance and prognosis, as well as with a particular treatment approach.

The patch stage is characterized by generally round to oval, flat or scaly cutaneous lesions – or patches – ranging from 1 to 10 cm in diameter, mainly appearing in areas of skin protected from sunlight (such as the buttocks, the groin and the breasts, in women). Patches may present slight erythema, wrinkling or telangiectases; areas of hypo- or hyperpigmentation may occur. Patches may be confined to a limited skin area or may become disseminated. Patients at this stage are often asymptomatic, and this phase may last indolently for several years or decades. Over time, cutaneous lesions may become infiltrated, and therefore flat or raised indurated lesions – or plaques – supervene. Plaques are red-to-brown in colour, often polycyclic and with a clear spot in the centre; sometimes they become ulcerated. An erythrodermic phase of the disease may be documented at either patch or plaque stage: the skin appears diffusely erythematous, infiltrated and desquamating; the face is also affected, although the elbow fold and the popliteal fossa seem to be characteristically spared. A thorough examination of peripheral blood is mandatory in this clinical context to correctly rule out the differential diagnosis between an erythrodermic mycosis fungoides and Sézary syndrome. The tumour stage is characterized by the appearance of nodules or tumours on pre-existent patches or plaques (Figure 28.5): tumours are raised and plum coloured, with a smooth surface that usually ulcerates, and display a 'fungal' growth, hence the name of 'mycosis fungoides', first attributed by Alibert and Bazin to this disease. Along with

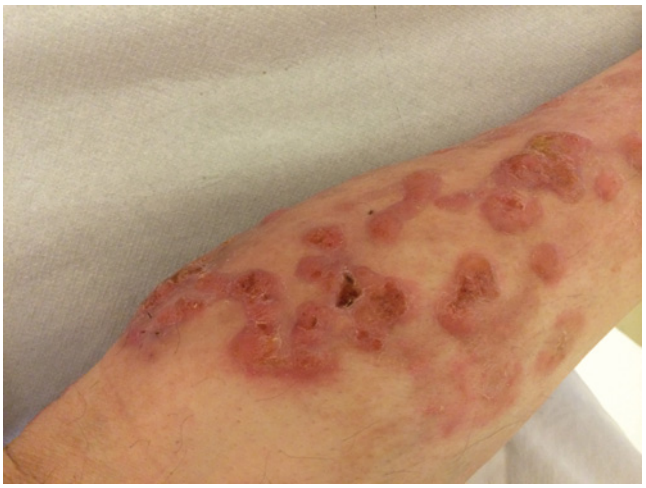


Figure 28.5 Mycosis fungoides, tumour phase.

the disease evolution to the tumour stage and the progressive appearance of new nodular lesions, the patient's general condition tends to worsen, and the prognosis becomes dismal.

A rare clinical variant (less than 10% of cases) of the disease is represented by folliculotropic mycosis fungoides (Figure 28.6), which is characterized by the presence of disseminated – and sometimes coalescent – papules, centred around hair follicles, or erythematous plaques, which generally involve the face (*facies leonina*) and the scalp, as well as the sternal and interscapular region. The prognosis seems less severe than for classical mycosis fungoides, with a 5-year overall survival of about 70%.

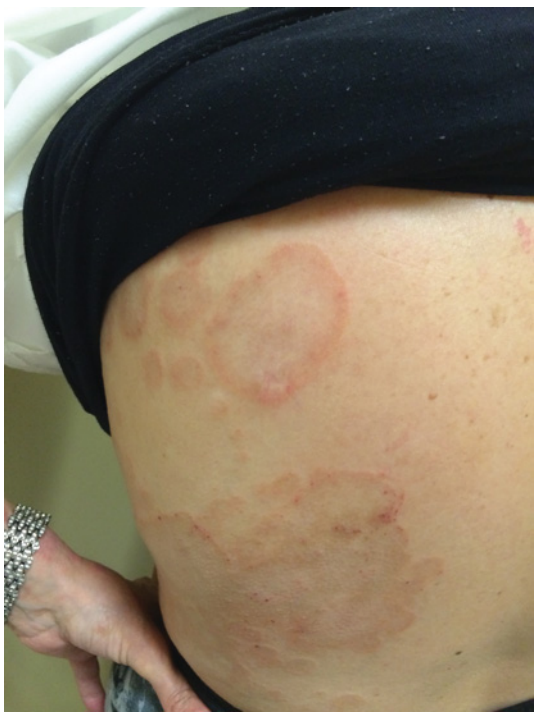


Figure 28.6 Folliculotropic mycosis fungoides.

Morphology

The histopathologic appearance of the disease is extremely variable, depending on the stage. Early mycosis fungoides is not histologically recognizable with certainty: early patches feature small, isolated and scattered – or sometimes clustered – lymphocytes within the papillary dermis, with just a few within the epidermis. As infiltrates become denser, lymphocytes acquire an irregular (cerebriform) nucleus, surrounded by a small halo, and tend to invade the epidermis (epidermotropism), eliciting spongiosis or oedema between keratynocytes. Plaques contain a great amount of atypical, cerebriform, CD3+, CD4+, CD8– lymphocytes, infiltrating the epidermis, the papillary and the reticular dermis (Figure 28.7). Prominent collections of lymphocytes within the epidermis are termed Pautrier's microabscesses; these are typical of plaques, and rarely found in patches. Mycosis fungoides tumours show a diffuse infiltration of the dermis, with lymphocytes displaying a large variety of cytological appearance, although cells with cerebriform nuclei tend to predominate (Figure 28.8). Epidermotropism in this phase is scarce or totally absent.

The CD30 antigen can be expressed in some plaques of mycosis fungoides, although it is mostly expressed in tumours harbouring larger amounts of anaplastic cells.

Staging system

Unlike the other B- and T-cell lymphomas, a revised TNM staging system is used for mycosis fungoides (Tables 28.2 and 28.3), which takes into account: the cutaneous involvement (T-stage), in terms of skin surface, presence of tumours or erythema; the

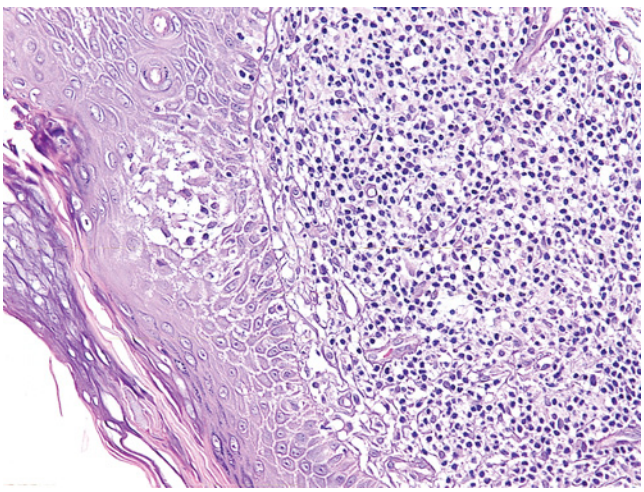


Figure 28.7 Histological appearance of plaques of mycosis fungoides (haematoxylin and eosin stain, magnification 200×).

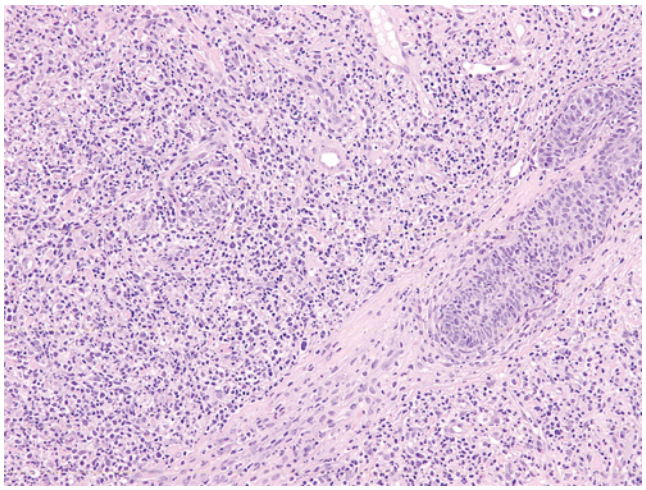


Figure 28.8 Histological appearance of tumours of mycosis fungoides (haematoxylin and eosin stain, magnification 100×).

involvement of peripheral lymph nodes (N-stage), which should be at least 1.5 cm in their major diameter, along with an evidence of T-cell receptor rearrangement (stages N1b and N2b) or not (stages N1a and N2a) and the involvement of visceral organs (M-stage), which requires pathological confirmation. The blood tumour burden is quantified by the B-parameter, and it reflects the amount of circulating Sézary cells (less or equal to 5% versus more than 5%, with at least 1000/μL Sézary cells being diagnostic for Sézary syndrome).

Therapy

The treatment of this peculiar kind of cutaneous T-cell lymphoma is specifically stage-adapted (see Table 28.3), which means that different treatment options are applied for each stage of the disease: patients with an early-stage mycosis fungoides, which is characterized by a slow progression and an indolent course, may take advantage of skin-directed therapies, although lacking a curative intent, whereas more advanced stages require systemic approaches, with either more conventional chemotherapeutic drugs or newer agents.

TNM stage I disease, consisting of patches and/or plaques, requires a skin-directed approach: photochemotherapy, using a psoralen compound as a photosensitizing agent and ultraviolet light type A (PUVA), is nowadays the most widely adopted treatment strategy, along with narrow-band ultraviolet light B (UVB) phototherapy (wavelength of 311–313 nm) if patches represent the most relevant cutaneous lesions. Recombinant α_2-interferon

Table 28.2 Description of TNM-B staging system categories for mycosis fungoides.

Skin	T_1	Limited patches, papules and/or plaques covering <10% of the skin surface. Further stratify into T_{1a} (patches only) and T_{1b} (plaques ± patches).
	T_2	Patches, papules or plaques covering ≥10% of the skin surface. Further stratify into T_{2a} (patches only) and T_{2b} (plaques ± patches).
	T_3	One or more tumours, ≥1 cm in diameter.
	T_4	Confluence of erythema, covering ≥80% of body surface area.
Nodes	N_0	No clinically abnormal peripheral lymph nodes: biopsy not required.
	N_1	Clinically abnormal peripheral lymph nodes (firm, irregular, clustered, fixed nodes, ≥1.5 cm in diameter)*. Further stratify into N_{1a} (clonally negative) and N_{1b} (clonally positive).
	N_2	Clinically abnormal peripheral lymph nodes**. Further stratify into N_{2a} (clonally negative) and N_{2b} (clonally positive).
	N_3	Clinically abnormal peripheral lymph nodes***. Either clonally positive or negative.
	N_X	Clinically abnormal peripheral lymph nodes; no histological confirmation.
Viscera	M_0	No visceral organ involvement.
	M_1	Visceral involvement (requires histological confirmation; involved organ should be specified).
Blood	B_0	Absence of significant blood involvement: ≤5% of peripheral blood lymphocytes are atypical (Sézary) cells. Further stratify into B_{0a} (clonally negative) and B_{0b} (clonally positive).
	B_1	Low blood tumour burden: >5% of peripheral blood lymphocytes are atypical (Sézary) cells, but criteria for B_2 are not met. Further stratify into B_{1a} (clonally negative) and B_{1b} (clonally positive).
	B_2	High blood tumour burden: ≥1000/μL Sézary cells clonally positive.

*Histopathology Dutch (lymph node effacement scale) grade 1 or National Cancer Institute (number of atypical lymphocytes within the involved node) LN_{0-2}.
**Histopathology Dutch grade 2 or National Cancer Institute LN_3.
***Histopathology Dutch grade 3–4 or National Cancer Institute LN_4.

Table 28.3 TNM-B staging system for mycosis fungoides.

	T	N	M	B
I A	1	0	0	0–1
I B	2	0	0	0–1
II A	1–2	1–2	0	0–1
II B	3	0–2	0	0–1
III A	4	0–2	0	0
III B	4	0–2	0	1
IV A_1	1–4	0–2	0	2
IV A_2	1–4	3	0	0–2
IV B	1–4	0–3	1	0–2

may be associated with PUVA therapy as a second-line therapy, in order to produce more durable and profound clinical responses; this association also represents a suitable treatment of stage IIA disease (with patches and/or plaques and lymph node involvement). Unilesional mycosis fungoides may also be treated with conventional external radiation therapy (roentgen therapy).

TNM stage IIB (tumour-stage disease) may be treated either with total skin electron beam irradiation (TSEBI), which yields only 40% satisfactory clinical response, or with systemic monochemotherapy, using gemcitabine or liposomal doxorubicin. Treatment alternatives are represented by bexarotene and dinileukin diftitox (the latter is approved in the United States, but is unavailable elsewhere). Bexarotene is an orally (but also topically, as a gel formulation) bioavailable retinoid that acts as a selective agonist of the retinoid X receptor, which consequently exerts its function through the recruitment of transcription factors, and induces apoptosis and cell cycle arrest in CTCL cells. Dinileukin diftitox is a hybrid molecule, containing the interleukin-2 (IL-2) receptor binding domain and the catalytically active fragment of diphtheria toxin, which targets the IL-2 receptor found on activated B- and T-cells. The toxin is then internalized, and induces the apoptosis of the targeted cell.

Stage III disease (erythroderma) can be first treated again with the combination of recombinant α_2-interferon and PUVA; possible alternatives or second-line options are represented by TSEBI, gemcitabine, liposomal doxorubicin or bexarotene.

Stage IV disease, which is characterized by extracutaneous involvement, requires an immediate systemic chemotherapy: gemcitabine or doxorubicin used as monotherapy should be preferred to polychemotherapy with cyclophosphamide, doxorubicin, vincristine and prednisone (CHOP) or CHOP-like regimens, which appear significantly more toxic and display poor disease control.

Histone deacetylase inhibitors (HDACi) represent the new frontier of mycosis fungoides treatment. Through the inhibition of the enzymatic removal of acetyl groups from histonic and non-histonic proteins, these molecules control the epigenetic regulation of gene transcription, ultimately regulating cell-cycle and apoptosis-related proteins. Vorinostat and romidepsin, the former being orally available and the latter being administered intravenously, have both shown a rate of response of 30–40% in Phase II studies, with a median duration of response varying from nearly 200 days to more than 1 year.

Allogeneic transplantation can be a suitable treatment option for younger patients failing at least two previous lines of standard therapy; durable remissions may be obtained as a consequence of a putative graft-versus-lymphoma immune effect, yielding 5-year progression-free and overall survival rates of 34% and 53%, respectively.

Sézary syndrome

Sézary syndrome, first described in 1938 by Sézary and Bouvrain, who identified a population of large lymphocytes in the peripheral blood, with grooved and lobulated 'cerebriform' nuclei (*cellules monstrueuses*), is regarded as the leukaemic form of mycosis fungoides, clinically characterized by the triad: erythroderma (diffuse reddening of the skin), generalized lymphadenopathy and atypical T-lymphocytes (Sézary cells) in peripheral blood, skin and lymph nodes. Like mycosis fungoides, it is a disease of the middle-aged and elderly, although much less common.

Clinical features

Erythroderma is the most relevant clinical aspect; it either tends to evolve gradually over months or even years from a pre-existing non-specific dermatitis, or it may abruptly develop *de novo*. Sometimes, erythroderma may reflect an evolution of a patch or plaque stage mycosis fungoides (see above). Skin is red and scaly, and may become doughy, as well as fissurated. Hyperkeratosis of palms and soles is also a distinguishing feature, along with onychodystrophy and sometimes alopecia.

Morphology

Sézary cells are clonal mature T-helper lymphocytes, generally CD3+, CD4+ and CD8–; an aberrant loss of T-cell antigens, such as CD2, CD3, CD4, CD5 and CD7 is frequently observed and T-cell receptor rearrangement found by polymerase chain reaction provides proof of clonality. Histologically, skin changes resemble those of mycosis fungoides, although epidermotropism is seen to a lesser extent. Sometimes, however, skin specimens can be non-diagnostic for mycosis fungoides at all, because there can be more spongiosis and lymphocytes are often small.

Diagnosis

The integration of clinical features, histology, cytology, immunophenotyping and molecular assays is necessary to make a correct diagnosis of Sézary syndrome. In the presence

of erythroderma, Sézary syndrome is diagnosed when the absolute count of Sézary cells is ≥1000/μL (or Sézary cells represent more than 20% of circulating lymphocytes). Clonal rearrangement of the T-cell receptor is required to establish the diagnosis. Alternatively, one of the following criteria should be satisfied, if a count of Sézary cells cannot be performed: increased CD4+ or CD3+ cells, with a CD4/CD8 ratio ≥10, aberrant immunophenotypes (CD4+/CD7– ratio ≥40% or CD4+/CD26– ratio ≥30%).

Therapy

Extracorporeal photopheresis (ECP) represents the treatment of choice, and it is performed for two consecutive days every 2–4 weeks. During ECP, pooled leucapheresis and plasmapheresis products are treated with 8-methoxypsoralen, then exposed to ultraviolet light type A (UVA) as they pass through a disposable narrow cassette, *ex vivo*. Irradiated leucocytes are subsequently re-infused: psoralens covalently bind to DNA after irradiation with UVA, therefore leading to the induction of apoptosis. In parallel, ECP seems also to induce dendritic cell differentiation, which in turn translates into an enhanced antigen presentation and the initiation of a host immune response against the tumour. Overall responses to ECP are around 60% and complete responses are seen in about 20% of the cases, with durable responses seen in patients lacking a significant nodal or visceral involvement. ECP can also be combined with TSEBI, α_2-interferon or bexarotene therapy. Systemic chemotherapy based on chlorambucil, methotrexate, fludarabine, gemcitabine or doxorubicin as single agents should be delivered in patients with high tumour burden (>2600 Sézary cells/μL), advanced disease with significant nodal or visceral involvement, or relapsed disease.

Primary cutaneous CD30+ lymphoproliferative disorders

These cover a wide spectrum of indolent T-cell lymphoproliferative disorders, primarily localized in the skin and account for nearly 30% of CTCL, thus representing one of the most frequent disease entities.

They are clinically characterized by a chronic clinical course, with cutaneous lesions sometimes showing a relapsing-remitting behaviour or spontaneous regressions, always featuring the presence of large and atypical CD30+ cells, variably represented within the histological specimen. Primary cutaneous anaplastic large-cell lymphoma (C-ALCL) and lymphomatoid papulosis (LyP) represent the two extremities of the spectrum, and are discussed in more detail.

Primary cutaneous anaplastic large-cell lymphoma

C-ALCL is the second most common type of CTCL, being mostly represented in middle-aged men (median age at presentation is 60 years, with a male/female ratio of 2–3:1). It clinically presents with solitary and localized papules, nodules or tumours, frequently ulcerated, involving the skin of the trunk, face, extremities and buttocks. Multifocal involvement is rather uncommon. Lesions may show a spontaneous complete or partial regression; however, skin relapses are frequent. Extracutaneous dissemination is rare, with regional lymph nodes being the most commonly affected organs other than the skin.

Tumour cells are large, with an anaplastic (abundant cytoplasm, round to oval or irregularly shaped nuclei with prominent nucleoli), pleomorphic or immunoblastic cytomorphology; the infiltrate involves the dermis, without epidermotropism. Cells mostly display a CD4+ T-cell phenotype, CD8+ in a minority of cases, with possible loss of CD2, CD5 or CD3; CD30 is by definition expressed by a majority (>30%) of the neoplastic cells, without CD15 expression. Cytotoxicity markers, such as TIA-1, perforin and granzyme B are frequently expressed.

The prognosis is good, with a 5-year disease-free survival higher than 90%, although relapses are frequent. Patients with multifocal skin lesions and patients with involvement of regional nodes show a similar prognosis to those presenting with a unique skin lesion. Localized lesions favourably respond to radiation therapy, and surgical excision may be curative in some cases. Systemic therapy is required for multifocal disease: low-dose methotrexate, α_2-interferon and bexarotene represent the treatments of choice; brentuximab vedotin has shown very promising results.

Lymphomatoid papulosis (LyP)

LyP is a chronic, recurrent, self-healing skin disease, which mainly affects middle-aged men (with a male/female ratio of 2–3:1). It is clinically characterized by the presence of papular, papulonecrotic or nodular skin lesions, mainly affecting the trunk and the extremities, at different stages of development and with a relapsing-and-remitting course over 3 to 12 weeks, leaving behind superficial scars. In about 20% of the patients, LyP is associated – or followed by – another malignant lymphoma, such as mycosis fungoides, C-ALCL or Hodgkin disease.

Three morphological subtypes are described: in type A, the most frequent, scattered or clustered large, Reed–Sternberg-like CD30+ cells infiltrate the superficial dermis, and are intermingled with numerous inflammatory cells; in type B, the most uncommon, cells with cerebriform nuclei involve the superficial dermis and show epidermotropism (CD30 expression is generally lacking); type C lesions demonstrate a monotonous population of large clusters of highly CD30-expressing cells, with a few admixed inflammatory cells.

The prognosis is excellent; no treatment is recommended for localized lesions, while topical steroids, low-dose methotrexate or PUVA phototherapy may be suitable in cases of disseminated disease.

Conclusions

T-cell lymphoproliferative disorders are classified on the basis of their clinical behaviour and presentation, thus distinguishing them into either indolent or aggressive nodal, extranodal and leukaemic diseases. Although infrequent diseases, they represent a big challenge for the treating haematologists, since their prognosis is rather disappointing when the current treatment strategies are applied, and relapse rates are still relevant. A precise clinical and histological diagnosis, as well as the presence of a multidisciplinary team, are the key elements to assure the best patient management. The application of molecular techniques and immunological markers has clarified some aspects of the pathogenesis of these diseases, and has provided the rationale for the development of new specific treatment approaches.

Acknowledgements

The Authors would like to thank Dr Claudio Agostinelli for providing Figures 28.2, 28.3, 28.4, 28.7 and 28.8.

Selected bibliography

Armitage JO (2013) The aggressive peripheral T-cell lymphomas: 2013. *American Journal of Hematology* **88**: 911–8.

Dearden C (2012) How I treat prolymphocytic leukemia. *Blood* **120**: 538–51.

Foss FM, Zinzani PL, Vose JM *et al.* (2011) Peripheral T-cell lymphoma. *Blood* **117**: 6756–67.

Lamy T, Loughran TP (2011) How I treat LGL leukemia. *Blood* **117**: 2764–74.

Swerdlow SH, Campo E, Harris NL *et al.* (eds) (2008) *World Health Organization Classification of Tumours of Haematopoietic and Lymphoid Tissues*. IARC press, Lyon.

Vose J, Armitage J, Weisenburger D (2008) International peripheral T-cell and natural killer/T-cell lymphoma study: pathology findings and clinical outcomes. *Journal of Clinical Oncology* **26**: 4124–30.

Wilcox RA (2011) Cutaneous T-cell lymphoma: 2011 update on diagnosis, risk-stratification and management. *American Journal of Hematology* **86**: 928–48.

Willemze R, Hodak E, Zinzani PL *et al.* (2013) Primary cutaneous lymphomas: ESMO clinical practice guidelines for diagnosis, treatment and follow-up. *Annals of Oncology* suppl 6: VI149–54.

CHAPTER 29

Multiple myeloma

29

Jesús San-Miguel[1] and Joan Bladé[2]

[1]Clinica Universidad de Navarra, Centro de Investigación Medica Aplicada (CIMA), Pamplona, Spain
[2]Hospital Clinic de Barcelona, Barcelona, Spain

Definition

Multiple myeloma (MM) is characterized by the proliferation of a single clone of plasma cells that produce a monoclonal protein. The plasma cell proliferation results in extensive skeletal involvement, with osteolytic lesions, hypercalcaemia, anaemia and/or soft tissue plasmacytomas. In addition, the excessive production of nephrotoxic monoclonal immunoglobulin can result in renal failure and an increased risk of developing potentially life-threatening infections due to the lack of functional immunoglobulins. The clinical and laboratory manifestations of the disease, including their management, are discussed in this chapter.

Epidemiology and aetiology

The annual incidence of MM is 4 per 100,000. It represents approximately 1% of all malignant diseases and 15% of all haematological malignancies. The incidence of MM is lower in Asian populations and in blacks is twice that in whites; MM is slightly more frequent in men than in women. The median age at diagnosis is 65–70 years. Only 15% and 2% of the patients are younger than 50 and 40 years, respectively.

The cause of MM is unknown. Radiation may play a role in some cases. An increased risk has been reported in farmers, particularly those who use herbicides and insecticides, and in people exposed to benzene and other organic solvents. However, the number of cases is small and more data are needed to establish a significant relationship. MM and monoclonal gammopathy of undetermined significance (MGUS) have been reported in familial clusters. A relationship between MM and pre-existing inflammatory diseases has been suggested, and plasma cell dyscrasias associated with protracted stimulation of the reticuloendothelial system have been reported in experimental studies. However, more recent case–control studies do not support a role for chronic antigenic stimulation in the aetiopathogenesis of MM. There is now clear evidence that most, if not all, myeloma cases are preceded by a previous MGUS.

Pathogenesis

MM is a B-cell malignancy characterized by the accumulation of terminally differentiated clonal plasma cells in the bone marrow, the production of a monoclonal immunoglobulin detectable in serum and/or urine and the presence of lytic bone lesions. In order to understand the pathogenesis of MM, it is important to review not only the molecular changes involved in the development of the malignant clone, but also the mechanisms responsible for the interaction between the malignant plasma cells and their microenvironment, since they play a relevant role in bone destruction, tumour cell growth, survival, migration and drug resistance.

Cellular origin of myeloma cells

Normal differentiation from early B cells to plasma cells is characterized by three B-cell-specific DNA remodelling mechanisms that modify immunoglobulin genes: VDJ rearrangement,

Postgraduate Haematology, Seventh Edition. Edited by A Victor Hoffbrand, Douglas R Higgs, David M Keeling and Atul B Mehta.

somatic mutation and class switch recombination. Rearrangements of the immunoglobulin genes of B-cell precursors to form a B-cell receptor (BCR) occur in the bone marrow, while antigen recognition, selection, somatic hypermutation and class switch recombination take place in the germinal centre lymph node. Sequence analysis of the immunoglobulin VH gene support the postgerminal origin of myeloma cells, which have successfully completed somatic hypermutation (without intraclonal variation) and IgH switching, before migrating to the bone marrow, where they will interact with stromal cells before finally differentiating into long-lived plasma cells.

Genomic abnormalities

Genome instability is a prominent feature of myeloma cells and in fact, almost all cases of MM are cytogenetically abnormal. Genomic abnormalities can be categorized as chromosomal translocations, mainly involving the IGH locus on chromosome 14q32, copy number abnormalities, mutations, methylation modifications, and gene and microRNA (miRNA) dysregulation.

IGH translocations

A primary event in many kinds of B-cell tumour is dysregulation of an oncogene that, as a result of translocation to the *IGH* locus (14q32) or, somewhat less often, the *IGL* locus (κ 2p11 or λ 22q11), is juxtaposed near one of the potent immunoglobulin enhancers. In MM, *IGH* translocation may be classified into primary or secondary. Primary *IGH* translocations occur as initiating events during the pathogenesis of MM, whereas secondary translocations are involved in progression. Most primary *IGH* translocations result from errors in B-cell-specific DNA modification processes, mostly *IGH* switch recombination or, less often, somatic hypermutation. The breakpoints occur mainly within or immediately adjacent to *IGH* switch regions or JH regions. In contrast, secondary translocations are mediated by other kinds of recombination mechanism that do not specifically target B-cell-specific DNA modification processes. Unlike other B-cell tumours, in MM there is a marked diversity of chromosomal loci involved in *IGH* translocations. About 40% of MM tumours have *IGH* translocations involving five recurrent chromosomal patterns (Figure 29.1): 11q13 (*CCND1*), 4p16 (*FGFR/MMSET*), 16q23 (*MAF*), 6p21 (*CCND3*) and 20q11 (*MAFB*).

The prevalence of t(11;14) according to interphase fluorescence *in situ* hybridization (FISH) analysis is 15–20% and is readily detectable by karyotyping. As a result of the translocation, *CCND1* is juxtaposed to the powerful *IGH* 3′ enhancer(s) on der(14), and its expression is dysregulated, as indicated by gene expression profiling and reverse transcriptase polymerase chain reaction (RT-PCR) in 100% of MM cases with t(11;14).

The t(4;14) translocation is identified in approximately 15% of MM cases using FISH analysis, but cannot be detected by karyotyping techniques. This translocation results in the simultaneous deregulation of the fibroblast growth factor receptor 3 (*FGFR3*) gene on der(14) and the multiple myeloma SET domain (*MMSET*) gene on der(4). FGFR3 is one of the high-affinity tyrosine kinase receptors for the FGF family of ligands. Both *FGFR3* and *MMSET* genes are not normally expressed in plasma cells, but are over-expressed as a result of t(4;14). However, gene expression profiling and RT-PCR analysis have shown that only 75% of MM cases with t(4;14) display

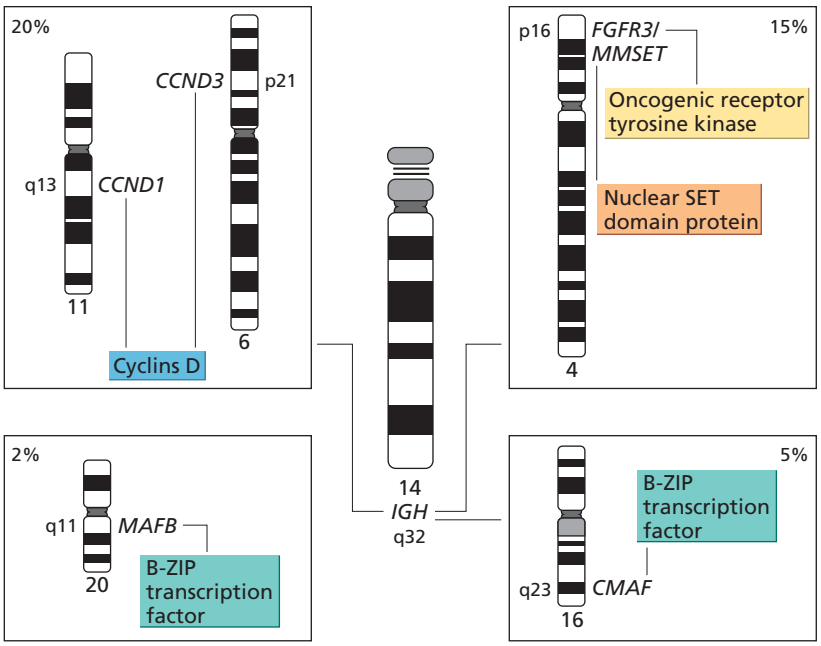


Figure 29.1 *IGH* translocations in multiple myeloma.

simultaneous over-expression of MMSET and FGFR3. In the remaining 25% of cases, only MMSET is upregulated and the lack of FGFR3 expression is linked in most cases to loss of the *FGFR3* gene on der(14). These data suggest that *MMSET* may be the critical transforming event in MM harbouring t(4;14), whereas *FGFR3* could be dispensable. In some cases (10%) the translocated *FGFR3* contains activating mutations that may be involved in MM progression.

The incidence of t(14;16) is 5–10%. The breakpoints on 16q23 occur over a region 550–1350 kb centromeric to *MAF*. Taking into account such a long distance, it is still an open question whether or not *IGH* may act as enhancer for *MAF* in this translocation. Moreover, over-expression of MAF is observed in half of myeloma cases, while the prevalence of t(14;16) is low. The t(6;14) translocation has been found in a low proportion (3%) of MM cases. Using microarray analyses, high levels of cyclin D3 mRNA have been shown in cases with t(6;14) detected by FISH. The t(14;20) translocation leads to deregulation of *MAFB* (20q23), which, like *MAF*, encodes a B-ZIP transcription factor, but in contrast to t(14;16), *MAFB* translocations have structural features that indicate they are secondary translocations.

Gains and losses of chromosomal material

Almost all MM cases are aneuploid, as evidenced by the measurement of DNA content by flow cytometry and cytogenetic techniques. Patients with MM may be grouped into two major categories, according to ploidy status assessed by karyotyping: the hyperdiploid group (more than 46/47 chromosomes) and the non-hyperdiploid group, composed of hypodiploid (up to 44/45 chromosomes), pseudodiploid (44/45 to 46/47) and near tetraploid (more than 74) cases. Non-hyperdiploid MM is characterized by a very high prevalence of *IGH* translocations involving the five recurrent partners. Likewise, monosomy/deletion 13 and gains on 1q occur predominantly in non-hyperdiploid MM. In contrast, the hyperdiploid group is associated with recurrent trisomies involving odd chromosomes (3, 5, 7, 9, 11, 15 and 19) and with a low incidence of structural chromosomal abnormalities. Similar associations have been observed on analysing DNA content by flow cytometry.

The loss of chromosome 13 is the most common monosomy in MM (40–50% of newly diagnosed patients). This abnormality shows a strong association with t(4;14) and t(14;16), deletion of 17p and gains on 1q. Chromosome 17p deletion, which includes loss of *TP53*, occurs at a lower frequency in newly diagnosed MM (5–10%), although the proportion is higher in advanced stages of the disease. Furthermore, 17p deletion is associated with extramedullary MM. Conventional cytogenetics, FISH and comparative genomic hybridization analysis have all demonstrated that lesions of chromosome 1 are the most common abnormalities in MM; mostly they are 1q gains, as the result of tandem duplications and jumping segmental duplications of the chromosome 1q band. Recently, a large FISH study has demonstrated that 1p losses (especially 1p22 and 1p32 deletions) are also frequent in MM patients.

Mutations detected by whole-genome sequencing

Whole-genome sequencing strategies have shown that there are approximately 35 non-synonymous mutations per myeloma sample. However, few recurrently mutated genes have been detected, apart from the well-known mutations in the ERK pathway. This is in agreement with other haematological malignancies, such as acute myeloid leukaemia, but is in contrast to hairy cell leukaemia and Waldenstrom's macroglobulinaemia, in which single unifying mutations are seen, *BRAF* and *MYD88*, respectively.

Epigenetic modifications

Little is known about the epigenetic changes involved in MM pathogenesis. The most relevant epigenetic change revealed so far is the global DNA hypomethylation and gene-specific DNA hypermethylation in MM as compared to MGUS. Interestingly, patients with the t(4;14) translocation have increased gene-specific DNA hypermethylation compared with myeloma samples of other cytogenetic subgroups.

Late genetic events

Some genetic changes in MM, such as secondary translocations, mutations, deletions and epigenetic abnormalities, are considered late oncogenic events and are associated with disease progression. Dysregulation of *MYC* is a paradigm for secondary translocations in MM. Most karyotypic abnormalities involving *MYC* correspond to complex translocations and insertions that often are non-reciprocal and frequently involve three different chromosomes. Activating *RAS* mutations are considered molecular markers of disease progression. Thus, the prevalence of activating *KRAS* and *NRAS* mutations is over 75% in MM cases at relapse. *TP53* inactivation, via either deletion or mutation, seems to be more frequently associated with disease progression. Methylation is an epigenetic change that has been described in MM and acts as an inactivating mechanism of the tumour-suppressor genes *CDKN2B* and *CDKN2A*. Although it has also been detected in MGUS, its prevalence is much higher in advanced MM and extramedullary forms of the disease.

Molecular classification of MM based on gene expression profiling

Gene expression analysis of MM has confirmed the huge genetic diversity of this tumour. Recently, the classification of MM into seven different groups has been proposed. Each group displays a specific genetic signature and some of them are associated with a particular *IGH* translocation or ploidy status and with a characteristic clinical behaviour. Table 29.1 summarizes this

Table 29.1 Molecular classification of multiple myeloma.

Group	Specific translocation	Frequency (%)	Cyclin D expression	Genetic signature	Prognosis	Other characteristics
1 PR	–	12	CCND2	↑*CCNB1*, ↑*CCNB2*, ↑*MCM2*, ↑*BUB1*, ↑*MAGEA6*, ↑*MAGEA3*, ↑*GAGE1*	Unfavourable	Normal karyotypes
2 LB	–	11	CCND2	↑*EDN1*, ↑*IL6R*, ↓*DKK1*, ↓*FRZB*	Favourable	Lower number of bone lesions
3 MS	t(4;14) *FGFR3/MMSET*	18	CCND2	↑*FGFR3*, ↑*MMSET*, ↑*PBX1*, ↑*PAX5*	Unfavourable	
4 HY	–	26	CCND1	↑*TRAIL*, ↑*DKK1*, ↑*FRZB*, ↓*CKS1B*	Favourable	Hyperdiploid karyotype, bone lesions
5 CD-1	t(11;14) *CCND1* or t(6;14) *CCND3*	8	CCND1 or CCND3	↑*CEBPB*, ↑*NID2*, ↑*SET7*	Favourable	
6 CD-2	t(11;14) *CCND1* or t(6;14) *CCND3*	17	CCND1 or CCND3	↑*MS4A1* (*CD20*), ↑*PAX5*, ↑*CD27*, ↑*CXCR4*	Favourable	
7 MF	t(14;16) *MAF* or t(14;20) *MAFB*	8	CCND2	↑*MAF*, ↑*MAFB*, ↑*CXCR1*, ↑*ITGB7*, ↓*DKK1*	Unfavourable	Lower number of bone lesions

PR, proliferation; LB, low bone disease; MS, *MMSET*; HY, hyperdiploid; CD-1, *CCND1/CCND3*; CD-2, *CCND1/CCND3*; MF, *MAF/MAFB*.

classification, which connects genetic abnormalities, cell transcriptome and clinical features of patients.

Dysregulation of cyclin D genes as a potential unifying event in MM pathogenesis

There is no common genetic mechanism to explain the pathogenesis of MM. However, it can be speculated that although *IGH* translocations induce upregulation of different oncogenes, it is possible that all *IGH* translocations involved in MM converge on a common pathway resulting in inhibition of differentiation and an increase in cell survival and proliferation. Gene expression profiling analysis has demonstrated that expression of *CCND1*, *CCND2* and *CCND3* is increased in virtually all MM patients, supporting the recent hypothesis of a potential unifying event in pathogenesis. Approximately 25% of MM cases display over-expression of one of these cyclins, which may be triggered directly by an *IGH* translocation such as t(11;14) and t(6;14) that dysregulates *CCND1* and *CCND3* respectively, or indirectly by an *IGH* translocation involving *MAF* and *MAFB* genes, which encode a transcription factor that targets cyclin D2. Nearly 40% of MM cases express increased cyclin D1 through biallelic dysregulation of *CCND1* and without apparent t(11;14); most of the remaining cases of MM, including those with t(4;14), have increased expression of cyclin D2. The expression level of cyclin D has also been incorporated into the molecular classification (Table 29.1).

MicroRNA expression

MicroRNAs (miRNAs) are small non-coding RNAs that regulate gene expression at the post-transcriptional level and are involved in critical biological processes, including cellular growth and differentiation. Different studies have shown that miRNA expression is deregulated in myeloma cells as compared to normal plasma cells, and that their expression profile is associated with genetic abnormalities. Moreover, several miRNAs have been involved in MM pathogenesis. In this sense, it has shown a mechanism of p53 regulation through miRNAs acting on MDM2 expression; thus, miR-192, 194 and 215 re-expression in myeloma cell lines induce degradation of MDM2 with the subsequent p53 upregulation and cell growth inhibition.

Multistep pathogenesis of multiple myeloma

The current pathogenic models assume that MM develops through a multistep transformation from normal plasma cells (PCs) to MGUS, which implies PC immortalization and, subsequently, the transformation to active MM, where clonal PCs cause end-organ damage. Cytogenetic studies using FISH have demonstrated that most genetic lesions typical of MM are already present at MGUS stage. We recently have shown that a major difference between these three entities is the number of PCs with genetic abnormalities, which increases from MGUS to SMM (smouldering multiple myeloma) and to MM, thus the progression from MGUS to SMM, and eventually to

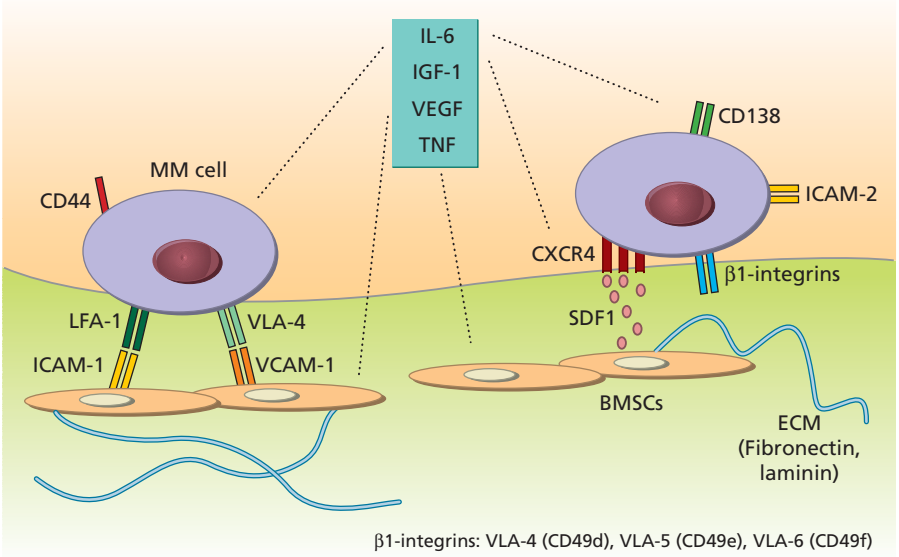


Figure 29.2 Interactions between plasma cells and the microenvironment. See text for definition of abbreviations.

MM, involves a clonal expansion of genetically abnormal PCs. These findings have also been confirmed by SNP-based mapping arrays and whole-genome sequencing. Similarly, the progression of overt MM implies a complex evolutionary process with intraclonal heterogeneity, where a Darwinian branching model drives a clonal competition with alternating dominance.

Interaction between plasma cells and their microenvironment

As far as the pathogenesis of MM is concerned, interactions between the myelomatous plasma cells and their microenvironment can be as important as the genetic lesions. In the bone marrow, MM cells adhere to extracellular matrix proteins and bone marrow stromal cells through a series of adhesion molecules, for example the β_1 integrin family (VLA-4, VLA-5, VLA-6; also called CD49d, CD49e and CD49f) (present in myeloma cells), and vascular cell adhesion molecule (VCAM)-1 and intercellular adhesion molecule (ICAM)-1 (present in stromal cells) (Figure 29.2). In addition, bone marrow stromal cells produce a stromal-cell-derived factor (SDF)-1 that binds to CXCR4 on the surface of myeloma cells, inducing both chemotaxis of plasma cells and upregulation of surface adhesion molecules such as VLA-4. Bone marrow homing of plasma cells is likely further facilitated through other adhesion molecules expressed by myeloma cells, such as CD138, CD38, CD44 and CD106.

Adhesion of myeloma cells to the bone marrow microenvironment induces a cell adhesion-mediated drug resistance phenotype via three mechanisms: (i) cell cycle arrest at G_1 (associated with upregulation of p27, an inhibitor of cyclin-dependent kinases), (ii) apoptosis inhibition via over-expression of FLIP-L, an endogenous inhibitor of FAS (CD95) and (iii) protection of tumour cells from initial drug-induced DNA damage (double-strand breaks) by reducing topoisomerase II activity. The binding of MM cells to the bone marrow microenvironment also induces the transcription and secretion of cytokines, such as tumour necrosis factor (TNF)-α, interleukin (IL)-6, insulin-like growth factor (IGF)-1, IL-21, SDF-1α and vascular endothelial growth factor (VEGF), by plasma cells and/or bone marrow stromal cells; this triggers signalling pathways (e.g. RAF/MEK/MAPK, PI3K/AKT, NF-κB and JAK/STAT) that promote cell proliferation and prevent apoptosis. These pathways are also potential targets for therapeutic intervention (Figure 29.3). In addition, cytokines modulate the production of additional adhesion molecules, which, in a vicious circle, further enhance cell adhesion.

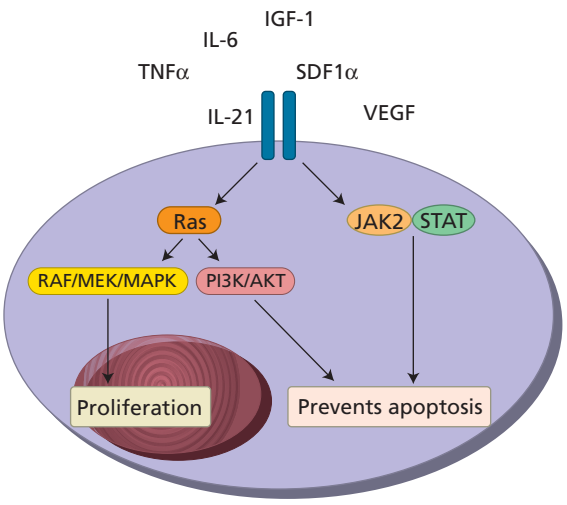


Figure 29.3 Signalling pathways involved in myeloma pathogenesis. See text for definition of abbreviations.

In summary, it appears that the bone marrow microenvironment provides a sanctuary for myeloma cells by both promoting proliferation and blocking apoptosis, thereby allowing tumour progression and eventual emergence of drug resistance. Interruption by downregulating the interactions between the tumour cell and its microenvironment can potentially halt cell growth and proliferation and be of benefit to patients.

Influence of pathogenesis on the clinical features of MM and the development of bone lesions

The interaction between MM cells and the microenvironment not only favours tumour growth, but is also responsible for the final myeloma portrait (see below and Figure 29.4). This

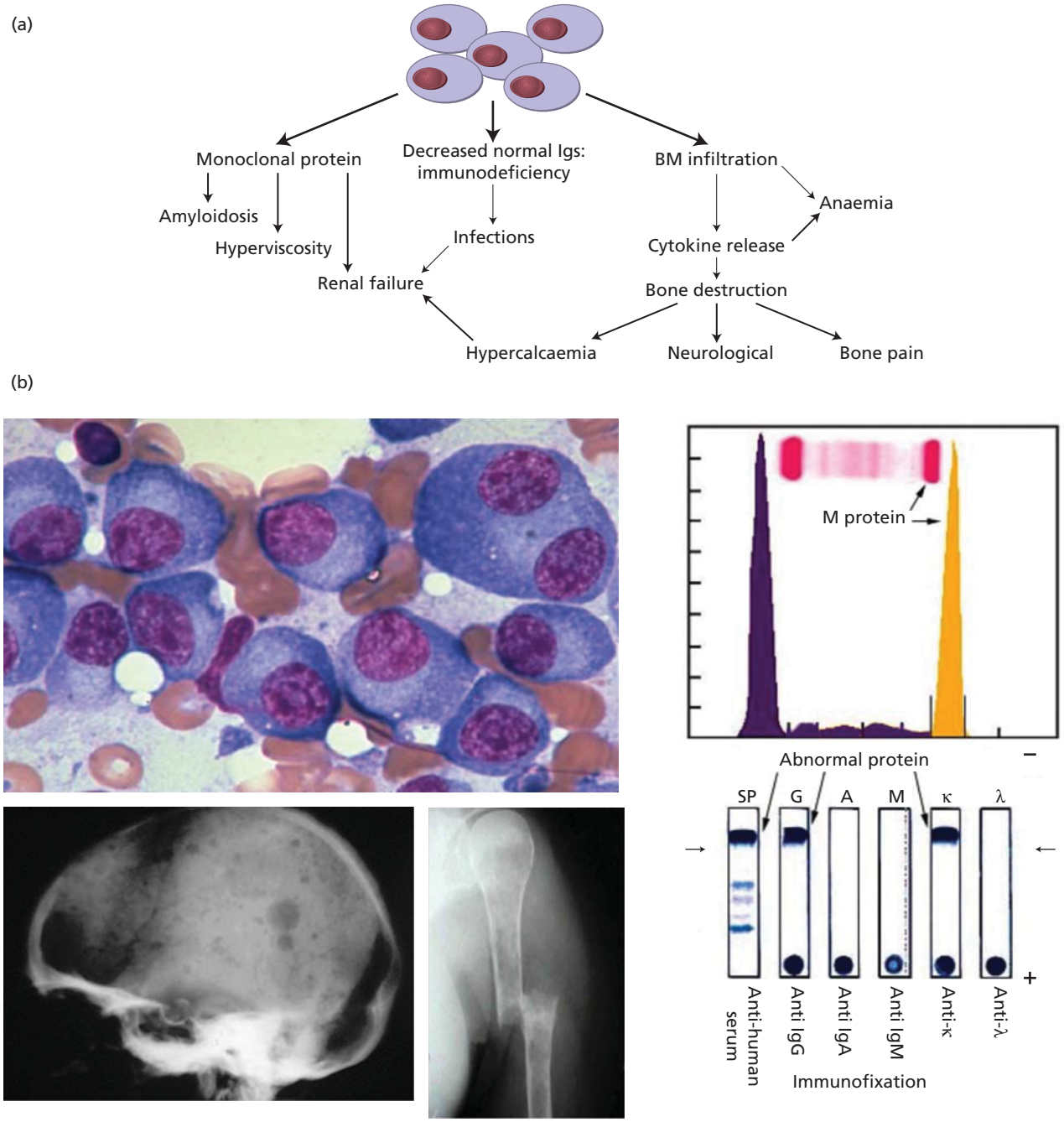


Figure 29.4 (a) Clinical manifestations in myeloma and (b) bone marrow infiltration, monoclonal band (IgGK by immunofixation) osteolytic lesions resulting in a pathological fracture.

interaction stimulates other cytokine cascades responsible for osteolytic lesions, which can result in bone pain, hypercalcaemia and neurological compression syndromes. Bone marrow infiltration also impairs normal haemopoiesis, leading to anaemia. MM cells secrete monoclonal protein (M-protein) that increases plasma viscosity. In addition, M-protein, and particularly light chains, are responsible for the impairment of renal function, leading ultimately to renal failure. Hypercalcaemia due to osteolytic disease may also contribute to renal failure. Patients are also at increased risk of developing potentially life-threatening infections due to the lack of functional immunoglobulins.

Bone disease is one of the hallmarks of MM (see below). Osteolysis is mediated by an imbalance between osteoclast activity (increased) and osteoblast activity (decreased). Interaction of MM cells with stromal cells and other cells in the microenvironment induces the secretion of numerous osteoclast-activating factors such as RANK-L, macrophage inflammatory protein (MIP)-1α (also known as CCL3), activin A, VEGF, hepatocyte growth factor (HGF), IL-3, IL-7, TNF-α, IL-6, IL-1β and MIP-3α. Two of the most important are RANK-L (receptor activator of NF-κB ligand) and MIP-1α.

RANK-L is a transmembrane molecule in stromal cells/osteoprogenitors, which is also called TRANCE (TNF-related activation-induced cytokine) or OPG-L (osteoprotegerin ligand). RANK-L binds to its functional receptor RANK (TNF receptor superfamily) on osteoclasts, stimulating osteoclastogenesis by inducing differentiation of osteoclast precursors and stimulating resorption. RANKL activity can be blocked by osteoprotegerin (OPG), a soluble-decoy receptor for RANK-L also produced by stromal/osteoprogenitor cells. Therefore osteoclastic activity is regulated by a delicate balance between RANK-L and OPG. In fact, under normal physiological conditions, the levels of OPG are significantly higher than those of RANK-L. In contrast, in MM this balance is disrupted by increased expression of RANK-L and decreased expression of OPG on stromal cells after interaction with myeloma cells. It has been suggested that the decline in OPG is also mediated by myeloma cell uptake by binding to CD138 and subsequent degradation. MIP-1α is a potent stimulator of osteoclast formation through a dual mechanism: (i) it enhances the activity of RANK-L and (ii) it directly stimulates osteoclast precursors to differentiate into mature forms. MIP-1α gene expression is abnormally regulated in MM due to unbalanced expression of the acute myeloid leukaemia (AML)-1A and AML-1B transcription factors in myeloma cells. This imbalance also induces IL-3, which stimulates osteoclast formation and resorption directly or by further augmenting that of RANKL and MIP-1α; in addition, as mentioned below, IL-3 inhibits osteoblast formation. Other chemokines such as IL-7, TNFα and IL1β indirectly stimulate osteolytic processes, inducing RANK-L expression. Other OAFs secreted by myeloma cells and/or stromal/or osteoclasts cells (e.g. HGF, IL-6, VEGF, activin A), further increase the gradient of osteoclastogenic factors in focal lesions and contribute to osteoclast production and activity.

In MM, in addition to the marked osteoclast activation, there is inhibition of osteoblast formation and function, which is mediated both by both soluble factors and direct cellular interactions of myeloma and stromal cells. The WNT/bone morphogenetic protein (BMP) signalling pathways are critical for the osteogenic differentiation of mesenchymal stem cells to mature bone-forming osteoblasts. Myeloma cells (and other BM microenvironmental cells) produce numerous soluble factors that inhibit osteoblast differentiation and/or function, such as Wnt signalling antagonists (e.g. DKK1, sclerostin, soluble frizzled related proteins (sFRP-2/3)), BMP inhibitors (e.g. activin A, TGF-β and HGF), and other cytokines and chemokines (e.g. IL-7, TNF-α and IL-3 (which indirectly inhibits osteoblast differentiation involving CD45+ cells in the BM)). In addition, osteoblastic cells from myeloma patients with lytic lesions show reduced Runx2 activity, Runx2 being the major transcription factor regulating osteoblast differentiation. Suppression of Runx2 activity is mediated, at least in part, by cell-to-cell contact of myeloma and mesenchymal osteoprogenitors, since blocking of VLA4-VCAM1 interactions with a neutralizing anti-VLA4 antibody partially restored Runx2 function in mesenchymal cells.

Differential diagnosis

The diagnostic criteria for the monoclonal gammopathies have been reviewed by the International Myeloma Working Group (IMWG). The main clinical entities are MGUS, primary systemic amyloidosis (see also Chapter 32), smouldering multiple myeloma and symptomatic multiple myeloma.

Monoclonal gammopathy of undetermined significance

MGUS has a high prevalence (3.2% and 5.8% in individuals over 50 and 70 years of age, respectively). It is characterized by the presence of a serum M-protein (<30 g/L) and less than 10% plasma cells in the bone marrow with no evidence of other B-cell lymphoproliferative disorder and no symptoms or organ or tissue impairment due to the monoclonal gammopathy. The transformation rate to a malignant plasma cell disorder is about 1% per year, with an actuarial probability of malignant evolution of 30% at 25 years of follow-up. When the different causes of death are considered, the actuarial probability of malignant transformation at 25 years of follow-up is only 11%, much lower than the actuarial prediction. The main factors associated with MGUS progression include M-protein size, IgA isotype, abnormal free light-chain ratio and the 'evolving type' (rising M-protein during the first years of follow-up), and the presence of more than

95% phenotypically aberrant plasma cells within the bone marrow compartment.

When the proportion of bone marrow plasma cells is consistent with MGUS, but the patient has a nephrotic syndrome, congestive heart failure, peripheral neuropathy, orthostatic hypotension or massive hepatomegaly, the most likely diagnosis is primary systemic amyloidosis resulting from the deposition of amyloidogenic light chains. On the other hand, in a patient with constitutional symptoms, lytic bone lesions, a small M-spike and less than 10% plasma cells in the bone marrow, the most likely diagnosis is metastatic carcinoma with coincidental MGUS.

Smouldering multiple myeloma

The term 'smouldering multiple myeloma' (SMM) was first defined by Kyle and Greipp as the presence of a serum M-protein (>30 g/L) and 10% or more plasma cells in the bone marrow in the absence of lytic bone lesions or clinical manifestations due to the monoclonal gammopathy. More recently, the IMWG considered that the term 'asymptomatic myeloma' could be more appropriate. This condition was defined as the presence of an M-protein (≥30 g/L) and/or 10% or greater bone marrow plasma cells in the absence of symptoms or organ or tissue impairment due to the monoclonal gammopathy. About 10% of patients diagnosed with MM have smouldering disease. This situation is clinically and biologically very close to that observed in MGUS. However, the plasma cell mass is much higher and most cases will eventually evolve into symptomatic MM.

The annual risk of progression to symptomatic disease is 10% per year for the first 5 years, and it significantly decreases thereafter, 5% per year during the following 5 years and only 1% per year from the 10th year. SMM is not a uniform entity, since it includes from indolent, low-risk forms (that behave as MGUS) to 'early myelomas' at high risk of developing symptomatic diseases. This heterogeneous outcome is supported by the identification of risk factors predicting progression to symptomatic MM. The Mayo Clinic Risk Classification proposes three different subgroups of SMM: group 1 with ≥3 g/dL of MC and ≥10% of plasma cells in bone marrow, in which the median time to progression (TTP) to symptomatic MM is 2 years; group 2 with <3 g/dL of MC and ≥10% bone marrow plasma cells M-protein with a median TTP of 8 years and group 3 with ≥3 g/dL of MC, but with <10% plasma cells bone marrow infiltration, translating into a median TTP of 19 years. The Spanish group has proposed a risk classification based on the percentage of PCs with aberrant phenotype (high risk if ≥95% of the total PCs are clonal) plus immunoparesis (decrease in one or two of the uninvolved immunoglobulins), with a median TTP of 23 months when the two risk factors were present, as compared with 73 months when only one risk factor was present and not reached when none of the risk factors was present. Other adverse risk factors include an abnormal serum free light chain (FLC), evolving levels of paraprotein, and cytogenetic abnormalities such as t(4;14), gain of 1q21 or hypodiploidy. Upon using either the Mayo or Spanish critera, the high-risk SMM patients have a 70% risk of progression at 3 years. Moreover, more sensitive criteria can be used to identify patients at 'ultra-high risk of progression' (70% transform at 2 years): presence of FLC ratio ≥100, circulating PCs ($>5 \times 10^6$/L), more than 60% of BMPCs, focal lesions on spinal magnetic resonance. These patients should probably be considered as early myeloma and they would be candidates for early intervention. These risk stratification classifications may allow a more individualized disease management.

Table 29.2 Myeloma-related organ or tissue impairment (end-organ damage) due to the plasma cell proliferative process.

- Increased serum calcium
- Renal insufficiency
- Anaemia: haemoglobin 20 g/L below the lowest normal limit
- Bone lesions: lytic lesions or osteoporosis with compression fractures (possibly confirmed by MRI or CT)
- Other: symptomatic hyperviscosity (rare), amyloidosis, recurrent bacterial infections (more than two episodes in 12 months), extramedullary plasmacytomas

Symptomatic multiple myeloma

The diagnosis of symptomatic MM requires the presence of an M-protein in serum and/or urine, increased plasma cells in the bone marrow or plasmacytoma, and related organ or tissue impairment (including bone lesions). The more common symptoms are fatigue from anaemia and bone pain due to the skeletal involvement. Some patients may have no symptoms, but they can have related organ or tissue impairment. Clinical and laboratory features may include anaemia, skeletal involvement (lytic lesions and/or severe osteoporosis, with or without compression fractures), renal failure, hypercalcaemia, recurrent bacterial infections, extramedullary plasmacytomas or associated amyloidosis (Table 29.2). The criteria agreed by the IMWG for the diagnosis of symptomatic MM are shown in Table 29.3.

Table 29.3 Symptomatic multiple myeloma*.

- M-protein in serum and/or urine
- Bone marrow (clonal) plasma cells or plasmacytoma†
- Related organ or tissue impairment (end-organ damage, including bone lesions)

*Some patients may have no symptoms, but have related organ or tissue impairment.
†If flow cytometry is performed, most plasma cells (>90%) will show a 'neoplastic' phenotype.

Table 29.4 Laboratory work-up for a patient with monoclonal gammopathy.

- History and physical examination
- Complete blood count and differential peripheral blood film
- Chemistry including calcium and creatinine
- Serum protein electrophoresis and immunofixation
- Nephelometric quantification of immunoglobulins
- 24-hour urine collection for electrophoresis and immunofixation
- Bone marrow aspirate (cytogenetics, immunophenotyping and plasma cell labelling index if available)
- Radiological skeletal bone survey: CT or MRI may be helpful
- β_2-Microglobulin, C-reactive protein and lactate dehydrogenase
- Measurement of free monoclonal light chains if available

Of note, no serum or urine M-protein values were included, since about 40% of patients with symptomatic MM have a serum M-protein level lower than 30 g/L and 3% have non-secretory myeloma. In the same sense, no minimal proportion of bone marrow plasma cells was required because about 5% of patients with well-documented symptomatic MM have less than 10% plasma cells in their bone marrow. Table 29.4 illustrates the laboratory work-up for patients with monoclonal gammopathies.

Other special forms of plasma cell dyscrasia

Plasma cell leukaemia

Plasma cell leukaemia was initially described by Kyle in 1974 as a plasma cell disorder characterized by a relative peripheral blood plasmacytosis of more than 20% of total nucleated cells, or an absolute number of plasma cells greater than 2×10^9/L. It is likely that lower levels of circulating plasma cells (i.e. $\geq$5% and/or $\geq 500 \times 10^9$/L has a similar prognostic meaning. There are two forms of plasma cell leukaemia: the *de novo* presentation in leukaemic phase, and secondary cases corresponding to already diagnosed MM that evolve into a leukaemic phase. The clinical course of plasma cell leukaemia is usually very aggressive and resistant to conventional treatment and therefore new agents should be urgently investigated in these patients. A consensus statement on the diagnostic criteria and treatment approach by the International Myeloma Working Group has been recently published.

Solitary plasmacytoma of bone

The existence of a solitary plasmacytoma has been recognized in up to 3% of patients with a plasma cell dyscrasia, usually on the vertebral column. The diagnostic criteria require the existence of a solitary plasma cell tumour in which the biopsy confirms plasma cell histology, a negative skeletal survey and absence of plasma cell infiltration in a random bone sample (<10%), as well as no evidence of anaemia, hypercalcaemia or renal impairment. Some groups suggest that patients in whom a paraprotein persists after the eradication of plasmacytoma with local treatment should undergo a review of the diagnosis. The treatment of choice is local radiotherapy, but about two-thirds of patients with solitary bone plasmacytoma develop MM at 10 years' follow-up, with a median time to progression of 2 years.

Extramedullary plasmacytoma

Extramedullary plasmacytoma is a plasma cell tumour that arises outside the bone marrow, most frequently in the upper respiratory tract (nose, paranasal sinuses, nasopharynx and tonsils). Other sites include parathyroid gland, orbit, lung, spleen, gastrointestinal tract, testes and skin. In most cases the lesion is unique, although the presence of more lesions (multiple plasmacytomas) has also been reported. Diagnosis is based on the detection of the plasma cell tumour in an extramedullary site, in the absence of bone marrow plasma cell infiltration, bone lytic lesions and other signs of MM (end-organ damage).

Non-secretory multiple myeloma

This specific type of MM requires particular attention, since it is very difficult to diagnose. The only way to make a definitive diagnosis is to demonstrate the presence of tissue infiltration (usually bone marrow) by cells with plasma cell morphology. However, plasma cell infiltration must be greater than 10% and clonality must be assessed by immunophenotyping (demonstration of cytoplasmic immunoglobulins with restricted light chain: positive production without excretion). In addition, the serum free light chains are abnormal and this is a most useful parameter for the follow-up. However, exceptional cases exist in which no monoclonal protein can be observed within the plasma cells. In these cases, it is mandatory to demonstrate clonality by the study of the rearrangement status of the immunoglobulin genes.

IgM multiple myeloma

This exceptional form of myeloma has been reported very rarely and must be distinguished from Waldenström macroglobulinaemia. The morphology and immunophenotype of the infiltrating cells will give the definitive diagnosis, as well as the existence of osteolytic lesions, which are absent in Waldenström macroglobulinaemia.

Osteosclerotic myeloma (POEMS syndrome)

POEMS syndrome is characterized by polyneuropathy, organomegaly, endocrinopathy, M-protein and skin changes. The clinical picture consists of a chronic inflammatory

demyelinating polyneuropathy, more motor than sensory, and osteosclerotic lesions. Hepatomegaly. hyperpigmentation, hypertrichosis, angiomatous lesions on the trunk, gynaecomastia, testicular atrophy and papilloedema or thrombocytosis may occur. The M-protein is commonly of IgA λ type and the bone marrow contains less than 5% plasma cells. Castleman disease can be associated and VEGF is universally increased. Biopsy of an osteosclerotic lesion may be necessary to confirm the diagnosis. Mandatory diagnostic criteria are the presence of an M-protein and polyneuropathy, and major criteria are the presence of osteosclerotic lesions, Castleman disease and elevated VEGF. At least the two major criteria plus one major finding and a minor clinical criteria (above mentioned) are necessary for diagnosis.

Amyloid disease is dealt with in Chapter 30.

Disease complications and their management

Figure 29.4 illustrates the clinical manifestations of multiple myeloma.

Bone involvement: assessment and treatment

Bone involvement is the most frequent clinical complication in patients with MM. About 70% of patients have lytic bone lesions, with or without osteoporosis, and another 20% have severe osteoporosis without lytic lesions. This frequency corresponds to conventional skeletal radiography assessment, a technique that is associated with low sensitivity (only demonstrating lytic disease when at least 30% of bone substance has been lost), particularly in some areas (ribs, sternum), low specificity (gas in colon) and long examination time. However, it has two major advantages: is widely available and remains as the international standard (for CRAB criteria). Newer imaging techniques have greater sensitivity compared with radiographic bone survey for detection of MM bone lesions (Figure 29.5). Computed tomography (CT) has the highest sensitivity for the detection of bone defects and with the whole-body low-dose modality the radiation exposure is much lower than with conventional CT, the scanning time is short and it may replace conventional X-ray in the near future. Magnetic resonance imaging has the highest resolution for soft tissue and bone marrow infiltration; it is particularly valuable for differentiation between benign and malignant fractures, but is inferior to CT for assessment of bone disease. Finally, positron emission tomography (PET) allows assessment of tumour metabolism and disease activity (versus inactive or necrosis), and may be of prognostic significance; however, it still requires great work on standardization.

These new techniques are already recommended to evaluate patients with SMM and solitary plasmacytoma, and although up until now the presence of one or more clear sites of osteolytic bone destruction seen on CT and/or PET-CT does fulfill the

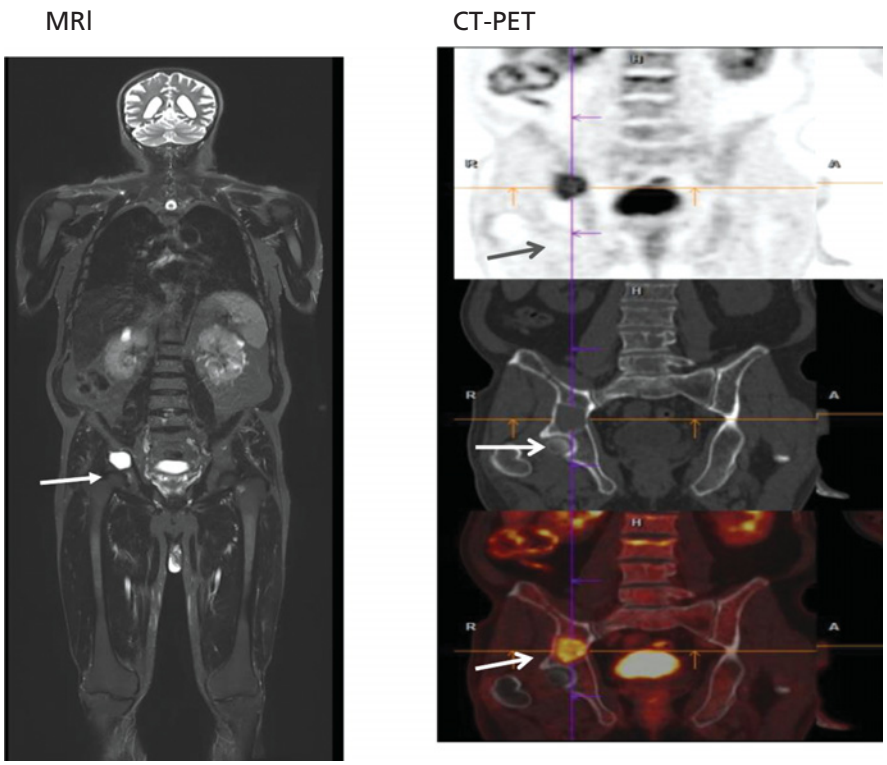


Figure 29.5 MRI and PET/CT in a myeloma patient. The arrows show a myeloma lesion in the lower part of the right iliac bone. The T2 weighted MRI shows a hyperintense image (left) and an active lesion at FDG PRT/CT is confirmed (right).

CRAB criteria, the situation will probably change soon. Nevertheless, great caution should be paid to avoid over-interpretation of equivocal or tiny lucencies seen only on CT or PET-CT.

From the clinical point of view, the skeletal involvement leads to bone pain and can result in pathological fractures. The pathophysiology of bone disease has been described above. Some patients develop pathological fractures of long bones and require orthopaedic surgery. In the event of extensive lesions, surgery can be followed by radiation therapy. On the other hand, prophylactic orthopaedic intervention must be considered in patients with large lytic lesions at high risk of fracture. It is important to consider that patients with severe back pain due to vertebral compression fractures can benefit from vertebroplasty or kyphoplasty. Spinal cord compression caused by a vertebral fracture is very rare in patients with MM. This complication is usually caused by a plasmacytoma arising from a vertebral body.

Between 15 and 20% of patients with MM have hypercalcaemia at the time of diagnosis. A common complication of hypercalcaemia is renal impairment caused by interstitial nephritis. Treatment of hypercalcaemia with hydration and bisphosphonates is a medical emergency. Zoledronic acid is the bisphosphonate of choice (quicker response and significantly longer time to recurrence compared with pamidronate).

The intravenous agents pamidronate and zoledronic acid are of clinical benefit in the treatment of bone disease in patients with MM. Pamidronate is administered at a monthly dose of 90 mg via a 2-hour intravenous infusion. Zoledronic acid, at a monthly dose of 4 mg, is at least as effective as pamidronate and has the advantage that it can be administered via a 15-min infusion. In patients with renal function impairment, the dose of zoledronic acid must be reduced to a maximum of 3 mg. It was suggested that bisphosphonates should be used indefinitely, once initiated. However, the appearance of severe late complications, such as osteonecrosis of the jaw, related to the duration of bisphosphonate exposure has resulted in a reconsideration of the initial recommendations. Osteonecrosis of the jaw is associated with the duration of bisphosphonate exposure, type of bisphosphonate (higher with zoledronic acid than with pamidronate) and history of recent dental procedure. The current recommendations for treatment with bisphosphonates in MM patients, based on consensus panels from both the IMWG and the ASCO, do not recommend the initial use of bisphosphonates for more than 2 years. In relapsed patients, treatment with bisphosphonates can be re-started and administered concomitantly with active therapy. Finally, in patients in whom the bone disease is a consequence of excess RANK-L activity, newer molecules such as denosumab might be of benefit. The pathogenesis of myeloma bone disease is summarized in Figure 29.6.

Renal failure

About 20% of patients with MM have a serum creatinine higher than 177 µmol/L (2 mg/dL) at diagnosis. The degree of renal failure is usually moderate, with a serum creatinine lower than 354 µmol/L (4 mg/dL). However, in some series up to 10% of patients with newly diagnosed MM have renal failure severe enough to require dialysis from the time of diagnosis. The main causes of renal failure in MM are: (i) light-chain excretion resulting in cast nephropathy (myeloma kidney) and (ii) glomerular deposition of immunoglobulin (light-chain amyloidosis or immunoglobulin deposition disease). In myeloma kidney, the typical feature consists of the presence of myeloma casts, mainly composed of light chains, in the distal tubules and

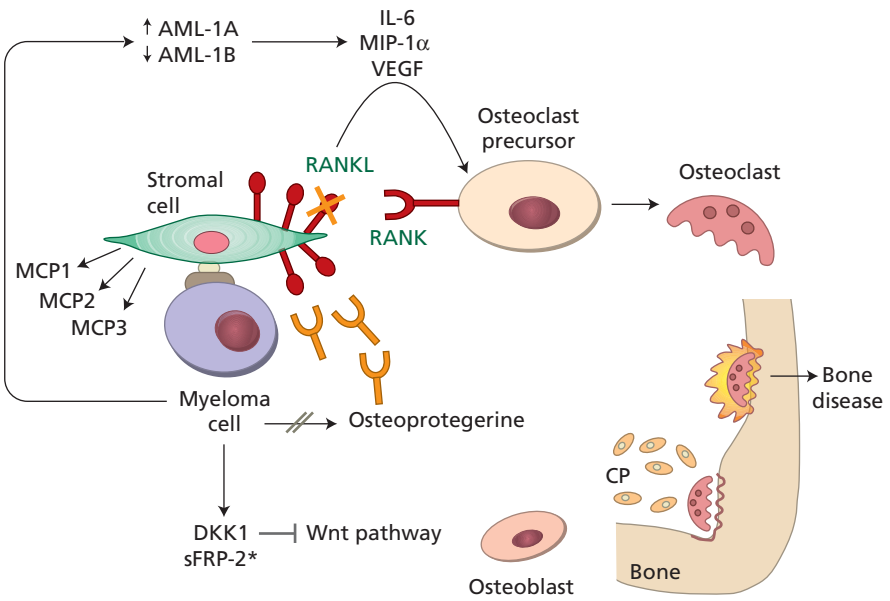


Figure 29.6 Pathogenesis of bone disease. See text for definition of abbreviations.

collecting ducts. There is a correlation between the degree of cast formation and the severity of renal failure. Light-chain tissue deposition usually consists of glomerular deposits of immunoglobulins, resulting in nephrotic syndrome. The amyloid deposits are fibrillar structures of light chains showing positive Congo red staining. In light-chain deposition disease, the deposit of light-chain immunoglobulins is non-fibrillar (Congo red negative). In contrast with amyloidosis, the light chain is usually of the κ type. The characteristic clinical feature is a nephrotic syndrome, but renal function can rapidly deteriorate, resembling glomerulonephritis (see also Chapter 30).

The median survival of patients with MM and renal insufficiency is less than 1 year. However, the prognosis mainly depends on the reversibility of renal function. Thus, the median survival of patients with reversible renal failure is similar to that of patients with normal renal function, whereas patients with non-reversible renal failure have a median survival of less than 6 months. The factors associated with renal function recovery include serum creatinine lower than 354 μmol/L (4 mg/dL), 24-hour urinary protein excretion lower than 1 g, and serum calcium higher than 2.875 mmol/L (11.5 mg/dL).

Vincristine, doxorubicin (Adriamycin) and high-dose dexamethasone (VAD), or cyclophosphamide and dexamethasone, or even dexamethasone alone, in very frail patients, appear to be better approaches than melphalan-containing regimens because of both lower myelosuppression and quicker action. The novel drugs introduced for myeloma treatment are of great value in patients with renal failure. Taking into account that the action of bortezomib is very quick, it is probably an ideal agent for rapidly decreasing light chains in order to prevent the development of irreversible renal failure by avoiding further tubular light-chain damage. In a retrospective series of 24 patients with relapsed/refractory MM and dialysis-dependent renal failure, the overall response rate (RR) was 75%, with 30% complete remissions (CR) or near-CR. Recent studies have confirmed the benefit of bortezomib-based therapies in patients with newly diagnosed myeloma and renal failure. The association of lenalidomide and dexamethasone could also be a good treatment option for patients with renal failure. However, the dose of lenalidomide must be adjusted to the degree of renal failure according to the guidelines for the use of lenalidomide in patients with renal function impairment.

With regard to the use of high-dose therapy/autologous stem cell transplantation (SCT) in patients with MM and renal failure, the largest experience comes from the Arkansas group, with a reversibility of renal failure of 43%, but higher morbidity and mortality (6% and 13% after a single or tandem transplant, respectively) than in patients with normal renal function. Chemoresistant disease, low serum albumin and older age were associated with a poorer outcome. In any event, the dose of melphalan must be reduced to 140 mg/m^2. In patients with no overt myeloma and low plasma cell mass in whom renal function impairment is due to glomerular light-chain deposition (light-chain deposition disease), the likelihood of response is higher than that in MM because of the low plasma cell mass at the time of transplantation. In this situation there is no need for tumour reduction with induction chemotherapy before stem cell mobilization and high-dose therapy.

Theoretically, the removal of nephrotoxic light chains with plasma exchange could avoid further renal failure and hopefully prevent irreversible renal failure. The Mayo Clinic group, in a small controlled trial, compared chemotherapy with chemotherapy plus plasma exchange and found only a trend in favour of the group including plasma exchange. Similarly, in a large randomized trial there was no conclusive evidence that plasma exchange improved the outcome of patients with MM and acute renal failure. When excluding the patients who die in this early period, the median survival of patients with MM and non-reversible renal failure needing chronic dialysis is almost 2 years and 30% of them survive for more than 3 years. Thus, long-term dialysis is a worthwhile supportive measure for patients with MM and end-stage renal failure. The use of high cut-off dialysis filters is very promising and hopefully prospective ongoing studies will confirm its benefit.

Anaemia and bone marrow failure

Approximately 35% of patients with newly diagnosed MM have a haemoglobin level lower than 90 g/L. In addition, severe anaemia is a frequent complication later in the course of the disease due to disease progression. Anaemia is associated with a significant loss in quality of life and poor prognosis. The main causes of anaemia in MM are bone marrow replacement by plasma cells, relative erythropoietin deficiency, renal insufficiency and chemotherapy with cytotoxic agents.

Severe granulocytopenia and thrombocytopenia at the time of diagnosis are unusual. About 10% of patients have a platelet count of less than 100×10^9/L, but platelet counts lower than 20×10^9/L with risk of severe bleeding are very unusual. The development of an unexplained pancytopenia in patients previously treated with alkylating agents, particularly melphalan, is suspicious of myelodysplasia.

A number of trials have shown the beneficial effect of recombinant human erythropoietins and darbepoetin-α in the treatment of myeloma-associated anaemia. The response to erythropoietin is associated with a significant improvement in quality of life. An important aspect is that the most significant improvement in quality of life is reached when the haemoglobin increases from 110 to 120 g/L, but levels above 140 g/L should be avoided due to its association with a higher risk of thrombosis. Thus, the goal should be maintenance of the haemoglobin around 120 g/L, with careful dose titration in order to achieve a good quality of life, while minimizing severe complications such as thrombotic events. The major cause of erythropoietin failure is iron deficiency. Iron repletion should be indicated when there is evidence of functional iron deficiency measured by an

increased soluble transferrin receptor. It seems that the best iron supplemental therapy is the administration of iron saccharate. Treatment with granulocyte colony-stimulating factor (G-CSF) may be required to treat chemotherapy-induced severe granulocytopenia. Patients treated with lenalidomide may require G-CSF therapy.

Infection

Infectious complications are the major cause of morbidity and mortality in patients with MM. The highest risk of infection is observed during the first 2 months of starting therapy, in patients with severe chemotherapy-induced granulocytopenia and in those with relapsed and refractory disease. The main cause of infection in MM is the impaired antibody production, leading to a decrease in the uninvolved immunoglobulins. Other important causes include chemotherapy-induced granulocytopenia, renal function impairment and glucocorticoid treatment, particularly high-dose dexamethasone. Most infections in newly diagnosed patients and during the first cycles of chemotherapy are caused by *Streptococcus pneumoniae*, *Staphylococcus aureus* and *Haemophilus influenzae*, while in patients with renal failure, as well as in those with relapsed and/or refractory advanced disease, more than 90% of the infectious episodes are caused by Gram-negative bacilli or *Staph. aureus*.

An infectious episode in a patient with MM should be managed as a potentially serious complication requiring immediate therapy. In case of suspected severe infection and before the identification of the causal agent, treatment against encapsulated bacteria and Gram-negative microorganisms should be initiated. Although prophylaxis of infection in patients with MM is a controversial issue, some general guidelines can be offered. Intravenous immunoglobulin prophylaxis is not recommended. Pneumococcal vaccination is recommended, particularly in patients with IgG myeloma with high-serum M-protein levels, which are usually associated with very low levels of uninvolved immunoglobulins. Antibiotic prophylaxis is likely of benefit within the first 2 months of initiation of therapy, especially in patients at high risk of infection (recent past history of serious infections, such as recurrent pneumonia, or renal failure). Patients treated with bortezomib should receive prophylaxis against varicella-zoster infections.

Nervous system involvement

Spinal cord compression from a plasmacytoma, which occurs in about 10% of patients, is the most frequent and serious neurological complication in MM. The dorsal spine is the most common site of involvement, followed by the lumbar region. The clinical picture of spinal cord compression consists of back pain and paraparesis. Although it can evolve for several days or even a few weeks, the onset can be abrupt, resulting in severe paraparesis or paraplegia in a few hours. The picture is usually accompanied by a high sensitivity level. Lumbar involvement can cause a cauda equina syndrome (low back pain with radicular distribution and leg weakness).

Spinal cord compression is an emergency requiring immediate medical action. When suspected, urgent magnetic resonance imaging (MRI) should be performed. If confirmed, treatment with high-dose dexamethasone must be started immediately at a loading dose of 100 mg, followed by 25 mg every 6 hours, followed by progressive tapering. Simultaneous local radiation therapy should be started as soon as possible. If the spinal cord compression is caused by a vertebral collapse or by spinal instability rather than a plasmacytoma (which is very rare), urgent surgical decompression followed by fixation of a prothesis of bone graft or methacrylate is required.

Nerve root compression can cause back pain that follows a radicular metameric distribution with or without radicular sensory involvement. Treatment with radiation therapy is helpful, allowing time for the action of systemic chemotherapy. Clinically significant peripheral neuropathy (PN) is very uncommon in newly diagnosed patients with MM. Leptomeningeal involvement in the central nervous system (CNS) with detection of plasma cells in the cerebrospinal fluid (CSF) is exceedingly rare. The frequency of CNS involvement is roughly 1%, the most frequent presenting features being paraparesis, symptoms from increased intracranial pressure, cranial nerves palsies and confusion. The CSF examination frequently shows plasma cells of plasmablastic morphology, as well as an increased protein concentration with positive immunofixation for the myeloma protein. Unfortunately, despite active treatment measures such as intrathecal therapy (methotrexate, cytarabine, glucocorticoids), cranial or even craniospinal irradiation, the prognosis is extremely poor with a median survival of 3 months from the diagnosis of CNS involvement.

Prognostic factors

The outcome for patients with MM is highly heterogeneous, with survival duration ranging from a few months to more than 10 years. This heterogeneity relates to: (i) specific characteristics of the tumour itself, (ii) host factors and (iii) a series of factors resulting from the interaction between the tumour clone and the host, which mainly reflect tumour burden and disease complications (Table 29.5).

Host factors

The favourable influence of a good performance status (ECOG $\geq$ 2) and young age (<65–70 years) is well established. In contrast, neither sex nor race has prognostic influence. Immune surveillance, by T and natural killer (NK) cells, plays a relevant role in most malignancies, and low numbers of

Table 29.5 Prognostic factors in MM.

Prognostic factors	Tumour related	Host related	Tumour burden
Essential	Cytogenetics/FISH: t(4;14), t(14;16), *17p deletions, 1q gains, 1p deletions*	Age Performance status	International Staging System (β_2-microglobulin/albumin)
Additional	Hypodiploidy LDH Plasma cell proliferation Immunophenotyping markers		
New and promising	DNA copy number abnormalities by SNP arrays	Immune status	Circulating clonotypic plasma cells sRANK-L

mature NK and CD4 cells have been reported in advanced-stage MM. In addition, patients who develop expanded T-cell clones (CD8+CD57+CD28−), which can recognize autologous idiotypic immunoglobulin structures, display favourable outcome.

Malignant clone factors

The second cohort of prognostic factors are those that reflect specific characteristics of myelomatous plasma cells and include morphology, immunophenotyping, cytogenetics, oncogenes, multidrug resistance and proliferative activity of plasma cells. Immature/plasmablastic morphology is associated with poor outcome. With regard to antigen expression, we have observed that downregulation of CD117 (c-Kit) and CD56, and expression of CD28 and CD19, are associated with poor prognosis. As occurs with acute leukaemia, cytogenetic abnormalities are the most important prognostic markers for MM. Nowadays, cytogenetic evaluation is mandatory in all patients with newly diagnosed MM, and should always include FISH in purified PCs or in combination with immunofluorescent detection of light-chain-restricted PCs (cIg-FISH). Among *IGH* translocations, patients with t(4;14) treated with either conventional or intensive chemotherapy display shorter event-free survival and overall survival. However, recent analysis supports the notion that patients with t(4;14) make up a heterogeneous group. Thus, the French group has discriminated a subgroup of these patients (approximately 45%) with both low β2-microglobulin and high haemoglobin levels at diagnosis, which displays prolonged survival after tandem transplant and benefits from high-dose therapy. The association of t(14;16) with prognosis has not been well supported because of its low frequency. According to Mayo Clinic data, t(14;16) was linked to poor outcome in the context of conventional chemotherapy. However, controversial results have been reported using autologous transplant: the IFM group did not confirm the poor prognostic value of t(14;16) in patients receiving a tandem-autologous transplantation approach, whereas a recent study from the MRC showed a significantly shorter survival among patients with t(14;16) treated with autologous transplant. In contrast to t(4;14) and t(14;16), the presence of t(11;14) was shown to have either a favourable or an irrelevant influence on prognosis. Initial studies indicated that monosomy 13/del(13) was associated with poor outcome; however, analyses based on large series of MM patients have revealed that the presence of monosomy 13/del(13) as a unique abnormality, without concomitant *IGH* abnormalities, is not indicative of unfavourable prognosis. In contrast, del(17p) (with deletion of the *TP53* gene) is one of the most, if not the most, adverse prognostic factor in MM. It seems that the higher the proportion of PCs bearing 17p deletions, the shorter the survival. Several studies have shown that 1q gains have a significant and independent poor prognostic factor; in addition, the IFM group has reported that 1p deletions are a major independent prognostic factor. Moreover, the presence of complex, as well as non-hyperdiploid karyotypes also predict treatment failure. In contrast, hyperdiploid tumours with multiple trisomies involving chromosomes 3, 5, 7, 9, 11, 15, 19 and 21 tend to have a favourable prognosis. In line with this latter observation, we have shown that patients with hyperdiploid DNA cell content (defined by flow cytometry) have a favourable outcome. Finally, the proliferative activity of the malignant plasma cells, as assessed either by the labelling index with bromodeoxyuridine or by flow cytometry, is one of the most important prognostic markers.

One of the aims of cytogenetic classification is to define a high-risk group that should be managed differently from standard-risk patients. According to IMWG recommendation, the term 'high-risk MM' should include those patients with at least one of the following features: deletion of 17p, t(4;14) or t(14;16), detected by FISH analysis. Hypodiploidy defined by karyotyping is also used by the Mayo Clinic to define high-risk patients. Recently, a prognostic model in MM based on the frequently associated genetic lesions has been proposed. Accordingly, three genetic risk groups were defined: a favourable risk group with no adverse FISH lesions, an intermediate group with one adverse genetic lesion – t(4;14), t(14;16) or t(14;20); 17p deletion or 1q gain – and a high-risk group with more than one adverse genetic lesion.

Tumour burden and disease complications

A high proportion of plasma cells in the bone marrow, diffuse bone marrow infiltration and the presence of circulating plasma cells reflect a high tumour burden, but their prognostic influence is modest. Similarly, the impact of skeletal lesions, evaluated by radiography or bone resorption markers, is not clear. In contrast, disease complications such as anaemia, thrombocytopenia and particularly renal insufficiency have a major influence. Nevertheless, the most important factor is the concentration of β2-microglobulin, which increases as a result of both growth of tumour burden and deterioration in renal function. The higher the value, the worse the prognosis. Nevertheless, β2-microglobulin is not helpful for disease monitoring. C-reactive protein is a surrogate marker for IL-6 (a major plasma cell growth factor), which also correlates with outcome. In contrast to hypoalbuminaemia, neither the amount of paraprotein nor its isotype influences prognosis.

Only a few of the prognostic factors have real independent value. A summary of the most important is shown below:

1 Two host factors (*age* and *performance status*) reflect the ability of the patient to tolerate chemotherapy

2 Two intrinsic characteristics of the malignant clone (*cytogenetics* and *plasma cell proliferative activity*)

3 One biochemical marker that reflects tumour burden (β_2-*microglobulin*).

The International Staging System (ISS), derived from more than 11,000 patients, has shown that β2-microglobulin and albumin are the best combination of easily available markers for discriminating prognostic subgroups: stage I (β2-microglobulin <3.5 mg/dL, albumin > 3.5 mg/dL); stage III (β2-microglobulin >5.5 mg/dL); stage II (the rest). This substitutes for the Durie and Salmon classification, which affords little prognostic information. In the next few years, improved staging systems using cytogenetics and S-phase analysis should be developed for use by reference centres and eventually for all patients with myeloma.

Response to therapy as a prognostic factor

In addition to all the variables that can be measured at diagnosis and which have already been mentioned, response to front-line therapy represents one of the most important prognostic factors in most haematological malignancies. In the case of MM there is still open debate about the importance of the depth of response on patient outcome. The reason for this controversy is probably that, historically, until the introduction of high-dose chemotherapy, CR was extremely rare and the only available comparison was between responding patients (achieving partial or minor responses) and non-responding patients, with the former having a better outcome. A recently published meta-analysis, including 21 reports with independent datasets for which outcome data were reported, has shown a very significant association between the maximal response obtained after treatment and the two main long-term-outcome parameters, i.e. overall survival (OS) and time to progression-related events. However, it is clear that more sensitive techniques (e.g. multiparametric immunophenotyping, molecular analysis of bone marrow plasma cells by PCR, serum free light-chain test and imaging techniques such as MRI or PET to detect myeloma activity outside the bone marrow) are needed to evaluate response in MM and these may contribute to a better assessment of the impact of response in final outcome. Both the Spanish and UK groups have shown that when transplanted and elderly MM patients show undetectable residual myelomatous plasma cells by immunophenotyping, progression-free survival (PFS) and OS are significantly longer, and this parameter is significantly more powerful than negative immunofixation.

Criteria of response

Response criteria were initially developed by the Chronic Leukaemia and Myeloma Task Force (CLMTF) in 1968. The main response parameter was a minimum 50% reduction in the M-protein. In 1972 the Southwest Oncology Group (SWOG) defined partial response (PR) as a reduction of at least 75% in the calculated serum paraprotein synthetic rate and/or a decrease of at least 90% in urinary light-chain urine protein excretion sustained for at least 2 months. The Medical Research Council (MRC) introduced the concept of plateau phase, defined as a period of stability after chemotherapy in which tumour progression does not occur, despite the presence of measurable disease. The minimum period of stability required for definition of plateau phase was 3 months. Since CRs were rarely observed with the classical conventional dose chemotherapy, neither the CLMTF nor the SWOG response criteria included a definition of CR. In addition there were no definitions for relapse and progression.

EBMT criteria for response, relapse and progression

With the introduction of high-dose therapy/SCT, the M-protein disappears in a significant number of patients, a fact that is associated with a significant prolongation of survival. In this context, the European Group for Blood and Marrow Transplantation (EBMT) developed new criteria defining CR (negative immunofixation in serum and urine and less than 5% bone marrow plasma cells), PR (≥50% reduction in serum M-component and ≥90% in 24-hour urine Bence–Jones proteinuria) and minimal response (MR) (25–49% decrease in M-component), as well as criteria for relapse (reappearance of M-protein by immunofixation in patient who had achieved CR, and progression from PR or MR). Any type of response should be maintained for a minimum of 6 weeks. These criteria have been shown to be useful and reproducible in both transplant and non-transplant series, as well as in prospective clinical trials.

Table 29.6 International uniform response criteria for multiple myeloma (IMWG response criteria).

Response category*	Criteria
Complete remission (CR)	Negative immunofixation (serum and urine) <5% bone marrow plasma cells No soft tissue plasmacytomas
Stringent CR	As above, *plus* Normal free light-chain ratio Absence of clonal plasma cells†
Very good partial response	≥90% decrease in serum M-protein Urine M-protein <100 mg per 24 hours
Partial response	≥50% decrease in serum M-protein ≥90% decrease in urine M-protein or <200 mg per 24 hours ≥50% decrease in soft tissue plasmacytomas

*All response categories require two consecutive measurements made at any time.
†Determined by immunohistochemistry or immunofluorescence.

Uniform response criteria

The IMWG has expanded the EBMT criteria by adding the categories of stringent CR (sCR) and very good partial response (VGPR) (Table 29.6). Patients with negative immunofixation in serum and urine and with a normal free light-chain ratio are considered in sCR. The free light-chain measurement has also been included for the evaluation of response in patients with non-secretory and oligosecretory disease. VGPR requires a decrease in the M-protein size of at least 90%. In addition, time to event, duration of response, clinical relapse and time to alternative therapy are emphasized as critical end points. More recently, the IMWG has considered the possibility of also adding the parameters 'CR by immunophenotyping' and 'CR by molecular techniques', as well as the reintroduction of minor responses for treatment evaluation of relapse/refractory patients.

Treatment

In this section we focus on the treatment of the malignant clone, since the management of disease complications (anaemia, renal insufficiency, bone disease) has been discussed above.

Melphalan–prednisone (MP) was introduced for the treatment of MM in the late 1960s. In the subsequent 30 years, treatment improvements remained stagnant, since more complex chemotherapy combinations (e.g. VCMP, VBAD, VAD) only led to small increases in the overall RR, but without differences in survival, as assessed in a large meta-analysis that included over 6000 patients. The next step forward was the use of high-dose melphalan followed by stem cell support (autologous SCT) for young myeloma patients, which resulted in a significant improvement in disease-free survival and OS. However, for elderly patients, MP remained the standard of care. From 2000, there was a revolution in the treatment of MM with the availability of new agents with distinct mechanisms of action: the immunomodulatory drugs thalidomide and lenalidomide (Revlimid) and the proteasome inhibitor bortezomib (Velcade).

In this section we first discuss the treatment of newly diagnosed patients stratified according to age (above or below 65–70 years), which categorizes the patients as transplant or non-transplant candidates, and then analyse the options for relapse/refractory patients, as well as emerging novel agents.

Should all myeloma patients be treated?

Currently, only myeloma patients with symptomatic disease (defined by CRAB criteria) should be treated, patients with smouldering myeloma are not treated until they develop symptomatic disease. Attempts at early intervention in SMM patients with alkylating agents, bisphosphonates, antagonists of the receptor of interleukin 1β or thalidomide failed to show a significant benefit; however, none of these studies discriminated between high- and standard- or low-risk SMM patients. By contrast, the Spanish group has conducted a Phase III randomized trial focusing on high-risk smouldering myeloma patients, comparing early treatment with lenalidomide–dexamethasone induction therapy followed by lenalidomide maintenance therapy versus observation. The results showed that the experimental arm was associated with a significant delay in progression to symptomatic myeloma (3 years after study entry, 77% of patients in Group A and 30% in Group B were progression-free; hazard ratio, 5.59; $P < 0.001$). This delay translated into a significant overall survival benefit (the proportions of patients alive at 3 years was 94% and 80%, respectively; hazard ratio, 3.24; $P = 0.03$). Although, these data suggest a benefit for early intervention in high-risk SMM patients, other trials are needed before this becomes a new standard of care. Moreover, ongoing efforts are trying to better define the high-risk population.

Treatment of newly diagnosed transplant candidate patients

Currently, treatment in this setting usually includes three to six cycles of induction therapy, followed by autologous SCT and the possibility of consolidation and maintenance.

Induction

The VAD combination has long been the gold standard as a preparatory regimen for young newly diagnosed MM patients

Table 29.7 Response to induction treatment in transplant candidate patients: results from Phase III front-line trials*.

Regimen	Patients	PR or better (%)	CR + n-CR (%)	Study
TD versus D	470	63 versus 46	7 versus 2.6	Rajkumar *et al*. (2008)
TD versus VAD	200	76 versus 52	10 versus 8	Cavo *et al*. (2005)
TAD versus VAD	400	–	35 versus 13	Macro *et al*. (ASH 2006)
TVAD versus VAD	230	81 versus 66	–	Zervas *et al*. (2007)
CTD versus CVAD	1161	87 versus 75	13 versus 8	Morgan *et al*. (ASH 2012)
BD versus VAD	482	80 versus 63	21 versus 8	Harousseau *et al*. (2010)
BTD versus TD	256	93 versus 74	36 versus 9	Cavo *et al*. (2010)
LD_{high} versus LD_{low} †	445	82 versus 71	4 versus 2	Rajkumar *et al*. (2010)
LD versus D	198	85 versus 51	22 versus 4	Zonder *et al*. (ASH 2007)
BTD versus TD	257		35 versus 14	Rosiñol *et al*. (2012)
PAD versus VAD	827		7 versus 2	Sonneveld *et al*. (2012)

*Response after autologous SCT (CR + n-CR): TAD versus VAD (16% versus 11%); CTD versus CVAD (51% versus 39%); BD versus VAD (35% versus 24%); BTD versus TD (57% versus 28%), BTD versus TD (46% versus 24%), PAD versus VAD (21% versus 5%), CTD versus CVAD (33% versus 25%)
*High-dose dexamethasone (three pulses), low-dose dexamethasone (one pulse).
† A, Adriamycin; B, bortezomib; D, dexamethasone; L, lenalidomide; T, thalidomide; V, vincristine; CR, complete remission; n-CR, near complete remission; ASH, American Society of Hematology (Annual Meeting).

who are candidates for autologous SCT, with PR rates ranging from 52 to 63% and CR rates from 3 to 13%. However, novel drug combinations are superior to VAD-like regimens for decreasing tumour burden before transplantation. As shown in Table 29.7, three randomized trials have compared thalidomide (T)-based regimens (TD or TAD or TVAD) versus either high-dose dexamethasone or VAD as initial therapy in transplant-eligible patients. In all studies, thalidomide combinations were superior to conventional induction treatment, although the response rate (PR or greater) obtained with thalidomide plus dexamethasone (63%) was lower than that achieved with TAD or TVAD (80%, with CR rates usually <10%). The MRC group has compared cyclophosphamide (C) plus thalidomide and dexamethasone with CVAD as an induction regimen before transplantation; the thalidomide arm was significantly superior (RR 87% versus 75%; CR 20% versus 12%). In studies evaluating bortezomib (B) combination therapy, data from a French randomized trial that compared bortezomib plus dexamethasone with VAD show the superiority of bortezomib plus dexamethasone, both before and after transplantation (Table 29.7). The Italian group has shown the superiority of bortezomib, thalidomide and dexamethasone over thalidomide and dexamethasone (Table 29.7). The high efficacy of bortezomib-based regimens as induction treatment is consistent with several pilot studies using either bortezomib plus dexamethasone alone or in combination with doxorubicin (Adriamycin/Doxil) or thalidomide, with RRs usually over 80% and CR rates of 18–32%. The Spanish group has also shown that BTD is superior in CR rate to TD both pre- and post-transplant and also in PFS. With regard to lenalidomide, two large randomized studies have shown that the majority of patients (>85%) respond to lenalidomide plus dexamethasone induction, but a minimum of four to six cycles would probably be required to achieve a substantial number of CRs. Thalidomide or bortezomib combinations did not affect stem cell collection or granulocyte and platelet recovery after transplantation. For lenalidomide, three recent reports indicate a decrease in CD34-positive cells collected and recommended harvesting early in the course of induction with lenalidomide and/or using cyclophosphamide along with G-CSF. It is evident that the pre-transplant induction regimen should be a triple combination based on bortezomib and dexamethasone. No data are available to draw conclusions regarding the superiority of one combination, such as BTD, BLD, BCD or BAD over the other.

Autologous stem cell transplantation

High-dose therapy (usually melphalan 200 mg/m^2) followed by autologous SCT prolonged OS compared with standard-dose therapy in prospective randomized trials conducted by the French (IFM) and English (MRC) groups and has provided evidence for more than 10-year survivorship in at least a subset of patients. Nevertheless, although the SWOG 9321 study in the USA, the French MAG91 study and the Spanish PETHEMA-94 trial confirmed the benefit of autologous SCT in terms of RR and event-free survival (EFS), they did not find superiority in terms of survival compared with standard-dose therapy. These discrepancies can be partly explained by differences in study design, differences in the conditioning regimens and, particularly, differences in the intensity and duration of the chemotherapy arm (the dose of alkylating agents and steroids was higher in the SWOG and Spanish trials, which may explain why OS for

conventionally treated patients was longer in these two studies compared with the IFM and MRC trials). Despite these discrepancies, high-dose therapy is currently considered the standard of care for younger patients with MM, mainly based on the benefit for RR and EFS.

In the setting of novel agents, it is also important to define whether autologous SCT enhances the RRs obtained with these new induction regimens. As mentioned above, studies based on bortezomib combinations, particularly those using BTD or PAD, have shown that the CR rate was improved following autologous SCT (Table 29.8), suggesting that induction with novel agents and autologous SCT are complementary rather than alternative treatment approaches. Nevertheless, the benefit in terms of EFS and OS remains to be seen. Data on lenalidomide are very encouraging, though still scanty. However, some investigators argue that this approach may be challenged by the optimal results obtained with 'long-term' treatment with novel combinations (i.e. lenalidomide plus dexamethasone).

With regard to tandem autologous SCT, its use will decrease since according to the French and Italian experience, only patients achieving less than a VGPR with the first transplant benefit from the second; nevertheless, recent data suggest that tandem transplant may be valuable in patients with high-risk cytogenetics. Second transplant at relapse may be increasingly used, providing that the duration of the response to the first transplant has lasted for more than 2–3 years.

Consolidation and maintenance

The concept of consolidation with short-term therapy consisting of two or three full-dose cycles in order to further decrease the tumour mass is most promising and there are international trials investigating the role of consolidation after ASCT.

The next step in the sequence of treatment is maintenance. Interferon and/or corticosteroids have shown little benefit and have been abandoned. The availability of novel agents (particularly thalidomide and lenalidomide, which are available in oral formulations) has transformed the concept of maintenance in an attempt to prolong the duration of responses after transplantation. Six randomized trials have investigated the role of maintenance with thalidomide. In all six there was a benefit in PFS (6 months prolongation according to a meta-analysis), but only in three was OS prolonged. This raises an important concern about whether the continuous use of novel agents may induce more resistant relapses with shorter survival after relapse in some studies. Two large randomized studies, by the IFM and the CALGB groups, of lenalidomide maintenance versus placebo have shown a significant benefit in PFS and in the CALGB trial also a significant prolongation of OS. However, there was an initial concern on the occurrence of second primary malignancies with lenalidomide in both trials and a short survival after relapse in the French study. A recent Italian trial also showed the benefit of lenalidmide maintenance. Regarding bortezomib maintenance, in two trials there was a benefit in PFS, but in OS only in one. Several prospective international trials on the role of maintenance are ongoing.

Allogeneic stem cell transplantation

Allogeneic SCT remains the only curative therapeutic approach in MM. However, it is associated with a high TRM (up to 30–50%) and high morbidity, mainly due to chronic graft-versus-host disease (GVHD). Accordingly, it should be used in carefully defined situations and preferably within the context of clinical trials. In order to decrease TRM, different reduced-intensity conditioning regimens (allo-RIC), mainly based on fludarabine and melphalan or fludarabine plus radiotherapy (2 Gy) have been introduced. The TRM decreases to 15 to 25%, but this is associated with a higher incidence of relapses. In a prospective randomized trial, the French group compared double autologous SCT with autologous SCT followed by allo-RIC among patients displaying poor prognostic features (high β_2-microglobulin and monosomy 13). The results of double autologous SCT or autologous SCT followed by allo-RIC were similar. In contrast, the Italian group, using a similar approach, has described an improvement in terms of OS among patients receiving autologous SCT followed by allo-RIC compared with double autologous SCT. The Spanish group also reported a comparison between double autologous SCT and autologous SCT followed by allo-RIC in patients failing to achieve at least near CR after a first autologous SCT. Although there was a higher increase in CR rate and a trend towards a longer PFS in favour of allo-RIC, there was a statistical difference in EFS and OS. The EBMT has reported the updated results of allo-RIC versus ASCT

Table 29.8 Result of randomized trials comparing autologous SCT with chemotherapy.

	Patients	CR (%)	EFS (months)	OS (months)
IFM90 (Attal *et al*. 1996)	200	22/5	28/18	57/44
MRC 03 (Child *et al*. 2003)	401	44/8	31/19	54/42
PETHEMA 95 (Bladé *et al*. 2005)	185	30/11	42/33	61/66
US-Intergroup (Barlogie *et al*. 2006)	516	11/11	25/21	58/53

CR, complete remission; EFS, event-free survival; OS, overall survival.

after a median follow-up of 7 years, showing a significant benefit in PFS and OS in favour of allo-RIC. Differences in patient characteristics, GVHD prophylaxis and conditioning regimens could contribute to these discrepant results. Moreover, unfortunately, a high proportion of patients developed extramedullary relapses without bone marrow involvement, indicating that although the disease may be under control in the bone marrow milieu, extramedullary spread may occur. In any event, about 10 to 15% of patients undergoing an allogeneic transplant are cured.

With regard to the use of allogeneic SCT as rescue therapy, a prerequisite is to obtain a CR or VGPR before the transplant, since most patients with active disease will not benefit from this procedure. Once again these transplants should be performed by experienced groups and within clinical trials. Donor lymphocyte infusions given for relapsed myeloma following allogeneic transplantation induce responses in 30–50% of patients, but unfortunately the long-term efficacy is limited.

Treatment of newly diagnosed elderly and non-transplant candidate patients

Melphalan + prednisone (MP) has been the gold standard for over 40 years; however, the scenario has completely changed with the introduction of novel agents such as thalidomide or bortezomib, and lenalidomide.

Six randomized trials have compared thalidomide (T) + MP (MPT) versus MP, showing significantly higher RR in the MPT arm (59% versus 37%) (CR 10% versus 2.5%), as well as longer PFS/TTP (prolongation in PFS of 5.4 months). However, only in the three studies, MPT treatment was associated with a significant prolongation in OS (median of 6 months benefit). Based on these data MPT has been approved as a standard of care. The toxicity associated with the high dose of thalidomide, used in some of these trials, may contribute to explain the survival discrepancies. Data from the MRC myeloma IX trial shows that the combination of cyclophosphamide + thalidomide + dexamethasone (CTD) is superior to MP in terms of RR (82% versus 49%), but not in OS. MP has also been compared with thalidomide + dexamethasone (TD) and although the RR was higher in the experimental arm, the OS was shorter. It should be noted that in this last trial, TD treatment was associated with a higher rate of early discontinuations due to toxicity, and higher mortality, particularly during the first year.

Lenalidomide (Len) has also been combined with MP. A randomized trial comparing MP versus Len + MP, using Len either only as part of the induction or also as maintenance, showed a significantly longer PFS for the maintenance approach (31 versus 14 and 12 months, respectively), but no significant differences in OS. A large clinical trial, including 1600 patients, has compared Len-dex (low dose dexamethasone 40 mg weekly) until progression versus Len-dex fixed time (18 cycles) versus MPT (9 cycles). Initial results show a significant advantage for continuous Len-dex treatment, both in terms of PFS (25.5 versus 20.7 versus 21.2 months, respectively) and OS (59.4%, 55.7% and 51.4% OS at 4 years, $P = 0.01$). Based on these data continuous LEN-dex could become a new standard (without alkylator) for newly diagnosed non-transplant candidate patients.

The proteasome inhibitor bortezomib (B) has been tested in combination with MP in a pilot study conducted by the Spanish group; the positive results were confirmed in a large randomized study, which compared BMP versus MP (9 cycles in each arm). The RR for BMP versus MP were 71% versus 35%, with 30% versus 4% CR. BMP treatment was associated with a longer TTP (24 versus 16.6 months) and OS at 3 years of 72% versus 59%. The results were updated after more than 5 years follow-up and confirmed a prolongation of 13 months in OS for BMP. This combination has been approved as a standard of care by EMEA and FDA. In an attempt to optimize the treatment of elderly untreated MM patients with VMP, the Spanish and Italian Myeloma groups (GEM/PETHEMA and GIMEMA) activated two randomized trials, exploring bortezomib only once weekly in instead of the standard twice weekly schedule. Results have indicated that the tolerability is increased substantially, and the efficacy is maintained with the reduced-dose bortezomib schedule (probably due to the better tolerability with fewer treatment discontinuations). Thus, grade 3/4 PN was only 5–7% with the reduced dose VMP regimen in the two studies. In line with this reduction in the frequency of PN, the rate of treatment discontinuations was low in both studies (8% and 10%). In addition, the French group has reported that when bortezomib is given by subcutaneous route of administration instead of the conventional intravenous route, the rate of grade 3/4 PN drops from 16% to 6%. New proteasome inhibitors such as carfilzomib or ixazomib (also called MLN9708, a twin brother of bortezomb in oral formulation) are being investigated in combination with either MP or Len-dex, with encouraging results. Bendamustine plus prednisone has been compared with MP and the former was associated with longer PFS (18 versus 11 months), but no longer OS.

Several trials have explored the value of maintenance treatment in the elderly population. Thalidomide was investigated in three studies and although some of them showed some benefit in PFS (ranging from 2 to 7 months), only one had benefit in OS and, accordingly, this approach has been abandoned. As mentioned above, continuous treatment with lenalidomde, both in the MPR and Len-dex trial have been associated with a significant prolongation in PFS (around 18 months benefit), which translated into longer OS in the second, but not in the first trial. As far as bortezomib maintenance is concerned, the Spanish group has investigated the value of 3 years of maintenance with either Btz-Thal or Btz-Pred (one pulse every three months) after a short course of 6 induction cycles with B-MP or B-Thal-P, and although this approach resulted in a long PFS of approximately 3 years, the overall benefit in OS remains to be seen. The Italian group has investigated the value of B-Thal

Table 29.9 Results from randomized Phase III trials in newly diagnosed elderly patients.

Study	Induction regimen	N	Maintenance regimen	CR (%)	ORR(%)	PFS/TTP (months)	Median OS (months or %)
Alkylator-based induction regimens							
Melphalan-based combinations							
Palumbo 2006, 2008	MPT vs.	129	T until DP vs.	16	76	22	48
	MP	126	None	2.4	48	15	45
Facon 2007	MPT vs.	125	None	13	76	28	52
	MP	196	None	2	35	18	33
Hulin 2009	MPT vs.	113	None	7	62	24	44
	MP	116	None	1	31	19	29
Wijermans 2010	MPT vs.	165	T until DP vs.	N/A	66	13	40
	MP	167	None	N/A	45	9	31
Beksac 2011	MPT vs.	60	None vs.	58	9	21	28
	MP	62	None	38	9	14	26
Waage 2010	MPT vs.	182	T until DP vs.	13	57	15	29
	MP	181	None	4	40	14	32
Palumbo 2012	MPR-R	153	R until DP	18	77	31	70% at 3 years
	MPR	152	Placebo until DP	13	67	15	62% at 3 years
	MP	154	Placebo until DP	5	49	12	66% at 3 years
San Miguel 2013	MPV	344	None	30	71	N/A	56
	MP	338	None	4	35	N/A	43
Mateos 2012	VMP vs.	130	Randomized to VT	20	80	37	60% at 5 years
	VTP	130	or VP up to 3 years	28	81	32	53% at 5 years
Palumbo 2010	VMP vs.	257	None	24	81	27	51% at 5 years
	VMPT	254	VT up to 2 years	38	89	37	61% at 5 years
Nievizsky 2011	VMP	167	V (five cycles) in all arms	32	69	N/A	N/A
Cyclophosphamide-based combinations							
Morgan 2011	CTDa	426	Randomized to T or	13	64	13	33
	MP	423	not until DP	2	33	12	31
Bendamustine-based combinations							
Ponisch 2006	BP	68	None	32	75	18	32
	MP	63	None	13	70	11	33
*Non-alkylator-based induction regimens**							
Ludwig 2009	TD	NS	Randomized to IFN	2	68	17	42
	MP	NS	or IFN-T	2	50	21	49
Nievizsky 2011	VD	168	V (five cycles) in all	24	73	NA	NA
	VTD	167	arms	36	80		
Rajkumar 2010	Len/Dex (RD)	214	None	5	81	19	75% at 2 years
	Len/dex (Rd)	208	None	4	70	25	87% at 2 years
Facon 2013	Len-dex (cont)	535	Continuous Len	15.1	75.1	25.5	59.4% at 4 years
	Len-dex (18 months)	541	None	14.2	73.4	20.7	55.7% at 4 years
	MPT	547	None	9.3	62.3	21.2	51.4% at 4 years

*Also VTP (from the randomized trial MVP versus VTP (Mateos 2012)).

maintenance after induction with B-MPT, as compared with 9 cycles of Btz-MP; continuous treatment was associated with a longer PFS (37 versus 27 months), as well as a longer OS (67% versus 55% at 4 years).

A controversial issue is whether there is any preference for one of the novel combinations. An individualized treatment approach would probably be valuable: 1) for patients with previous history or risk of deep venous thrombosis, BMP could be the preferable option; 2) in patients with pre-existing peripheral neuropathy, Len-MP or better Len-dex (continuous) should be the choice; 3) in patients with renal insufficiency, BMP is the most safe approach, although Len-dex can also be used; 4) in patients living long distances from hospital, oral treatment (MPT or even better Len-dex) would be preferable; 5) in patients with poor compliance with treatment, BzMP could be better; 6) in fragile patients probably Len-dex is better tolerated, finally, 7) if cost is an issue MPT or Cyclo-TD are the cheapest options.

In patients >75 years or with frail condition it would be recommended to use modified regimens, with a lower dose of thalidomide (100 mg); or bortezomib (weekly schedule or 1 mg/m2). One additional possibility in these patients is to substitute melphalan by cyclophosphamide (50 mg/day or 1g/21days), since this latter agent is less myelotoxic. Lenalidomide do not require dose reduction if combined with only low dose dexamethasone, but 15 mg should be used if combined with MP. In very elderly patients, special attention must be paid to infectious episodes (require active treatment) and renal function (appropriate hydration), particularly during the first three months of treatment when they are responsible for the high incidence of early deaths. Ongoing studies will establish optimal dosing and treatment schedules for different populations with the aim of maximizing efficacy and improving tolerability.

Treatment at relapse

It is important to separate the young (<65 years) from the elderly (>65 years) patients. In young patients relapsing after transplantation, we discriminate three cohorts of patients: early relapse (<1 year), intermediate relapse (1–3 years) and late relapse (>3 years). If the relapse occurs within the first year after transplantation, patients should be immediately considered high risk and, in order to overcome drug resistance, rescued with either a combination of all potentially effective drugs (e.g. BTD plus cisplatin, Adriamycin, cyclophosphamide and etoposide; or bortezomib, lenalidomide and dexamethasone) or alternating cycles of two combinations of non-cross-resistant agents (BCD alternating with lenalidomide/dexamethasone). If ≥PR is achieved, the patient could proceed to allogeneic SCT with RIC, although this must still be considered an investigational approach.

If relapse occurs 1–3 years after autologous SCT, we would favour rescue with novel agents used in a sequential (not simultaneous) manner, starting with one line of treatment (different from the one used as induction) and shifting to the second and subsequent lines only when disease progression occurs. Within this category of patients, those under 60 years old with an HLA-identical donor and a suboptimal response to the first line of treatment should be considered for allogeneic SCT with RIC. Finally, if relapse occurs more than 3 years after the first autologous SCT, an attractive possibility is re-induction with the initial treatment or other novel-agent combination, followed by a second autologous SCT.

In elderly patients, treatment decisions at relapse must take into account the general condition of the patient. Once the patient relapses, after up-front treatment, the durations of subsequent responses to rescue therapies are progressively shortened. Therefore, the current goal in relapsed MM is to optimize the efficacy of novel drugs through their most appropriate combinations, to establish optimal sequences of treatment and to promote active clinical research on experimental agents that have already shown promising activity in *in vitro* and animal models.

At first relapse, a regimen based on novel drugs (lenalidomide or bortezomib) and different from that used as induction should be instituted. Table 29.10 summarizes the most relevant results and combinations reported in relapsed patients. Bortezomib as a single agent has been shown to be significantly superior to high-dose dexamethasone in relapsed/refractory patients (RR 43% versus 18%; CR 16% versus 1%; TTP 6 versus 3 months). The addition of pegylated doxorubicin increased EFS to 9.3 months. With regard to lenalidomide, a large randomized trial has shown that lenalidomide plus dexamethasone is significantly superior to dexamethasone alone (RR 60% versus 20%; CR 15% versus 2%; TTP 11.1 versus 4.7 months). The most widely used regimens at relapse include:

- lenalidomide/dexamethasone
- bortezomib/dexamethasone or bortezomib/liposomal doxorubicin with or without dexamethasone or bortezomib/cyclophosphamide/dexamethasone (Table 29.10).

At second or subsequent relapse, usually after the patient has already failed bortezomib and at least one immunomodulator, a clinical trial with experimental agents should be encouraged (see section on 'promising new drugs'). If the patient is not a candidate for active therapy, palliative treatment with oral cyclophosphamide (50 mg/day) and prednisone (30 mg on alternating days) can be considered.

Side-effects associated with novel agents

Because of the previous history of thalidomide, a major concern was its toxicity profile. The side-effects are dose related and the most common include constipation, weakness, somnolence and neuropathy. Peripheral neuropathy is a common adverse event with thalidomide therapy and often limits the dose and duration of treatment. The use of combination therapy has raised concerns about an increased risk of DVT.

Table 29.10 Relapse/refractory patients: response to thalidomide, lenalidomide and bortezomib; single agents and combinations.

Drug/combination	No. of patients	Response rate		Study
		PR or better (%)	CR (%)	
Thalidomide-based combinations				
T monotherapy (200–800 mg)	169	24	2	Singhal (*N Engl J Med* 1999)
T monotherapy (meta-analysis)	1629	28	1.6	Glasmacher (*Br J Haematol* 2005)
TD	>400	42–58	3–13	Several*
TD vs. placebo-D	116	65 vs. 28	NR	Fermand *et al.* (ASH 2006)
TCD	>200	56–76	5–20	Several†
Lenalidomide-based combinations				
L monotherapy	104	14	4	Richardson (*Blood* 2006)
LD vs. D	175 vs. 176	60 vs. 22	15 vs. 2	Weber and Dimopoulos (*N Engl J Med* 2007)
LAD	69	87	23	Knop (ASH 2007)
Pegylated liposomal doxorubicin + VDL	62	75	29 (CR + n-CR)	Baz (*Ann Oncol* 2007)
LCD	21	65	NR	Morgan (*Br J Haematol* 2007)
Bortezomib-based combinations				
B monotherapy vs. D	669	43 vs. 18	16 vs. 1	Richardson (*N Engl J Med* 2005)
Pegylated liposomal doxorubicin + B vs. B monotherapy	646	48 vs. 43	14 vs. 11	Orlowski (*J Clin Oncol* 2007)
BM	26	47	11	Berenson (*J Clin Oncol* 2006)
BM	21	68	34 (CR + n-CR)	Popat (ASH 2007)
BCP	37	88	40	Reece (*J Clin Oncol* 2008)
BCD	50	82	16	Kropff (*Br J Haematol* 2007)
BCD	47	75	31	Davies (*Haematologica* 2007)
Immunomodulators (thalidomide/lenalidomide) plus bortezomib				
BT ± D	85	55	16 (CR + n-CR)	Pineda-Román (*Leukemia* 2008)
BT + pegylated liposomal doxorubicin	21	56	22 (CR + n-CR)	Chanan-Khan (*Leukemia and Lymphoma* 2005)
BTAD	20	63	24 (CR + n-CR)	Hollmig (ASH 2006)
BTMP	30	67%	14 (CR), 7 (n-CR)	Palumbo (*Blood* 2007)
BTMD	53	60	11	Terpos (*Leukemia* 2008)
BL	27	79	33	Richardson (ASH 2007)

*Weber (*J Clin Oncol* 2003), Dimopoulos (*Ann Oncol* 2001) and Palumbo (*Haematologica* 2001).
†Kropff (*Haematol J* 2003), García-Sanz (*Leukemia* 2004) and Dimopoulos (*Haematol J* 2004).
A: Adriamycin, B: bortezomib, C: cyclophosphamide, D: dexamethasone, L: lenalidomide, M: melphalan, P: prednisone, T: thalidomide, V: vincristine, CR: complete remission, n-CR: near complete remission, ASH: American Society of Hematology (Annual Meeting).

Apparently the major risk of DVT occurs when tumour load is high and thalidomide is combined with chemotherapy, especially Adriamycin. Accordingly, in this setting, anticoagulant prophylaxis with low-molecular-weight heparin (LMWH) or aspirin is mandatory. Current data suggest that lenalidomide is better tolerated than thalidomide in several aspects: it does not usually produce clinically significant somnolence, constipation or neuropathy, although the incidence of myelosuppression is higher, mainly neutropenia (grade 3 in 17–30%) and thrombocytopenia, which are manageable with dose reduction and growth factor support. Similarly to thalidomide, lenalidomide is associated with a higher risk of DVT (5–25%) and the risk increases in patients with comorbidities, previous history of DVT, concomitant use of erythropoietin, high-dose dexamethasone, anthracyclines or high tumour mass. For these reasons, anticoagulant prophylaxis with LMWH or aspirin is mandatory. The most frequent toxicities of bortezomib include fatigue, gastrointestinal symptoms, cyclical thrombocytopenia and,

particularly, peripheral neuropathy. This latter side-effect, classified as grade 3 in 9–20% of patients, is the main reason for treatment discontinuation, and the early detection of peripheral neuropathy is most important in order to reduce the dose or frequency of injections. The use of a weekly schedule of bortezomib and particularly the subcutaneous administration has resulted in a significant decrease in peripheral neuropathy. Concerning panobinostat thrombocytopenic and gastrointestinal effects is the most relevant toxicity. Elotuzumab and daratumumab can produce infusion reactions.

Promising new drugs

Undoubtedly, the development of proteasome inhibitors and immunomodulatory drugs (IMiDs) and the success of these agents in multiple myeloma has revolutionized the way we think about tumour biology and treatment of this disease. However, as mentioned previously, most patients eventually become resistant to these drugs, indicating that novel agents and combination strategies are clearly needed in the setting of relapsed and refractory disease.

The third-generation IMiD, pomalidomide, in combination with dexamethasone has demonstrated substantial efficacy in patients with pretreated multiple myeloma, including those refractory to lenalidomide or lenalidomide + bortezomib. In a cohort of patients with relapsed multiple myeloma in which 62% of patients had received prior IMiDs, pomalidomide + dexamethasone produced a response rate of 65% (≥PR) and a PFS of 13 months. In patients with lenalidomide-refractory disease, approximately 25 to 35% responded to pomalidomide. This was confirmed in the Phase III randomized MM-003 trial, comparing pomalidomide plus low-dose dexamethasone versus high-dose dexamethasone in patients who had failed prior bortezomib and lenalidomide. The combination significantly improved the primary endpoint of median PFS (4 versus 1.9 months, HR 0.50; $P < .001$). In addition, a benefit in median OS was also observed for pomalidomide-dexamethasone compared to high-dose dexamethasone (13.1 versus 8.1 months, HR 0.72; $P = 0.009$). This data suggests pomalidomide/dexamethasone may become a new standard for patients who are refractory to lenalidomide and bortezomib. Triple combinations of pomalidomide with cyclophosphamide-prednisone or with bortezomib-dexamethasone are being investigated and apparently duplicate the PFS.

The second-generation proteasome inhibitor, carfilzomib, has also shown encouraging efficacy in heavily pretreated MM patients. Thus, a response rate of 50% (≥PR) with PFS of >8 months is achieved when used as single agent, with 16% responses in Btz-refractory patients. In a Phase II trial, including lenalidomide refractory patients, carfilzomib in combination with Len-Dex showed 69% ≥PR with a PFS of 11.8 months. Of note, the incidence of PN is very low (<3%). Combinations with pomalidomide, cyclophosphamide or panobinostat are also being tested. Moreover, three Phase III trials have been completed. The FOCUS trial has compared carfilzomib single agent with best supportive care in patients with very advanced disease showing similar outcome. The ENDEAVOR trial has compared carfilzomib + dexamethasone versus bortezomib + dexamethasone in relapsing patients; the first option achieved higher response rate (77% versus 63%) and longer PFS (18.7 versus 9.4 months). The ASPIRE study has investigated the addition of carfilzomib to the combination of lenalidomide and dexamethasone (CRd versus Rd). CRd was associated with longer PFS (26.3 versus 17.6 months) and OS (HR, 0.79). This CRd scheme has shown impressive preliminary results in newly diagnosed patients. Furthermore, a new oral proteasome inhibitor Ixazomib (MLN9708) is already in clinical trials. In relapse/refractory patients (most of them previously exposed to bortezomib) ≥PR was achieved in 15% of cases, with very low frequency of G3 PN. Based on these encouraging results Ixazomib is being studied in a Phase III trial in combination with Rd, not only in relapsed patients, but also in newly diagnosed patients in which this later combination has shown in a pilot study 92% ≥PR. Other oral proteasome inhibitors, such as oprozomib or marizomib are also at early phase of development.

Another drug currently under investigation in relapsed/refractory multiple myeloma is bendamustine, which is a hybrid between an alkylating agent and a purine analogue. As a single agent, it produced an ORR of 31% in patients with relapsed disease following high-dose therapy. Combinations of bendamustine with bortezomib, thalidomide or lenalidomide increase response rates substantially. Numerous ongoing clinical trials are investigating the role of bendamustine in patients with relapsed and/or refractory multiple myeloma.

Agents with novel mechanisms of action are also emerging, including monoclonal antibodies. The monoclonal antibody most advanced in clinical development for multiple myeloma is elotuzumab, a humanized IgG1 antibody targeting the CS1 glycoprotein. This cell surface glycoprotein is highly expressed in myeloma cells with little to no expression in normal tissues. The mechanism of action of elotuzumab is primarily through natural killer-cell-mediated ADCC. While elotuzumab monotherapy only elicited modest activity in patients with multiple myeloma, the addition of lenalidomide and low-dose dexamethasone resulted in an ORR of 82% in a Phase I trial of relapsed or refractory myeloma. The second type of monoclonal antibody under investigation is anti-CD38 (daratumumab, SAR650984). Results from phase I/II dose escalation studies have demonstrated activity in monotherapy, with 30 to 40% responses at the optimized doses, and only mild infusion reactions, which were well controlled with steroids. In the SIRIUS trial that included 112 heavily pre-treated and double refractory MM patients, daratumumab single agent induced 29% RR, with a DOR of 7.4 months and a PFS of 3.7 months. These positive results has prompted the investigation of anti-CD38 in combination with lenalidomide plus dexamethasone, with RR close to 90% in

lenalidomide sensitive patients and 50% in patients previously refractory to lenalidomide. Combinations with bortezomib are also under investigation.

The positive results of this combination in a phase II trial, with a PFS of 26.9 months, supported the activation of a large randomized trial (ELOQUENT) in relapsed/refractory patients after 1–3 prior lines of therapy comparing lenalidomide+dexamethasone plus/minus elotuzumab; the triplet combination was associatted wiht longer PFS (19.4 versus 14.9 months) but not differences in OS have been so far observed.

Deacetylase (DAC) inhibitors also have a novel mechanism of action in multiple myeloma. DAC enzymes remove the acetyl group from client proteins, including histones, p53, HIF-1α and Hsp90. Single-agent DAC inhibitors have demonstrated only modest activity and minor responses or disease stabilization in multiple myeloma. However, there is a clear rationale for the combination of DAC inhibitors with proteasome inhibitors. Two unique mechanisms exist to remove unfolded or misfolded proteins within a cell; the ubiquitin-proteasome cascade that degrades ubiquitinated proteins and the aggresome. Aggresomal degradation can be inhibited by histone deacetylase 6 (HDAC 6) inhibitors, as HDAC 6 binds ubiquitinated protein complexes and microtubule complexes, shuttling proteins into the aggresome for degradation. In preclinical studies, simultaneous targeting of proteasome activity and aggresome activity triggered substantial accumulation of misfolded proteins and synergistic multiple myeloma cytotoxicity. Additional *in vitro* and *in vivo* studies showed that the pan-DAC inhibitor panobinostat was highly synergistic with bortezomib-dexamethasone in multiple myeloma cell lines and murine models.

This led to the VANTAGE 088 trial investigating bortezomib in combination with the broad-acting HDAC inhibitor vorinostat or placebo in relapsed multiple myeloma. Although the combination significantly improved ORR (54% versus 41%; $P < 0.0001$), the PFS advantage was only 24 days compared to placebo (7.63 versus 6.83 months). The reasons for this are unclear and could be related to the adverse event profile or the omission of corticosteroids, which may play a role in the synergy of these combinations. In the Phase II PANORAMA 2 trial, the combination of panobinostat with bortezomib-dexamethasone produced a partial response in 35% of patients with relapsed and bortezomib-refractory multiple myeloma. An encouraging PFS of 4.9 months was also observed. The Phase III PANORAMA 1 trial of bortezomib/dexamethasone with or without panobinostat in relapsed/refractory multiple myeloma has recently be completed and the results show a superiority for PFS in the experimental arm. (PFS, 12 versus 8 months). The improvement in the outcomes appeared to be particularly of interest among patients who had been previously exposed to proteasome inhibitors and IMiDs (PFS, 12.5 versus 4.7 months).

Novel agents are also targeting signalling pathways important for multiple myeloma pathogenesis. One of the most important signalling pathways for tumour cell survival and proliferation is the PI3K/mTOR pathway. Several agents have been developed to target these pathways, including the Akt inhibitor perifosine and the mTOR inhibitors everolimus and temsirolimus. These agents have demonstrated modest efficacy as single agents and are being evaluated in combination with proteasome inhibitors and IMiDs (Table 29.3). It may be advantageous to consider simultaneous inhibition of more than one molecule along the PI3K/mTOR pathway, as targeting only the TORC1 complex with a rapamycin analogue can lead to compensatory upregulation of PI3K and TORC2 activation.

The RAS/MEK/ERK pathway is also important in multiple myeloma for tumour growth and differentiation. In an *in vivo* mouse model evaluating IMiD activity and resistance, decreased ERK activation was observed when IMiDs were administered. In contrast, activation of ERK and MEK signalling occurred when tumorus became resistant to IMiD therapy. The addition of a MEK inhibitor potentiated the activity of lenalidomide-dexamethasone and pomalidomide-dexamethasone in multiple myeloma cell lines, suggesting a rationale for combining ERK pathway inhibitors with IMiDs.

The kinase spindle protein inhibitor ARRY-520 is currently under investigation and demonstrating preliminary efficacy in multiple myeloma. Kinase spindle protein is required for multiple myeloma cell-cycle progression, and inhibition with ARRY-520 induces mitotic arrest and subsequent apoptosis. Phase II trial data demonstrated activity for this novel agent, including a ≥PR rate of 22% when ARRY-250 was combined with low-dose dexamethasone in patients refractory to bortezomib, lenalidomide and dexamethasone.

Selinexor, and oral selective inhibitor of Nuclear Export, that inhibit XPO 1 and as consequence activate tumour suppressor proteins and reduces oncoproteins is another promising drug. Finally, active research on immunotherapy, particularly targeting check point inhibitors (PD1, PDL1 …) is underway.

The evolution of therapy for multiple myeloma is far from over, with continued discovery of new drugs and progress in our understanding of myeloma cell biology and prognostic factors. The future success of multiple myeloma treatment depends on the use of rationally designed drug combinations. Myeloma needs to be approached as a heterogeneous disease with distinct disease subtypes. This will allow individualization of therapy with newer agents and potentially make multiple myeloma a chronic disease.

Acknowledgements

This work was supported by the Cooperative Research Thematic Network grants RD12/0036/0058 and RD12/0036/0046 of the Red de Cancer (Cancer Network of Excellence); Instituto de Salud Carlos III, Spain (FIS: PI060339; 06/1354; 02/0905;

01/0089/01-02; PI12/01293, PS09/01, PS09/01897/01370; G03/136); and Asociación Española Contra el Cáncer (GCB120981SAN), Spain.

Selected bibliography

Benboubker L, Dimopoulos MA, Dispenzieri A *et al.* (FIRST Trial Team) (2014) Lenalidomide and dexamethasone in transplant-ineligible patients with myeloma. *New England Journal of Medicine* **371**(10): 906–17.

Bladé J, Rosiñol L (2007) Complications of multiple myeloma. *Hematology/Oncology Clinics of North America* **21**: 1231–46.

Cavo M, Rajkumar SV, Palumbo A *et al.* (2011) International myeloma working group consensus approach to the treatment of multiple myeloma patients who are candidates for autologous stem cell transplantation. *Blood* **117**(23): 6063–73.

Dimopoulos M, Spencer A, Attal M *et al.* (2007) Lenalidomide plus dexamethasone for relapsed or refractory multiple myeloma. *New England Journal of Medicine* **357**: 2123–32.

Durie BGM, Harousseau JL, San Miguel JF *et al.* (2006) International uniform response criteria for multiple myeloma. *Leukemia* **20**: 1467–73.

Fernández de Larrea C, Kyle RA *et al.* (2013) Plasma cell leukemia: consensus statement on diagnostic requirements, response criteria and treatment recommendations by the International Myeloma Working Group. *Leukemia* **27**(4):780–91.

Lohr JG, Stojanov P, Carter SL *et al.* (2014) Widespread genetic heterogeneity in multiple myeloma: implications for targeted therapy. *Cancer Cell* **25**(1): 91–101.

Lokhorst H, Einsele H, Vesole D, Bruno B (2010) International Myeloma Working group consensus statement regarding the current status of allogeneic stem cell transplantation for multiple myeloma. *Journal of Clinical Oncology* **28**(29): 4521–30.

Ludwig H, Sonneveld P, Davies F *et al.* (2014) European perspective on multiple myeloma treatment strategies in 2014. *Oncologist* **19**(9): 829–44.

Mateos MV, Hernández MT, Giraldo P *et al.* (2013) Lenalidomide plus dexamethasone for high-risk smoldering multiple myeloma. *New England Journal of Medicine* **369**(5): 438–47.

Mateos MV, Ocio EM, Paiva B *et al.* (2015) Treatment for patients with newly diagnosed multiple myeloma in 2015. *Blood Review* [epub ahead of print] Available at: http://www.ncbi.nlm.nih.gov/pubmed/26094881. Accessed 5th August 2015.

Morgan GJ, Walker BA, Davies FE (2012) The genetic architecture of multiple myeloma. *Nature Reviews Cancer* **12**(5): 335–48.

Ocio EM, Richardson PG, Rajkumar SV *et al.* (2014) New drugs and novel mechanisms of action in multiple myeloma in 2013: a report from the International Myeloma Working Group (IMWG). *Leukemia* **28**(3): 525–42.

Paiva B, Puig N, García-Sanz R *et al.* (2015) Is this the time to introduce minimal residual disease in multiple myeloma clinical practice?. *Clinical Cancer Research* **21**(9): 2001–8.

Rajkumar SV, Dimopoulos MA, Palumbo A *et al.* (2014) International Myeloma Working Group updated criteria for the diagnosis of multiple myeloma. *Lancet Oncol* **15**(12): e538–e48.

San Miguel JF, Schlag R, Khuageva NK *et al.* (2008) Bortezomib plus melphalan and prednisone for initial treatment of multiple myeloma. *New England Journal of Medicine* **359**: 906–17.

San Miguel JF (2014) Multiple myeloma: a model for scientific and clinical progress. *Hematology* ASH Education Program **5**; 1–7.

Sonneveld P, Goldschmidt H, Rosinol L *et al.* (2013) Bortezomib-based versus nonbortezomib-based induction treatment before autologous stem-cell transplantation in patients with previously untreated multiple myeloma: a meta-analysis of phase III randomized, controlled trials. *Journal of Clinical Oncology* **31**(26): 3279–87.

CHAPTER 30

Amyloidosis

30

Simon DJ Gibbs and Philip N Hawkins

National Amyloidosis Centre, Royal Free and University College London Medical School, London, UK

Introduction

Amyloidosis is a disorder of protein folding in which normally soluble proteins are deposited in the extracellular space as insoluble fibrils that progressively disrupt tissue structure and function. Some 25 unrelated proteins can form amyloid *in vivo*, and the clinical amyloidosis syndromes are classified according to the fibril protein (Table 30.1).

Amyloid deposition is remarkable in its diversity; it can be systemic or localized, acquired or hereditary, life-threatening or merely incidental. Clinical consequences occur when accumulation of amyloid fibrils is sufficiently substantial to interfere with normal function. The pattern of organ involvement varies within and between fibril types and clinical phenotypes overlap greatly. In systemic amyloidosis virtually any tissue may be involved and the disease is often fatal, although prognosis has improved with increasingly effective treatments of many underlying conditions. Greater understanding of the pathogenesis of the disease, allowing improved subtyping and appropriate targeted therapies, with improved supportive care, including haemodialysis and solid organ transplantation in selected patients have also influenced the prognosis of this disorder. In localized amyloidosis, deposits are confined to a particular organ or tissue, but may have serious consequences such as haemorrhage. Localized amyloid deposition is also a pathological hallmark of uncertain significance in some important common diseases including Alzheimer's disease and type 2 diabetes mellitus.

This chapter describes the clinical features, differential diagnosis and management of systemic AL (monoclonal immunoglobulin light chain) amyloidosis, which is the most common and serious form of the disease.

Pathogenesis of amyloid

Proteins that can form amyloid can exist in two completely different three-dimensional conformations, the transformation involving massive re-folding of the normal form into one that can autoaggregate in a highly ordered manner to produce the characteristic predominantly β-sheet, rigid, non-branching fibrils of 10–15 nm in diameter and indeterminate length. Acquired biophysical properties common to all amyloid fibrils include insolubility in physiological solutions, relative resistance to proteolysis, and ability to bind Congo red dye in an ordered manner that gives the diagnostic green birefringence under cross-polarized light.

Amyloid deposition can occur in three circumstances:

1 When there is a sustained abnormally high concentration of certain normal proteins, such as serum amyloid A protein (SAA) in chronic inflammation or β_2-microglobulin in renal failure, which underlie susceptibility to AA and β_2-microglobulin amyloidosis, respectively.

2 When there is a normal concentration of a normal, but inherently amyloidogenic, protein over a prolonged period, such as transthyretin (TTR) in senile amyloidosis and β-protein in Alzheimer's disease.

3 When there is production of an acquired or inherited variant protein with an abnormal amyloidogenic structure, such as monoclonal immunoglobulin light chains in AL amyloidosis or some hereditary variants of TTR, lysozyme, apolipoprotein A-I and fibrinogen Aα chain.

Genetic and environmental factors that may influence susceptibility and timing of amyloid deposition are unclear, but once underway, amyloid deposition is unremitting so long as the supply of the precursor protein continues.

Postgraduate Haematology, Seventh Edition. Edited by A Victor Hoffbrand, Douglas R Higgs, David M Keeling and Atul B Mehta.

Table 30.1 Classification of amyloidosis.*

Type	Fibril precursor protein	Clinical syndrome
AA	Serum amyloid A protein	Systemic amyloidosis associated with acquired or hereditary chronic inflammatory diseases. Formerly known as secondary or reactive amyloidosis
AL	Monoclonal immunoglobulin light chains	Systemic amyloidosis associated with myeloma, monoclonal gammopathy, occult B-cell dyscrasia. Formerly known as primary amyloidosis
ATTR	Normal plasma transthyretin	Senile systemic amyloidosis with predominant cardiac involvement
ATTR	Genetic variants of transthyretin (e.g. ATTR Met30, Ala60, Ile122)	Familial amyloid polyneuropathy (FAP), with systemic amyloidosis and often prominent amyloid cardiomyopathy
$A\beta_2M$	β_2-Microglobulin	Dialysis-related amyloidosis (DRA) associated with renal failure and long-term dialysis. Predominantly musculoskeletal symptoms
Aβ	β-Protein precursor (and rare genetic variants)	Cerebrovascular and intracerebral plaque amyloid in Alzheimer's disease. Occasional familial cases
AApoAI	Genetic variants of apolipoprotein A-I (e.g. AApoAI Arg26, Arg60)	Autosomal dominant systemic amyloidosis. Predominantly non-neuropathic with prominent visceral involvement, especially nephropathy. Minor wild-type ApoAI amyloid deposits may occur in the aorta
AApoAII	Genetic variants of apolipoprotein A-II	Autosomal dominant systemic amyloidosis with predominant renal involvement
AFib	Genetic variants of fibrinogen α chain (e.g. AFib Val526)	Autosomal dominant systemic amyloidosis. Non-neuropathic usually with prominent nephropathy
ALys	Genetic variants of lysozyme (e.g. ALys His67)	Autosomal dominant systemic amyloidosis. Non-neuropathic with prominent renal and gastrointestinal involvement
ACys	Genetic variant of cystatin C (Gln68)	Hereditary cerebral haemorrhage with cerebral and systemic amyloidosis
AGel	Genetic variants of gelsolin (e.g. Asn187)	Autosomal dominant systemic amyloidosis. Predominant cranial nerve involvement with lattice corneal dystrophy
AIAPP	Islet amyloid polypeptide	Amyloid in islets of Langerhans in type 2 diabetes mellitus and insulinoma
ALECT2	Leucocyte chemotactic factor II	Systemic amyloidosis with predominant renal involvement

*Amyloid composed of peptide hormones, prion protein and unknown proteins not included.

Amyloid deposits consist mainly of amyloid fibrils, but also contain some common minor constituents, including certain glycosaminoglycans (GAGs) and the normal circulating plasma protein serum amyloid P component (SAP), as well as various other trace proteins. SAP binds in a specific calcium-dependent manner to all amyloid fibrils, but not their precursor proteins, which is the basis for the use of SAP scintigraphy to image and monitor amyloid deposits. Studies in knockout mice indicate that SAP contributes to amyloidogenesis.

Amyloid fibril-associated GAGs mainly comprise heparan and dermatan sulfates. Their universal presence, restricted heterogeneity and intimate relationship with the fibrils suggest they may also contribute to the development or stability of amyloid deposits, a possibility currently being investigated in a clinical trial of a GAG analogue in AA amyloidosis.

Many of the pathological effects of amyloid can reasonably be attributed to its substantial physical presence or precise location, for example within glomeruli or nerves. Aggregated amyloid proteins may also be directly cytotoxic, but they appear to evoke little or no local reaction in the tissues. The relationship between amyloid load and associated organ dysfunction differs greatly between locations and individuals, and the rate of new amyloid deposition may be an important factor.

Treatments that substantially reduce the supply of amyloidogenic precursor protein frequently result in stabilization or regression of existing amyloid deposits, and are often associated with preservation or improvement in the function of amyloidotic organs.

Systemic AL amyloidosis

Systemic AL ('primary') amyloidosis occurs in 1–2% of individuals with monoclonal B-cell dyscrasias. AL fibrils are derived

from monoclonal immunoglobulin light chains, which are unique in each patient. The uniqueness of an individual's light chain may explain the considerable heterogeneity in organ involvement and outcome in AL amyloidosis. Virtually any organ other than the brain may be directly affected, most commonly the kidneys, heart, liver and peripheral nerves. Early symptoms are non-specific, and routine screening techniques often fail to detect the subtle underlying monoclonal gammopathy, causing delay in diagnosis and detrimental disease progression. Awareness of AL amyloidosis is thus essential, with early recourse to more informative investigations such as serum free light chain (SFLC) analysis and target organ or screening tissue biopsies.

AL fibrils and monoclonal light chains

AL amyloid fibrils are derived from monoclonal immunoglobulin light chains and consist of the whole or part of the variable (V_L) domain with molecular mass between about 8 and 30 kDa. Monoclonal light chains are unique in structure in each individual and in their inherent propensity to form amyloid fibrils. Overall, about 1–2% of patients with monoclonal gammopathies eventually develop AL amyloidosis, but this is not possible to predict. The inherent 'amyloidogenicity' of certain monoclonal light chains has been demonstrated in an *in vivo* model in which purified Bence–Jones proteins were injected into mice. Light chains from patients with AL amyloidosis formed amyloid, whereas those from myeloma patients did not. AL fibrils are more commonly derived from λ than κ light chains, despite κ isotypes predominating among both normal immunoglobulins and monoclonal gammopathies. Some amyloidogenic light chains have distinctive amino acid replacements or insertions compared with non-amyloid monoclonal light chains, including replacement of hydrophilic framework residues by hydrophobic ones, changes that can promote aggregation and insolubility. Certain light-chain isotypes, notably $V_{\lambda VI}$, are especially amyloidogenic, and there is a degree of concordance between some isotypes and their tropism for being deposited as amyloid in particular organ systems. For example, the $V_{\lambda VI}$ isotype often presents with dominant renal involvement, whereas the $V_{\lambda II}$ isotype frequently involves the heart.

The plasma cell dyscrasia

Any B-cell dyscrasia, including multiple myeloma, Waldenström macroglobulinaemia and other malignant lymphomas/leukaemias, producing a monoclonal immunoglobulin may be complicated by AL amyloid deposition, but over 80% of patients with AL amyloidosis have a low-grade and otherwise 'benign' monoclonal gammopathy of undetermined significance (MGUS). Conversely, some minor and clinically insignificant amyloid deposits can be found histologically in up to 15% of patients with myeloma. The cytogenetic abnormalities that commonly occur in multiple myeloma and MGUS, such as 14q translocations and 13q deletion, have also been observed in AL amyloidosis, but their prognostic significance has not been fully elucidated.

Clinical features

AL amyloidosis accounts for 1 in 1500 deaths in the UK and occurs equally in men and women, equating to an incidence of about 8 per million persons per year. The median age at presentation is about 65 years, but it can occur in young adults and most likely remains underdiagnosed in the elderly.

Commonly, presenting features are non-specific, such as decreased exercise tolerance, fatigue, anorexia, weight loss and malaise. Proteinuric renal dysfunction and cardiac failure are common, and the combination is strongly suggestive of amyloidosis. The heart is affected in more than 50% of patients and in 30% a restrictive cardiomyopathy is a presenting feature. Rarer cardiac presentations include arrhythmias and angina, the latter sometimes due to coronary amyloid angiopathy. Dominant renal amyloid is the presenting feature in one-third of patients, typically presenting with nephrotic syndrome and/or renal impairment. Gut involvement may cause motility disturbances, which can also be secondary to autonomic neuropathy, and malabsorption, perforation, haemorrhage or obstruction. Macroglossia occurs in 5–10%, but is almost pathognomonic of AL amyloidosis (Figure 30.1). Hyposplenism is not

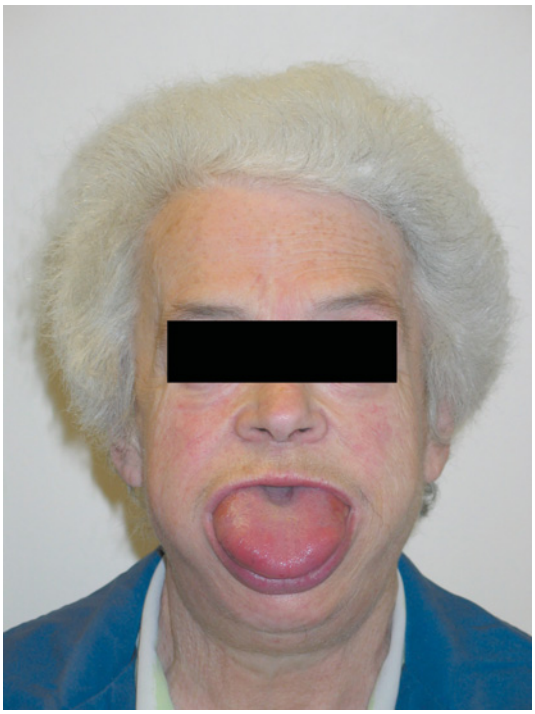


Figure 30.1 Macroglossia in AL amyloidosis.

infrequent, but is rarely documented. Painful sensory polyneuropathy with changes in pain and temperature sensation followed later by motor deficits occur in 10–20% of cases and carpal tunnel syndrome occurs in 20%. Autonomic neuropathy leading to impotence, orthostatic hypotension and gastrointestinal disturbances may occur alone or together with peripheral neuropathy, and has a poor prognosis. Involvement of dermal blood vessels is common and may cause purpura, most distinctively in a periorbital distribution ('raccoon eyes'). Direct skin involvement takes the form of papules, nodules and plaques, usually on the face and upper trunk. Articular amyloid is rare, but the symptoms can be severe and superficially mimic an inflammatory polyarthritis. Soft-tissue infiltration may occur, characteristically involving the submandibular region or the glenohumeral joints and surrounding tissues to produce the 'shoulder pad' sign, or lymph nodes themselves can be infiltrated. Thyroid infiltration with amyloid, occasionally resulting in hypothyroidism, has been reported. An uncommon, but serious manifestation peculiar to AL amyloidosis is an acquired bleeding diathesis that may be associated with deficiency of factor X and sometimes also factor IX.

Diagnosis and investigation of AL amyloidosis

Amyloid should be considered in the differential diagnosis of renal failure, nephrotic syndrome, restrictive cardiomyopathy, peripheral or autonomic neuropathy and hepatomegaly, but in the presence of any symptoms at all in patients known to have clonal B-cell dyscrasias. In practice, the diagnosis of amyloidosis is usually an unexpected finding following biopsy of an organ with disturbed function. An approach to diagnosis is outlined in Table 30.2, and comprising confirmation of amyloid, determination of fibril type, characterization of the underlying plasma cell dyscrasia, and evaluation of the extent, distribution and function of involved organs.

Confirming the presence of amyloid

Histology

Systemic amyloid deposits occur in blood vessels and small interstitial foci throughout the body, providing the basis for 'screening' biopsies of the GI tract and abdominal fat aspiration, which are both diagnostic in 50–80% of cases. The presence of amyloid can usefully be sought in bone marrow biopsies. Biopsy of a clinically affected organ is diagnostic in 95% of cases. Pink amorphous material on haematoxylin and eosin-stained tissue should raise suspicion of amyloid and prompt Congo red staining, which produces green birefringence under cross-polarized light and is the diagnostic gold standard (Figure 30.2). False-positive and false-negative interpretation of the Congo red stain is not rare, but can be minimized by using the alkaline-alcohol method (described by Puchtler and colleagues), fresh reagents, tissue sections of optimal 5–10 μm thickness, inclusion of positive control tissue and high-quality polarizing filters. Some Congo red techniques stain connective tissues quite strongly and produce white or very pale-green birefringence that cause diagnostic errors.

The differential diagnosis includes light chain deposition disease (LCDD), but these non-fibrillar deposits do not stain with Congo red.

Electron microscopy

Electron microscopy is helpful in confirming the presence of amyloid and excluding LCDD. A diagnosis of amyloidosis made through electron microscopy alone should be regarded with caution, since deposition of fibrillar material occurs in other pathological processes.

SAP scintigraphy

Radiolabelled SAP scintigraphy is a specific nuclear medicine imaging technique that demonstrates the presence and distribution of amyloid deposits in a quantitative manner. It was developed and is used routinely at the UK National Amyloidosis Centre. ^{123}I-labelled SAP localizes rapidly and specifically to amyloid deposits of all fibril types, in proportion to the amount of amyloid present, and is particularly effective for imaging amyloid in the liver, spleen, kidneys, bone marrow and adrenal glands. Bone uptake is very specific for AL amyloidosis. Serial SAP scans enable whole body and organ loads of amyloid to be tracked over time, for example following chemotherapy in AL amyloidosis (Figure 30.3). Unfortunately, amyloid in the moving heart and in small or diffuse hollow structures such as nerves, the gastrointestinal tract and the lungs is not adequately visualized by SAP scintigraphy.

Identifying fibril type

A monoclonal gammopathy may be incidental to the presence of amyloid and be gravely misleading. AL amyloidosis must be confirmed through immunohistochemical staining, proteomic analysis or by comprehensive exclusion of other types, including hereditary forms through DNA analysis.

Immunohistochemistry

Immunohistochemical staining of amyloid-containing tissue sections using a panel of antibodies against known amyloid proteins is the most accessible method for characterizing fibril type. However, even when positive and negative controls are used, and the specificity of staining has been sought through absorption with appropriate antigens, results are often not definitive in AL amyloidosis due to background immunoglobulin and failure of antibodies to bind to light chains in the amyloid conformation. In contrast, immunohistochemical staining can usually exclude AA and hereditary amyloidosis.

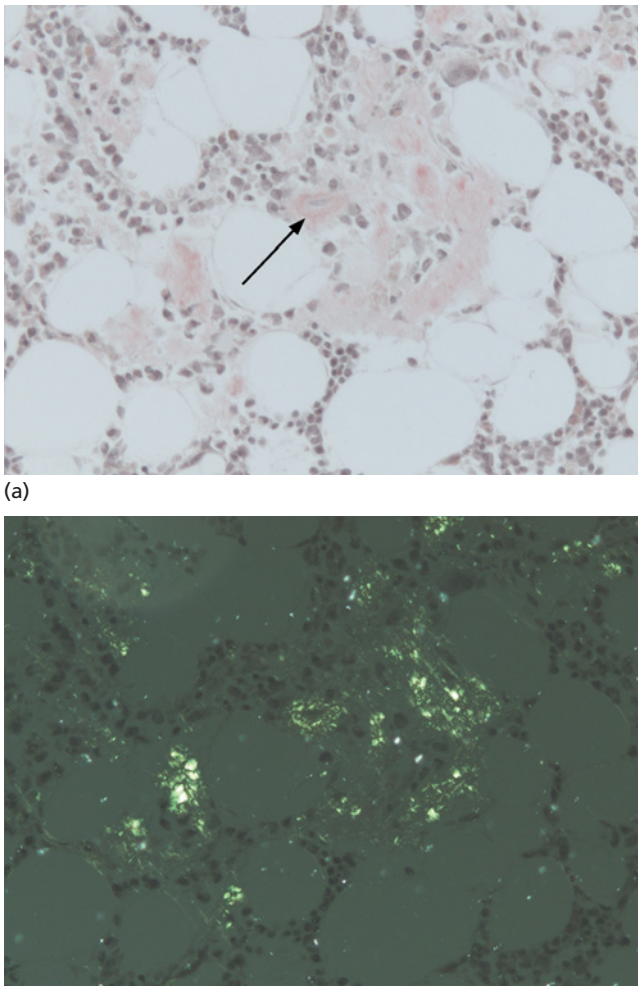
Proteomics: mass spectrometry

Proteomic identification of amyloid type through mass spectrometry, using amyloidotic material cut out from tissue

Table 30.2 Suggested approach to the investigation and monitoring of suspected AL amyloidosis.

	Confirmation of amyloid	Determination of amyloid type	Evaluation of organ involvement	Investigation of plasma cell dyscrasia	Monitoring
Pathology	Biopsy and Congo red histology of affected organ, screening tissue (e.g. rectum or fat aspirate) or any available specimen	Immunohistochemical staining of tissue sections with a panel of antibodies to amyloid fibril proteins (often not definitive in AL amyloidosis)	Biopsy of affected organ (but subsequent biopsies merely to determine the extent of amyloid involvement are not recommended)	Bone marrow aspirate and trephine biopsy with immunophenotyping	Follow-up biopsies usually not helpful in monitoring amyloid load
Haematology, biochemistry, immunology		Identification of a monoclonal gammopathy supports AL type, but may be an incidental finding	Serum creatinine and creatinine clearance, albumin, 24-hour urine protein Liver function tests Coagulation screen NT-pro-BNP, troponin, thyroid function tests	Full blood count, urea and electrolytes, creatinine, calcium Immunoglobulins Electrophoresis and immunofixation of serum and urine Quantitative serum free light chain assay	Quantitative serum free light chain assay (every month or two) NT-pro-BNP (3–6 monthly) Electrophoresis and immunofixation of serum and urine (3–6 monthly)
Imaging	SAP scintigraphy	SAP scintigraphy (evidence of marrow involvement is indicative of AL type)	Echocardiogram, ECG, SAP scintigraphy	Skeletal survey	SAP scintigraphy (6–12 monthly)
Other		DNA analysis for hereditary forms of amyloidosis Amyloid fibril protein sequencing	As otherwise indicated, e.g. nerve conduction studies, cardiac MRI		Serial assessment of organ function, e.g. liver and renal function tests, including 24-hour proteinuria estimations, echocardiography and other investigations as indicated

NT-pro-BNP, N-terminal pro-brain natriuretic peptide.



(a)

(b)

Figure 30.2 Appearance of amyloid in a bone marrow biopsy (×40, 6-μm section). (a) Congo red stain showing amorphous pink material in the interstitium and small blood vessel (arrow). (b) Same section under high-intensity cross-polarized light showing diagnostic apple-green birefringence.

sections by laser dissection, is increasingly being used in specialist amyloidosis practice with a substantial improvement in diagnostic yield. We currently regard immunohistochemistry and mass spectrometry as complementary and often mutually informative investigations at the National Amyloidosis Centre.

DNA analysis

Sequencing of the genes associated with hereditary amyloidosis should be performed routinely to exclude hereditary amyloidosis when the AL type cannot be confirmed with absolute certainty by immunohistochemistry and/or proteomics. Of patients attending the UK National Amyloidosis Centre, 5% have hereditary amyloidosis, those with TTR mutations usually presenting with neuropathy and/or cardiac amyloidosis and those with fibrinogen Aα chain mutations presenting with renal dysfunction. It is necessary to confirm that identified mutations are indeed the cause of the amyloid with fibril typing as described above.

Assessment of the plasma cell dyscrasia

A serum paraprotein or urinary free light chains are often not identified by electrophoresis and immunofixation. However, SFLC analysis (Freelite, The Binding Site, Birmingham, UK) demonstrates a clonal excess of SFLCs in well over 90% of cases and can be used to quantitatively monitor the effects of chemotherapy on the underlying plasma cell disorder.

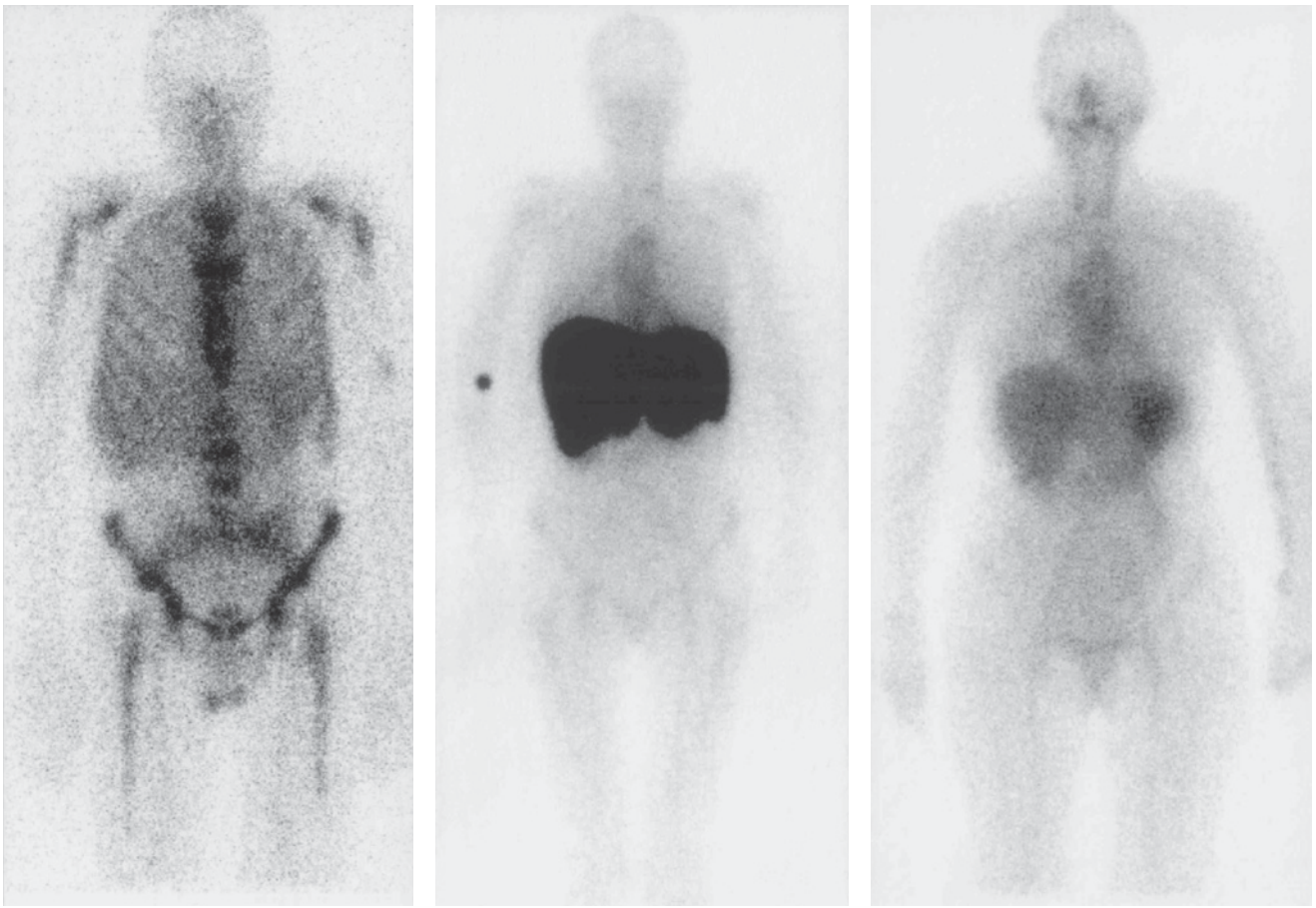
Bone marrow aspiration and trephine biopsy are required to characterize the underlying B-cell dyscrasia, which comprises a subtle plasma cell infiltrate in about 80% of cases. A skeletal X-ray survey should be performed to further exclude myeloma.

Assessment of organ involvement

International consensus criteria produced by Gertz and colleagues in 2005 define organ involvement in AL amyloidosis (Table 30.3). ECG and two-dimensional Doppler echocardiography are vital tools for evaluating cardiac involvement. In cardiac amyloidosis, ECG may demonstrate reduced QRS voltages in the limb leads and poor R-wave progression in the chest leads ('pseudo-infarct' pattern). Echocardiography typically reveals concentrically thickened and echogenic heart valves. Cardiac amyloidosis is a restrictive cardiomyopathy and diastolic dysfunction is easily missed and difficult to quantify. Cardiac magnetic resonance imaging (MRI) is emerging as a very useful tool for identifying cardiac amyloidosis and distinguishing it from other restrictive cardiomyopathies. The status of other organs can be assessed either through routine tests of organ function, such as liver function tests, serum creatinine and 24-hour urine measurements, or by specialist investigations as indicated, for example autonomic function and nerve conduction tests and high-resolution pulmonary computed tomography (CT).

Glomerular proteinuria (predominantly albuminuria) is present in about 90% of patients with renal involvement. Abnormalities in liver function tests are unusual and do not occur until liver amyloidosis is substantial, and are most commonly obstructive in nature. Right-sided heart failure due to amyloid cardiomyopathy may cause obstructive liver function tests in the absence of liver involvement by amyloid. Anaemia is uncommon unless amyloidosis is associated with myeloma, bleeding or chronic kidney disease. An abnormal clotting screen is relatively common. A prolonged thrombin time is the most frequent abnormality, but is generally of no clinical consequence. A prolonged prothrombin time due to acquired factor X deficiency is the most likely coagulation abnormality to be associated with clinically significant bleeding.

Elevation of N-terminal probrain natriuretic peptide (NT-pro-BNP) and cardiac troponin concentrations occur in a wide



(a) (b) (c)

Figure 30.3 Radiolabelled ^{123}I-SAP whole-body scintigraphy, anterior images. (a) The tracer has localized virtually exclusively to amyloid deposits in the bones in this particular patient, a distribution that is almost pathognomonic for amyloid of AL type. (b, c) Serial scintigraphs in a 56-year-old woman with AL amyloidosis. At presentation in 1998 (b) she had massive uptake in the spleen and liver, obscuring any renal signal. She underwent high-dose chemotherapy, and the follow-up scan in 2002 (c) shows that the deposits have regressed substantially.

variety of cardiac conditions and in chronic kidney disease. However, significant myocardial AL amyloidosis is excluded by NT-pro-BNP concentrations below 30 pmol/L. By contrast, cardiac troponin and NT-pro-BNP concentrations appear to be powerful predictors of prognosis and survival after chemotherapy. Elevation of cardiac troponin and/or NT-pro-BNP provide the basis for the powerful Mayo Clinic staging system in patients with newly diagnosed AL amyloidosis. Stage 1, with cardiac biomarkers below threshold values (cTnT <0.035 μg/L and NT-proBNP <332 ng/L) is associated with the lowest risk of death and a median survival of 26.4 months, stage 2, with either biomarker above the threshold values has a median survival of 10.5 months, and Stage 3, with both biomarkers above the threshold values, is associated with a median survival of only 3.5 months.

Differential diagnosis

Alternative diagnoses that frequently need to be considered include amyloidosis of non-AL types (principally AA and hereditary forms), localized forms of AL amyloidosis, and non-amyloid diseases such as LCDD and paraprotein-associated neuropathies, including POEMS syndrome and treatment-related toxicities.

Natural history

Systemic AL amyloidosis is a progressive disease with a very poor prognosis. Median survival without treatment in historical series is only 12–15 months. At diagnosis, 40% of patients have Mayo Stage III disease (elevated cardiac troponin and

Table 30.3 Non-invasive diagnostic criteria for amyloid-related major organ involvement.

Heart
Echocardiogram demonstrates mean wall thickness >12 mm, no other cardiac cause found
Kidney
Proteinuria >0.5 g per 24 hours, predominantly albumin
Liver
Total liver span >15 cm in the absence of heart failure or alkaline phosphatase >1.5 times the institutional upper limit of normal
Nerve
Peripheral: clinical; symmetric lower extremity sensorimotor peripheral neuropathy
Autonomic: gastric emptying disorder, pseudo-obstruction, voiding dysfunction not related to direct organ infiltration
Gastrointestinal tract
Direct biopsy verification with symptoms
Lung
Direct biopsy verification with symptoms
Interstitial radiographic pattern
Soft tissue
Tongue enlargement, clinical
Arthropathy
Claudication, presumed vascular amyloid
Skin
Myopathy by biopsy or psudohypertrophy
Lymph node (may be localized)
Carpal tunnel syndrome

Source: Gertz *et al*., 2005. Reproduced with permission of John Wiley & Sons Ltd.x

NT-pro-BNP), and median survival remains on the order of 6 months despite chemotherapy. However, amyloidosis is not, as previously believed, inexorably progressive and irreversible and it is now clear that amyloid deposits often gradually regress when the supply of amyloid fibril precursor protein is reduced.

Management

Specific therapies that inhibit formation of amyloid or enhance its clearance are not yet available, although are in development. Currently, the objective of treatment in AL amyloidosis is to suppress production of amyloidogenic monoclonal light chains with the aim that progression of the disease will be slowed, halted or reversed. However, many patients have advanced multisystem disease at diagnosis and tolerate chemotherapy poorly.

Quantitative measurements of SFLCs are usually the most effective means for guiding ongoing treatment in individual patients. Although it is generally desirable to suppress the underlying clonal disease as rapidly and as extensively as possible, reduction in the concentration of the amyloidogenic free light chain by just 50–75% can be sufficient to halt disease progression in many patients and confer substantial survival benefit, regardless of the chemotherapy used. However, more intensive suppression of the clonal disease, defined by 90% or greater suppression of free light production, is associated with the best outcomes. Although treatment of AL amyloidosis is derived from that of myeloma, it is tolerated much more poorly in amyloidosis, and serious treatment-related toxicity is frequent. Chemotherapy must be tailored to the individual patient, balancing the need for deep remission with the need to avoid undue toxicity.

Certain organs affected by amyloid tend to fare better than others, following treatment. Proteinuria and liver function often improve when the clonal disease is adequately suppressed, whereas cardiac disease, macroglossia and peripheral nerve function tend to improve extremely slowly if at all. International consensus criteria on the definitions of clonal and organ responses and progression were described by Gertz and colleagues in 2005.

Autologous peripheral blood stem cell transplantation

Use of high-dose melphalan therapy and autologous stem cell transplantation (SCT) in AL amyloidosis was first reported in 1996, and several series have reported clinical benefit in up to 80% of patients who survived the procedure. In the largest reported series, consisting of 394 patients, a complete clonal response occurred in 41%. Of those alive and in clonal response 1 year after autologous SCT, median survival has not been reached after 10 years of follow-up. However, treatment-related mortality (TRM) with autologous SCT has been consistently and substantially higher than in multiple myeloma, ranging from 5 to 43%, reflecting the compromised function of multiple organ systems by amyloid. Even stem cell mobilization has significant risks in AL amyloidosis. Causes of death include cardiac arrhythmias, intractable hypotension, multiple organ failure and gastrointestinal bleeding. Patient selection has since been refined and criteria for eligibility for autologous SCT developed, with the main exclusion criteria being overt cardiac involvement, more than two vital organ systems involved, a history of gastrointestinal bleeding and age over 65 years. Improvements in TRM have gradually resulted from better patient selection with improved supportive care and physician experience.

The exact role of SCT in AL amyloidosis remains controversial and its apparently good outcome may reflect selection of fitter good-prognosis patients. Regardless, because of its special problems in AL amyloidosis, it is recommended that such patients are treated in units with expertise in this particular

disease. In eligible patients, autologous SCT with melphalan 200 mg/m^2 is widely used as first-line therapy without induction chemotherapy except in those with coexisting myeloma. Lower doses of melphalan compromise the efficacy of transplantation. Stem cell mobilization is often achieved with the use of granulocyte colony-stimulating factor (G-CSF) alone, as the underlying plasma cell burden is usually relatively small. The French group compared autologous SCT with oral melphalan and dexamethasone in a randomized study and found no difference in response rates between the two arms, but a significantly higher TRM for the autologous SCT arm (24%) and a longer overall survival in the melphalan/dexamethasone arm, throwing into question the routine use of autologous SCT in AL amyloidosis. However, the trial has attracted some controversy, with higher than expected rates of TRM in autologous SCT compared with other case series and the use of lower doses of melphalan (140 mg/m^2) in some of the patients undergoing SCT, prompting calls for further study in this area.

Melphalan and dexamethasone

Since 1999, cyclic melphalan and dexamethasone has been widely used in Europe and has been found to be a well-tolerated and effective oral chemotherapy, with overall clonal response rates of 67%, including complete responses in 37%; amyloidotic organ responses have been reported in up to 50% of cases along with median overall survival of up to 5 years. However, these promising figures do not apply to patients with significant cardiac amyloidosis. Median time to haematological response is slow at 4.5 months, and the issue of stem cell depletion makes it unattractive for patients who may subsequently be eligible for autologous SCT. Renal dose adjustment is required and the long-term leukaemogenesis of this regimen has yet to be established.

Cyclophosphamide, thalidomide and dexamethasone

Thalidomide has been used extensively in combination therapies for plasma-cell disorders, but doses above 200 mg daily are poorly tolerated in AL amyloidosis. Amyloid-related cardiac involvement or peripheral neuropathy results in increased sensitivity to the adverse effects of thalidomide, such as fluid retention, fatigue, constipation and neuropathy. Lower doses of single-agent thalidomide reduce toxicity, but also markedly reduce responses rates. The use of thalidomide as maintenance treatment in AL amyloidosis has not yet been established. Addition of dexamethasone to intermediate-dose thalidomide improves the response rate to about 50%, with a median time to response of less than 4 months, but is associated with significant toxicity in 60% of patients. The cyclophosphamide, thalidomide and dexamethasone (CTD) regimen has been used extensively in the treatment of AL amyloidosis in the UK since 2002 and incorporates the addition of weekly oral cyclophosphamide 500 mg to low-dose thalidomide (100–200 mg) and two pulses of dexamethasone over a 21-day cycle. A risk-adapted approach using even lower doses of both thalidomide (50–100 mg) and dexamethasone in high-risk patients maintains good clonal response rates, while decreasing the incidence of severe toxicity to around 20%. Overall SFLC response rates are around 75%, with an SFLC clonal response rate of 37%. Organ responses occur in approximately 50%. Median duration of response is 32 months. Fatigue and fluid retention requiring increased diuretic use remain the most troublesome toxicities of this regimen. Patients with severe nephrotic syndrome are at particular risk of venous and arterial thrombosis and should receive thromboprophylaxis. However, CTD has the advantage of being a stem-cell-sparing oral regimen that does not require dose adjustment in renal impairment and is rapid-acting, the majority of patients achieving at least a partial (>50%) SFLC response after only 1 month of treatment.

Bortezomib, cyclophosphamide and dexamethasone

The proteasome inhibitor bortezomib (Velcade) has emerged as one of the most effective and rapidly acting therapies in the treatment of AL amyloidosis. In a recent multicentre study of 94 patients, most of whom had relapsed following previous treatments, clonal responses with bortezomib-based therapy occurred in 70%, including complete remission (CR) in 25%, and the median time to clonal response was only 1 month. Organ responses occurred in 30%, and median time to progression was 14 months; major toxicities included peripheral and autonomic neuropathy. However, unlike in myeloma, cytopenias were rare with bortezomib, most likely due to the smaller plasma cell burden in the bone marrow.

Most new AL amyloidosis patients in the UK now receive a cyclic combination of oral cyclophosphamide, subcutaneous bortezomib (Velcade) and oral dexamethasone (CVD). Common adverse effects include peripheral neuropathy, cytopenias, fatigue, GI disturbances, rashes and serious fluid retention in patients with cardiac involvement.

Lenalidomide

The oral thalidomide analogue lenalidomide (Revlimid) is also now widely used in patients with AL amyloidosis, although usually not first-line other than in some patients with severe neuropathies. Although few studies have been performed, accrued clinical experience indicates that the maximum tolerated dose is typically about 15 mg daily. Efficacy is markedly improved when administered with dexamethasone, and is associated with clonal response rates of about 50%, including complete remission in 20%. Organ responses were reported as approximately 22%. Toxicity from infections, even without neutropenia, appears to be surprisingly high, especially in patients with severe renal or cardiac impairment. Time to response is relatively slow, with a median of up to 6 months. Lenalidomide treatment is associated with a lower incidence of neuropathy than thalidomide or bortezomib, but adverse effects include cytopenias, venous thromboembolism, fatigue, GI disturbances and rashes.

Melphalan and prednisolone

In the early 1990s, cyclical oral melphalan with or without prednisolone was the first chemotherapy to be demonstrated in a randomized controlled trial as effective in the treatment of AL amyloidosis. However, benefits of treatment were typically not observed for at least 6 months and were substantial in only 20% of cases. As such, the use of melphalan was a breakthrough in the treatment of AL amyloidosis, but it is now no longer routinely recommended in low doses either as a single agent or with prednisolone. Although it may be tempting to offer this treatment to older patients, the small likelihood of benefit favours using novel drug regimens described above, despite their greater risks.

Allogeneic bone marrow transplantation

A case of successful allogeneic bone marrow transplantation in AL amyloidosis was reported in 1998 and since then a European study has reported on allogeneic and syngeneic haemopoietic cell transplantation in 19 patients with AL amyloidosis. Overall TRM was 40% and haematological response occurred in eight patients, with only one relapse after 31 months of follow-up. This treatment option is rarely exercised.

Supportive treatment

Supportive therapy remains a critical component of management of AL amyloidosis. For cardiac amyloidosis, the mainstay of treatment is diuretics. Vasodilating drugs and β-blockers are generally best avoided as they often worsen symptoms significantly. Fluid overload should be treated with fluid restriction, loop diuretics and a low-salt diet. Refractory oedema may respond well to the addition of spironolactone. Salt-poor albumin infusions can occasionally be helpful. Dysrhythmias may respond to amiodarone or to pacing. Implantable cardioverter defibrillators have been used very effectively in some patients.

In renal amyloidosis, rigorous control of hypertension is vital, and inhibition of angiotensin-converting enzyme and/or angiotensin II is often recommended in the context of proteinuria. Renal dialysis may be necessary, and is usually both feasible and acceptably tolerated. In autonomic neuropathy, fludrocortisone 100–200 μg/day can be helpful in some patients, but may cause or exacerbate fluid retention. Midodrine is an effective pressor agent, starting at 2.5–15 mg thrice daily. Midodrine causes activation of the α-adrenergic receptors of the arteriolar and venous vasculature, producing an increase in vascular tone and elevation of blood pressure. Its chief adverse effect is supine hypertension, and other pressor agents must be coadministered with caution. Gastroparesis causing symptoms of early satiety and nausea can be managed with prokinetic agents such as metaclopramide, along with advice about small frequent meals that are of soft consistency. Diarrhoea due to amyloid gut involvement or autonomic neuropathy may respond to loperamide and codeine phosphate. Malnutrition is not uncommon and is often underestimated. Significant weight loss should be treated with protein and vitamin supplementation and the patient should be reviewed by an experienced dietitian. Feeding via a percutaneous endoscopic gastrostomy may occasionally be required. Amyloid- and treatment-related peripheral neuropathy can be disabling and difficult to treat. Analgesia including opioids and non-steroidal anti-inflammatory drugs, amitriptyline, venlafaxine, antiepileptics, TENS and/or gabapentin/pregabalin have all been used, although anecdotal reports suggest limited efficacy. Withdrawal or dose reduction of any neurotoxic chemotherapies, such as thalidomide or bortezomib, should be considered.

Surgical resection of amyloidotic tissue is occasionally beneficial but, in general, a conservative approach to surgery, anaesthesia and other invasive procedures is advised. Should any such procedure be undertaken, meticulous attention to blood pressure and fluid balance is essential, especially in patients with renal and/or cardiac involvement. Amyloidotic tissues may heal poorly and are liable to haemorrhage.

Solid organ transplantation

Renal and cardiac transplantation have been used in selected patients with AL amyloidosis for over 25 years. Renal transplantation should be considered in relatively young patients with otherwise well-preserved extrarenal organ function who have chemotherapy-responsive disease. Cardiac transplantation should be considered in patients with isolated end-stage cardiac failure who would otherwise be eligible for autologous SCT. In such patients, cardiac transplantation must be followed by chemotherapy to prevent recurrence of cardiac amyloid or its accumulation in other organ systems. Liver transplantation has been performed in only a very small number of patients, with generally disappointing results, reflecting the fact that liver failure in AL amyloidosis is always associated with substantial amyloid deposition in other vital organs.

Localized AL amyloidosis

Localized deposits of AL amyloid can occur almost anywhere in the body, with characteristic sites including the skin, airways, conjunctiva and urogenital tract. They may be nodular or confluent and are associated with a usually inconspicuous focal infiltrate of clonal B-cell-producing amyloidogenic light chains. Progression of localized AL amyloid into a truly systemic disease is exceedingly rare, and conservative management is usually appropriate. Orbital AL amyloid presents as mass lesions that can disrupt eye movement and the structure of the orbit. Localized laryngeal AL amyloidosis is a well-recognized syndrome that is often amenable to direct or laser excision, but hereditary systemic apolipoprotein A-I amyloidosis can also present in this manner. Amyloidosis in the bronchial tree is virtually always of localized AL type, as are solitary or multiple amyloid nodules within the lung tissue. If accessible, bronchoscopic lasering of stenosing lesions can bring symptom relief. In contrast,

diffuse alveolar septal parenchymal deposition is commonly a manifestation of systemic AL amyloidosis. There are anectodal reports that inhaled steroids may benefit pulmonary symptoms. Breast amyloidosis has been associated with Sjögren disease in some cases, and mucosa-associated lymphoid tissue (MALT) lymphoma should be excluded. Lichenoid and macular forms of cutaneous amyloid are thought to be derived from keratin or related proteins, whereas nodular cutaneous amyloid deposits are generally of AL type and can sometimes be a manifestation of systemic AL amyloidosis. Localized urogenital AL amyloid deposits are often incidental findings, but may present with haematuria or, less commonly, obstruction. They can occur anywhere from the renal collecting system to the urethra, although are most usually identified within the bladder. Management is conservative or with transurethral laser resection when symptoms occur; cystectomy is rarely required.

Other forms of systemic amyloidosis

Among more than 5000 patients with systemic amyloidosis who have been referred for evaluation at the National Amyloidosis Centre, approximately 55% have had AL type, 15% AA (reactive systemic) type, and the remainder have had a variety of hereditary, localized and other types. Although some features of systemic AL amyloidosis are very characteristic of this particular type, such as macroglossia, periorbital purpura and certain permutations of organ involvement, the clinical phenotypes of AA, AL and hereditary systemic amyloidosis can be indistinguishable. Furthermore, the underlying chronic inflammatory disease process is clinically covert in 5–10% of patients with AA amyloidosis, and a family history is often absent in patients with hereditary ATTR and fibrinogen Aα-chain amyloidosis, which are the most common familial forms.

AA amyloidosis

Reactive systemic AA (secondary) amyloidosis occurs in 1–5% of patients with chronic inflammatory diseases that evoke a substantial acute-phase response, after a median latency of about 15–20 years. AA amyloid fibrils are derived from the circulating acute-phase reactant SAA, the serum concentration of which can increase from the healthy reference range of less than 10 mg/L to over 1000 mg/L during active inflammation. The commonest associated diseases in the developed world include rheumatoid arthritis, juvenile idiopathic arthritis and Crohn's disease. Familial Mediterranean fever and other rare periodic fever syndromes can result in AA amyloidosis. Chronic infections remain important causes in some parts of world. Castleman disease of the solitary plasma cell type should be considered among the underlying conditions that remain clinically covert.

Most patients present with nephropathy, particularly proteinuria, but liver and gastrointestinal involvement may occur at a late stage. Clinical involvement of the heart and nerves occurs very rarely. Diagnosis of AA amyloid is usually achieved by renal (or rectal) biopsy, and the AA fibril type can be confirmed immunohistochemically using anti-SAA antibodies in almost all cases. SAP scintigraphy virtually always shows involvement of the spleen and kidneys; hepatic involvement is a late feature associated with a poor prognosis. Treatment in AA amyloidosis should decrease the underlying inflammation so that SAA levels remain well controlled, preferably below 10 mg/L. The exact nature of the treatment will depend on the underlying inflammatory disease, and ranges from inhibitors of tumour necrosis factor in rheumatoid arthritis and colchicine in familial Mediterranean fever to surgical resection of Castleman disease tumours. Any therapy that reduces SAA production to healthy baseline levels prevents further deposition of AA amyloid, frequently leads to the regression of existing amyloid deposits, with improvement in amyloid-related organ dysfunction, and significantly improves long-term survival. Kidney transplantation generally has excellent outcomes in AA amyloidosis.

β_2-Microglobulin amyloidosis

β_2-Microglobulin amyloidosis, also known as dialysis-related amyloidosis, occurs because of the accumulation of β_2-microglobulin in renal failure and predominantly affects articular and periarticular structures in patients with end-stage renal failure who have been on dialysis for at least 7–10 years. Susceptibility factors include older age and the use of non-biocompatible dialysis membranes. Carpal tunnel syndrome is often the first clinical manifestation, and large-joint arthralgias, tenosynovitis, spondyloarthropathies and periarticular bone cysts are common. Although β_2-microglobulin amyloidosis is a systemic form of amyloidosis, deposits outside the musculoskeletal system are seldom of clinical significance. The disabling arthralgia may respond partially to non-steroidal anti-inflammatory drugs or corticosteroids, but the only really effective treatment for this condition is normalization of β_2-microglobulin levels through renal transplantation. Carpal tunnel syndrome is amenable to surgery, but may recur.

Transthyretin amyloidosis (ATTR amyloidosis)

Normal TTR is inherently, but weakly, amyloidogenic, and minor ATTR amyloid deposits are common in elderly individuals. Clinically significant involvement is seemingly very rare and almost completely restricted to the heart and carpal tunnel, but non-hereditary ATTR amyloid deposits occur in some measure in up to 25% of subjects over 80 years of age. It is likely that ATTR cardiac amyloidosis remains under-recognized by physicians, although new cardiac MRI techniques and scintigraphy with the re-purposed bone tracer ^{99m}Tc-DPD are greatly

improving the diagnostic yield. Non-hereditary cardiac ATTR amyloidosis is extremely rare before 65 years of age. Patients typically survive for several years with reasonably good quality of life managed with diuretics alone. A number of specific novel therapies are currently in late-stage clinic trial, including small interfering RNA and anti-sense oligonucleotide therapies that inhibit TTR protein production, and agents that are designed to stabilize TTR in the blood in its normal non-amyloid conformation.

Hereditary systemic amyloidoses

Hereditary systemic amyloidosis is caused by deposition of genetically variant proteins as amyloid fibrils, and is associated with mutations in the genes for TTR, fibrinogen Aα chain, cystatin C, gelsolin, apolipoprotein A-I, apolipoprotein A-II and lysozyme. These disorders are all inherited in an autosomal dominant manner with variable penetrance, and usually present in adult life.

Familial amyloidotic polyneuropathy (FAP) associated with mutations in the gene for TTR is the most common type of hereditary amyloidosis. It is characterized by progressive and disabling peripheral and autonomic neuropathy, often with cardiac involvement. Vitreous amyloid deposits may also occur and are virtually pathognomonic of the syndrome. Symptoms typically present between the third and seventh decades. More than 100 TTR variants are associated with FAP, the most frequent of which is the substitution of methionine for valine at residue 30 (TTR Met30). There are well-recognized foci of this in Portugal, Japan and Sweden, but FAP has been reported in most ethnic groups. TTR Ala60 is the most frequent cause of FAP in the British and Irish population, typically presenting after 50 years of age and usually with marked cardiac involvement. TTR Ile122 occurs in 3–4% of black Africans and is associated with a phenotype indistinguishable from senile (wild-type) cardiac amyloidosis. The majority of TTR is produced by hepatocytes, and liver transplantation is the only effective treatment for this disorder. Although the visceral amyloid deposits frequently regress following liver transplantation, the neuropathy is often irreversible and established cardiac amyloidosis may paradoxically progress due to ongoing fibril formation by wild-type TTR in this particular organ. Combined heart and liver transplantation has been performed successfully in a small number of cases.

The syndrome of non-neuropathic hereditary systemic amyloidosis is caused by mutations in the genes for fibrinogen Aα chain, lysozyme, apolipoprotein A-I and apolipoprotein A-II. Most such patients present with renal impairment and/or proteinuria, but substantial deposits in the liver and spleen are frequent in hereditary lysozyme and apolipoprotein A-I amyloidosis, and the heart may be involved in hereditary apolipoprotein A-I amyloidosis. Prominent neuropathy occurs in some patients with apolipoprotein A-I Arg26. Although kindreds with hereditary amyloidosis are rare, 5–10% of patients referred to the National Amyloidosis Centre with apparently sporadic amyloidosis do in fact have hereditary forms of the disease. About half of these are associated with TTR mutations, and most of the remainder are associated with variant fibrinogen Aα chain Val526. Penetrance of this particular mutation is extremely low in most families, thus obscuring the genetic aetiology, but the renal histology is characteristic, showing substantial accumulation of amyloid within enlarged glomeruli, but none in blood vessels or the interstitium. Renal impairment is often noted in the fourth or fifth decade, and end-stage renal failure is managed with dialysis or renal transplantation. Variant fibrinogen Aα-chain amyloidosis can recur in the transplanted graft and median survival of kidney transplants in patients attending our centre has been 7 years, though this can be much longer. Combined liver–kidney transplants have been performed in this condition though, with notable mortality. Hereditary cystatin C amyloidosis manifests in Icelandic families as cerebral amyloid angiopathy, with recurrent cerebral haemorrhage, and gelsolin variants are associated with corneal lattice dystrophy and cranial neuropathy, most often in Finnish patients.

DNA analysis should now be performed routinely on patients with systemic amyloidosis in whom AA or AL fibril type cannot be definitively verified. The newly described leucocyte chemotactic factor II (LECT2) form of amyloidosis can also mimic AL, AA and hereditary forms of amyloid-related kidney disease, but usually manifests as isolated, stable or slowly progressive renal impairment, often with positive kidney and adrenal gland signalling on SAP scintigraphy; currently it can only be reliably identified with mass spectrometry. The pathogenesis of ALECT2 amyloidosis is not known, but it is currently not thought to have a genetic aetiology.

Conclusion and future directions

Improved understanding of the aetiology and pathogenesis of amyloid has led to many recent advances in the characterization and management of amyloidosis. Chemotherapy in systemic AL amyloidosis can now be guided by its early effect on SFLC concentration, and DNA analysis can prevent patients with otherwise unrecognized hereditary amyloidosis from receiving inappropriate cytotoxic treatment. Clinical improvement following successful treatment of the various conditions that underlie amyloidosis is always gradual, and supportive measures are of great importance, but the new generation of myeloma therapeutics, including proteasome inhibitors and thalidomide analogues, have already had major impacts on outcomes in AL amyloidosis.

Novel therapies specifically for amyloidosis now in development include small molecules, monoclonal antibodies, peptides and GAG analogues that variously interfere with amyloidogenesis. Immunotherapy approaches are in development, with the

aim of promoting regression of amyloid. Several of these new therapeutic approaches are already being tested in patients with the hope that they may be effective in a diverse range of amyloid-related disorders in the near future.

Selected bibliography

Gillmore JD, Wechalekar A, Bird J *et al.* (2015) Guidelines on the diagnosis and investigation of AL amyloidosis. *British Journal of Haematology* **168**: 186–206.

Jaccard A, Moreau P, Leblond V *et al.* (2007) High-dose melphalan versus melphalan plus dexamethasone for AL amyloidosis. *New England Journal of Medicine* **357**: 1083–93.

Kastritis E, Wechalekar AD, Dimopoulos MA *et al.* (2010) Bortezomib with or without dexamethasone in primary systemic (light chain) amyloidosis. *Journal of Clinical Oncology* **20**: 1031–7.

Kaufman GP, Dispenzieri A, Gertz MA *et al.* (2015) Kinetics of organ response and survival following normalization of the serum free light chain ratio in AL amyloidosis. *American Journal of Hematology* **90**(3): 181–6.

Lachmann HJ, Booth DR, Booth SE *et al.* (2002) Misdiagnosis of hereditary amyloidosis as AL (primary) amyloidosis. *New England Journal of Medicine* **346**: 1786–91.

Mahmood S, Venner CP, Sachchithanantham S *et al.* (2014) Lenalidomide and dexamethasone for systemic AL amyloidosis following prior treatment with thalidomide or bortezomib regimens. *British Journal of Haematology* **166**(6): 842–8.

Sanchorawala V (2012) Role of high-dose melphalan and autologous peripheral blood stem cell transplantation in AL amyloidosis. *American Journal of Blood Research* **2**(1): 9–17.

Venner CP, Gillmore JD, Sachchithanantham S (2014) A matched comparison of cyclophosphamide, bortezomib and dexamethasone (CVD) versus risk-adapted cyclophosphamide, thalidomide and dexamethasone (CTD) in AL amyloidosis. *Leukemia* **28**(12): 2304–10.

Wechalekar AD, Gillmore JD, Bird J (2015) Guidelines on the management of AL amyloidosis. *British Journal of Haematology* **168**: 207–18.

CHAPTER 31

The classification of lymphomas: updating the WHO classification

31

Elias Campo[1] and Stefano A Pileri[2]

[1]Hospital Clinic, University of Barcelona, Barcelona, Spain
[2]Bologna University School of Medicine, St Orsola Hospital, Bologna, Italy

Introduction

The classification of lymphoid neoplasms in the World Health Organization (WHO) proposal is based on two major principles: stratification of the neoplasms according to their cell lineage and the definition of non-overlapping distinct diseases that are clinically relevant. The identification of these diseases is based on a combination of morphology, immunophenotype, genetic and molecular features and clinical manifestations.

The 2001 WHO classification of lymphoid neoplasms was based on the Revised European–American Classification of Lymphoid Neoplasms (REAL) published by the International Lymphoma Study Group (ILSG) in 1994. The validation of this proposal in a large series of tumours and the publication of the third edition of the WHO classification ended a long history of controversy between different classification schemes that used different terminologies and concepts among pathologists and clinicians in different parts of the world.

The current WHO classification published in 2008 was the result of an international effort that integrated criteria developed by different working groups, refined definitions of well-established diseases, incorporated new entities and developed new concepts and ideas related to the biology of lymphomas. However, the classification still has contentious aspects, such as provisional entities corresponding to categories for which the WHO Working Group felt there was insufficient evidence to recognize as distinct diseases at this time (Table 31.1).

Mature B-cell neoplasms

This is a heterogeneous group of neoplasias that in most cases may be related to different steps in the normal B-cell differentiation process. The combination of morphological, phenotypic and molecular features are essential for the diagnosis of the differen entities (Table 31.2).

Chronic lymphocytic leukaemia/small lymphocytic lymphoma (see also Chapter 27)

Chronic lymphocytic leukaemia/small lymphocytic lymphoma (CLL/SLL) is a neoplasm of mature small B-cell lymphocytes that commonly express CD5 and CD23 and have dim expression of surface IgM/IgD. The tumour usually involves the peripheral blood and bone marrow, but also lymph nodes, spleen, liver and other extranodal sites. The diagnosis of the disease requires, in the absence of extramedullary tissue involvement, the presence of 5×10^9/L monoclonal B lymphocytes with a CLL phenotype in the blood. The diagnosis may be established with lower cell counts when the patient has cytopenias or disease-related symptoms. The term 'small lymphocytic lymphoma' refers to non-leukaemic neoplasms with the same morphology and phenotype. The presence in blood of a clonal population of B-cells with the CLL phenotype below this number of lymphocytes is called monoclonal B-cell lymphocytosis (MBL). MBL is considered a precursor of CLL, although not all MBL will progress to an overt CLL.

Postgraduate Haematology, Seventh Edition. Edited by A Victor Hoffbrand, Douglas R Higgs, David M Keeling and Atul B Mehta.

Table 31.1 WHO classification of mature lymphoid neoplasms.

Mature B-cell neoplasms
Chronic lymphocytic leukaemia/small lymphocytic lymphoma
B-cell prolymphocytic leukaemia
Splenic B-cell marginal zone lymphoma
Hairy-cell leukaemia
Splenic lymphoma/leukaemia unclassifiable
Splenic diffuse red pulp small B-cell lymphoma
Hairy-cell leukaemia variant
Lymphoplasmacytic lymphoma/Waldenström macroglobulinaemia
Heavy-chain diseases
Plasma-cell myeloma
Solitary plasmacytoma of bone
Extraosseous plasmacytoma
Extranodal marginal zone lymphoma of mucosa-associated lymphoid tissue (MALT lymphoma)
Nodal marginal zone lymphoma
Paediatric nodal marginal zone lymphoma
Follicular lymphoma
Paediatric follicular lymphoma
Primary cutaneous follicle centre lymphoma
Mantle-cell lymphoma
Diffuse large B-cell lymphoma (DLBCL), not otherwise specified
T-cell/histiocyte-rich large B-cell lymphoma
Primary DLBCL of the CNS
Primary cutaneous DLBCL, leg type
EBV-positive DLBCL of the elderly
DLBCL associated with chronic inflammation
Lymphomatoid granulomatosis
Primary mediastinal (thymic) large B-cell lymphoma
Intravascular large B-cell lymphoma
ALK-positive large B-cell lymphoma
Plasmablastic lymphoma
Large B-cell lymphomas arising in HHV8-associated multicentric Castleman disease
Primary effusion lymphoma
Burkitt lymphoma
B-cell lymphoma, unclassifiable, with features intermediate between DLBCL and Burkitt lymphoma
B-cell lymphoma, unclassifiable, with features intermediate between DLBCL and classical Hodgkin lymphoma
Mature T-cell and NK-cell neoplasms
T-cell prolymphocytic leukaemia
T-cell large granular lymphocytic leukaemia
Chronic lymphoproliferative disorder of NK cells
Aggressive NK-cell leukaemia
Systemic EBV-positive T-cell lymphoproliferative disease of childhood
Hydroa vacciniforme-like lymphoma
Adult T-cell leukaemia/lymphoma
Extranodal NK/T-cell lymphoma, nasal type

Table 31.1 (*Continued*)

Mature T-cell and NK-cell neoplasms (Continued)
Enteropathy-associated T-cell lymphoma
Hepatosplenic T-cell lymphoma
Subcutaneous panniculitis-like T-cell lymphoma
Mycosis fungoides
Sézary syndrome
Primary cutaneous CD30-positive T-cell lymphoproliferative disorders
Lymphomatoid papulosis
Primary cutaneous anaplastic large-cell lymphoma
Primary cutaneous γδ T-cell lymphoma
Primary cutaneous CD8-positive aggressive epidermotropic cytotoxic T-cell lymphoma
Primary cutaneous CD4-positive small/medium T-cell lymphoma
Peripheral T-cell lymphoma, unspecified
Angioimmunoblastic T-cell lymphoma
Anaplastic large-cell lymphoma, ALK positive
Anaplastic large-cell lymphoma, ALK negative
Hodgkin lymphoma
Nodular lymphocyte predominant Hodgkin lymphoma
Classical Hodgkin lymphoma
Nodular sclerosis classical Hodgkin lymphoma
Lymphocyte-rich classical Hodgkin lymphoma
Mixed cellularity classical Hodgkin lymphoma
Lymphocyte-depleted classical Hodgkin lymphoma

Source: Swerdlow *et al*., 2008. Reproduced with permission of WHO.

The bone marrow is infiltrated in virtually all cases with an interstitial, nodular or diffuse pattern. The lymph nodes show effacement of the architecture by a diffuse infiltration. Aggregates of prolymphocytes and paraimmunoblasts, called proliferation centres, are an almost constant feature in the lymph nodes (Figure 31.1). The proliferative cells tend to accumulate in these areas and are also associated with some follicular dendritic cells and increased numbers of $CD4^+$ T cells. The detection of ZAP-70 by flow cytometry is an important prognostic marker of the disease that correlates well with the mutational status of the immunoglobulin genes. CD38 expression is also considered a prognostic marker (see also Chapter 27).

The mutational status of the immunoglobulin genes (IG) distinguishes two major subtypes of the disease. Thus, 40–50% of cases have unmutated IG (>98% identity with the germline), whereas 50–60% have hypermutated genes. Patients with unmutated-IG CLL have a more aggressive disease. Analysis of immunoglobulin gene sequences has revealed a marked bias in use of the different VH families and the identification of subsets of cases with quasi-identical IGHV sequences called stereotypes, suggesting the influence of certain antigens in the

Table 31.2 Immunophenotypic features of common B-cell neoplasms.

Neoplasm	SIG; CIg	CD20	CD5	CD10	BCL6	MUM1	CD23	CD43	CD103	Cyclin D	CD38/CD138	CD30	EBV
CLL	+; –/+	+weak	+	–	–	–/+	+	+	–	–	–	–	–
Lymphoplasmacytic lymphoma	+; +	+	–	–	–	+	–	+/–	–	–	+	–	–
Hairy-cell leukaemia	+; –	+	–	–	–	–	–	+	++	+/–	–	–	–
Splenic marginal-zone lymphoma	+; –/+	+	–	–	–	–/+	–	–	+	–	–	–	–
Follicular lymphoma	+; –	+	–	+/–	+	–/+	–/+	–	–	–	–	–	–
Mantle-cell lymphoma	+; –	+	+	–	–	–	–	+	–	+	–	–	–
MALT lymphoma	+; +/–	+	–	–	–	–/+	–/+	–/+	–	–	–	–	–
DLBCL-GCB	+/–; –/+	+	–	+	+	–		–/+		–	–	–	–
DLBCL-ABC	+/–; –/+	+	–	–	–	+		–	–	–	–	–/+	–/+*
PMBL	–	+	–	–	+/–	+/–	+	–	–	–	–	+	–
Burkitt lymphoma	+; –	+	–	+	+	–	–	–		–	–	–	–/+
Plasmablastic lymphoma	+	–	–	–/+	–	+	–	–	–	–	+	+/–	+
Plasma cell myeloma	+	–/+	–	+/–	–	+	–	–/+	–	–/+	+	–	–

*EBV+ DLBCL.
ABC, activated B-cell type; cIG, cytoplasmic immunoglobulin; DLBCL, diffuse large B-Cell lymphoma; EBV, Epstein–Barr virus; GCB, Germinal-centre B cell type; PMBL, primary mediastinal B-cell lymphoma; sIg, surface immunoglobulin.

selection of tumour cells. The most common chromosomal alterations are deletions of chromosome 13q (50%), trisomy 12 (20%), 11q deletions (15%) and 17p deletions (10%); 13q deletions are more common in mutated CLL, whereas 11q and 17p deletions are more frequent in unmutated CLL and are associated with a worse prognosis. Next-generation sequencing (NGS) studies have identified the complex mutational landscape of the disease that in addition to the frequent *TP53* (3–15%) and *ATM* (10–15%) alterations also include frequent mutations in *NOTCH1* (4–15%), *SF3B1* (5–17%) *BIRC3* (3%), *MYD88* (3%), among many others. The distribution of these mutations varies in different subgroups of patients and in the evolution of the disease

Transformation into a more aggressive tumour occurs in 2–10% of patients (Richter syndrome). The two most common forms of transformation are diffuse large B-cell lymphoma (DLBCL) and, less frequently, Hodgkin lymphoma (HL). DLBCL arising in unmutated-IG CLL usually corresponds to the clonal evolution of the preceding CLL, whereas in mutated CLL it frequently corresponds to a different lymphoid neoplasm. DLBCL transformation is frequently associated with *TP53* or *NOTCH1* mutations.

B-cell prolymphocytic leukaemia (see also Chapter 27)

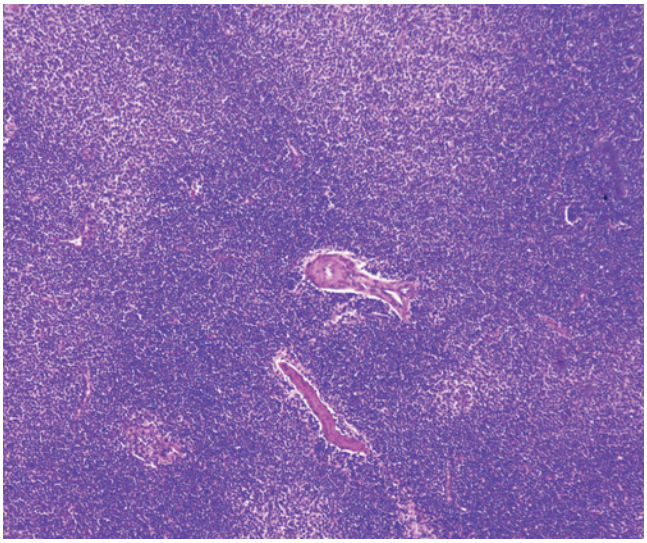
B-cell prolymphocytic leukaemia (B-PLL) is a malignancy of B prolymphocytes that affects the blood, bone marrow and spleen and is characterized by more than 55% prolymphocytes in the blood. B-PLL does not include transformed CLL, CLL with increased prolymphocytes, and blastoid mantle cell lymphoma with t(11;14)(q13;q32), which should be excluded. B-PLL is an uncommon disease of old patients (median age 65–69 years) and similar male/female distribution. Patients have 'B' symptoms, massive splenomegaly, absent or minimal lymphadenopathy and a rapidly rising lymphocyte count, usually over 100×10^9/L. Anaemia and thrombocytopenia are seen in 50%.

The cells strongly express surface IgM with or without IgD and mature B-cell antigens. CD5 and CD23 are only positive in 20–30% and 10–20% of cases, respectively. Complex karyotypes are common and 17p deletions associated with *TP53* gene mutations are detected in 50% of cases.

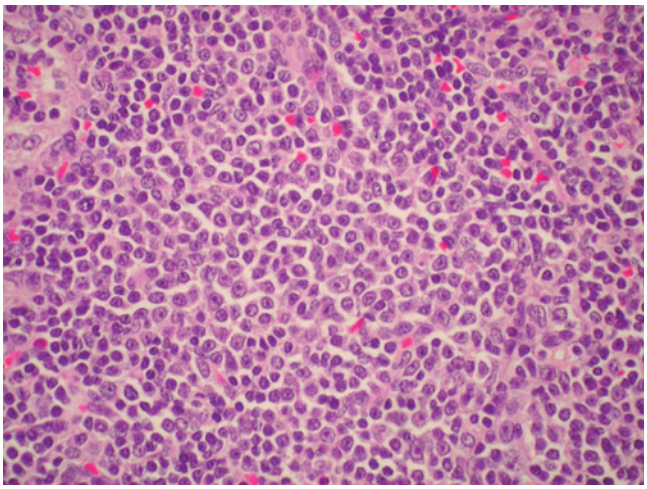
Median survival is 30–50 months. *IGHV* mutations, ZAP-70 and CD38 expression are heterogeneous and, contrary to CLL, lack precise clinical prognostic value.

Splenic marginal zone lymphoma

Splenic marginal zone lymphoma (SMZL) is a B-cell neoplasm composed of small lymphocytes that surround and replace the germinal centres and mantle of the reactive follicles in the white pulp, and merge with a peripheral (marginal) zone of larger cells, including scattered transformed blasts. Both small and larger cells infiltrate the red pulp. The patients usually have a leukaemic and splenomegalic presentation with villous lymphocytes in the blood. The disease also involves the splenic hilar lymph nodes



(a)



(b)

Figure 31.1 Lymph node with chronic lymphocytic leukaemia. (a) At low power the lymph node shows diffuse effacement of the architecture with vaguely nodular pale areas that correspond to the proliferation centres. (b) At higher magnification the proliferation centre contains larger cells with nuclei that have a central nucleolus (prolymphocytes and paraimmunoblasts).

and bone marrow, but extension to peripheral lymph nodes is uncommon (Figure 31.2).

There are no distinctive phenotypic markers for SMZL, whose diagnosis requires the exclusion of other lymphoma types. Most cases exhibit IgM/IgD and a mature B-cell phenotype, but other markers are usually negative (Figure 31.2). The immunoglobulin genes are unmutated in approximately half of the cases, which also tend to have deletions of chromosome 7q31–32 (45%) and a more unfavourable evolution. NGS studies have identified frequent mutations in *NOTCH2* (20–25%), genes of the NF-κB pathway (*BIRC3*, *TNFAIP3*), and KLF2. Some cases are positive for hepatitis C virus and may respond to antiviral treatment. The clinical course is usually indolent, but some patients may have progressive disease and transformation to a large-cell lymphoma may occur. The hairy cell leukaemia variant and lymphomas with a diffuse infiltrate of the red pulp are included in a provisional category termed 'splenic B-cell lymphoma/leukaemia, unclassifiable'.

Lymphoplasmacytic lymphoma/Waldenström macroglobulinaemia

Lymphoplasmacytic lymphoma (LPL) is a B-cell neoplasm composed of small lymphocytes, plasmacytoid lymphocytes and plasma cells, usually involving the bone marrow and sometimes lymph nodes and spleen that does not fulfil the criteria for any other B-cell neoplasm which may have plasmacytic differentiation. Although the detection of a paraprotein is common, it is not required for the diagnosis. Waldenström macroglobulinaemia is an LPL with bone marrow involvement and an IgM monoclonal gammopathy of any concentration. Because these entities do not have specific markers, it is essential to exclude the presence of any other B-cell neoplasm that may have a plasmacytic differentiation. NGS have identified activating mutations of *MYD88* in 90–100% of the cases and *CXCR4* in 28%. These findings may help in the differential diagnosis of the disease.

Plasma cell neoplasms (see also Chapter 29)

Plasma cell neoplasms encompass a spectrum of lesions characterized by the clonal expansion of terminally differentiated B cells that usually secrete a class-switched immunoglobulin (Table 31.3).

Monoclonal gammopathy of undetermined significance

Monoclonal gammopathy of undetermined significance (MGUS) is defined by less than 10% clonal plasma cells in the bone marrow and less than 30 g/L of an M-protein and absence of end-organ damage. This condition is considered a preneoplastic process since evolution to overt plasma cell malignancy does not always occur.

Plasma cell myeloma (see also Chapter 29)

Plasma cell myeloma is a bone marrow neoplasm characterized by multifocal proliferation of plasma cells associated with a serum M-protein and symptomatic disease related to different organ dysfunction or lytic bone lesions. The plasma cells in the bone marrow expand in small or large clusters. Some cases may present asymptomatically, whereas in others the tumour cells spread to the peripheral blood in the form of a plasma-cell leukaemia. The phenotype of these cells is aberrant, with expression of CD19 and CD56 in 75% of cases and other markers such

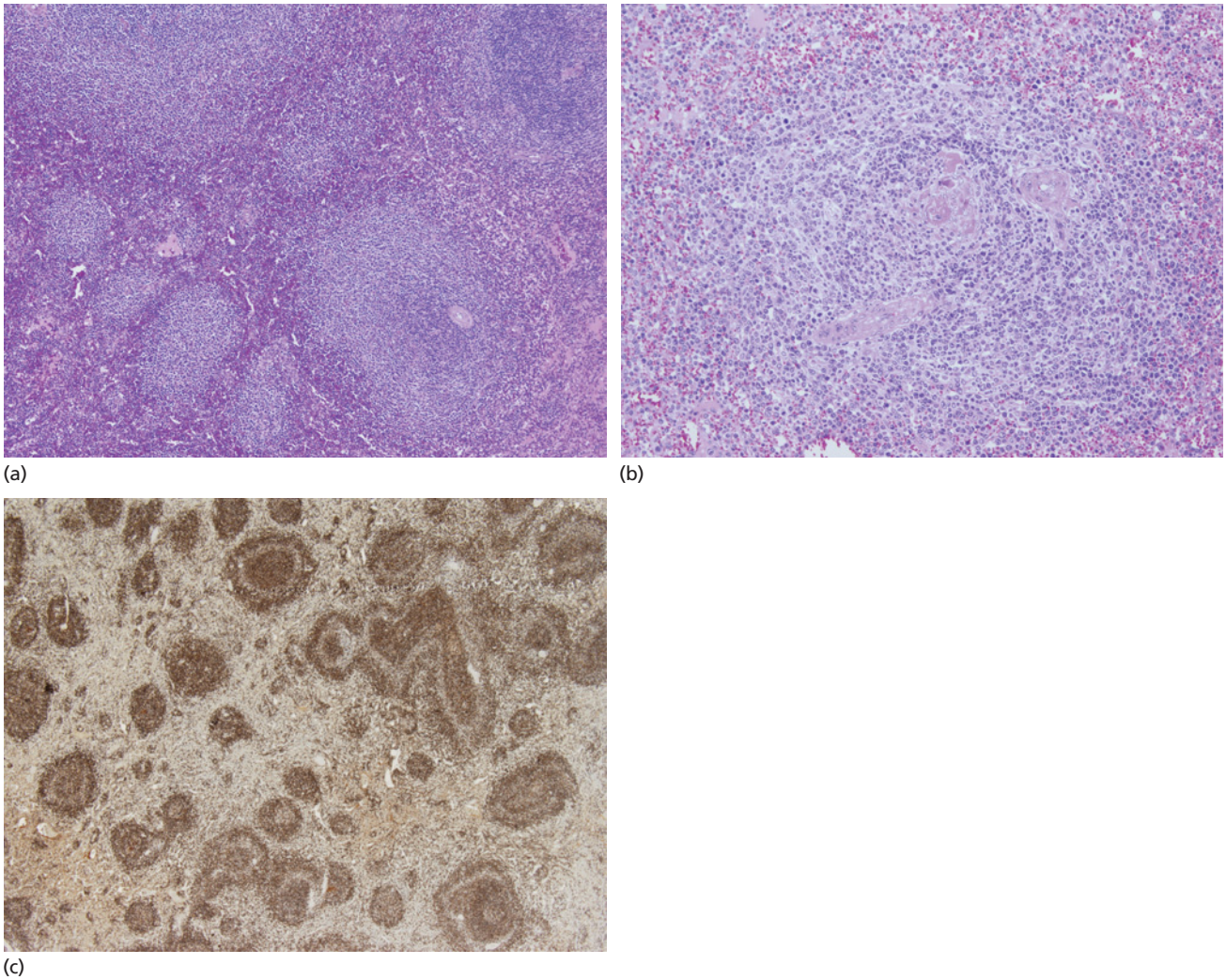


Figure 31.2 Splenic marginal zone lymphoma. (a) The spleen shows multiple nodular proliferations in the red pulp and expanding the white pulp areas. (b) At higher magnification the nodules are composed of small lymphoid cells in the centre and larger cells at the periphery. (c) The tumour cells are IgD positive.

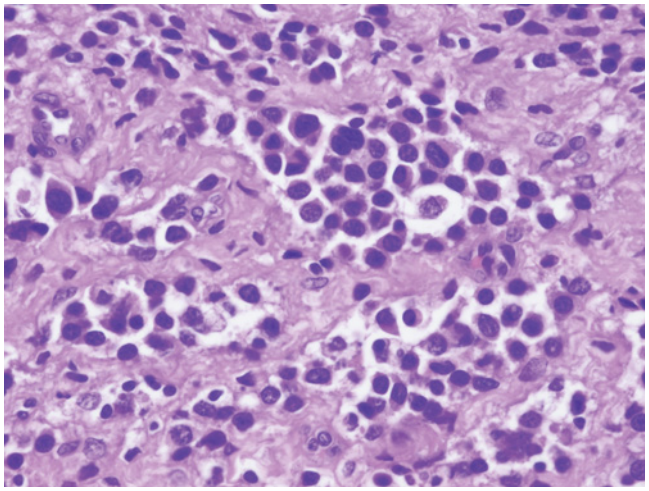
Table 31.3 Plasma cell neoplasms.

Monoclonal gammopathy of undetermined significance
Plasma cell myeloma
Asymptomatic (smouldering) myeloma
Non-secretory myeloma
Plasma-cell leukaemia
Plasmacytoma
Solitary plasmacytoma of bone
Extraosseous (extramedullary) plasmacytoma
Immunoglobulin deposition diseases
Osteosclerotic myeloma (POEMS syndrome)

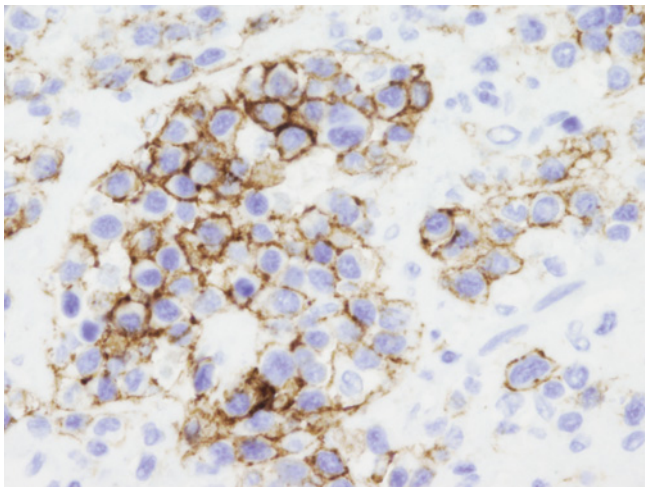
as CD117, CD20, CD52 and CD10. Chromosomal translocations involving the 14q32 region occur in 40% of cases; the partner translocations include *CCND1* (cyclin D1) at 11q13, *MAF* at 16q23, *FGFR3* and *MMSET* at 4p16, *CCND3* (cyclin D3) at 6p21 and *MAFB* at 20q11. Tumours lacking these alterations are usually hyperdiploid. *MYC* translocations have been recently identified in around 20% of cases.

Plasmacytoma

Plasmacytoma is a solitary tumour lesion composed of clonal plasma cells that may occur in bones or extraosseous tissues (Figure 31.3). Bone plasmacytomas evolve to plasma cell myeloma in two-thirds of patients. In contrast, extraosseous



(a)



(b)

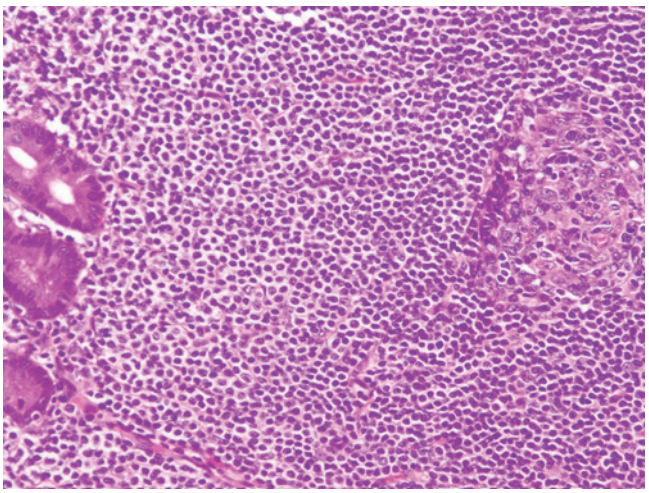
Figure 31.3 Plasmacytoma of the nasal cavity. (a) Atypical plasma cells embedded in fibrous stroma. (b) Plasma cells are CD138 positive.

localized lesions have a relatively indolent behaviour without bone marrow involvement, suggesting a closer relationship to MALT lymphomas than bone marrow plasma-cell neoplasms.

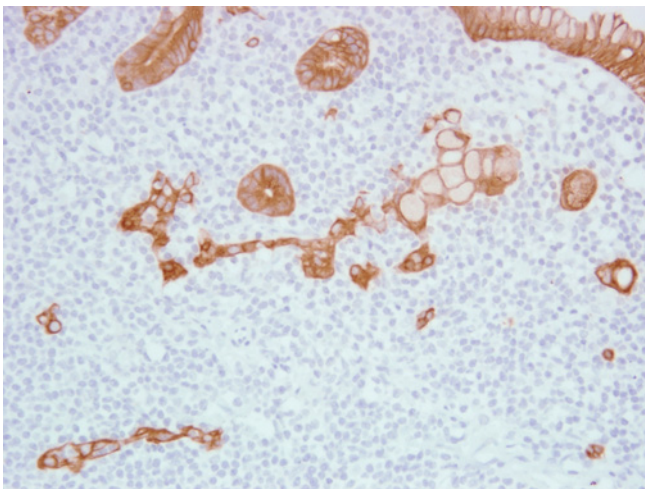
Extranodal marginal zone lymphoma of mucosa-associated lymphoid tissue (MALT lymphoma)

MALT lymphomas are extranodal B-cell neoplasms composed of a heterogeneous population of small lymphocytes, marginal zone lymphocytes with cleaved nuclei (centrocyte-like), cells with clear cytoplasm resembling monocytoid B cells, cells with plasmacytic differentiation and occasional large transformed cells. All these cells expand the marginal zone of reactive follicles and in some cases may colonize the germinal centres, but preserve the mantle zone area. In epithelial tissues the neoplastic cells infiltrate the epithelium forming lymphoepithelial lesions (Figure 31.4). These tumours occur most commonly in the gastrointestinal tract, salivary gland, lung, head and neck, ocular adnexa, skin and, less frequently, thyroid and breast.

The lymphoma cells express mature B-cell markers, IgM and less often IgG or IgA, but IgD is usually negative. CD5, CD10 and CD23 are negative, although some CD5-positive cases have been described. Four major translocations have been associated with these lymphomas, t(11;18), t(1;14), t(14;18) and t(3;14), that generate the chimeric gene *BIRC3–MALT1* or activate



(a)



(b)

Figure 31.4 Gastric marginal zone lymphoma of the mucosa-associated lymphoid tissue (MALT lymphoma). The gastric mucosa is infiltrated by a monotonous population of lymphoid cells that surrounds a residual germinal centre (a) and destroys the glands, creating lymphoepithelial lesions clearly observed with cytokeratin staining (Cam 5.2 immunoperoxidase staining) (b).

BCL10, *MALT1* and *FOXP1*, respectively. The t(11;18) translocation tends to occur more frequently in gastric and lung lymphomas, whereas the t(14;18) translocation is more often seen in salivary gland and ocular adnexa tumours. Interestingly, *MALT1* and *BCL10* participate in activation of the same NF-κB pathway.

These lymphomas arise in topographic sites with a preexisting chronic inflammatory lesion induced by infectious, immunological or unknown stimuli. Thus gastric lesions are associated with *Helicobacter pylori*, *Campylobacter jejuni* is detected in the immunoproliferative small intestinal disease, *Chlamydia psittaci* in ocular adnexa tumours of certain geographic regions and *Borrelia burgdoferi* in some cutaneous MALT lymphomas. Autoimmune disorders of the salivary gland (Sjögren syndrome) and thyroid (Hashimoto disease) are the preceding lesions of MALT lymphomas in these locations.

The lymphoma responds to control of the underlying infectious disease and local treatments, but may relapse in extranodal territories after many years. Nodal dissemination may precede clinically the detection of extranodal involvement. Tumours with the translocations described appear to be resistant to antibiotic therapy. Transformation to large-cell lymphomas may occur.

Nodal marginal zone lymphoma (NMZL)

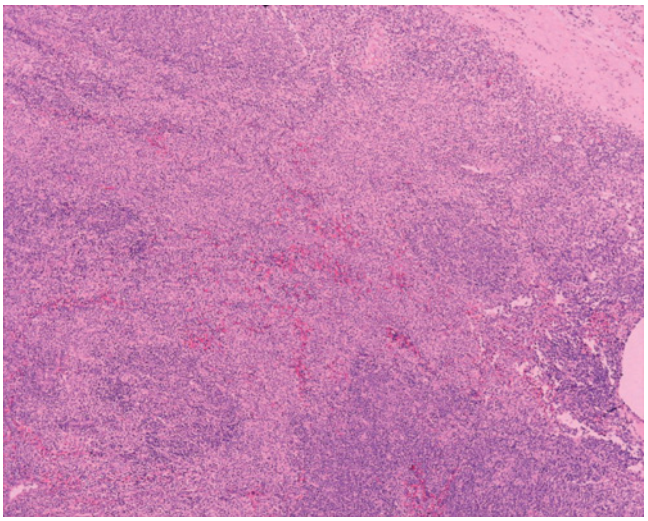
NMZL is a primary neoplasm of the lymph nodes that resembles extranodal MALT lymphomas (Figure 31.5). An extranodal or splenic lymphoma should be ruled out before this diagnosis is established. Although the morphology and phenotype is similar to MALT lymphoma, primary NMZL does not exhibit the typical translocations of these tumours. Patients may present with disseminated disease and 60–80% survive more than 5 years. Some cases may evolve to large B-cell lymphomas.

NMZLs in the paediatric age group seem to have distinctive clinical and morphological features. Thus, they present predominantly in males as asymptomatic localized tumours and morphologically the residual follicles may have features of progressive transformed germinal centres with eroded mantle zones by the tumour cell infiltration. Molecular and genetic studies detect clonal rearrangements of the IG or chromosomal alterations in virtually all cases and these findings are useful in the differential diagnosis with reactive hyperplasias. The prognosis of these patients seems excellent with conservative therapy.

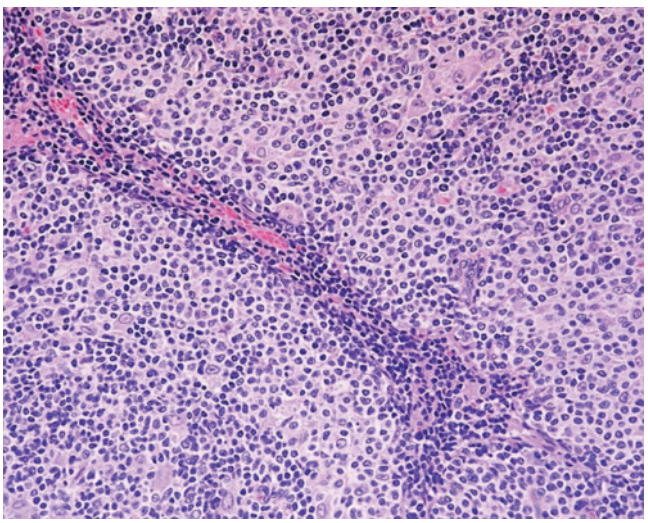
Follicular lymphoma (FL)

FL is a neoplasm composed of cells of the germinal centres, with different proportions of small centrocytes and large centroblasts, that usually has a follicular growth pattern (Figure 31.6). These tumours are common in Western countries, accounting for around 20–30% of all lymphomas.

These tumours have been graded according to the number of large cells, as follows: grade 1 (0–5 large cells per high-power field), grade 2 (6–15 large cells) and grade 3 (>15 large cells). Grades 1 and 2 represent a morphological continuum that is not associated with relevant clinical or biological differences and therefore this distinction is not encouraged. However, grade 3 tumours are further distinguished into grades 3a and 3b according to the presence of intermingled small centrocytes in grade 3a, whereas grade 3b is composed entirely of large cells. According to the genetic and phenotypic features, FL grade 3a seems more related to FL grade 1–2, whereas FL grade 3b may be closer to DLBCL. Any diffuse area composed of more than 15 large cells should be reported as DLBCL, and the percentage of the respective DLBCL and FL components indicated.



(a)



(b)

Figure 31.5 Nodal marginal zone lymphoma. (a) Tumour cells with a clear appearance infiltrate the node, with partial preservation of the architecture. (b) Higher magnification reveals the clear cytoplasm of the tumour cells.

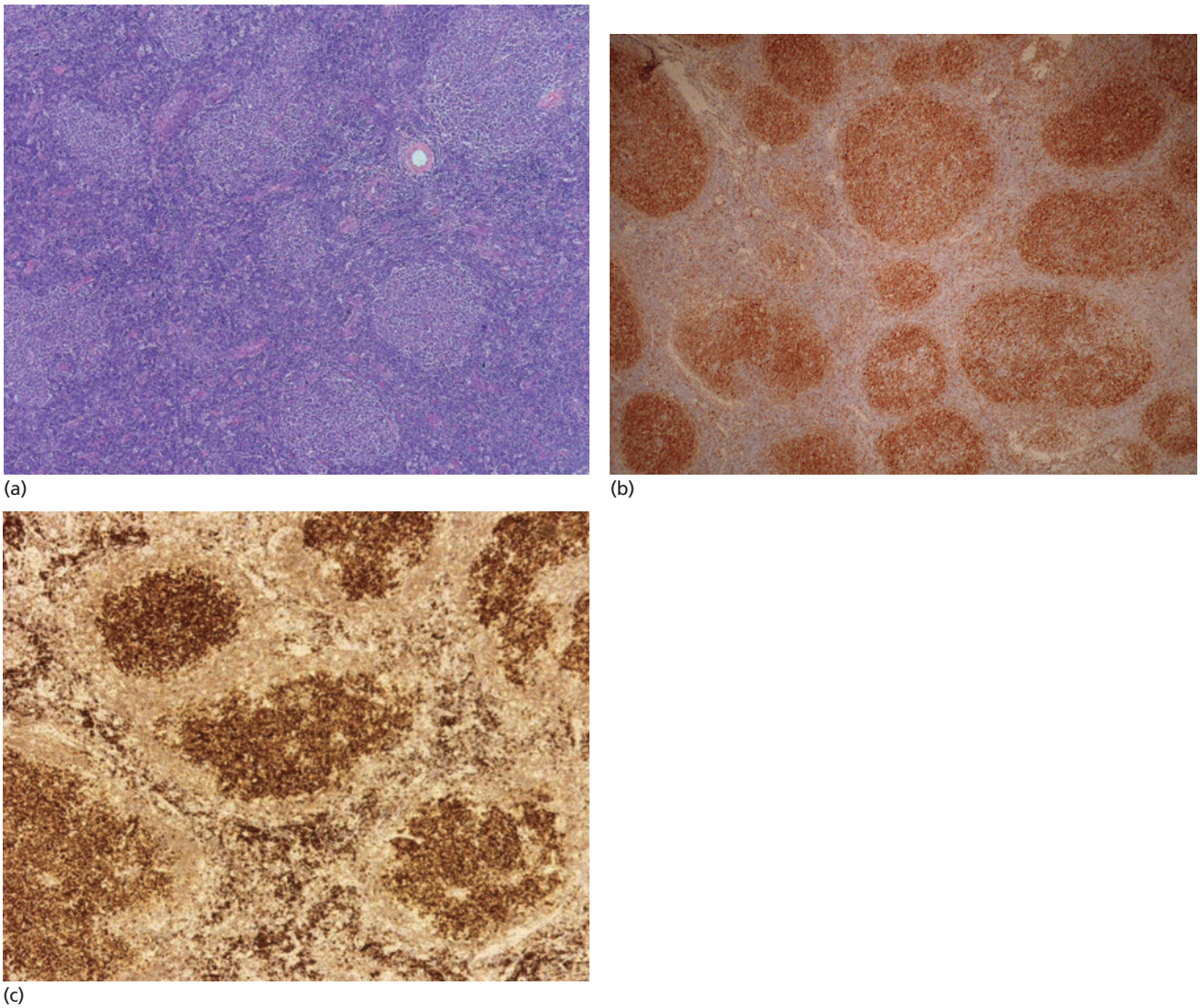


Figure 31.6 Follicular lymphoma. (a) The tumour grows in a follicular pattern with expanded germinal centres without macrophages. The tumour cells are positive for CD10 (b) and BCL-2 (c).

FL has a mature B-cell phenotype with coexpression of the germinal-centre markers CD10 and BCL-6 (Figure 31.6). CD5, CD43 and CD23 are negative. IRF4/MUM1, a transcription factor related to plasma cell differentiation, is also usually negative. BCL-2 is positive in 85–90% of FL grade 1–2, but only in 50% of grade 3. BCL-2 staining is very useful because reactive germinal centres are negative (Figure 31.6). This expression reflects the presence of the t(14;18) translocation that is the genetic hallmark of this lymphoma and which targets the *BCL2* gene. FL grade 3b is less frequently positive for CD10 and BCL-2 protein and the t(14;18) translocation is only detected in 5–40% of cases. In contrast, these lymphomas express IRF4/MUM1 in 40% of the cases and carry 3q27 and *BCL6* rearrangements in 30–50% of cases, whereas these aberrations are rare in FL grades 1–3a. A predominant diffuse variant of FL that also lacks the t(14;18) has been described. The presentation in these patients is usually inguinal or pelvic with large masses. The tumour expresses CD10, CD23 and carries del 1p36. Some cases may progress to a DLBCL and this transformation may be associated with *TP53* mutations, *CDKN2A/B* homozygous deletions or t(8;14) translocations involving *MYC* and mutations of *MYD88* and *TNFAIP3*.

The tumour cells in FL are associated with a rich microenvironment of different types of T cells and histiocytes that seem to play a major role in determining the biological behaviour of the lymphoma. NGS studies in FL have identified frequent mutations in genes involved in chromatin remodelling, particularly in *MLL2* (89%), *CREBB* (36%) and *EZH2* (22%).

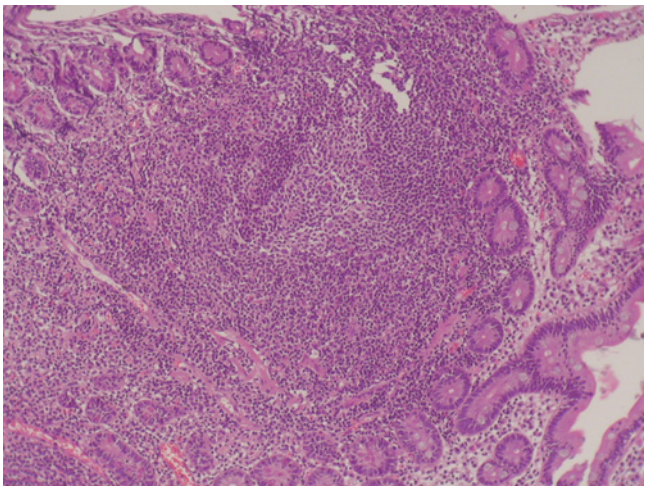


Figure 31.7 Follicular lymphoma of the duodenum.

In situ follicular lymphoma/intrafollicular neoplasia

In situ FL or intrafollicular neoplasia is a variant in which the tumour cells are limited to the germinal centres of the follicles. Usually, only a number of the lymph node follicles are involved, whereas others are reactive. The tumour cells are recognized by their strong BCL-2 expression inside the germinal centres and the absence of CD10- and BCL-6-positive tumour cells outside the follicles. The BCL-2-positive cells in the follicle carry the t(14;18) translocation. Some of these cases may correspond to early involvement of a lymph node by a disseminated tumour. A second group of patients may develop overt FL during follow-up, but most patients (94%) remain free of lymphoma elsewhere after many years of follow-up.

Paediatric follicular lymphoma

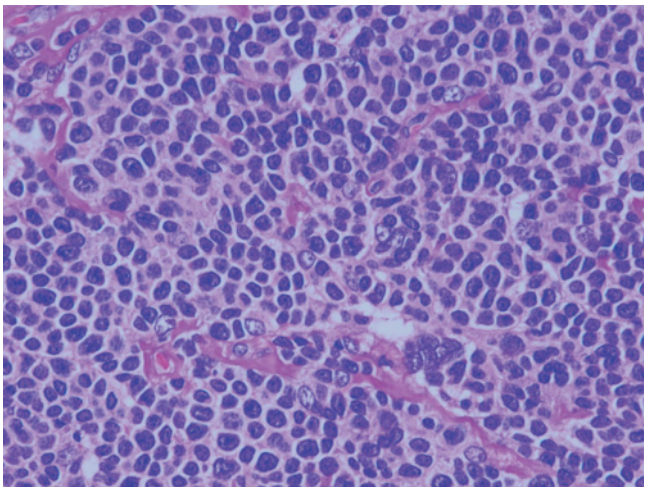
Paediatric FL occurs in children and young adults, usually involving the head and neck region, but also the testis with localized disease. The lymphomas are BCL-2 negative, do not carry the t(14;18) translocation and are grade 3. These cases have a relatively good evolution, with local therapy. A subset of tumours presenting in the Waldeyer's ring are frequently associated with a DLBCL component, express strong IRF4 and carry translocations of this gene. These cases may require systemic therapy.

Primary intestinal follicular lymphoma

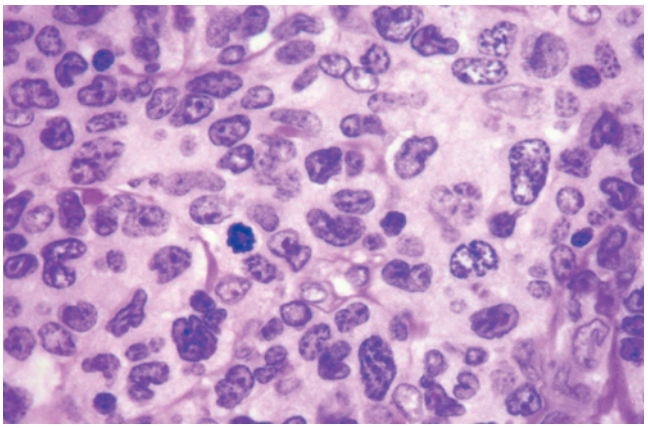
Primary intestinal FL occurs frequently in the duodenum as an incidental finding (Figure 31.7). These cases have a conventional morphology, phenotype and genetic findings. However, the tumours express IgA and remain localized, with an excellent prognosis, even without treatment.

Primary cutaneous follicle centre lymphoma

Primary cutaneous follicle centre lymphoma generally presents in the skin of the head and trunk. The tumour cells may grow with a follicular or more diffuse pattern. The cells express B-cell markers and BCL-6, whereas CD10 is positive in cases with a follicular pattern, but tend to be lost in the diffuse component. BCL-2 expression and the t(14;18) translocation are usually negative. The tumours may relapse, but have a good outcome without extracutaneous dissemination, even with only localized therapy.



(a)



(b)

Figure 31.8 Mantle-cell lymphoma with classical (a) and pleomorphic (b) morphology.

Mantle-cell lymphoma

Mantle-cell lymphoma (MCL) is a B-cell neoplasm generally composed of monomorphous small- to medium-sized lymphoid cells with irregular nuclear contours and cyclin D1 overexpression secondary to *CCND1* translocation (Figures 31.8 and 31.9). This tumour comprises 3–10% of non-Hodgkin lymphomas and occurs mainly in males, with a median age around 60 years. Most MCLs are diagnosed as disseminated nodal disease, but a number of patients are initially seen with

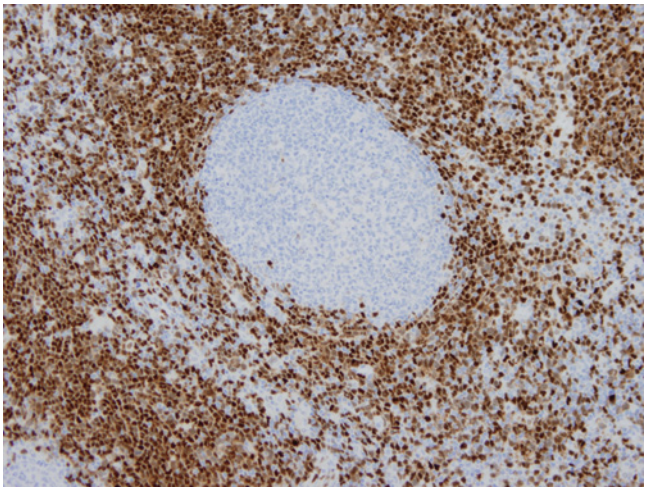


Figure 31.9 Mantle-cell lymphoma is positive for cyclin D1.

leukaemic disease, bone marrow involvement and frequently splenomegaly, but without lymphadenopathy. Extranodal involvement is very common. Some tumours may have a blastoid or pleomorphic cytology resembling lymphoblasts or DLBCL, respectively. These cytological variants are associated with higher proliferative activity, more complex karyotypes and worse prognosis. The tumour cells grow with a vaguely nodular, diffuse or mantle-zone pattern. Some cases may show restricted infiltration to the inner mantle zones of otherwise reactive follicles, a situation that has been named *in situ* MCL. The tumour cells express mature B-cell markers, with intense CD20 and surface immunoglobulin, and coexpress CD5 and CD43. CD23 and germinal-centre markers are negative.

MCL exhibits the t(11;14) translocation that targets *CCND1*. SOX11 is a transcriptional factor highly expressed in MCL but not in other mature B-cell lymphomas, with the exception of Burkitt lymphoma (30%). SOX11 seems to play an oncogenic role in MCL by regulating a broad transcriptional program that includes blocking the B-cell differentiation program and promoting angiogenesis. SOX11 positive tumours usually have unmutated *IGHV* and accumulate complex karyotypes. A subset of MCL does not express SOX11. These cases tend to have hypermutated *IGHV* and simple karyotypes. Some cases of cyclin-D1-negative MCL have been recognized with similar morphological and phenotypic features to those of cyclin-D1-positive tumours. Approximately 50% of these cases have *CCND2* translocations. SOX11 is also expressed in cyclin-D1-negative tumours and its detection is a useful tool to recognize this variant. The clinical presentation and behaviour is similar to *CCND1*-positive tumours. NGS have identified new mutated genes that in addition to the frequently mutated *ATM* (41%) and *TP53* (28%) include *NOTCH1/2* (10%), genes involved in chromatin modification (*MLL2* (15%) *WSHC1* (10%), *MEF2B* (3%)) and activation of the NF-κB pathway (*BIRC3*, *TRAF2/3*) (6%). SOX11-negative tumours may have *TP53* mutations associated with worse prognosis but usually do not have mutations in the other genes.

In general MCL has an aggressive clinical course, with frequent relapses after responding to chemotherapy and a median survival of 3–5 years. Proliferation is considered the best biological prognostic parameter. Patients presenting with disease limited to the blood, bone marrow and sometimes the spleen, without lymphadenopathy, have been reported to have a better prognosis and may be candidates for watch and wait management. These tumours have simple karyotypes and a distinct expression profile, including a negative expression of SOX11.

Diffuse large B-cell lymphoma not otherwise specified (see also Chapter 34)

DLBCL is a neoplasm of large B lymphocytes with a diffuse growth pattern. These lymphomas are heterogeneous and several morphological variants, phenotypical and molecular subtypes, and different entities have been recognized (Table 31.4). DLBCL is very common and accounts for 25–30% of adult non-Hodgkin lymphomas. Usually, these cases present as primary tumours, but they can also represent the transformation of a less aggressive B-cell lymphoma.

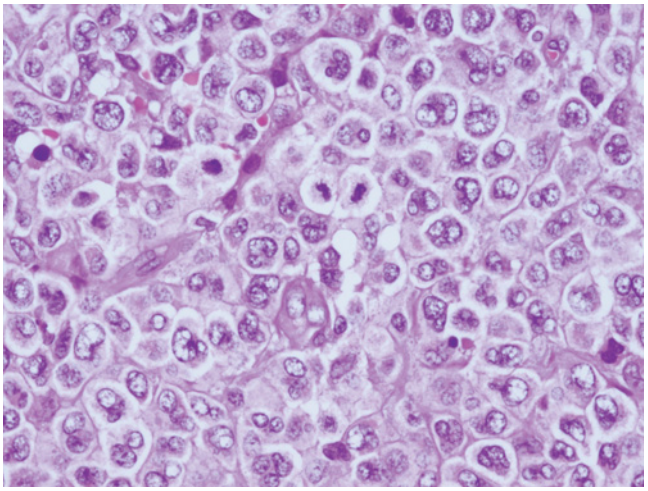
Morphologically, DLBCL not otherwise specified (NOS) may have centroblastic, immunoblastic or anaplastic cytology (Figures 31.10 and 31.11). These variants are related to biological and genetic features, but the low reproducibility prevents their use as major classifiers. Phenotypically, DLBCL, NOS expresses mature B-cell markers. CD5 is detected in a subset of these tumours that has been associated with a more aggressive behaviour. The expression of the germinal-centre markers CD10, BCL-6 and LMO2 has been related to a germinal-centre origin of the tumours, whereas the expression of IRF4/MUM1 and FOXP1 has been associated with an origin in non-germinal-centre activated B cells (Figures 31.10 and 31.11). These two immunophenotypic subgroups of DLBCL have a clear correlation with the two molecular subtypes of germinal-centre (GCB) and activated B-cell (ABC) DLBCL, recognized by gene expression profiling. However, the conflicting results obtained in different studies on the prognostic significance of the phenotypic subgroups make difficult their use in clinical practice.

Genetically, 20–30% of DLBCL, NOS exhibit the t(14;18) translocation and *BCL2* gene rearrangement. These cases are associated with CD10 expression and a germinal-centre origin. Translocation of the 3q27 region and *BCL6* rearrangements are found up to 30% of cases. *MYC* translocations have been observed in up to 10–15% of cases and are usually associated with complex karyotypes and with *BCL2* or, less frequently, *BCL6* rearrangements. These DLBCL carrying *MYC* and *BCL2* or *BCL6* translocations have been called DLBCL with a genetic

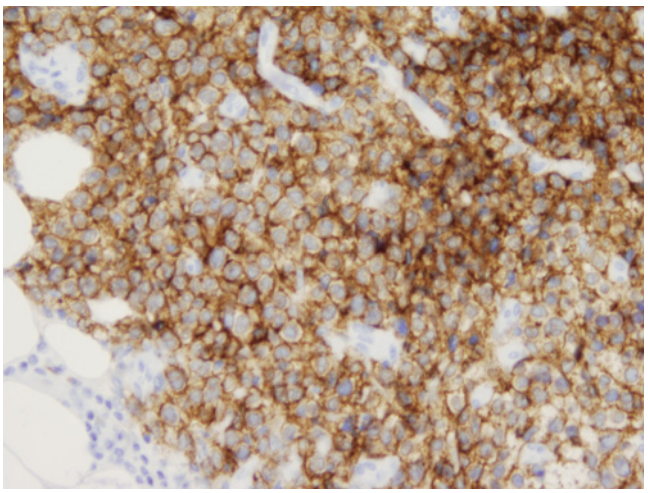
Table 31.4 Diffuse large B-cell lymphoma (DLBCL): variants, subgroups, subtypes and entities.

DLBCL, not otherwise specified
Common morphological variants
Centroblastic
Immunoblastic
Anaplastic
Rare morphological variants
Molecular subgroups
Germinal centre B-cell-like
Activated B-cell-like
Immunohistochemical subgroups
CD5-positive DLBCL
Germinal centre B-cell-like
Non-germinal centre B-cell-like
DLBCL subtypes
T-cell/histiocyte-rich large B-cell lymphoma
Primary DLBCL of the central nervous system
Primary cutaneous DLBCL, leg type
EBV-positive DLBCL of the eldery
Other lymphomas of large B cells
Primary mediastinal (thymic) large B-cell lymphoma
Intravascular large B-cell lymphoma
DLBCL associated with chronic inflammation
Lymphomatoid granulomatosis
ALK-positive large B-cell lymphoma
Plasmablastic lymphoma
Large B-cell lymphoma arising in HHV8-associated multicentric Castleman disease
Primary effusion lymphoma
Borderline cases
B-cell lymphoma, unclassifiable, with features intermediate between DLBCL and Burkitt lymphoma
B-cell lymphoma, unclassifiable, with features intermediate between DLBCL and classical Hodgkin lymphoma

Source: Swerdlow *et al*., 2008. Reproduced with permission of WHO.



(a)

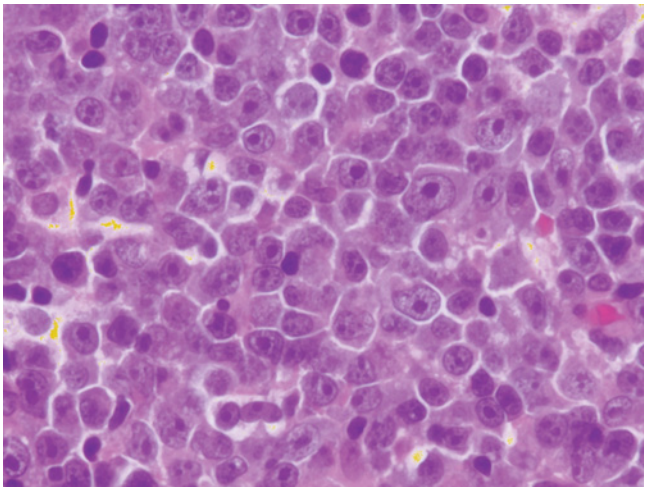


(b)

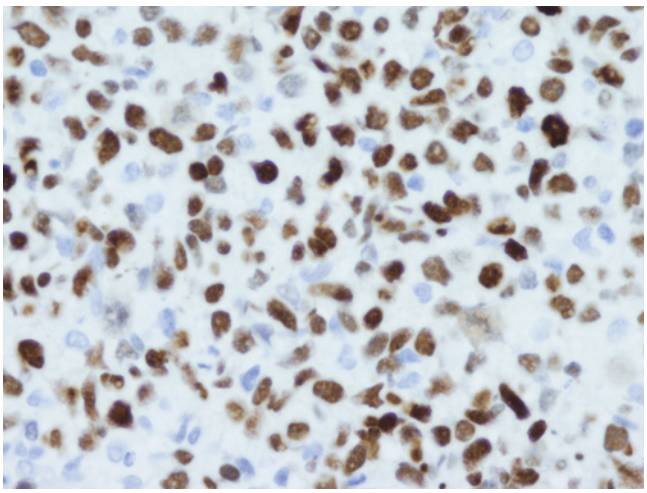
Figure 31.10 Diffuse large B-cell lymphoma (a) positive for CD10 (b) suggests a germinal-centre cell origin.

'double hit' and have a very poor prognosis. The recognition of this subset of tumours has been recommended to investigate alternative therapeutic strategies. DLBCL with double expression of MYC and BCL2 proteins have also poor prognosis, but the evolution is not as aggressive as the 'double genetic hit' tumours and it is not clear whether these patients require different treatments.

Gene expression profiling of DLBCL, NOS has identified two major molecular subtypes that express genes related to germinal-centre cells (GCB) or activated B cells (ABC), respectively. These subtypes of DLBCL differ in genetic, molecular and clinical aspects, indicating that they correspond to different biological entities. ABC, but not GCB, DLBCL exhibits constitutive activation of the NF-κB pathway. The clinical outcome is also different, with 5-year survival rates of 59% and 30% in GCB and ABC DLBCL, respectively. The prognostic value of these molecular subtypes has been retained in patients treated with rituximab-containing regimens. The recognition of these molecular subtypes of DLBCL may be relevant to implement new therapeutic strategies targeting the pathogenic differences. NGS has identified a large number of mutated genes involved in chromatin remodelling, such as *CREBB/EP300* (39%), *MML2* (24-32%), *EZH2* (32%) and *MEF2B* (16%), and genes activating the NF-κB pathway, such as *MYD88* (29%), or *CARD11*. The first group predominate in GC-DLBCL whereas the later are more common in ABC-DLBCL.



(a)



(b)

Figure 31.11 Diffuse large B-cell lymphoma with immunoblastic morphology (a) and positivity for MUM1/IRF4 (b) suggesting a non-germinal-centre cell origin.

T-cell/histiocyte-rich large B-cell lymphoma

T-cell/histiocyte-rich large B-cell lymphoma is characterized by a limited number of scattered large B cells immersed in a rich background of T cells and frequently histiocytes. The tumour cells express mature B-cell markers and frequently BCL-2 and EMA, although CD30, CD15 and CD138 are negative. The background is composed of CD3 and CD4 T cells and CD68-positive histiocytes. T-cell rosettes surrounding tumour B cells are not seen. Cases positive for Epstein–Barr virus (EBV) and a similar morphology would be better classified as EBV-positive DLBCL. This tumour presents with disseminated disease involving lymph nodes but also spleen, liver and bone marrow. Failure of therapy and International Prognostic Index (IPI) score are predictors of survival.

DLBCL with a predominant extranodal location

The WHO classification recognizes a series of DLBCL subtypes and entities characterized by predominantly extranodal presentation (Table 31.4).

Primary mediastinal (thymic) large B-cell lymphoma

Primary mediastinal (thymic) large B-cell lymphoma (PMBL) seems to originate in a thymic B cell and predominates in young women presenting with a large mediastinal mass that frequently invades adjacent structures. Progression outside the mediastinum frequently involves extranodal sites such as kidney, liver, adrenal or central nervous system (CNS). The tumour is mainly composed of large cells with abundant pale cytoplasm, which express mature B-cell markers, but usually lack surface immunoglobulin. CD30 is positive in 80% of cases, although not as uniformly as in HL. CD15 is negative. Genetically, PMBL has frequent gains and amplifications of 9p24 and inactivating mutations of *SOCS1*. These tumours also carry translocations of the MHC class II transactivator *CIITA* and the immunoresponse modulator *PDL1* and activating mutations of the phosphatase *PTPN1*. Intriguingly, all these alterations may also be seen in a proportion of Hodgkin lymphomas highlighting the genetic relationship between these entities. These tumours have a distinctive expression profile, relatively similar to that of HL, confirming the pathological evidence of the relationship between these two tumours. In fact, some patients may present with combined HL and PMBL at diagnosis or relapse, or have tumours with intermediate features of both (see below). The outcome of these tumours is more favourable than other DLBCL, with a 5-year survival of 65%.

Intravascular large B-cell lymphoma

Intravascular large B-cell lymphoma is a distinctive lymphoma characterized by the growth of large B cells within the lumina of small- to medium-sized vessels and capillaries (Figure 31.12). It is relatively rare in Western countries, but more common in Eastern populations. This lymphoma is very aggressive and frequently diagnosed only at post mortem. An isolated cutaneous variant with better prognosis has been identified, mainly in women.

Primary cutaneous DLBCL, leg type

Primary cutaneous DLBCL, leg type, is a primary cutaneous lymphoma composed almost exclusively of atypical large B cells that commonly presents in the lower extremities. These tumours express an activated B-cell phenotype with positivity for IRF4/MUM1 and CD10 negativity. BCL-6 is frequently positive and BCL-2 is strongly expressed. The genetic and molecular profiles are relatively similar to those found in ABC DLBCL.

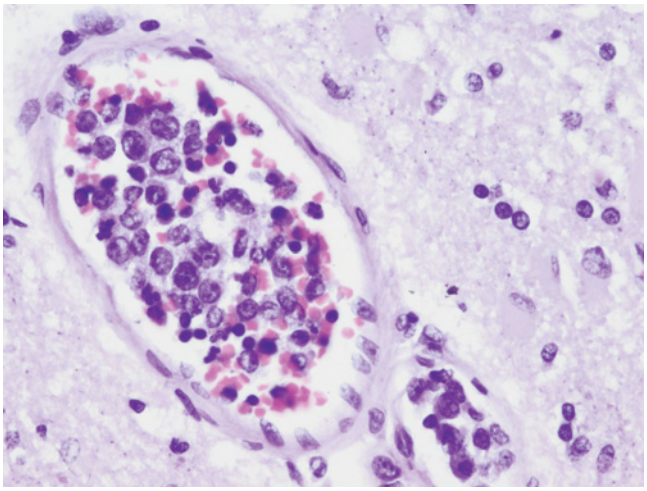


Figure 31.12 Intravascular lymphoma in the brain.

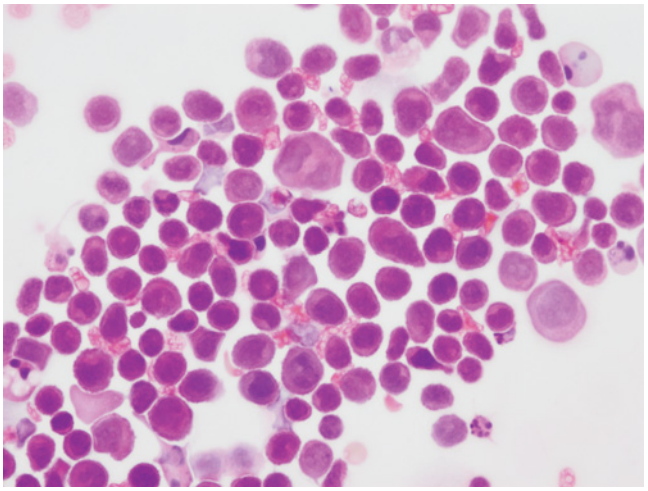


Figure 31.13 Primary effusion lymphoma. The tumour cells are present in the pleura and have a large plasmablastic morphology.

The prognosis is poor and the tumour tends to disseminate to extracutaneous sites.

Large-cell lymphomas of terminally differentiated B cells

The updated WHO classification has included different lymphoma entities that have in common the proliferation of large lymphoid cells with a terminally differentiated B-cell phenotype characterized by variable or total lack of CD20 expression, and less frequently CD79a, but which express plasma-cell-associated antigens such as CD38 and CD138. Most of these tumours are infected with EBV or human herpesvirus (HHV)-8 and occur in patients with a certain immunodeficiency status.

ALK-positive large B-cell lymphoma

ALK-positive large B-cell lymphoma is a distinctive disease characterized by proliferation of large B cells with plasmablastic differentiation that express ALK and are EBV negative. These tumours present in young immunocompetent patients, usually with nodal involvement and aggressive behaviour. ALK expression is due to *ALK* rearrangements, particularly with clathrin [t(2;17)] or less frequently nucleophosmin [t(2;5)]. Contrary to other lymphomas in this group, viral infection and immunodeficiency are absent.

Plasmablastic lymphoma

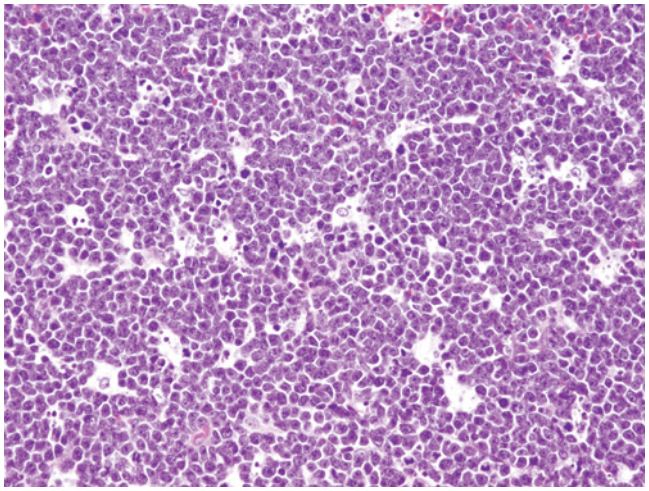
Plasmablastic lymphoma is a large B-cell lymphoma with immunoblastic morphology and plasma-cell immunophenotype that presents mainly in extranodal sites such as oral mucosa and gastrointestinal tract of immunosuppressed patients. HIV infection is the main cause, but post-transplant or immunosuppressive treatments are also common causes. Most of the tumours are EBV positive and latent membrane protein (LMP)-1 negative. The clinical behaviour is aggressive, with poor response to therapy. *MYC* is rearranged in around 50% of the cases.

Primary effusion lymphoma

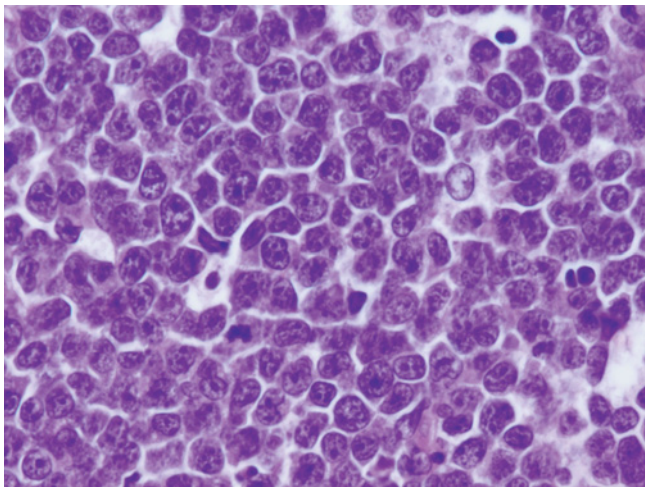
Primary effusion lymphoma occurs mainly in immunosuppressed patients associated with HHV-8 and, frequently, EBV infection. It is characterized by proliferation of large pleomorphic B cells that lack expression of B-cell markers and immunoglobulin, but which express plasma-cell-associated antigens and CD30 (Figure 31.13). These tumours usually present as effusion lymphomas, with no involvement of tissues. However, some patients may present with, or later develop, a solid tumour in extranodal sites.

Burkitt lymphoma

Burkitt lymphoma (BL) is mainly a tumour of children and young adults and is characterized by monotonous proliferation of medium-sized B cells with a mature B and germinal-centre phenotype, negative or very weak BCL-2 expression, high proliferation (Ki-67 > 95%) and the t(8;14) translocation with *MYC* rearrangement. These tumours occur endemically in equatorial Africa and other geographic areas and sporadically throughout the world. BL is also seen associated with immunodeficiency, particularly HIV infection. The tumour mainly involves extranodal sites, particularly the abdominal cavity, and, in endemic areas, jaws and facial bones. The tumour may infiltrate the CNS. A leukaemic phase may be seen in patients with bulky disease and occasional patients may present with a pure leukaemic disease (Burkitt leukaemia variant). The tumour



(a)



(b)

Figure 31.14 Burkitt lymphoma. (a) 'Starry sky' pattern due to the abundant histiocytes with apoptotic bodies. (b) Monotonous medium-sized cells with high proliferation.

is clinically aggressive, but potentially curable with current protocols.

Morphologically, the tumour cells are very monomorphic. The high proliferation is typically associated with a 'starry sky' pattern because of the high number of histiocyte phagocytosing apoptotic bodies (Figure 31.14). Genetically, the t(8;14) translocation is found in most tumours with very few if any additional genetic alterations. NGS studies have identified a distinctive profile of somatic mutations targeting the *ID3* (11%)/TCF3 (40–70%) pathway not detected in DLBCL and CCND3 (38%). These mutations promote survival and proliferation of tumour cells. The gene expression profile of BL differs from that observed in DLBCL with over-expression of *MYC* target genes, high expression of germinal-centre-related genes and low expression of NF-κB target genes and MHC class I genes.

B-cell lymphoma, unclassifiable, with features intermediate between DLBCL and Burkitt lymphoma

Borderline cases between DLBCL and BL comprise tumours with a morphology that resembles BL, but which have an atypical phenotype, such as strong BCL-2 protein expression, tumours with blastic nuclear morphology, but negative for terminal deoxynucleotidyl transferase and cyclin D1, or tumours with a BL phenotype, but with more irregular nuclei than acceptable for this lymphoma. Some of these cases were diagnosed previously as Burkitt-like lymphomas. These tumours express a mature B-cell phenotype with a high proliferative index. Strong expression of BCL-2 protein should suggest the presence of t(14;18) in addition to t(8;14) ('double hit'). *MYC* rearrangements may occur with a non-immunoglobulin gene and may be associated with complex karyotypes. Lymphomas with typical morphology of DLBCL with high proliferative activity and/or t(8;14) should not be included in this category. Clinically, the patients are older than those with conventional BL, the disease is usually disseminated and the bone marrow and peripheral blood may be involved. The clinical evolution is very aggressive, with poor response to conventional therapy. These cases are not considered a specific entity but a working category. Some cases may represent transformed FL, evolving DLBCL activating particular molecular pathways. However, independent of the possible biological significance, these patients should be recognized and studied separately from conventional DLBCL and BL.

B-cell lymphoma, unclassifiable, with features intermediate between DLBCL and classical Hodgkin lymphoma

These tumours are B-cell lymphomas that have clinical, morphological and/or immunophenotypic features that overlap those of DLBCL, particularly PMBL, and HL. These cases are diagnosed frequently in young men presenting with a large mediastinal mass that is sometimes associated with surpraclavicular lymph nodes. Most of these cases are composed of sheets of large cells with a pleomorphic aspect, sometimes lacunar appearance, associated with some inflammatory infiltrate. The phenotype combines strong expression of a complete B-cell programme (CD20, CD79a, BOB-1, OCT2, PAX5) with CD30 and CD15 positivity. Some cases resembling classical nodular sclerosis HL with uniform expression of CD20 and other B-cell markers and lack of CD15 may also be included in this category. In contrast, cases resembling PMBL that lack CD20 and are positive for CD30 and CD15 may also be diagnosed in this category. Composite lymphomas with clear areas of PMBL and HL should not be included in this group and both components should be mentioned in the diagnosis. The overlapping morphological and phenotypical features of these tumours have been

Table 31.5 Immunophenotypic features of common T-cell neoplasms.

Neoplasm	CD3S; C	CD5	CD7	CD4	CD8	CD30	CXCL13	TCR	CD56	Cytotoxic granules	EBV
T-PLL	+	–	+, +	+/–	–/+	–	–	αβ	–	–	–
T-LGL	+	–	+, +	–	+	–	–	αβ	+, –	+	–
NK-LGL	–	–	+, –	–	+/–	–	–	–	–, +	+	+
Extranodal NK/T-cell lymphoma	–; +	–	–/+	–	–	–	–	–	NA, +	+	++
Hepatosplenic T-cell lymphoma	+	–	+	–	–	–	–	γδ > αβ	+, –/+	+	–
Enteropathy-type T-cell lymphoma	+	+	+	–	+/–	+/–	–	αβ >> γδ	–	+	–
Mycosis fungoides	+	+	–/+	+	–	–	–	αβ	–	–	–
Subcutaneous panniculitis like T-cell lymphoma	+	+/–	+/–	+/–	–	++	–	αβ	–	–/+	–
Primary cutaneous αβ T-cell lymphoma	+	+	+	–	+	–/+	–	γδ	–, +/–	+	–
PTCL-NOS	+/–	+/–	+/–	+/–	–/+	–/+	–	αβ>γδ	–/+	–/+	–/+
Angioimmunoblastic T-cell lymphoma	+	+	+	+/–	–/+	–	+	αβ	–	NA	+/–
ALCL	+/–	+/–	NA	–/+	–/+	++	–	αβ	–	+	–

TCR, T-cell receptor gene; NA, not available; Cytotoxic granules, TIA-1, perforin and/or granzyme.

validated as authentic biological characteristics by microarray expression profiling studies that have demonstrated the similarities between PMBL and HL. These tumours probably represent evidence of plasticity in the differentiation pathways of B cells.

Mature NK-cell/T-cell neoplasms (see also Chapter 28)

These tumours derive from natural killer (NK) cells and T cells of peripheral lymphoid organs and can be roughly subdivided into three groups: leukaemic, extranodal and nodal. The combination of morphological, phenotypic and molecular features are essential for the diagnosis of the different entities (Table 31.5) These are discussed in the following sections, with the exception of T-cell prolymphocytic leukaemia, T-cell large granular lymphocyte leukaemia and chronic lymphoproliferative disorders of NK cells, which are considered in Chapter 28.

Aggressive NK-cell leukaemia

This is a rare neoplasm, mainly observed among Asians, and exhibits systemic proliferation of NK cells almost always associated with EBV and an aggressive clinical course. Patients are usually middle-aged of either sex. The most commonly involved sites are the peripheral blood, bone marrow, liver and spleen (Figure 31.15). Fever and constitutional symptoms are common. Frequent complications are coagulopathy, haemophagocytic syndrome (HPS) or multiorgan failure. A specific overlap with extranodal NK/T-cell lymphoma may exist. Circulating leukaemic cells exhibit a morphological spectrum, from 'normal LGL' to elements with atypical infolded nuclei, open chromatin and distinct nucleoli. They express CD2, CD3ε, CD56 and cytotoxic molecules. Surface CD3 is absent. T-cell receptor (TCR) genes are in germline configuration. Most cases pursue a fulminant clinical course.

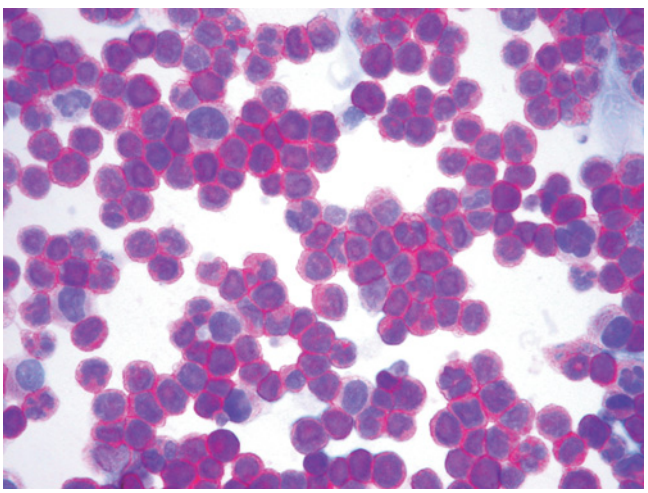


Figure 31.15 Cytospin from the peripheral blood of a patient with aggressive NK-cell leukaemia, stained for CD56 (alkaline phosphatase anti-alkaline phosphatase technique).

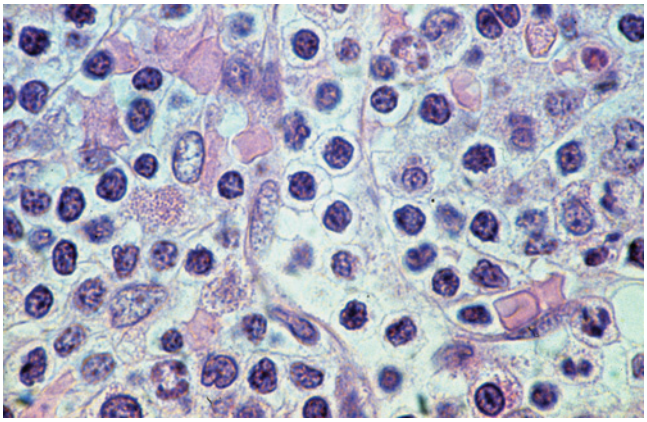


Figure 31.16 Bone-marrow involvement by EBV-positive lymphoproliferative disorder in a 16-year-old boy.

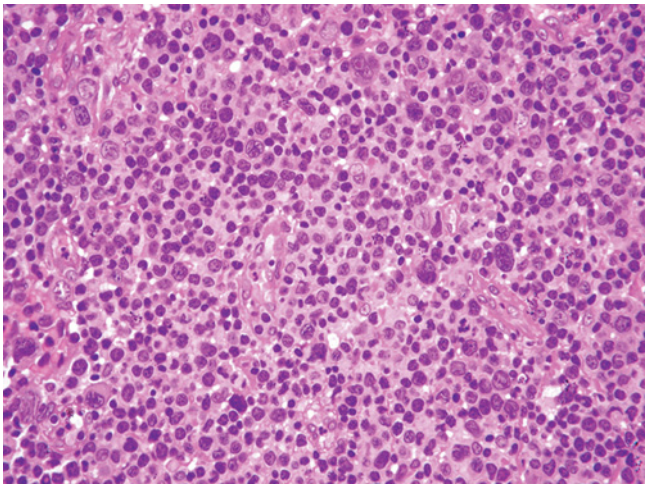


Figure 31.17 Lymph node infiltration in a patient with adult T-cell leukaemia/lymphoma.

EBV-positive T-cell lymphoproliferative disorders of childhood

Systemic EBV-positive T-cell lymphoproliferative disease of childhood and Hydroa vacciniforme-like lymphoma are the prototypes of these diseases. Both are prevalent in Far East Asia and people with native-American ancestry, with no sex predilection. EBV-positive T-cell lymphoproliferative disease of childhood can develop shortly after primary EBV infection or in the setting of chronic active EBV infection. It may have a rapid evolution with fatal outcome. Hydroa vacciniforme-like lymphoma is associated with sensitivity to insect bites and sun exposure and behaves indolently. In both conditions, there is strong suggestion of a genetic defect in the host immune response to EBV. The infiltrating T cells are usually small and lack significant cytological atypia, but at times are pleomorphic, medium/large with irregular nuclei and frequent mitoses (Figure 31.16). They are positive for CD2, CD3, TIA and EBV, but negative for CD56, and have monoclonally rearranged *TCR* genes. Occasional cases with γδ-TCR or NK characteristics have been reported.

Adult T-cell leukaemia/lymphoma (see also Chapter 34)

Adult T-cell leukaemia/lymphoma (ATLL) is endemic in south-western Japan, the Caribbean basin, Iran and parts of Central Africa, its distribution resembling that of human T-cell leukaemia/lymphoma virus (HTLV)-I. The disease has a long latency with exposure of affected individuals to the virus very early in life. Although causally linked to ATLL, HTLV-I itself is insufficient to cause neoplastic transformation, additional genetic events being required. Most patients present with widespread lymph node and peripheral blood involvement, the skin being affected in more than 50%. ATLL is characterized by variable clinical and histological features. In particular, the behaviour varies from acute (with usual hypercalcaemia and possible lytic bone lesions) to lymphomatous, chronic or smouldering, while the growth may consist of small- to large-sized pleomorphic cells or display anaplastic or even angioimmunoblastic-like morphology (Figure 31.17). Tumour elements are usually $CD2^+CD3^+CD5^+CD25^+CD7^-$. Most cases are $CD4^+CD8^-$, a few $CD4^-CD8^+$ or $CD4^+CD8^+$. CCR4 and FoxP3 are frequently expressed, while cytotoxic molecules are absent. CD30 positivity is at times detected, especially in cases lacking FoxP3, by heralding a more aggressive clinical course. *TCR* genes are clonally rearranged. Clinical subtypes, age, performance status, serum calcium and lactate dehydrogenase levels are major prognostic indicators. The acute and lymphomatous variants have a short survival (2 weeks to more than 1 year), while the chronic and smouldering forms behave more indolently, but can progress to an acute phase (see also Chapter 30).

Extranodal NK/T-cell lymphoma, nasal type (see also Chapter 34)

Extranodal NK/T-cell lymphoma, nasal type, more often occurs in adult males from Far East Asia, and from native-American ancestry. The tumour is EBV positive with type II latency, although LMP-1 expression is variable. Usually, these lymphomas affect the nasal cavity, nasopharynx, paranasal sinuses and palate. Extranasal locations include the skin, soft tissue, gastrointestinal tract and testis. Patients with nasal disease suffer from nasal obstruction or epistaxis, or extensive mid-facial destruction (so-called lethal midline granuloma). Irrespective of the site of presentation, the lymphomatous infiltrate is diffuse with angiocentric and angiodestructive features and frequent

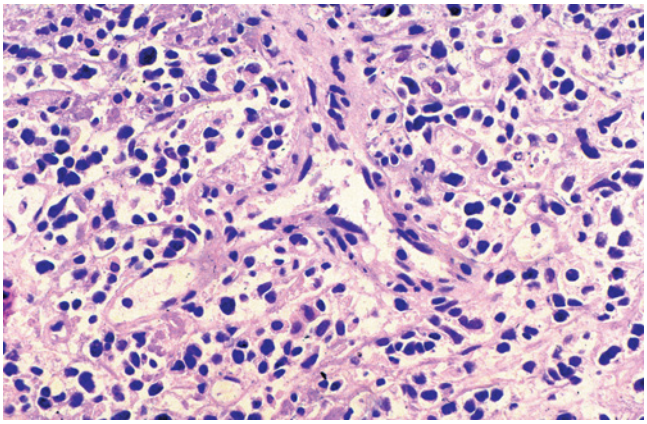


Figure 31.18 Destructive lesion of the nasal cavities in the course of extranodal NK/T-cell lymphoma, nasal type.

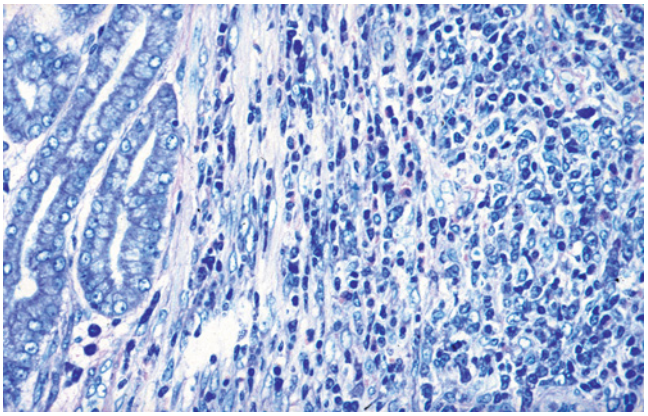


Figure 31.19 Enteropathy-associated T-cell lymphoma. Remnants of the mucosa are seen on the left-hand side.

necrosis. Neoplastic cells may be small, medium-sized, large or anaplastic with numerous mitotic figures and azurophilic granules in Giemsa-stained touch preparations (Figure 31.18). Phenotypically, they are positive for CD2, CD56, CD3ε and cytotoxic markers, but negative for CD3. Occasional cases are $CD7^+$ or $CD30^+$. It is noteworthy that CD56 is not specific for these lymphomas, being detected in other peripheral T-cell lymphomas, mainly carrying γδ-TCR. A diagnosis of extranodal NK/T-cell lymphoma, nasal type should be accepted with scepticism if EBV-encoded RNA *in situ* hybridization is negative. *TCR* and *IGHV* genes are usually in germline configuration, the former being clonally rearranged in a few instances. The prognosis of extranodal NK/T-cell lymphoma, nasal type is variable, being always poor for cases occurring outside the nasal cavity. It is noteworthy that survival has recently been improved with more intensive regimens, including up-front radiotherapy and L-asparaginase.

Enteropathy-associated T-cell lymphoma (see also Chapter 28)

In its classical form, enteropathy-associated T-cell lymphoma (EATL) is an intestinal tumour of intraepithelial T lymphocytes, usually consisting of large lymphoid cells admixed with numerous histiocytes and eosinophils and associated with necrosis. The adjacent small intestinal mucosa shows villous atrophy, crypt hyperplasia, increased lamina propria lymphocytes and plasma cells, and intraepithelial lymphocytosis (Figure 31.19). The disease occurs more often in areas with a high prevalence of coeliac disease (i.e. northern Europe) and has a strong association with HLA DQ2 or D8 phenotype. Patients (usually adults) present with abdominal pain, frequently associated with intestinal perforation. The tumour forms one or more ulcerating mucosal masses that invade the wall of the intestine (frequently jejunum or ileum). Neoplastic cells are positive for CD3, CD7, CD103, TCR-β and cytotoxic molecules, negative for CD4, CD5, and MATK, with variable expression of CD8 and CD30. *TCR*β and γ genes are always clonally rearranged. Cytogenetics reveals +9q31-qter or -16q12 and frequent +1q32-q41 and +5q34-q35. The prognosis is poor, with death frequently resulting from abdominal complications. An indolent form of T-cell lymphoproliferative disease of the gastrointestinal tract has recently been described that should be differentiated from EATL. The tumour cells are small and non-destructive, usually limited to the mucosa and submucosa and most cases do not have intraepithelial infiltration.

In the fourth edition of the WHO Classification, a variant of the process is recognized (termed type II), which corresponds to 10–20% of EATL cases. It may occur sporadically, without risk factors for coeliac disease or specific HLA phenotype, and has a broader geographic distribution, representing the only form observed among Asians, Hispanics and American indigenous populations. This is composed of monomorphic medium-sized cells with no inflammatory background and rare necrosis. Enteropathic changes in the surrounding or adjacent mucosa consists of focal or extensive epitheliotropism in the absence of villous atrophy and crypt hyperplasia. 'Type II EATL' has a distinctive immunophenotype ($CD3^+CD4^-CD8^+CD56^+$, $MATK^+$). In a significant proportion of cases, neoplastic cells carry γδ-TCR. On cytogenetic grounds, besides +9q31-qter or 16q12, +8q24 (with amplification of *MYC* locus) does frequently occur. At the present time, there is a great deal of discussion of whether to maintain the name type II EATL as well as the inclusion of this tumour in the same category as classical EATL.

Hepatosplenic T-cell lymphoma

Hepatosplenic T-cell lymphoma is a rare neoplasm derived from cytotoxic T cells, mostly of the γδ type. It affects adolescents

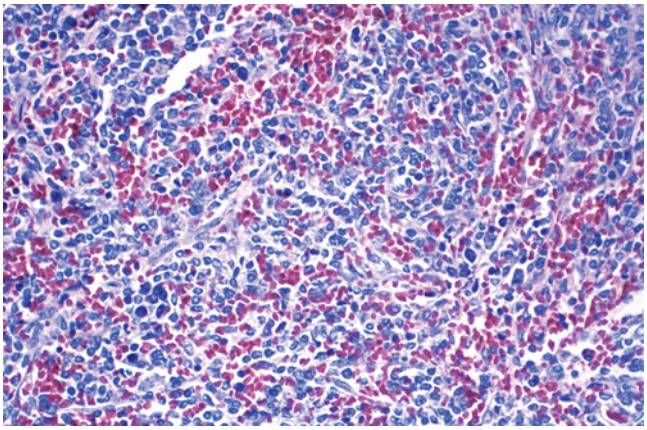


Figure 31.20 Splenic red pulp involvement by hepatosplenic T-cell lymphoma.

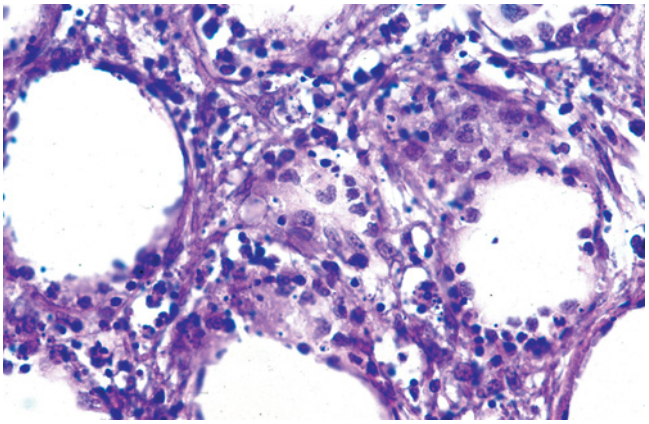


Figure 31.21 Example of subcutaneous panniculitis-like T-cell lymphoma. Features of 'rimming' are easily seen.

and young adults, with male predominance. Up to 20% of these lymphomas arise in the setting of chronic immuno-suppression, usually for solid organ transplantation or prolonged antigenic stimulation. It may also occur in patients, especially children, treated with azathioprine and infliximab for Crohn's disease. Clinically, it is characterized by hepatosplenomegaly, systemic symptoms and marked thrombocytopenia. Morphologically, neoplastic cells are monotonous, with medium-sized nuclei, loosely condensed nuclear chromatin, inconspicuous nucleoli and a rim of pale cytoplasm. In the spleen, they infiltrate the red pulp cords and sinuses with atrophy of the white pulp, while in the liver and bone marrow they show intrasinusoidal diffusion (Figure 31.20). Phenotypically, most cases are positive for CD3 and TCR-δ1, partly positive for CD56 and CD8, and negative for TCR-αβ, CD4 and CD5. A minority carries the α/β dimer. The cells express TIA1 and granzyme-M, but not granzyme-B and perforin. They have rearranged TCR-γ genes. Cases of γδ origin show biallelic rearrangement of TCR-δ genes. Isochromosome 7q is present in most instances. Trisomy 8 and loss of a sex chromosome may also be present. The course is aggressive, with initial response to chemotherapy, but rapid relapse. The median survival is less than 2 years.

Subcutaneous panniculitis-like T-cell lymphoma

Subcutaneous panniculitis-like T-cell lymphoma is a rare cytotoxic T-cell lymphoma that preferentially infiltrates subcutaneous tissue, is slightly commoner in females and has a broad age spectrum. In the new WHO classification, cases expressing TCR-γδ are classified as primary cutaneous γδ T-cell lymphoma. About 20% of patients have an associated autoimmune disease, mostly systemic lupus erythematosus. Patients present with multiple subcutaneous nodules, usually in the absence of other involved sites. Systemic symptoms are recorded in more than 50% of patients. Cytopenias and elevated liver function tests are common, and frank HPS is seen in 15–20%. The infiltrate involves the fat lobules with riming around individual fat cells and usually spares septa as well as the overlying dermis and epidermis (Figure 31.21). Vascular invasion is seen occasionally, while necrosis and karyorrhexis are common. The cells have a mature αβ T-cell phenotype ($\beta F1^+$), usually $CD8^+$ with expression of cytotoxic molecules and negative for CD56. The neoplastic cells show rearrangement of TCR genes, and are negative for EBV. Dissemination to lymph nodes and other organs is rare. The 5-year overall survival (OS) is 80%; however, if HPS occurs, the prognosis is poor.

Mycosis fungoides and Sézary syndrome (see also Chapter 28)

Mycosis fungoides is the commonest primary cutaneous T-cell lymphoma (CTCL) and is characterized by epidermotropic infiltrates of small to medium-sized T lymphocytes with cerebriform nuclei (Figure 31.22). It usually affects adults or the elderly with a male predominance, but can also occur in children and adolescents. The disease is limited to the skin, with classical evolution of patches, plaques and tumours. Extracutaneous dissemination occurs in advanced stages. Mycosis fungoides slowly progresses over years or sometimes decades. Histological transformation (i.e. >25% blasts in the dermal infiltrates) occurs mainly in the tumour stages. The typical phenotype is $CD2^+CD3^+$TCR-$\beta^+CD5^+CD4^+CD8^-CD7^-$. Recently, some reports have highlighted occasional cases expressing follicular T-helper (FTH) markers or γδ T-cell phenotype. *TCR* genes are clonally rearranged in a proportion of patients. The single most important prognostic factor is the clinical stage. In the more advanced stages, the prognosis is poor, as it is in the case of histological transformation. Rare clinical–morphological variants

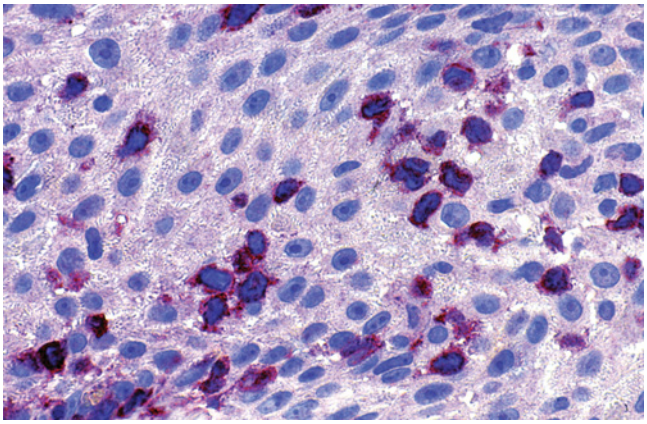


Figure 31.22 Mycosis fungoides. The epidermis is infiltrated by neoplastic T cells with cerebriform nuclei and strong CD3 expression.

include folliculotropic mycosis fungoides, pagetoid reticulosis and granulomatous slack skin.

Sézary syndrome is defined by erythroderma, generalized lymphadenopathy and presence of neoplastic T cells with cerebriform nuclei (so-called Sézary cells) in skin, lymph nodes and blood. This is a rare disease accounting for less than 5% of CTCLs. It occurs in adults, with a male predominance. The histological and phenotypic features in Sézary syndrome may be similar to those of mycosis fungoides. TCR genes are clonally rearranged. The 5-year OS is 10–20%.

Primary cutaneous CD30-positive T-cell lymphoproliferative disorders

These lesions include primary cutaneous anaplastic large-cell lymphoma (C-ALCL) and lymphomatoid papulosis, which form a spectrum with overlapping histopathological and phenotypic features. Their appearance and course are therefore critical for the diagnosis.

Primary cutaneous anaplastic large-cell lymphoma

C-ALCL is composed of large cells with an anaplastic, pleomorphic or immunoblastic morphology, the vast majority of which express CD30. By definition, C-ALCL should not be preceded by or coincide with mycosis fungoides. It must also be distinguished from systemic ALCL, which is a different entity. Most patients present with solitary or localized nodules or tumours, and sometimes papules located on the trunk, face, extremities or buttocks that often undergo ulceration and sometimes show partial or complete spontaneous regression, as in lymphomatoid papulosis. Besides CD30 positivity, the neoplastic cells display an activated $CD4^+$ T-cell phenotype, with variable loss of CD2, CD5 and/or CD3. Unlike systemic ALCL, C-ALCL expresses CLA, but not EMA or ALK. The latter molecule has recently been reported in some exceptional cases, which run the same course as the negative ones. Of course, under these circumstances it is pivotal to exclude a systemic ALCL ALK+. Unlike Hodgkin and Reed–Sternberg cells, staining for CD15 is generally negative. Most cases show clonal *TCR* gene rearrangement. The prognosis is favourable, with a 10-year OS of approximately 90%, although local relapses are frequent.

Lymphomatoid papulosis

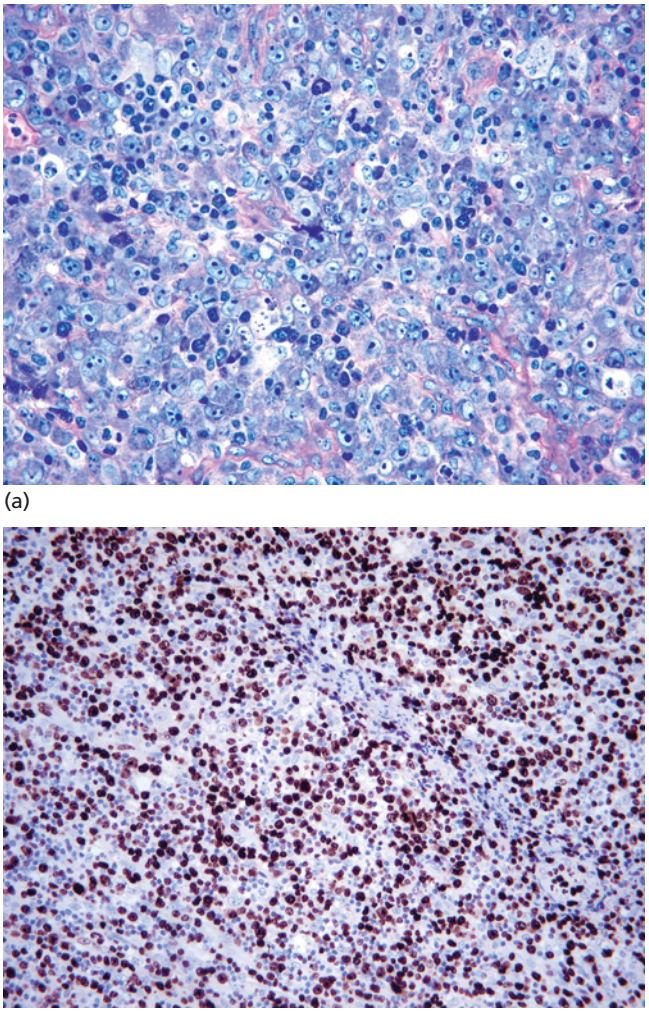
Lymphomatoid papulosis (LyP) is a chronic, recurrent, self-healing skin disease that most often occurs in adults, with male preponderance, affects trunk and extremities and is characterized by paradoxical eruptions of papular, papulonecrotic and/or nodular lesions that tend to disappear within 3–12 weeks, possibly leaving a scar. The duration of the disease may vary from several months to over 40 years. The prognosis is excellent. Morphologically, the large CD30-positive cells may be intermingled with a variable number of inflammatory cells or predominate as a monotonous population. The phenotype is similar to that of C-ALCL. Depending on the amount and distribution of atypical cells, five variants of LyP (A–E) are recognized, which can mimic other conditions, including mycosis fungoides and primary cutaneous CD8-positive aggressive epidermotropic cytotoxic T-cell lymphoma. More recently, a further variant has been reported, characterized by 6p25.3 rearrangement, as observed in ALCL ALK^- (see below). Clonally rearranged *TCR* genes are detected in about 60% of lesions.

Primary cutaneous γδ T-cell lymphoma

This tumour consists of a clonal proliferation of activated γδ T cells with a cytotoxic phenotype. A possibly related condition may affect mucosal sites. Primary cutaneous γδ T-cell lymphoma is rare (1% of CTCLs) and more often observed in the extremities of adults with no sex predilection. It may be predominantly epidermotropic with patches/plaques or may give rise to deep dermal or subcutaneous tumours, with or without epidermal necrosis and ulceration. Dissemination to mucosal and other extranodal sites is frequent. HPS may occur. 'B' symptoms are seen in most patients. The tumour cells characteristically exhibit a $\beta F1^-CD3^+CD2^+CD5^-CD7^{+/-}CD56^+$ phenotype with strong expression of cytotoxic proteins. The cells are strongly positive for TCR-δ with appropriate detection methods. The cells show clonal rearrangement of the TCR-γ and TCR-δ genes. TCR-β may be rearranged or deleted, but is not expressed. Primary cutaneous γδ T-cell lymphoma is resistant to multiagent chemotherapy and/or radiation and has a poor prognosis, with a median survival of approximately 30 months.

Peripheral T-cell lymphoma not otherwise specified (see also Chapter 28)

Peripheral T-cell lymphoma not otherwise specified (PTCL/NOS) is a heterogeneous group of tumours that corresponds with none of the T-cell entities defined in the WHO classification. It accounts for approximately 30% of peripheral T-cell lymphomas in Western countries. Most patients are adults, with a slight male predominance, and show peripheral lymphadenopathy and 'B' symptoms. Generalized disease is often encountered, with bone marrow, liver, spleen and extranodal tissue involvement. Leukaemic spread is uncommon. The cytological spectrum is extremely broad, from highly polymorphous to monotonous. Clear cells and Reed–Sternberg-like elements can be seen. An inflammatory background is often present (Figure 31.23). The differential diagnosis with angioimmunoblastic T-cell lymphoma (AITL) and ALCL ALK$^-$ may at times be difficult. In fact, some PTCLs/NOS share with AITL the relationship to FTH cells, as shown by specific marker expression, as well as recurrent gene mutations (see below). These cases, however, lack hyperplasia of both follicular dendritic cells (FDCs) and high endothelial venules (HEVs), typical of AITL. On the other hand, some PTCL, NOS tumours almost exclusively consist of CD30$^+$ elements, although they lack hallmark cells and cytotoxic profile. PTCL/NOS usually displays an aberrant T-cell phenotype, with frequent loss of CD5 and CD7. A CD4$^+$CD8$^-$ profile predominates in nodal cases. CD4$^+$CD8$^+$ or CD4$^-$CD8$^-$ is sometimes seen, as is CD8, CD56 and/or cytotoxic granule expression. CD52 is absent in 60% of cases. Aberrant CD20 and/or CD79a expression is occasionally encountered. In half of cases, 25–100% of lymphomatous elements carry CD30. Proliferation is high and Ki-67 rates over 80% are associated with poor outcome. *TCR* genes are clonally rearranged. Gene and miRNA expression profiling studies have revealed signatures distinct from those of other PTCLs. In addition, they have recently allowed the identification of morphologically unrecognizable subtypes, respectively related to FTH, T_H1, T_H2 and cytotoxic T-cells. Some of the deregulated genes might have prognostic and/or therapeutic implications. EBV integration is found in some cases. PTCL/NOS is highly aggressive with poor response to therapy, frequent relapses and 5-year OS of 20–30%, unless autologous bone-marrow transplantation is used in the first line.


(a)

(b)

Figure 31.23 Peripheral T-cell lymphoma not otherwise specified. (a) The tumour consists of large cells with immunoblastic-like appearance. (b) The proliferative rate as shown by Ki-67 marking is very high.

Angioimmunoblastic T-cell lymphoma

Angioimmunoblastic T-cell lymphoma (AITL) is characterized by a polymorphous infiltrate constantly involving lymph nodes with prominent proliferation of HEVs and FDCs (Figure 31.24). Spleen, liver, skin and bone marrow are also frequently affected. It occurs in middle-aged/elderly individuals of both sex and accounts for approximately 15–20% of peripheral T-cell lymphomas. Virtually all patients present with advanced-stage disease, systemic symptoms, polyclonal hypergammaglobulinaemia, circulating immune complexes, cold agglutinins with haemolytic anaemia, positive rheumatoid factor and anti-smooth-muscle antibodies. Skin rash is frequently observed. Other common findings are pleural effusion, arthritis and ascites. AITL is characterized by partial effacement of the lymph node architecture with sparing of the peripheral cortical sinuses. The neoplastic infiltrate is composed of small- to medium-sized lymphocytes, with clear to pale cytoplasm and mild cytological atypia. They often form small clusters around hyperplastic HEVs and FDCs, and are admixed with small reactive lymphocytes, eosinophils, plasma cells and histiocytes. An

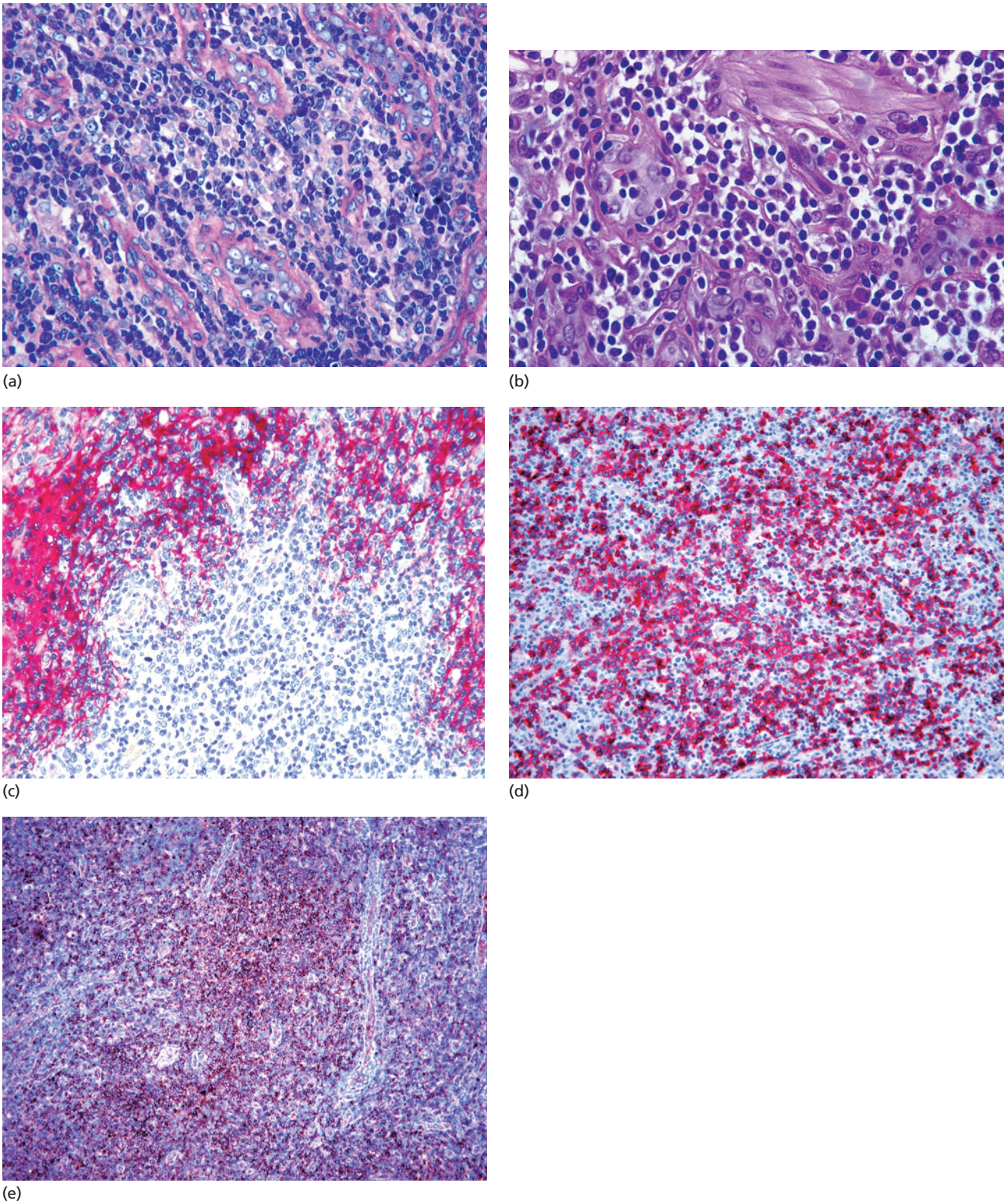
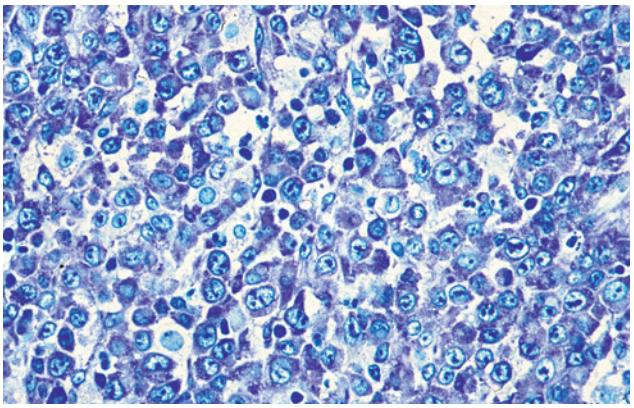


Figure 31.24 Angioimmunoblastic T-cell lymphoma. (a) Low-power view: note the amount of high endothelial venules as well as the cellular pleomorphism. (b) At higher magnification, neoplastic cells display irregular nuclear contours and a rim of clear cytoplasm. (c) Hyperplasia of follicular dendritic cells recognized with CD21 staining. The tumour cells are positive for CD10 (d) and CXCL13 (e).

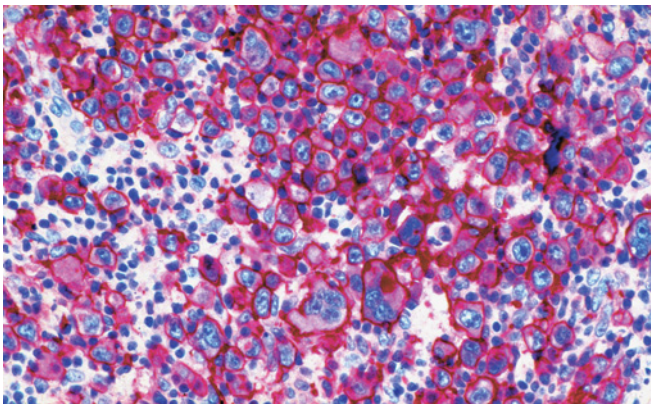
expansion of EBV-positive B immunoblasts is usually present. The neoplastic cells have the phenotype of FTH cells positive for mature T-cell markers, i.e. CD4, CD10, CXCL13, PD1, SAP and ICOS (Figure 31.24). *TCR* genes are clonally rearranged in most cases. Clonal *IGHV* gene rearrangements are found in about 25–30% of cases, and correlate with expanded EBV-positive B cells. Gene expression profiling (GEP) studies provided evidence of the FTH cell correlation of the tumour and proposed potential prognostic signatures. More recently, next generation sequencing (NGS) has identified a series of aberrations affecting the *RHOA*, *TET2*, *IDH2*, *DNM3TA*, *FYN*, and *ITK/SYK* genes. Notably, these GEP and WES findings have also been encountered in tumours that are at present classified as PTCL/NOS (see above); this questions whether or not a new category of FTH-related should be envisaged. The clinical course is aggressive with a median survival of less than 3 years. Patients often succumb to infectious complications. Supervening large B-cell lymphoma (often but not invariably EBV positive) can occur.

Anaplastic large-cell lymphoma, ALK$^+$

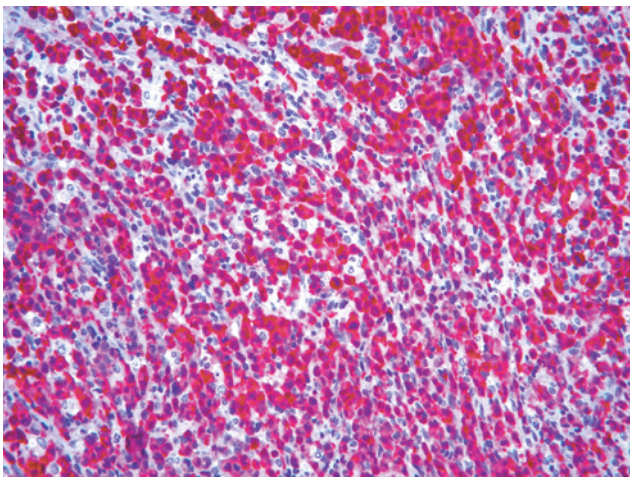
ALCL ALK$^+$ is characterized by large neoplastic cells with abundant cytoplasm and kidney/horseshoe-shaped nuclei (so-called hallmark cells), with a translocation involving the *ALK* gene and expression of both ALK protein and CD30. ALCL with similar morphology and phenotype, but lacking *ALK* rearrangement and ALK protein, is considered as a separate category (ALCL ALK$^-$). ALCL ALK$^+$ shows a higher prevalence in the first three decades of life and a male predominance, and frequently involves both lymph nodes and extranodal sites. Most patients present with stage III–IV disease and 'B' symptoms. The morphology may be variable. The 'common' appearance (60% of cases) consists of large tumour cells, frequently with hallmark appearance, that tend to grow within the sinuses, resembling a metastatic tumour (Figure 31.25). Some cases may have a lymphohistiocytic pattern (10%) characterized by few tumour cells overwhelmed by reactive histiocytes. A small-cell variant (5–10%) has been recognized that seems to have a worse prognosis. Some



(a)



(b)



(c)

Figure 31.25 Anaplastic large-cell lymphoma, ALK$^+$. (a) Some hallmark cells can be easily seen. (b, c) Expression of CD30 (b) and ALK protein (c) (alkaline phosphatase anti-alkaline phosphatase, Gill's counterstain, ×500 and ×250, respectively).

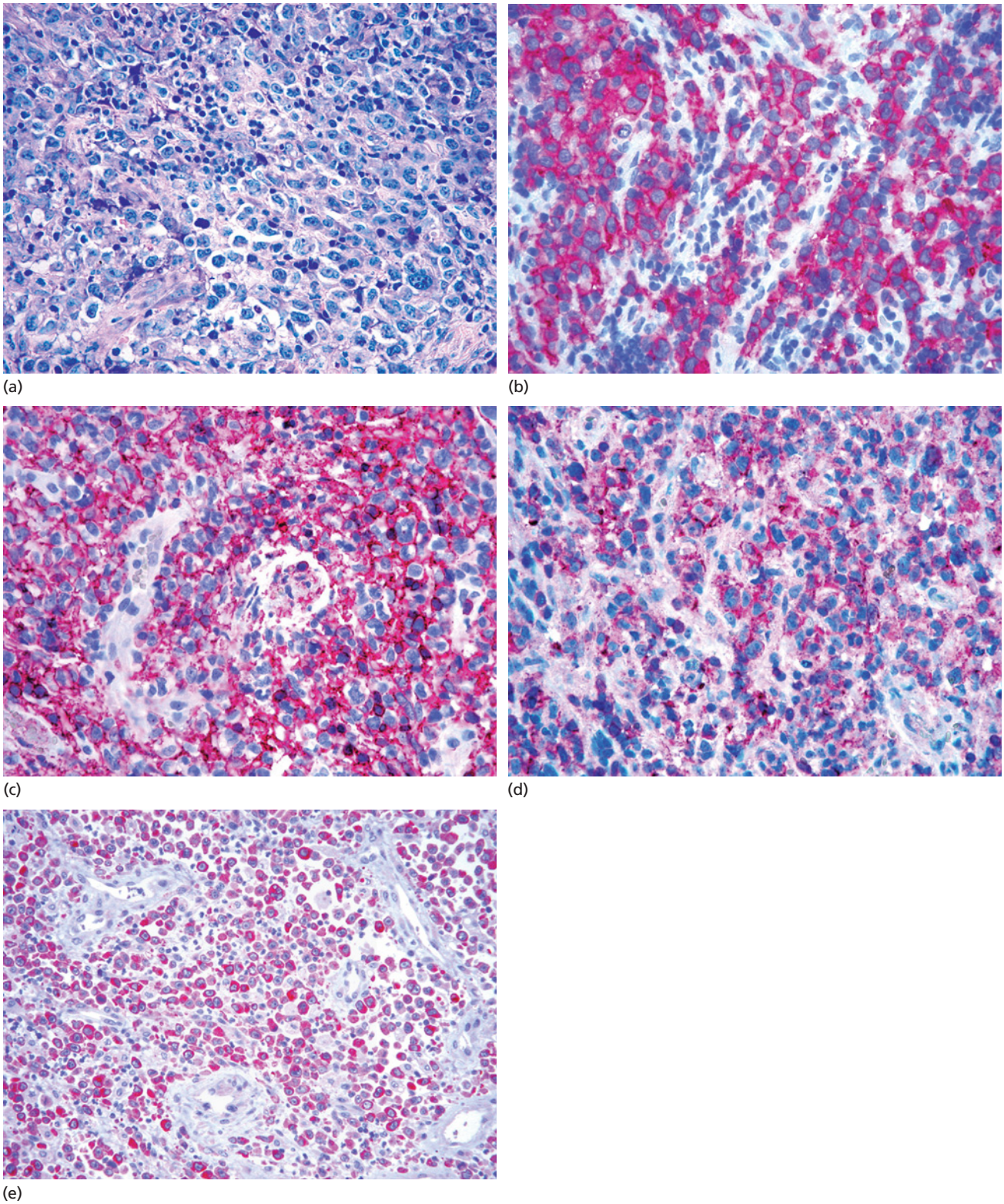


Figure 31.26 Anaplastic large-cell lymphoma, ALK$^-$. (a) The tumour consists of large cells partly showing hallmark appearance, and strong positivity for CD30 (b), CD45 (c), CD3 (d) and granzyme B (e).

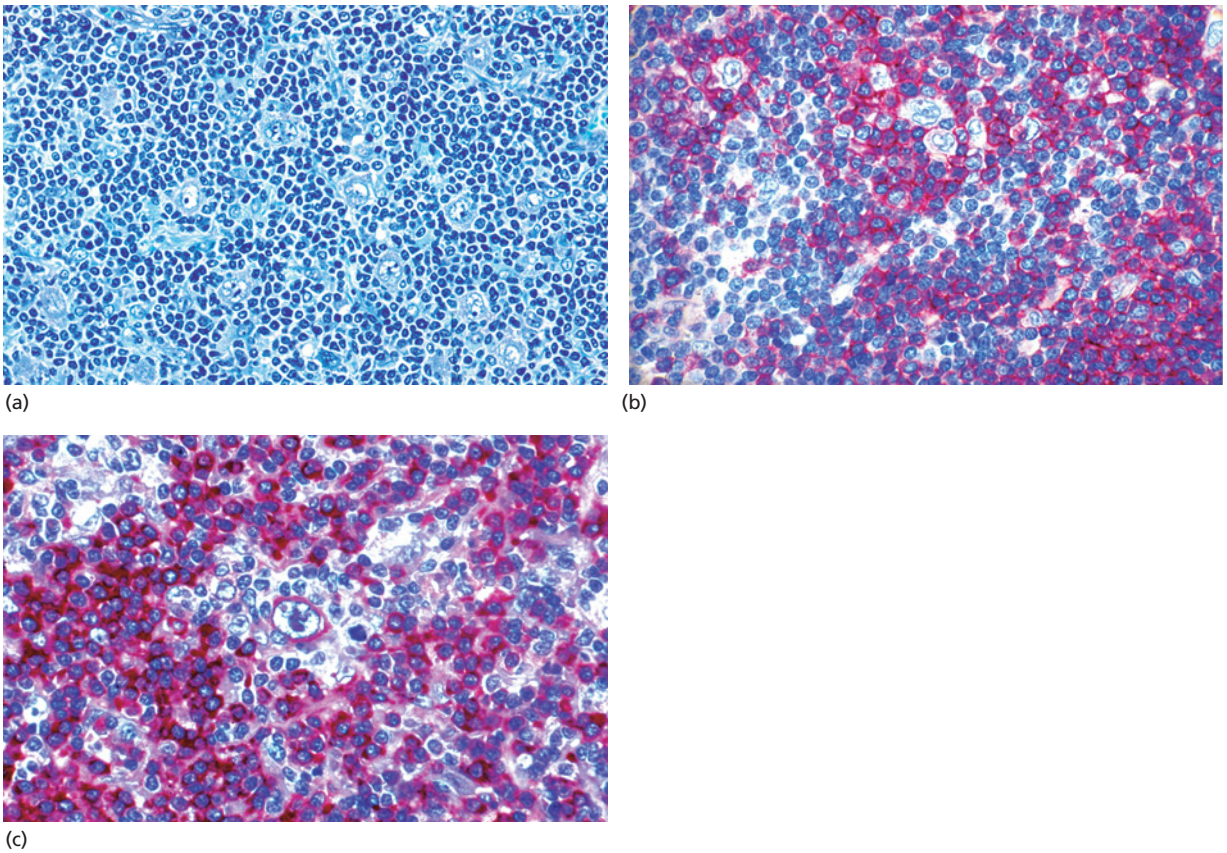


Figure 31.27 Nodular lymphocyte predominant (LP) Hodgkin lymphoma. (a) Typical LP cells within a milieu of small lymphocytes. (b) The LP tumour cells are surrounded by rosettes of $CD3^+$ T lymphocytes. (c) LP cells and most small lymphocytes express CD79a.

cases (3%) may mimic HL, but the cells are ALK positive and CD15 negative. Relapses may reveal morphological features different from those seen initially. The tumour cells are CD30 positive. In cases with t(2;5)/*NPM–ALK* translocation, ALK staining is both cytoplasmic and nuclear. In cases with variant translocations, ALK staining may be membranous or cytoplasmic. Most ALCL ALK^+ are positive for EMA, express cytotoxic markers and bear one or more T-cell antigens, although some may have null phenotype. EBV is negative and the *TCR* gene is rearranged. The most frequent genetic alteration is t(2;5)(p23;q35) involving the *ALK* and nucleophosmin (*NPM*) genes. Variant translocations lead to fusion of the *ALK* gene with other partners on chromosomes 1, 2, 3, 17, 19, 22 and X. All these aberrations result in upregulation and different subcellular distribution of ALK protein, but do not have clinical relevance. ALCL ALK^+, ALCL ALK^- and PTCL/NOS have different chromosomal and gene expression profiles. The IPI predicts the clinical outcome and the 5-year OS is about 80%.

Anaplastic large-cell lymphoma, ALK^-

In the new WHO Classification, ALCL ALK^- is a provisional entity that is morphologically and phenotypically indistinguishable from ALCL ALK^+ (Figure 31.26). The tumour affects adults, with a slight male predominance, involves lymph nodes and, less frequently, extranodal sites. Most patients are staged III–IV, with 'B' symptoms. Usually, the organ architecture is effaced by solid cohesive sheets of neoplastic cells. In the lymph node, the latter tend to diffuse through sinuses, mimicking metastatic carcinoma. Most cases show clonal *TCR* gene rearrangement and EBV is negative. The main differential diagnosis is with PTCL/NOS and classical HL. In this regard, BSAP/PAX5 staining is useful: classical HL shows weak positivity in most cases, a finding not observed in ALCL ALK^-. In contrast, the distinction between ALCL ALK^- and $CD30^+$PTCL/NOS is not always easy on conventional grounds, especially when no careful search for hallmark cells and cytotoxic markers is performed. Newly developed molecular tools based on gene expression and miRNA profiling can assist in this differential diagnosis, even when applied to formalin-fixed, paraffin-embedded (FFPE) material. Such distinction may be relevant. In fact, ALCL ALK^-, although more aggressive than ALCL ALK^+ (5-year OS 49% versus 80%) has a more favourable course than $CD30^+$ PTCL/NOS (5-year OS 32% versus 49%). In this respect, genomic studies have recently identified lesions allowing the prognostic stratification of ALCLs ALK^-

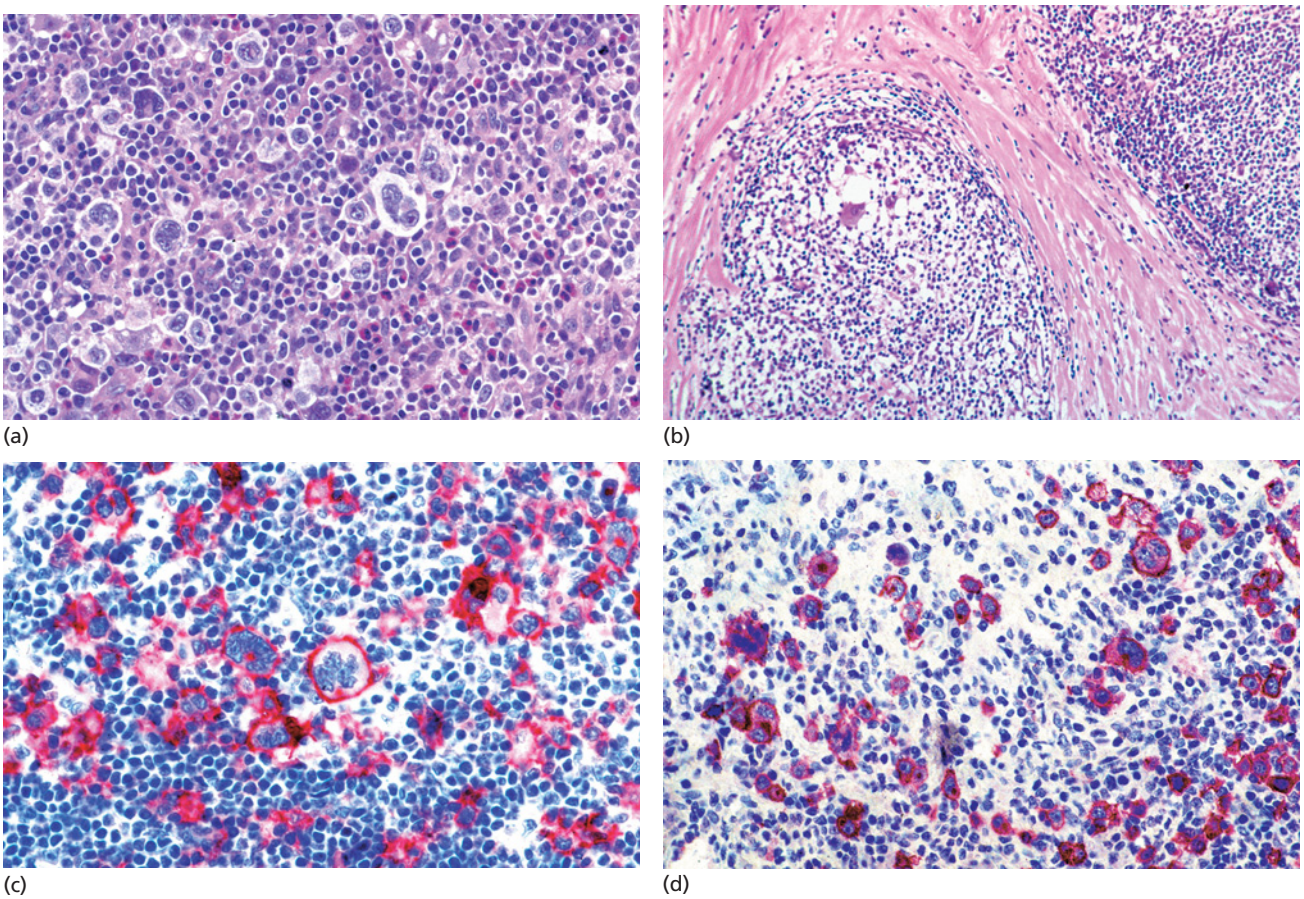


Figure 31.28 Classical Hodgkin lymphoma. (a) Mixed cellularity type. (b) Nodular sclerosing type (H&E, ×100). (c, d) Neoplastic cells positive for CD30 (c) and CD15 (d).

(e.g. 6p25.3 rearrangement or *BLIMP1* gene loss confers an indolent behaviour, while *TP63* rearrangement or*TP53* loss is associated with high aggressiveness). Finally, yet importantly, a special form of ALCL ALK$^-$has recently been described in patients who have received breast implants: if limited to the periprosthesis capsule it can be cured by surgery.

Hodgkin lymphoma (see also Chapter 32)

HL accounts for about 10% of lymphoid tumours. Irrespective of the histotype, HL is characterized by the following features: (i) usual onset in the lymph nodes, preferentially cervical ones; (ii) predilection for young adults and (iii) small number of mononucleated and multinucleated tumour cells (Hodgkin and Reed–Sternberg cells) within an exuberant inflammatory milieu and often ringed by rosettes of T lymphocytes. Biological and clinical studies have shown that HL comprises two disease entities: nodular lymphocyte-predominant HL (NLPHL) and classical HL.

Nodular lymphocyte-predominant Hodgkin lymphoma

NLPHL is a monoclonal B-cell tumour characterized by a nodular, or a nodular and diffuse, proliferation of scattered large multilobated neoplastic cells, known as 'popcorn' or LP cells (Figure 31.27). These reside in a follicular background with large FDC networks filled with non-neoplastic lymphocytes and histiocytes. One-third of the cases diagnosed in the past as NLPHL were lymphocyte-rich classical HL. It is currently unclear whether a diffuse NLPHL exists, its borders with T-cell/histiocyte-rich large B-cell lymphoma not always being sharp. NLPHL represents approximately 5% of HL cases. Patients are predominantly males in the fourth decade of life. LP cells are regularly positive for CD20, CD79a, EMA, BCL-6 and CD45. In contrast to Hodgkin and Reed–Sternberg cells of classical HL, tumour cells are positive for PAX5, Oct-2, BOB.1 and immunoglobulin light and/or heavy chains but – with a few exceptions – lack CD15, CD30 and EBV infection. The tumour cells are ringed by FTH cells that are positive for CD3, CD4, PD-1 and (to a lesser extent) CD57. LP cells show clonal *IGHV* gene

rearrangement, with a high load of somatic mutations that may be ongoing. Aberrant somatic hypermutations of *PAX5*, *PIM1*, *RhoH/TTF* (now called *RHOH*) and *MYC* genes are recorded in 80% of cases. NLPHL is an indolent disease with frequent relapses that usually remains responsive to therapy. Advanced stages have an unfavourable prognosis. Progression to large B-cell lymphoma-like lesions is seen in 1–14% of cases.

Classical Hodgkin lymphoma

Classical Hodgkin lymphoma is a monoclonal lymphoid neoplasm (mostly derived from B-lymphocytes) composed of Hodgkin and Reed–Sternberg cells (HRSCs) within a variable mixture of reactive lymphocytes, eosinophils, neutrophils, histiocytes, plasma cells, fibroblasts, and collagen fibres (Figure 31.28). Based on the microenvironmental composition, four histological subtypes are distinguished: lymphocyte-rich, mixed cellularity, lymphocyte-depleted and nodular sclerosis (characterized by HRSC variants, known as 'lacunar cells'). The immunophenotype and genetic features of HRSCs are identical in these histological subtypes, whereas clinical features and EBV association vary. Classical HL accounts for 95% of HLs, with a bimodal age curve in resource-rich countries, showing a peak at 15–35 years and a second peak in later life. Classical HL most often involves cervical lymph nodes and in approximately 60% of patients (usually with the nodular sclerosis subtype) causes a mediastinal mass. HRSCs typically represent 0.1–10% of the cellular infiltrate and are CD30 and CD15 positive in nearly all and most cases, respectively. In 30–40% of cases, CD20 may be detected, usually of varied intensity and in a minority of HRSCs. The B-cell nature of neoplastic cells is sustained in about 95% of cases by their expression of BSAP/PAX5, though weaker than that of reactive B cells. EBV-positive HRSCs show a latency II pattern ($LMP1^+$, $EBNA\text{-}2^-$). Characteristically, the transcription factors Oct-2 and BOB.1 are absent in about 90% of cases, while PU.1 is consistently absent. RSCs contain clonal *IGHV* gene rearrangements in more than 98% of cases, and clonal *TCR* gene rearrangements in rare instances. The rearranged *IGHV* genes harbour a high load of somatic hypermutations, usually not ongoing. GEP has recently produced some interesting data from both microdissected neoplastic cells and FFPE tissue samples, which may be of value in risk-class prediction. In this respect, the determination of microenvironmental components expressing CD68, CD163, and FoxP3 may assist in better defining the prognosis of patients with classical HL.

Selected bibliography

Campo E, Swerdlow SH, Harris NL *et al.* (2011). The 2008 WHO classification of lymphoid neoplasms and beyond: Evolving concepts and practical applications. *Blood* **117**(19): 5019–32.

Iqbal J, Wright G, Wang C *et al.* (2014) Gene expression signatures delineate biological and prognostic subgroups in peripheral T-cell lymphoma. *Blood* **123**: 2915–23.

Jares P, Colomer D, Campo E (2012) Molecular pathogenesis of mantle cell lymphoma. *Journal of Clinical Oncology* **122**(10): 3416–23.

Küppers R (2009) The biology of Hodgkin's lymphoma. *Nature Reviews Cancer* **9**: 15–27.

Kridel R, Sehn LH, Gascoyne RD (2012) Pathogenesis of follicular lymphoma. *Journal of Clinical Oncology* **122**(10): 3424–31.

Lenz G, Wright G, Dave SS *et al.* (2008) Stromal gene signatures in large-B-cell lymphomas. *New England Journal of Medicine* **359**: 2313–23.

Ott G, Rosenwald A, Campo E (2013). Understanding MYC driven aggressive B-cell lymphomas: pathogenesis and classification. *Blood* **122**(24): 3884–91.

Palomero T, Couronné L, Khiabanian H *et al.* (2014) Recurrent mutations in epigenetic regulators, RHOA and FYN kinase in peripheral T cell lymphomas. *Nature Genetics* **46**: 166–70.

Parrilla-Castellar ER, Jaffe ES, Said JW *et al.* (2014) ALK-negative anaplastic large cell lymphoma is a genetically heterogeneous disease with widely disparate clinical outcomes. *Blood* **124**(9): 1473–80.

Piccaluga PP, Fuligni F, De Leo A *et al.* (2013) Molecular profiling improves classification and prognostication of nodal peripheral T-cell lymphomas: results of a phase III diagnostic accuracy study. *Journal of Clinical Oncology* **31**: 3019–25.

Pileri SA, Piccaluga PP (2012) New molecular insights into peripheral T cell lymphomas. *Journal of Clinical Investigation* **122**: 3448–55.

Swerdlow SH, Jaffe ES, Brousset P *et al.*; on behalf of the International Lymphoma Study Group. (2014) Cytotoxic T-cell and NK-cell lymphomas: current questions and controversies. *American Journal of Surgical Pathology* **38**(10): e60–71.

Swerdlow SH, Campo E, Harris NL *et al.* (eds) (2008) *WHO Classification of Tumours of Haematopoietic and Lymphoid Tissues*. IARC, Lyon.

Shaffer AL 3rd, Young RM, Staudt LM (2012) Pathogenesis of human B cell lymphomas. *Annual Review of Immunology* **30**: 565–610.

Vaqué JP, Martínez N, Batlle-López A *et al.* (2014). B-cell lymphoma mutations: improving diagnostics and enabling targeted therapies. *Haematologica* **99**: 222–31.

CHAPTER 32

Hodgkin lymphoma

32

Piers Blombery and David Linch

UCL Cancer Institute, School of Life and Medical Sciences, University College London, London, UK

Introduction

Hodgkin lymphoma (HL) is a clinicopathologically unique, aggressive B-cell lymphoma, which is one of the most curable of all haematological malignancies. The excellent outcomes that are seen in HL represent the culmination of over 50 years of refining the combination of multiagent chemotherapy and radiotherapy. The evolution of the management of HL has resulted in new paradigms for combined modality therapy, demonstrated the advantages of risk-adapted treatment and has significantly advanced our understanding of the late-effects of chemoradiotherapy. Despite its relative rarity (approximately 20–30 new cases per million people in the United Kingdom per year), numerous randomized controlled trials in HL have been conducted, which have provided a large and robust evidence base to inform treatment decisions.

Despite these advances and the overall excellent outcomes, there are still subgroups of patients whose prognosis remains suboptimal (including the elderly and those with relapsed disease). Moreover, many aspects of the pathobiology of HL still remain unclear. This chapter will discuss the clinicopathological features, investigational approach and both current and future treatment strategies in patients with HL.

Pathological features

Classification

HL is classified into two broad groups according to the current *World Health Organisation Classification of Tumours of Haematopoietic and Lymphoid Tissues* (2008); classical HL (cHL) and nodular lymphocyte-predominant HL (NLPHL). Classical HL accounts for the vast majority of cases of HL (95%) compared to NLPHL. Despite being categorized together within the WHO (2008) classification, these two malignancies are fundamentally different with regard to pathogenesis, prognosis, clinical features and management. NLPHL is best considered separately, and within the same investigation and treatment paradigm as the low-grade B-cell lymphoproliferative disorders. The majority of this chapter focuses on cHL with NLPHL discussed as a separate entity further in Table 32.1.

There are four major histological subtypes of cHL; nodular sclerosis, lymphocyte-rich, lymphocyte-depleted and mixed-cellularity (see Histological features). Whilst the four histological subtypes have different epidemiological/aetiological implications and are associated with slightly different clinical features, the management approach for each type does not differ and therefore the distinction is currently not of major therapeutic relevance.

Pathogenesis

Cell of origin

The malignant lymphocyte in cHL is known as the Hodgkin and Reed–Sternberg (HRS) cell. Whilst the exact cell of origin of the HRS cell was the subject of research and dispute for many years, the demonstration of clonal rearrangements of the immunoglobulin heavy- and light-chain loci has confirmed their B-cell origin. Furthermore, the presence of somatic hypermutation within the immunoglobulin heavy and light chain loci suggest that HRS cells are derived from B-lymphocytes with germinal centre exposure. Despite their B-cell ontogenesis, HRS cells have lost most of the normal B-cell lineage gene expression program (including the expression of immunoglobulin)

Postgraduate Haematology, Seventh Edition. Edited by A Victor Hoffbrand, Douglas R Higgs, David M Keeling and Atul B Mehta.

Table 32.1 Nodular lymphocyte predominant Hodgkin lymphoma (NLPHL).

Clinical features

Male predominance. Patients typically presents with chronic asymptomatic peripheral lymphadenopathy. Majority present as early stage (stage I-II) disease. B-symptoms are rare. May transform to diffuse large B-cell lymphoma (DLBCL).

Lymph nodes in NLPHL may not be as reliably FDG-avid on PET scan.

Pathology

Lymph node involvement by neoplastic cells known as lymphocyte-predominant ("LP") cells which are large with folded nuclei and multiple nucleoli (also known as "popcorn" cells) in a nodular background consisting of expanded follicular dendritic cell meshworks and small B-lymphocytes.

LP cells are CD20+, CD45+, Oct-2+, BOB.1+, CD30− and CD15−.

Treatment – early stage (I–II)

Radiotherapy alone is associated with excellent outcomes.

Observation may be considered after complete surgical excision of single node.

Combined modality therapy (chemoradiotherapy) is often used in those with B-symptoms or risk factors for poorer outcomes (see further on).

Treatment – advanced stage (III–IV)

Chemotherapy with a variety of regimens is used including "cHL-type" (e.g. ABVD) as well as "NHL-type" (e.g. CVP/CHOP). Rituximab is usually incorporated into these regimens (due to CD20 expression on tumour cells and its observed activity as a single-agent) however the exact benefit remains uncertain.

Prognosis

Excellent outcomes in both early and advanced stage disease (similar to cHL) are seen. Risk factors for poorer outcomes include advanced age, higher stage, anaemia and lymphopenia.

Continued observation is necessary due to late relapses and large-cell transformation (typically to T-cell/histiocyte-rich large B-cell lymphoma).

through numerous aberrant genetic mechanisms, such as epigenetic silencing at B-cell gene promoter regions.

Molecular pathogenesis

Genes involved in normal B-lymphocyte growth and differentiation are suppressed in HRS cells. Instead, numerous aberrant intracellular signalling pathways contribute to the malignant phenotype of HRS cells including the following:

- HRS cells show constitutive activation of the NF-κB pathway, which is associated with apoptosis resistance. The basis for constitutive NF-κB activation in at least a proportion of cases is the result of inactivating mutations in *TNFAIP3* and *NFKBIA*, which encode inhibitors of the NF-κB pathway. Other mechanisms of NF-κB over-activity include genomic amplification of REL, expression and stimulation of CD40 by HRS cells and EBV infection of HRS cells (resulting in LMP-1 expression, which can mimic activation of CD40).
- The JAK-STAT signalling pathway is overactive in HRS cells, resulting in uncontrolled growth and proliferation. Mechanisms of JAK-STAT over-activity include chromosomal gains at 9p24.1-24.3 (which includes the JAK2 locus) and inactivating mutations in *PTPN1* (leading to increased phosphorylation of JAK-STAT pathway members).
- HRS cells have been shown to have deacetylated histones (H3), increased H3K27 trimethylation and DNA methylation patterns, leading to silencing of tumour-suppressor genes and the extinction of the normal B-lymphocyte expression profile.
- High-throughput sequencing studies have detected deleterious mutations in β2M which could potentially contribute to immune evasion by HRS cells through decreased β2M expression.

Microenvironment

The HRS cells have a complex and symbiotic relationship with the tumour microenvironment. Numerous cytokines and chemokines are produced by HRS cells, which attract a range of immune cells including T-lymphocytes (predominantly CD4+ (TH2 and Treg)), macrophages, plasma cells and eosinophils. The cytokines and chemokines produced by HRS cells induce the surrounding inflammatory milieu to produce their own cytokines, leading to propagation of the microenvironment. The immune cells within the microenvironment in turn provide signalling to the HRS cells via surface receptors that activate intracellular pathways, including NF-κB, which contribute to the malignant phenotype of the HRS cells.

Role of EBV

EBV can be detected in HRS cells in a variable proportion of cases of HL, depending on age, sex, race and histological subtype. HRS cells that have destructive or 'crippling' immunoglobulin loci rearrangements are almost always EBV positive. This

likely represents a mechanism by which the HRS cell bypasses the need for a functional B-cell receptor by using EBV-encoded proteins. EBV infection and expression of latent EBV-associated proteins (LMP1) is also associated with activation of NF-κB as well as JAK/STAT pathways.

Histological features

An accurate histological diagnosis of cHL is made by recognizing the morphological and immunophenotypic characteristics of the HRS cell within the appropriate cellular background. Systematic immunohistochemical evaluation of both the tumour cell compartment (HRS cells and its variants) and the associated cellular background is important to accurately distinguish cHL from the main competing differential diagnoses; diffuse large B-cell lymphoma (DLBCL) (in particular T-cell/histiocyte-rich diffuse large B-cell lymphoma and primary mediastinal B-cell lymphoma), anaplastic large-cell lymphoma (ALCL) and peripheral T-cell lymphoma not otherwise specified (PTCL-NOS).

HRS cells

Classic HRS cells are very large (20–50 μm) and can either have one nucleus or be multinucleate. HRS nuclei typically contain large, eosinophilic, inclusion-like nucleoli (Figures 32.1 and 32.2). The classic 'owl's eye' appearance describes a binucleate HRS cell with two mirror-image nuclei containing eosinophillic nucleoli and a thick nuclear membrane. In addition to the classic HRS cell, other HRS cell variants may also be seen, including mononuclear forms (Hodgkin ('H') cells), lacunar HRS cells (particularly in the nodular sclerosis (NS) subtype) and mummified HRS cells. In most histological subtypes the HRS cells make up only a small proportion of the cellularity of the tumour (typically 1% or less). High levels of circulating tumour DNA can be detected in most patients with Hodgkin lymphoma at diagnosis and this suggests that there is a much higher proliferation rate of tumour cells than previously appreciated, with a very high rate of apoptosis accounting for the relative scarcity of malignant Hodgkin cells seen histopathologically.

The immunohistochemical hallmark of the HRS cell is positive membrane staining for CD30 (with a paranuclear focus of staining within the Golgi) (Figure 32.3) and CD15. Due to the disruption of the normal B-cell gene expression program me,

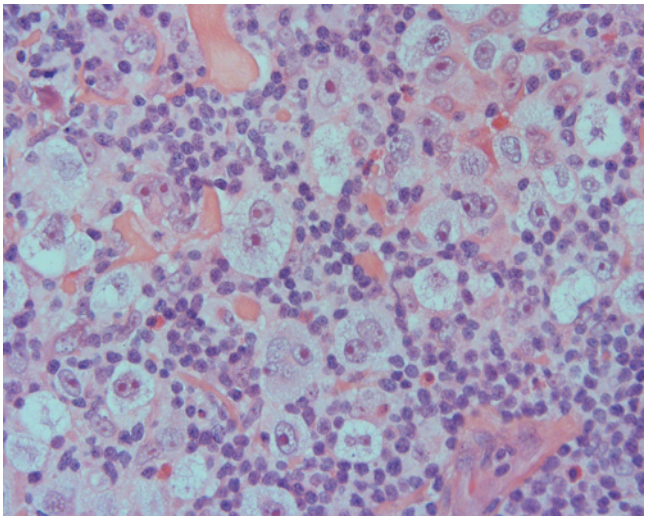


Figure 32.2 High-power view of Hodgkin Reed–Sternberg cells.

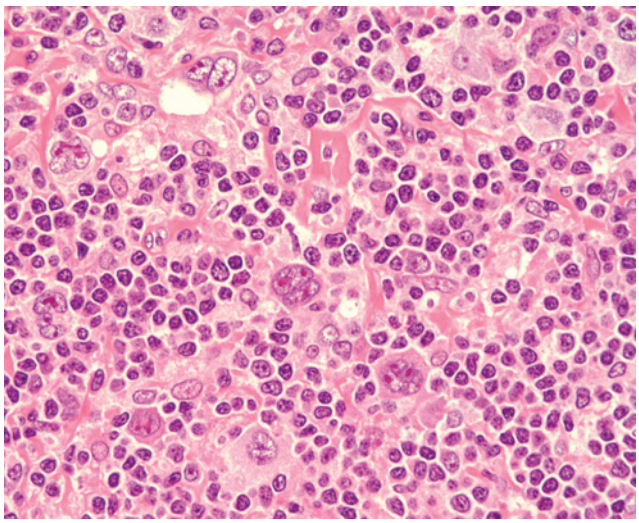


Figure 32.1 Classical Hodgkin lymphoma with numerous multinucleate Hodgkin Reed–Sternberg cells with prominent nucleoli.

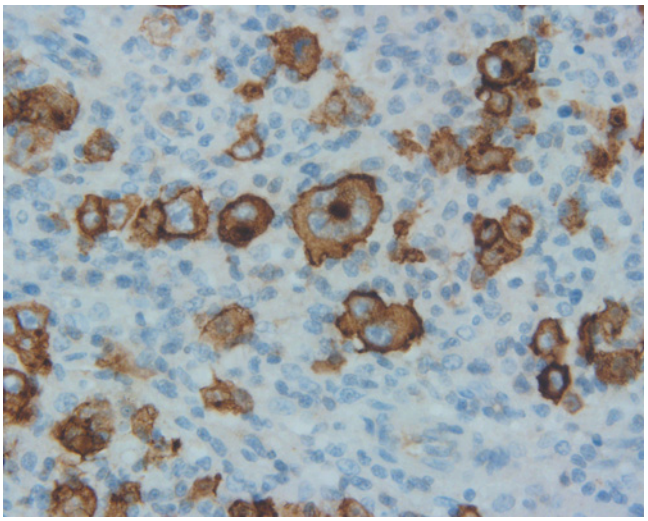


Figure 32.3 CD30 immunohistochemistry of classical Hodgkin lymphoma showing membrane and Golgi apparatus staining of the Hodgkin Reed–Sternberg cells.

HRS cells show different immunohistochemical staining characteristics to both normal B-lymphocytes and malignant B-lymphocytes of other B-lymphoproliferative disorders. Therefore, whilst HRS cells usually stain strongly for MUM1 they show weak expression of the master B-cell transcription factor PAX5. CD20 may be positive in approximately 20–50% of cases, but this is usually patchy and weak (as opposed to the strong uniform staining seen in the malignant B-cells of T-cell/histiocyte-rich DLBCL). CD45 expression is negative in HRS cells; however, due to the surrounding CD45-positive haemopoietic compartment, this feature can sometimes be difficult to appreciate.

Associated histological features

The associated cellular milieu seen in cHL is an important histological feature that helps to confirm the diagnosis of cHL (as well as differentiating the histological subtypes). The non-malignant associated cellular component differs depending on the histological subtype as follows:

- Nodular sclerosis (NS) (approx. 70–75% of cases) – the NS subtype is characterized by fibrous bands arising from a thickened capsule that run throughout the node and compartmentalize the tumour into nodules (Figure 32.4). Lacunar HRS cells are typical of the NS subtype and are present within a heterogenous cellular background that consists predominantly of CD4+ T-lymphocytes, but also includes histiocytes, eosinophils, plasma cells and neutrophils (Figure 32.5).
- Mixed cellularity (MC) (approx. 20–25% of cases) – this subtype is named after the heterogenous inflammatory infiltrate that surrounds the HRS cells, consisting of lymphocytes, plasma cells, epithelioid histiocytes and eosinophils. Whilst there may be interstitial fibrosis in the MC subtype, broad bands of compartmentalizing fibrosis are not present. The majority of MC cases are EBV positive.

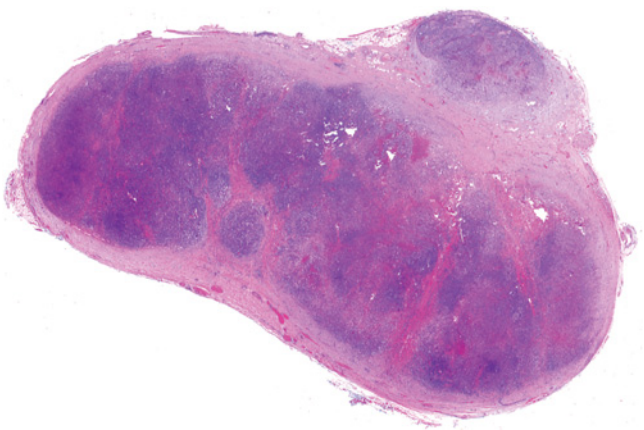


Figure 32.4 Low-power view of lymph node from patient with nodular sclerosis subtype classical Hodgkin lymphoma showing broad bands of fibrosis and thickened capsule.

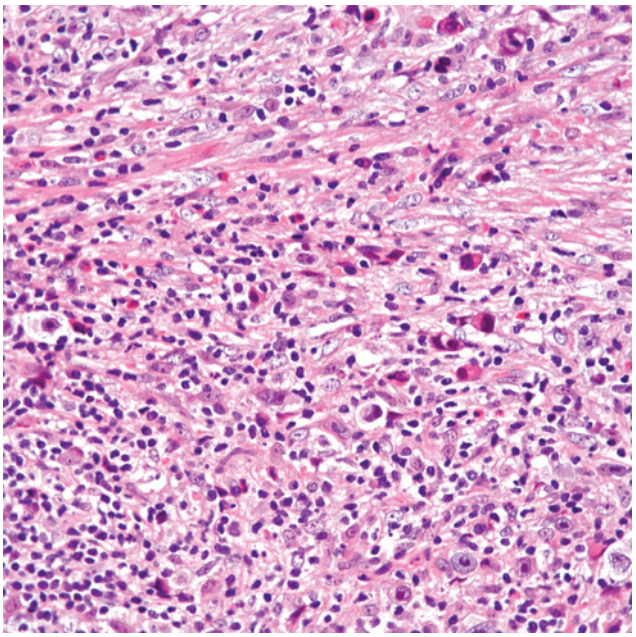


Figure 32.5 Nodular sclerosis classical Hodgkin lymphoma showing two lacunar HRS cells.

- Lymphocyte-rich (LR) (approx 5% of cases) – compared to the T-cell-predominant lymphocytic infiltrate in NS and MC cHL, this subtype contains a cellular background containing both B and T lymphocytes. In addition there is usually a relative paucity of neutrophils and eosinophils. This form is distinguished from NLPHL by the presence of typical HRS cells, which show membrane staining for CD30 and CD15 (which are not seen in NLPHL) (see Table 32.1).
- Lymphocyte-depleted (LD) (<1% of cases) – true LD cHL is a very rare entity that histologically may be characterized by either large numbers of HRS cells or rare HRS cells in a densely fibrotic background stroma. The LD histological subtype was reported to be associated with a poorer prognosis; however, many cases were probably misclassified and included cases of ALCL.

Clinical features

Presentation

Clinically, cHL usually presents as the result of one (or more) of three main mechanisms:

1 *Enlarging lymphadenopathy* – lymphadenopathy in cHL is typically painless. The rate of enlargement reported by the patient is variable, but it is not usually rapid; rather a slow progressive growth over months is more common. The most frequent site of lymphadenopathy is in the cervical and supraclavicular nodal regions, with isolated infradiaphragmatic

Table 32.2 Ann Arbor staging system with Cotswolds modifications.

Ann Arbor stage	Criteria
Stage I	Involvement of a single lymph node region (see Figure 32.6) or lymphoid structure*
Stage II	Involvement of two or more lymph node regions or lymph node structures on the same side of the diaphragm
Stage III	Involvement of lymph node regions or lymphoid structures on both sides of the diaphragm
Stage IV	Diffuse or disseminated involvement of one or more extralymphatic organs or tissues with or without associated lymph node enlargement. Involvement of liver and bone marrow always considered stage IV
Additional classifiers	
B	Presence of B symptoms defined as • Unexplained weight loss of more than 10% of the body weight during the 6 months before initial staging investigation • Unexplained, persistent, or recurrent fever with temperatures above 38 °C during the previous month • Recurrent drenching night sweats during the previous month
E	Limited extranodal extension from adjacent nodal site or apparent discrete single extranodal deposit (excluding liver and bone marrow)
X	A node or nodal mass greater than 10 cm (by largest dimensions of a single node or conglomerate nodal mass). Maximum width is equal to or greater than one-third of the internal transverse diameter of the thorax at the level of T5/6 on chest X-ray.

*e.g. spleen, thymus, Waldeyer's ring.

presentations being significantly more rare. On examination, nodes involved by HL are firm, mobile and non-tender.

2 *Complications of lymphadenopathy* – compression of vital structures may occur as the result of pathologically enlarged lymph nodes or lymph node masses. A typical presentation is from complications related to a mediastinal lymph node mass. These may present with a spectrum of symptoms from mild chest pain and cough to overt superior vena cava obstruction with headaches, facial suffusion, positive Pemberton's sign and engorged neck veins. Other complications of lymphadenopathy include abnormal LFTs and jaundice from obstruction of the biliary tree by porta hepatis lymphadenopathy.

3 *Systemic symptoms* – cHL is a highly inflammatory tumour and may be associated with numerous systemic symptoms. The systemic 'B' symptoms that are important to elicit on history due to their implications for Ann Arbor stage (Table 32.2) are loss of weight, fever and drenching night sweats. Other symptoms (which are not considered B symptoms according to the Ann Arbor staging system) that may be present include itch and alcohol-induced lymph node pain. Other very rare systemic associations include vanishing bile duct syndrome and paraneoplastic cerebellar degeneration.

Investigation

As outlined above, cHL is diagnosed by the morphological recognition of the HRS cell within the appropriate cellular background. The optimal tissue for diagnosis of cHL (and indeed almost all subtypes of lymphoma) is a complete lymph node taken as an excisional lymph node biopsy. There is an increasing tendency to perform core biopsies on lymph nodes for the investigation of lymphoma and whilst these are sufficient to make a diagnosis in some cases, there is a significant risk of either a false-negative biopsy due to sampling of the reactive component, or a potentially misleading or inadequate sampling of the tumour compartment leading to an incorrect diagnosis. Fine needle aspirates are never acceptable for the diagnosis of cHL and should not be performed if cHL is the suspected diagnosis.

Radiological investigations and bone marrow biopsy are performed to determine the extent of disease in order to define the Ann Arbor stage (see Staging and risk stratification). In addition, other investigations that are performed in the diagnostic work-up of patients with cHL (and suspected cHL) include:

- *Full blood count* – typical findings include mild anaemia, lymphopenia and mild microcytosis (rare)
- *Renal and liver function tests* – usually performed as a baseline before chemotherapy and as a screen for hepatic dysfunction as a result of cHL
- *HIV serology* – the incidence of cHL is increased in patients with HIV
- *Echocardiogram/nuclear medicine assessment of cardiac function* – treatment of cHL usually involves treatment with anthracycline containing chemotherapy and therefore consideration of cardiac function is important (see Late effects).

Staging and risk stratification

Ann Arbor staging

The staging system that is used in patients with HL is the Ann Arbor staging system (Table 32.2, Figure 32.6). Despite originally being described over 40 years ago, accurate staging using the Ann Arbor staging system (as well as integration of clinical risk factors) is still relevant to contemporary practice and is central to directing management strategies in patients with cHL. Staging via the Ann Arbor classification involves a thorough history to detect presence of B-symptoms (Table 32.3), examination of affected lymph node groups and radiological investigations (+/– bone marrow biopsy) to accurately determine the extent of nodal and extranodal involvement.

The most sensitive method for detecting the involvement of a nodal or extranodal structure by cHL is by 'functional' imaging with ^{18}F-FDG-PET scan with a correlative 'structural' computed tomography (CT) scan. cHL is highly FDG-avid and pathologically involved nodes (and extranodal sites) are able to be detected on PET scan that would not be detected using conventional modalities (i.e. CT alone) (Figure 32.7). Up to 15% of patients will be 'up-staged' using PET (when compared to CT alone) to a higher stage of disease with potentially significant implications for frontline treatment. Baseline PET scanning in patients with HL is now considered standard of care.

Table 32.3 Outcomes in patients with advanced-stage Hodgkin lymphoma using the International Prognostic Score (patients treated in modern era with ABVD or ABVD-like regimen).

International prognostic score	5 year FFP	5 year OS
0	88%	98%
1	84%	97%
2	80%	91%
3	74%	88%
4	67%	85%
≥5	62%	67%

Score one point for each of albumin <4 g/dL, haemoglobin <105 g/L, male sex, age ≥45 years, stage IV disease, leucocytosis (≥15 × 10^9/L), lymphopenia (<8% of WCC and/or <0.6 × 10^9/L).

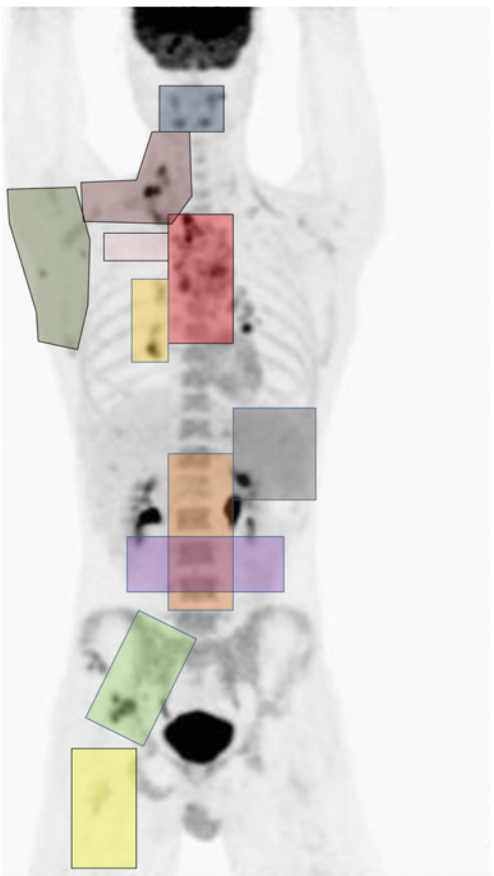


Figure 32.6 Ann Arbor lymph node groups.

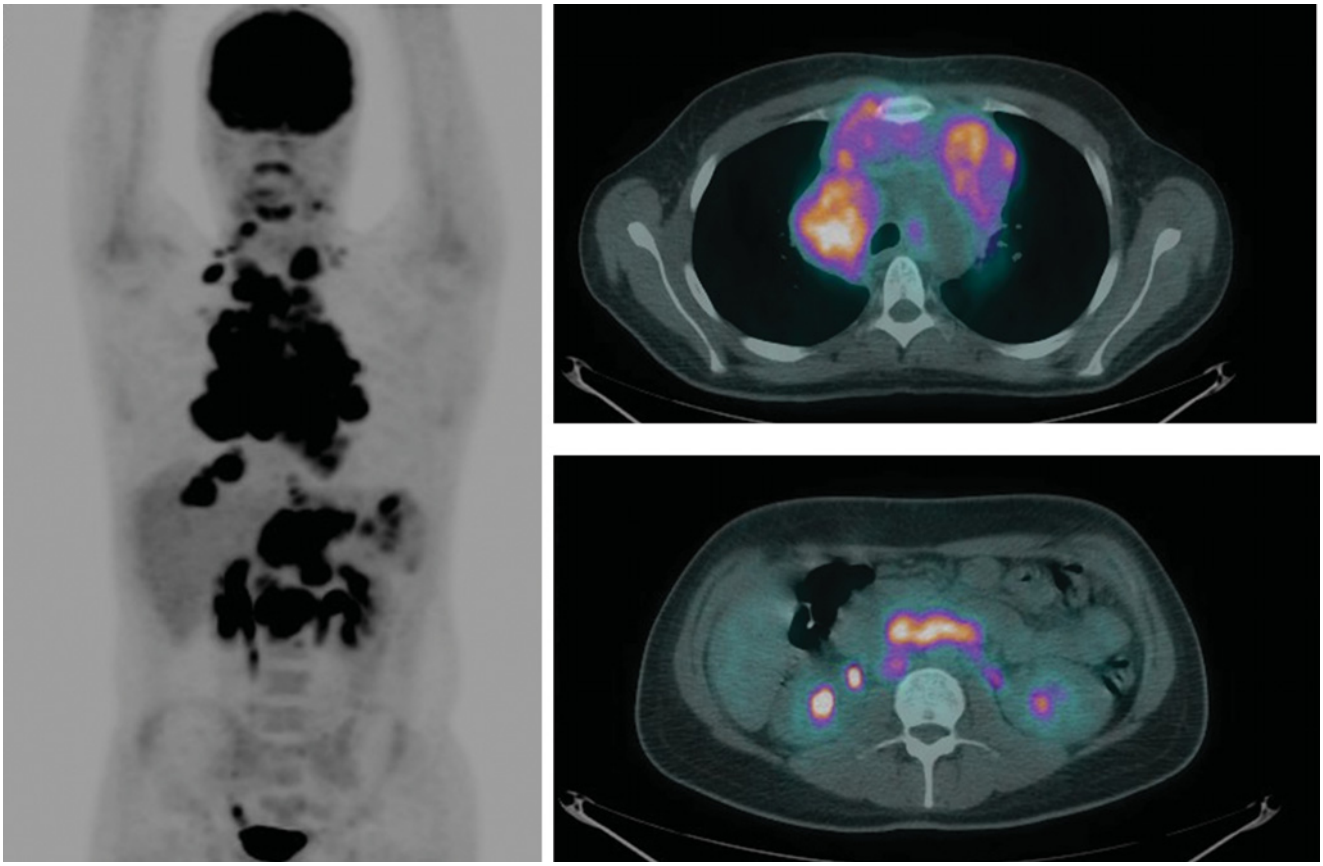


Figure 32.7 PET scan images of patient with stage III Hodgkin lymphoma showing FDG-avid cervical, mediastinal, mesenteric and retroperitoneal lymphadenopathy.

PET scans are also useful in the follow-up of patients. PET scan assessment at the end of treatment has been adopted in the *Revised Response Criteria for Malignant Lymphoma (2007)*. In these criteria a complete remission can still be assigned to the patient with a mass of any size, as long as it is PET negative. Negative end-of-treatment PET scans have a 94% negative predictive value for progression or early relapse in advanced-stage disease.

PET scan is a sensitive modality for detecting bone marrow involvement in patients with HL. In addition, the incidence of bone marrow involvement in patients with early-stage (stages I and II) disease determined by PET is low (<1%). Therefore, routine bone marrow biopsy in patients with early stage disease without evidence of marrow involvement on PET is a low-yield investigation. However, some centres still perform bone marrow biopsy in these cases due to the significant implications if bone marrow involvement is identified (e.g. change from stage I or II disease to stage IV).

Based on the Ann Arbor stage, patients are usually divided into either early-stage HL (ESHL) or advanced-stage HL (ASHL).

Early-stage Hodgkin lymphoma

ESHL broadly refers to Ann Arbor stage I or II disease. The presence of B symptoms does not preclude the designation of early-stage disease; however, some large trials in ESHL which inform practice have excluded patients with B symptoms. Stage IIB disease (especially those patients with clinical risk factors) is considered to be a borderline entity and many groups treat stage IIB as advanced-stage disease.

Advanced-stage Hodgkin lymphoma

ASHL refers to any patient with Ann Arbor stage III or IV disease. In addition to stage III and IV disease, many prospective trials in advanced-stage disease have included patients with earlier Ann Arbor stages of disease if they have high-risk features (e.g. a large mediastinal mass or B symptoms). The rationale for including these earlier-stage patients in the same trials as those with stage III and IV disease is that both groups of patients share a common high risk of disseminated systemic disease and therefore a reliance on systemically directed therapy (i.e. chemotherapy) for disease cure.

Risk stratification

Central to the treatment decisions in cHL are the competing risks of: (i) treatment minimization in order to reduce the risk of acute and chronic treatment-related morbidity and mortality, but leading to an increased incidence of disease recurrence, versus (ii) treatment intensification leading to superior disease control but an increased short and long-term treatment related morbidity and mortality.

The general approach in both early-stage cHL (ESHL) and advanced-stage cHL (ASHL) is an attempt to de-intensify treatment in patients who can be defined as being at low risk of disease recurrence and to intensify treatment in patients who can be identified as being at a high risk of disease recurrence. The identification of low and high risk patients is different for ESHL and ASHL.

Early-stage Hodgkin lymphoma

ESHL can be divided into favourable and unfavourable subgroups based on the presence of various risk factors (e.g. large mediastinal mass [≥ 1/3 transthoracic diameter], extranodal disease, ≥ 3 involved nodal areas or elevated erythrocyte sedimentation rate). This distinction has been made in many of the large prospective randomized trials that guide practice in ESHL. Importantly, the criteria used to define favourable and unfavourable are not standardized between groups, and whilst the criteria are similar, there are important differences.

Advanced-stage Hodgkin lymphoma

The international prognostic score (IPS) is the most commonly used method to determine upfront risk in ASHL. This risk score was determined after study of a large international cohort of patients with ASHL (mostly with ABVD-like chemotherapy). Seven factors were identified that predicted outcome (Table 32.3). A common way to divide patients is to include those with an IPS of 0, 1 or 2 as 'low-risk' and IPS 3–7 as 'high-risk'.

Of note, both of these risk stratifications are assessed on baseline features of disease. A strategy that is currently under evaluation is the assessment of risk based on the early response to therapy (e.g. appearance on PET scan after two cycles of chemotherapy) (see Risk-adapted frontline therapy).

Management

There are three main groups of therapeutic modalities used to treat cHL: conventional chemotherapeutic agents, radiotherapy and novel agents.

Conventional chemotherapy regimens

cHL is a chemosensitive malignancy. Chemotherapy regimens used in upfront treatment usually contain an alkylating agent (e.g. dacarbazine or procarbazine) and an anthracycline (e.g. adriamycin). The two most commonly used chemotherapy regimens used for upfront treatment of cHL are ABVD (adriamycin, bleomycin, vinblastine and darcarbazine), and the more intense escalated BEACOPP regimen (bleomycin, etoposide, adriamycin, cyclophosphamide, vincristine, procarbazine and prednisolone). Other chemotherapy regimens such as COPP and ChlVPP (chlorambucil, vinblastine, procarbazine, prednisolone) are sometimes used in frail patients and also are of value in the palliative setting. The classic MOPP regimen (containing the nitrogen mustard mechlorethamine) is infrequently used due to unacceptable gonadal toxicity and risk of secondary myeloid malignancy.

Radiotherapy

Radiotherapy is a highly effective modality for the treatment of cHL. Indeed, radiotherapy alone used to be the basis (in the 1960s) of curative therapy for cHL. The introduction of effective chemotherapy regimens such as ABVD have meant that both field sizes and radiotherapy doses have been able to be significantly decreased without a compromise in disease outcomes. Older style radiotherapy fields (e.g. the mantle field), which encompassed large numbers of nodes, as well as normal tissue, are rarely used nowadays. Instead, involved field radiotherapy (IFRT) and, more typically, involved node radiotherapy (INRT) are now used, resulting in a dramatically lower volume of normal tissue irradiated compared to the more extensive historical fields. In addition, there have been improvements in staging techniques with more accurate radiological delineation of involved sites, more effective chemotherapy and other improvements in radiotherapy techniques (including shielding and advances in planning technology).

Novel agents

One of the single most effective new anti-HL therapies to emerge is brentuximab vedotin (BV). BV is an antibody–drug conjugate, which consists of an anti-CD30 antibody conjugated to an antimicrotubule agent (monomethyl auristatin E (MMAE)). The specificity of the anti-CD30 antibody is used to guide the cytotoxin (MMAE) to the HRS cells, where binding of the anti-CD30 antibody–drug conjugate is followed by receptor mediated endocytosis and delivery of the cytotoxic agent upon hydrolysis of the linker molecule by lysosomal enzymes. Anti-CD30 antibodies that were not linked to cytotoxic agents had been investigated previously, but were found to not be effective. Other agents that have shown promise in cHL include the immunomodulatory drug lenalidomide, the histone deacetylase inhibitors (e.g. panobinostat) and the anti-PD-1 antibody nivolumab.

Conventional frontline treatment

Early-stage favourable

One of the largest prospective trials using contemporary radiotherapy techniques and chemotherapy regimens in patients with ESHL was the HD10 trial performed by the German Hodgkin Study Group (GHSG). In this trial, 1370 patients with favourable ESHL were randomized to either two or four cycles of ABVD, followed by either 20 Gy or 30 Gy of involved field radiotherapy (IFRT). No difference in freedom from treatment failure (FFTF) or OS between any of the arms was observed, and therefore 2 × ABVD and 20 Gy IFRT was concluded to be the optimal treatment from this trial.

The HD.6 trial performed by the NCIC Clinical Trials Group and the Eastern Cooperative Oncology Group (ECOG) randomized patients to either ABVD alone (for four to six cycles, depending on CT scan response) or subtotal nodal irradiation (an extensive field size, which is no longer used for early stage disease) +/– ABVD. Patients with a favourable risk profile that were treated with ABVD alone had similar disease control and overall survival outcomes to those treated in the HD10 trial (Table 32.4).

Based on these data, 2 × ABVD with 20 Gy IFRT should be considered the current standard treatment of favourable ESHL. ABVD alone for four to six cycles remains an option.

Given the excellent results of two cycles of ABVD + IFRT in favourable ESHL, the GHSG HD13 trial was performed to determine whether chemotherapy could be further reduced without an unacceptable loss of tumour control. HD13 evaluated removing either dacarbazine (ABV), bleomycin (AVD) or both (AV) from ABVD. In each arm, chemotherapy was followed

Table 32.4 Outcomes of selected trials in early-stage and advanced-stage Hodgkin lymphoma.

Trial	Treatment	Disease control outcome	Overall survival
Early stage			
Favourable			
GHSG HD10	ABVD × 2 + IFRT 20Gy	8-year FFTF 85.9%	8-year OS 95.1%
NCIC HD.6 (chemotherapy only arm)	ABVD × 4 – 6	12-year FFDP 89%	12-year OS 98%
GHSG HD13	ABVD × 2 + IFRT30Gy	5-year FFTF 93.1%	5-year OS 97.6%
GHSG HD13	AVD × 2 + IFRT 30Gy	5-year FFTF 89.2%	5-year OS 97.6%
Unfavourable			
GHSG HD11	ABVD × 4 + IFRT 30Gy	5-year FFTF 85.3%	5-year OS 94.3%
GHSG HD14	ABVD × 4 + IFRT 30Gy	5-year FFTF 87.7%	5-year OS 96.8%
	BEACOPP(esc) × 2 + ABVD × 2 + IFRT 30Gy	5-year FFTF 94.8%	5-year OS 97.2%
NCIC HD.6 (chemotherapy only arm)	ABVD × 4 – 6	12-year FFDP 86%	12-year OS 92%
Advanced stage			
Outcomes regardless of IPS (0–7)			
GHSG HD15	BEACOPP(esc) × 6	5-year PFS 90.3%	5-year OS 95.3%
	BEACOPP(esc) × 8	5-year PFS 85.6%	5-year OS 91.9%
HD2000	BEACOPP(esc) × 4, BEACOPP (baseline) × 2	5-year PFS 81%	5-year OS 92%
	ABVD × 6	5-year PFS 68%	5-year OS 84%
'Low-risk' – IPS (0–2)			
H34	BEACOPP(esc) × 4, BEACOPP (baseline) × 4	5-year PFS 93%	5-year OS 99%
	ABVD × 8	5-year PFS 75%	5-year OS 92%
E2496	ABVD × 6 – 8	5-year FFS 77%	5-year OS 91%
'High-risk' – IPS (3–7)			
EORTC 20012	BEACOPP(esc) × 4, BEACOPP (baseline) × 4	4-year PFS 83.4%	4-year OS 90.3%
	ABVD × 8	4-year PFS 72.8%	4-year OS 86.7%
E2496	ABVD × 6–8	5-year FFS 67%	5-year OS 84%

by 30 Gy IFRT. This trial showed that ABVD provided superior disease control when compared to AV, AVD or ABV. However, the gain in disease control from the addition of bleomycin (i.e. AVD versus ABVD) was modest (approximately 4% in 5-year FFTF) and not associated with any difference in overall survival.

Another potential strategy for minimizing chemotherapy toxicity in early-stage HL is to include novel agents in the place of conventional cytotoxic agents (e.g. brentuximab, vedotin, lenalidomide). As discussed below, this is a particularly attractive strategy in the elderly, where the morbidity and mortality from ABVD is substantial.

Early-stage unfavourable

In patients with unfavourable ESHL, the GHSG HD11 trial randomized 1395 patients to either 4 × ABVD + 30 Gy IFRT, 4 × ABVD + 20 Gy IFRT, 4 × BEACOPP (baseline doses) + 30 Gy IFRT and 4 × BEACOPP (baseline doses) + 20 Gy IFRT. This trial showed both a loss of tumour control in the 4 × ABVD + 20 Gy IFRT and no net benefit of using the more intense BEACOPP chemotherapy regimen. The 5-year FFTF in the 4 × ABVD and IFRT 30 Gy was 85.3% (5-year overall survival 94.3%). A similar finding was observed when 4 × ABVD + 30 Gy IFRT was compared to two cycles of escalated BEACOPP, followed by two cycles of ABVD + 30 Gy IFRT in HD14. Similarly again, patients with unfavourable early-stage disease treated with ABVD (four to six cycles) alone from the HD.6 trial had a 12-year freedom from progression of 86% and an OS of 92%.

Overall, it is generally regarded that 4 × ABVD + 30 Gy IFRT represents the best balance between efficacy and toxicity for unfavourable ESHL. Despite superior disease control, escalated BEACOPP is associated with excessive toxicity and is not typically used in unfavourable ESHL. ABVD alone (four to six cycles) appears to give comparable disease control and overall survival; however, there are less data supporting this approach.

Advanced-stage disease

Multiple prospective studies have shown that escalated BEACOPP results in better disease control than ABVD in the frontline treatment of advanced-stage cHL. Variation of the scheduling, doses and number of cycles of BEACOPP in these trials has resulted in six cycles of escalated BEACOPP being defined as the optimal balance between efficacy and toxicity. Despite superior disease control, the overall survival benefit of escalated BEACOPP over ABVD is controversial and has not yet been directly demonstrated. The observed superior disease control without benefit in overall survival is due to two reasons:

1 Escalated BEACOPP is associated with significantly higher haematological toxicity and secondary myeloid malignancy when compared to ABVD (see Late effects) and this may result in higher treatment-related mortality.

2 cHL is a highly salvageable malignancy. Therefore, despite the lower complete remission rates upfront with ABVD, a significant proportion of patients can still achieve long-term survival and cure with salvage chemotherapy and autologous stem-cell transplantation (see Relapsed/refractory disease). However, this treatment pathway then assumes a similar toxicity profile to treatment with escalated BEACOPP up front, but with the additional risk of non-response to salvage therapy.

Due to these competing factors, ABVD and escalated BEACOPP are both still used in the frontline setting, and despite numerous prospective trials involving thousands of patients, the optimal frontline treatment of advanced-stage cHL remains an issue of considerable controversy. Some authorities recommended limiting escalated BEACOPP to those patients with the poorest prognosis, as defined by the international prognostic score, but the long-term follow-up of the GHSG HD9 trial indicated that the benefit of escalated BEACOPP was similar in all risk groups.

Brentuximab vedotin has been combined with ABVD (ABVD-BV) and AVD (AVD-BV) up front in advanced-stage cHL, with encouraging preliminary efficacy. In a Phase I study of 51 patients in advanced-stage cHL, all patients were PET negative after two cycles of ABVD-BV (92% after two cycles of AVD-BV) and the CR rate at the end of treatment was 95%. In contrast, the expected CR rate for ABVD in advanced-stage patients, based on the literature to date, is approximately 80%. The main adverse effects observed included an approximately 80% incidence of grade III–IV neutropenia and an approximately 70% incidence of grade I–II sensory peripheral neuropathy. Importantly, pulmonary toxicity was observed in approximately 40% of the ABVD-BV patients. Therefore, the use of bleomycin with BV should be avoided. This approach is being further evaluated in a large prospective Phase III trial of AVD-BV versus ABVD.

Risk-adapted frontline treatment

In addition to stratifying therapy based on early- or advanced-stage disease and the presence or absence of baseline risk factors, patients may also be stratified by their early response to therapy. This was based on observations that patients with residual disease present on PET scan after two cycles of therapy had a significantly worse prognosis than those with no evidence of disease on PET.

Early-stage disease

The general approach of risk-adapted therapy in ESHL so far has been with PET scanning after initial chemotherapy (i.e. two

Table 32.5 Deauville score for the standardized reporting of PET scans in lymphoma.

Deauville Score 1	No uptake
Deauville Score 2	Uptake less than mediastinum
Deauville Score 3	Uptake greater than mediastinum, but less than liver
Deauville Score 4	Uptake greater than liver
Deauville Score 5	Uptake greater than liver and new sites of disease
Deauville Score X	New areas of uptake unlikely to be lymphoma

to three cycles of ABVD), which then guides the use of further cycles of chemotherapy and/or radiotherapy. The aim of this strategy is to offer escalated therapy to those patients with suboptimal initial disease response, whilst sparing those patients with good early disease control potentially unnecessary further treatment/toxicity. One approach being evaluated is to treat with chemotherapy (ABVD) alone and to omit radiotherapy in patients who have a negative PET scan (as defined on a standardized centralized PET assessment, Table 32.5). Initial results from these trials suggest that interim PET scanning in ESHL is potentially able to identify a group of patients that have an excellent outcome after chemotherapy alone.

Advanced-stage disease

There are three general approaches to risk-adapted therapy in advanced stage cHL:

1 Commence treatment with ABVD in all patients and then intensify treatment to escalated BEACOPP if interim PET appearance is unfavourable. Preliminary data from this approach suggests that many patients who are still PET positive after ABVD can be rendered PET negative with intensified treatment (e.g. escalated BEACOPP).

2 Commence treatment with escalated BEACOPP and subsequently de-intensify treatment to ABVD (or a reduced number of cycles of escalated BEACOPP) if interim PET is favourable.

3 Stratify patients based on IPS upfront to decide between ABVD or escalated BEACOPP as frontline therapy. Treatment is then intensified or de-intensified to the alternative regimen, based on interim PET scan.

These strategies are currently being evaluated in ongoing prospective trials.

Relapsed/refractory disease

For patients that relapse or who are refractory to primary therapy, the traditional treatment paradigm has been to administer salvage chemotherapy, choosing agents that were not administered in frontline treatment. A commonly used salvage regimen is ICE (ifosfamide, carboplatin, etoposide), but others include ESHAP (etoposide, methylprednisolone, cytarabine and cisplatin) and DHAP (dexamethasone, cytarabine, cisplatin).

Patients who demonstrate chemotherapy-responsive disease (usually assessed on PET scan performed after two cycles of salvage chemotherapy) undergo stem cell mobilization followed by high-dose therapy and autologous stem cell transplantation. Data from early trials suggest that this strategy results in approximately 50% disease-free survival at 3 years. The most potent determinant of outcome after autologous stem cell transplantation is the degree of chemoresponsiveness to salvage treatment. Patients that undergo autologous stem cell transplantation with a complete metabolic response on PET scan have a far superior outcome to those achieving only a partial metabolic response (PMR). Patients that are refractory to salvage chemotherapy (i.e. PET scan showing stable disease, disease progression or a mixed response) should not undergo autologous stem cell transplantation due to the very low chance of long-term disease control. Instead they should be considered for alternative therapies.

BV has been shown to be highly efficacious in the relapsed/refractory setting. Studies have shown an overall response rate of approximately 75% and approximately one-third of patients achieving complete remission in a heavily pretreated patient cohort. BV is currently used both in patients who have not responded adequately to initial conventional salvage therapy and also in first-line salvage. It is also useful in the palliative setting for relapsed patients post autologous stem cell transplantation.

Another promising novel approach in relapsed/refractory disease is monoclonal antibody blockade of PD-1. HRS cells over-express PD-1 ligand as a result of amplifications of chromosome 9p24.1. This results in increased stimulation of the PD-1 receptor on tumour-infiltrating T-lymphocytes and subsequent immune evasion through T-cell exhaustion. A Phase I study using the anti-PD-1 antibody nivolumab in patients with relapsed/refractory Hodgkin lymphoma (including patients previously treated with BV) has shown promising efficacy with a high overall response rate and some longer-term remissions.

In addition to the incorporation of novel therapeutic agents, numerous other strategies are being evaluated to try and improve outcomes in the relapsed/refractory setting, including PET-adapted sequential salvage therapy, tandem autologous transplant for high-risk patients and allogeneic transplantation in place of autologous transplantation.

A graft-versus-HL effect exists, and whilst allogeneic transplantation has typically been reserved for multirefractory disease (usually relapsed post autologous transplantation), there is emerging data to suggest that it may be a useful treatment in place of autologous transplantation in high-risk patients (i.e. those that do not achieve a CR to salvage chemotherapy or those with primary refractory disease). Whilst the progression-free

survival (PFS) in patients undergoing autologous transplantation with a PMR on PET to conventional salvage is 25–30% at 3–5 years, allogeneic transplant in this group may achieve a 3-year PFS of approximately 70% in some centres. This approach continues to be evaluated.

Treatment of older patients

Compared to the management of younger patients, there is a paucity of evidence to guide the management of patients over the age of 60 with cHL. However, a consistent feature that emerges from the data that are available is that the use of ABVD in patients over the age of 60 is associated with significant toxicity, dose reductions and treatment delays. In an analysis of patients over the age of 60 treated on the HD10 and HD11 trials, the use of ABVD was associated with a 68% incidence of grade III/IV toxicity and a 5% mortality rate from acute toxicity. In addition, elderly patients may be at particular risk of specific chemotherapy side-effects such as bleomycin lung toxicity.

Chemotherapy regimens utilizing conventional cytotoxic agents have been specifically developed for use in this group (e.g. VEPEMB) and whilst they may possibly represent an improvement over ABVD, they are still associated with suboptimal disease outcomes and significant toxicity. A standardized geriatric comorbidity assessment is one way to determine the likelihood of tolerance of chemotherapy and patients that 'fail' such a tool have a dismal outcome.

Given the poor outcomes, novel approaches are particularly attractive in this group of patients. Substitution of conventional chemotherapy agents with novel agents in frontline treatment (e.g. lenalidomide or BV) are currently being evaluated to potentially offer a more tolerable and efficacious treatment.

Late effects

Due to the high cure rates achieved by the treatment of both early stage and advanced stage cHL, patients treated for cHL are at risk of numerous late toxicities related to treatment. During the first 10 years after diagnosis, most deaths are the result of relapse, however, after this period most deaths are a result of late effects of treatment. For early-stage disease, the cause of death for the majority of patients overall is related to treatment rather than cHL.

The chronic toxicity profile of cHL treatment depends on numerous factors, including the intensity of chemotherapy regimen, types of chemotherapeutic agents to which the patient is exposed, site of radiotherapy fields and age of patient at treatment. Late effects of therapy are discussed in the following sections.

Secondary solid organ malignancy

Patients treated for HL have an approximately threefold increased risk of solid organ malignancy compared to the general population. The most common secondary malignancies are breast, lung and gastrointestinal tract cancers. Radiotherapy increases the risk of secondary solid organ malignancy at exposed sites, whereas chemotherapy has a more complex association with secondary solid organ malignancy, resulting in an increased risk of some types (e.g. colorectal carcinoma associated with procarbazine use), but also acting to abrogate the risk of others (e.g. reduction of breast cancer risk in patients receiving radiotherapy through reducing ovarian function and inducing premature menopause).

One of the most potent determinants of the risk of developing a secondary solid organ malignancy is the age at which the patient received treatment. This is particularly the case for the risk of breast cancer due to radiotherapy. If significant amounts of breast tissue are irradiated in patients under the age of 20, the relative risk of subsequent breast is approximately 15–25 fold. In contrast, the risk of breast cancer in patients treated at age 40 and over is similar to that of the general population. The explanation for this is likely to be a mixture of increased vulnerability of developing breast tissue to the transforming effects of radiotherapy, and a longer duration of ovarian function post treatment. Of note, however, these estimates of risk are based on older-style radiation fields and radiotherapy techniques.

The risk of subsequent breast cancer (and indeed most cancers) is related to the radiotherapy dose and field size. Patients are offered screening with yearly mammography or MRI from 8 years after treatment or at age 25 years (whichever is later). The risk of lung cancer is increased in those receiving lung exposure and patients should be counselled to cease smoking.

Secondary myeloid malignancy

The risk of secondary MDS/AML is approximately 1% or less with ABVD, but may be up to 4% with escalated BEACOPP. The risk of secondary myeloid malignancy was a particular concern in the early trials using BEACOPP (e.g. HD9); however, the incidence has decreased in subsequent GHSG trials (HD12 and HD15). The reason for this is not clear, but it has been hypothesized to be due to the lower use of radiation in later trials, as well as exposure to fewer cycles of chemotherapy.

Cardiovascular disease

Patients treated for cHL are at an increased risk of late cardiovascular disease. This is contributed to by exposure of the heart to radiotherapy, as well as treatment with anthracycline-based chemotherapy regimens. The types of cardiovascular disease that may occur include premature coronary-artery disease, myocardial infarction, congestive cardiac failure, valvular heart

disease, conduction abnormalities (e.g. atrial fibrillation) and pericardial disease. The latency of coronary artery disease as a result of radiotherapy is approximately 10 years. Importantly, concomitant cardiovascular risk factors including hypertension, hypercholesterolemia and smoking all compound this risk and should be aggressively managed.

Infertility

Women (regardless of age) receiving escalated BEACOPP have severely reduced ovarian reserve and significantly higher rates of sustained amenorrhea post treatment. Women over the age of 30 treated with BEACOPP regimens have a 50% rate of sustained amenorrhea (4 years) post treatment and those over 35 treated with escalated BEACOPP have a less than 5% rate of return to a regular menstrual cycle. In contrast, premature menopause after ABVD alone is not significantly increased.

For male patients, sperm storage can usually be offered before treatment, whereas fertility-sparing measures in female patients are time consuming and, because of treatment delays, not always advisable.

Conclusion

The treatment of cHL has now progressed to the stage where the significant majority of patients will be cured of their disease. A truly enviable situation when compared to many other haematological and non-haematological malignancies. The onus is now on the managing clinical team to deliver this curative therapy in a manner commensurate with individual patient risk in order to minimize the risk of acute and long-term morbidity and mortality. This will be best achieved by a multidisciplinary approach, including both radiation oncologists as well as haemato-oncologists. Moreover, the deepening understanding of the molecular basis of the HRS cell, as well as the development of new therapeutics hold the promise of an even further improvement in outcomes for patients with cHL.

Selected bibliography

Biggi A, Gallamini A, Chauvie S *et al*. (2013) International validation study for interim PET in ABVD-treated, advanced-stage Hodgkin lymphoma: interpretation criteria and concordance rate among reviewers. *Journal of Nuclear Medicine* **54**: 683–90.

Borchmann P, Haverkamp H, Diehl V *et al*. (2011) Eight cycles of escalated-dose BEACOPP compared with four cycles of escalated-dose BEACOPP followed by four cycles of baseline-dose BEACOPP with or without radiotherapy in patients with advanced-stage hodgkin's lymphoma: final analysis of the HD12 trial of the German Hodgkin Study Group. *Journal of Clinical Oncology* **29**: 4234–42.

Engert A, Diehl V, Franklin J *et al*. (2009) Escalated-dose BEACOPP in the treatment of patients with advanced-stage Hodgkin's lymphoma: 10 years of follow-up of the GHSG HD9 study. *Journal of Clinical Oncology* **27**: 4548–554.

Gallamini A, Hutchings M, Rigacci L *et al*. (2007) Early interim 2-[18F]fluoro-2-deoxy-D-glucose positron emission tomography is prognostically superior to international prognostic score in advanced-stage Hodgkin's lymphoma: a report from a joint Italian-Danish study. *Journal of Clinical Oncology* **25**: 3746–52.

Hasenclever D, Diehl V (1998) A prognostic score for advanced Hodgkin's disease: International Prognostic Factors Project on Advanced Hodgkin's Disease. *New England Journal of Medicine* **339**: 1506–14.

Lister TA, Crowther D, Sutcliffe SB *et al*. (1989) Report of a committee convened to discuss the evaluation and staging of patients with Hodgkin's disease: Cotswolds meeting. *Journal of Clinical Oncology* **7**: 1630–6.

Proctor SJ, Wilkinson J, Jones G *et al*. (2012) Evaluation of treatment outcome in 175 patients with Hodgkin lymphoma aged 60 years or over: the SHIELD study. *Blood* **119**: 6005–15.

Younes A, Gopal AK, Smith SE *et al*. (2012) Results of a pivotal phase II study of brentuximab vedotin for patients with relapsed or refractory Hodgkin's lymphoma. *Journal of Clinical Oncology* **30**: 2183–9.

CHAPTER 33

Non-Hodgkin lymphoma: low grade

33

William Townsend[1] and Robert Marcus[2]

[1]University College Hospitals NHS Foundation Trust and King's College London, London, UK
[2]Department of Clinical Haematology, King's College Hospital, London, UK

Introduction

The low-grade non-Hodgkin lymphomas (NHL) included in this chapter are follicular lymphoma, the marginal-zone lymphomas, Waldenström macroglobulinaemia/lymphoplasmacytic lymphoma and mantle-cell lymphoma. Chronic lymphocytic leukaemia (CLL)/small lymphocytic lymphoma (SLL) and hairy-cell leukaemia are also mature B-cell malignancies with some similar characteristics to the low-grade NHLs discussed in this chapter; they are covered in Chapter 27 and T-lymphoproliferative disorders are covered in Chapter 28.

Whilst the low-grade lymphomas have distinct epidemiology, pathogenesis, morphologic, immunophenotypic and clinical features, they share a broadly similar disease course characterized by slow rate of growth and long median survival, with the possible exception of mantle-cell lymphoma, which frequently has a more aggressive course. In the majority of cases, these are incurable malignancies and whilst the progress is usually slow, the development of refractory disease, or transformation to high-grade disease can occur in all subtypes.

Since the previous edition of this book there have been advances in the understanding of the pathogenesis of lymphoma and new therapies are emerging that show great promise. Major improvements in survival have been observed since the introduction of anti-CD20 monoclonal antibodies and the aims of current research are to reduce the toxicity of treatment, while continuing to improve outcomes (Table 33.1).

Epidemiology

Registry data from the Surveillance Epidemiology and End Results (SEER) database in the United States shows that there was a marked increase in the number of cases of NHL reported over several decades at the end of the twentieth century, but the rate of increase appears to have slowed since 2000. It is not clear why there was a rise in incidence, but it is likely that changes in medical practice, reporting conventions, and changes to the classification of lymphomas have had an impact. Numerous studies investigated possible links to environmental agents such as pesticides or industrial exposures, but no firm associations have been identified. The increase in incidence of low-grade lymphomas is less closely linked to the AIDS epidemic as high-grade lymphoma. There is marked geographical variation in incidence of some lymphoma subtypes, which is suggestive of an infectious aetiology. Although low grade NHLs can present at any age they are all more common with increasing age.

Histology and classification of low-grade NHL

The histology and classification of low-grade NHL are covered in Chapter 31. Salient features will be covered in the relevant sections below.

Postgraduate Haematology, Seventh Edition. Edited by A Victor Hoffbrand, Douglas R Higgs, David M Keeling and Atul B Mehta.

Table 33.1 Immunophenotypic profiles of 'low-grade' B-lymphomas.

CD5/10 category	Disease	SmIg	CD19	CD23	CD79b	FMC7	CD200	CD11c	CD25	CD103	CD123	*CyD1
CD 5+/10–	*CLL/SLL*	+wk	+	+	– or wk	–	+	–/+	–	–	–	–
	MCL	+	+	–	+	+	–	–	–	–	–	+
CD 5–/10+	*FL*	+	+	–	+	+	–	–/+	–	–	–	–
CD 5–/10–	*B-PLL*	+	+	–	+	+	–	–	–	–	–	–
	HCL	+	+	–	–	+	+	+	+	+	+	–
	HCLv	+	+	–	–	+	–	+	–	–/+	–	
	S-MZL	+	+	–/+	+	+	–	–+/	+/–	–	–	–
	LPL/WM	+	+	–	+	–/+	–	+/–	+/–	–	–	–

The diseases are initially categorised into groups according to CD5 and CD10 status.
CD 5+/10+ lymphoproliferative disorders are rare, but include FL and MCL with aberrant expression of CD5 and CD10, respectively.
The immunophenotypes indicated are the typical phenotypic profiles; however, phenotypic variation with aberrant over- or under-expression is common.
–/+ indicates that disease is more commonly negative and +/– indicates that it is more commonly positive.
CL/SLLL, chronic lymphocytic leukaemia/small lymphocytic lymphoma; *CyD1, Cyclin D1 (assessed by immunohistochemistry); FL, follicular lymphoma; HCL, hairy-cell leukaemia; HCLv, hairy-cell leukaemia variant; LPL/WM, lymphoplasmacytic lymphoma/Waldenstrom's macroglobulinaemia; MCL, mantle-cell lymphoma; SmIg, Surface-marking immunoglobulin; S-MZL, splenic marginal-zone lymphoma; wk, weak.
The pink shaded boxes indicate the typical profile and most discriminatory CD markers to distinguish CLL from MCL.
The blue shaded boxes indicate the typical HCL phenotype and the green boxes indicate the most discriminatory markers to distinguish HC from HCLv.
(Source: Robin Ireland, King's College London.)

Follicular lymphoma

Follicular lymphoma (FL) is the commonest low-grade NHL and the second commonest of all lymphomas after diffuse large B-cell lymphoma (DLBCL). FL, like other low-grade NHLs, is characterized by a slow rate of progression and a relatively indolent course in most patients, but it is heterogeneous and can progress rapidly, become refractory to treatment or transform to high-grade disease. Like other B-cell lymphoproliferative disorders, the outlook has been much improved by the introduction of anti-CD20 monoclonal antibodies and this, alongside improvements in supportive care, is one of the principal reasons for the improved long-term survival reported from registry data. There are also many novel, non-cytotoxic agents becoming available for the treatment of FL and, it is likely that they will radically change the management of FL in the near future.

Epidemiology of FL

The incidence of FL increases with age with a median age of diagnosis of 64 years, and whilst it can present in childhood, this is rare. The sex distribution is approximately equal, but there is considerable geographical variation in incidence: in Europe and North America, it has an annual incidence of approximately 4 per 100,000 people, but in Asia the incidence is 10 times lower. The long natural history of the disease and increasing survival give rise to a high prevalence, which is predicted to increase further with an ageing population.

Pathology of FL

FL is a malignancy of germinal-centre B-cells, predominantly affecting the lymph nodes, but the spleen, bone marrow, peripheral blood and other extranodal sites can also be involved. The architecture of involved lymph nodes is usually replaced by uniform large follicles with some similarities to normal germinal centres (see Chapter 31).

FL is histologically classified into grades depending on the number of centroblasts per high-power field. Grades 1–2 account for 80–90% of cases, are considered the same and are treated identically. Grade 3 is separated into grade 3a and grade 3b. The outcome of grade 3a FL is similar to that of grade 1–2 disease, whereas grade 3b FL has an aggressive course analogous to DLBCL and is clinically managed as such. Bone marrow is involved in approximately 50% of patients, and the pattern of infiltration is characteristically paratrabecular. The proliferation index, as measured by Ki67, usually reflects the histological grade and is typically <20% in grade I–2 FL and >20% in grade 3 disease. The malignant B cells do not exist in isolation, but are admixed with various T-cell subsets, macrophages and follicular dendritic cells that make up the tumour microenvironment.

Pathophysiology of FL

The characteristic cytogenetic abnormality in FL is the chromosomal translocation t(14;18)(q32;q21), resulting in constitutive expression of the antiapoptotic proto-oncogene *BCL2* (B-cell lymphoma-2) by bringing it into proximity with enhancer sequences of the immunoglobulin heavy-chain gene on chromosome 14. It is over-expressed in approximately 80–90% of grade 1–2 FL, is less commonly found in grade 3a disease and is infrequent in grade 3b. In up to 50% of BCL2-negative cases, 3q27 *BCL6* rearrangements can instead be identified. Numerous additional genetic events have been identified in FL, including mutations in the epigenetic modifiers *EZH2*, *CREBBP* and *MLL2*, loss of function mutations in *TNFRSF14* and mutations in the apoptosis regulator *FAS*. Transformation and progression have been attributed to clonal selection and the accumulation of further genetic events such as upregulation of *MYC* or mutation of *TP53*.

The t(14;18) translocation leads to failure of apoptosis of B cells, but it is clear that this is insufficient to lead to the development of FL. Lymphocytes carrying the t(14;18) translocation can be identified by PCR in the peripheral blood of 50 to 70% of healthy individuals with no evidence of FL. Until recently, an increased predisposition to FL was not apparent in healthy people with this translocation, but a recent study involving 520,000 healthy individuals found that a high frequency of circulating t(14;18) cells (greater than 1 in 10^{-4}) was associated with a 23-fold increased risk of developing FL, sometimes many years later. The clinical significance of this finding is not fully understood, and screening for people carrying t(14;18) is neither practical or desirable at present. The finding of the same *BCL2-IGH* break points in the original peripheral blood samples of t(14;18) cells as in the tumour biopsies of individuals who subsequently developed FL many years later, indicates the existence of a founder clone of premalignant B cells.

In the normal germinal-centre reaction, the B-cell receptor is rearranged and undergoes somatic hypermutation under the influence of the enzyme activation-induced cytidine deaminase (AID). It is postulated that further genetic events leading to the development of FL occur through the off-target influence of AID during repeated germinal-centre reactions of B cells that are resistant to apoptosis as a consequence of the t(14;18) translocation.

The microenvironment is important in the pathophysiology of FL and has an impact on prognosis, with evidence that the gene expression profile of the non-malignant cells infiltrating the tumour is predictive of prognosis. An immune signature consistent with increased T-cell infiltration has been associated with improved outcome compared to an immune signature consistent with increased macrophage and dendritic cell infiltration. These findings have been replicated in subsequent studies, but this has not yet been applied prospectively in the clinical setting.

A growing number of studies, mainly using single-parameter immunohistochemistry, have investigated the role of the microenvironment in the pathogenesis and prognosis of FL, but to date these have yielded inconsistent and sometimes contradictory findings. It is likely that there is a complex interplay between protumour cells such as T follicular-helper cells, antitumour immune-effector cells and T regulatory cells that inhibit the antitumour immune response. It is also evident that the FL B cells themselves are also able to influence the composition of the microenvironment.

Clinical features of FL

FL is often described as 'indolent'; this reflects the slow rate of progression observed in the majority of cases and the fact that many patients are asymptomatic and some do not require treatment for many years after diagnosis. However, the disease progresses rapidly in some patients and, in most cases, FL is an incurable condition characterized by responses to initial therapy with inevitable relapses. Both the response rate and length of remissions were observed to decline with subsequent lines of therapy in the prerituximab era. Studies of untreated individuals have revealed that a significant proportion of patients with asymptomatic, advanced-stage disease never require therapy, and prolonged, spontaneous remissions in untreated patients are also well recognized.

Presentation of FL is most commonly with lymphadenopathy, which is often asymptomatic, but may cause local symptoms or organ compromise. Constitutional symptoms are less frequently reported than in high-grade NHL. Extranodal disease may affect the bone marrow, liver, spleen, gastrointestinal tract or skin; CNS involvement is rare. In a large retrospective review of patients, which included 4167 patients enrolled in clinical trials, 78% had advanced stage, 22% early stage, 19% had B symptoms (weight loss, night sweats or fever), 52% had bone marrow involvement and 38% had areas of extranodal disease other than bone marrow. These clinical characteristics were broadly verified in a prospective review of 942 patients, but may not be wholly representative of all patients with FL, as only patients in clinical trials were included.

There is an associated risk of transformation to high-grade lymphoma, most commonly to DLBCL, but occasionally Burkitt lymphoma. The risk is approximately 3% per year or 15–28% at 10 years. There is considerable variation in the reported rates of transformation, in part due to differences in the definition of transformation (whether histologically confirmed or clinically suspected) and variations in clinical practice with regard to performing biopsies at relapse. Transformation has a very poor prognosis with a median time from histological transformation to death of 1.2–1.8 years.

Rare entities including *in situ* FL, paediatric FL, primary intestinal FL and primary cutaneous FL are described on page 583 of Chapter 31.

Table 33.2 Risk factors for the calculation of FLIPI and FLIPI2 scores.

Feature	Adverse risk factor	
	FLIPI	FLIPI2
Age	>60 years	>60 years
Stage	III–IV	Bone marrow infiltration
Haemoglobin	<120 g/L	<120 g/L
Disease burden	>4 nodal sites of disease	Maximal diameter of largest nodal mass >6 cm
LDH or β2MCG	LDH > upper limit of normal	β2MCG > upper limit of normal

Staging and baseline investigations in FL

FL is clinically staged according to the Ann Arbor staging system which was originally devised for the staging of Hodgkin lymphoma. Stage I–II_A is early-stage disease, while II_B–IV is characterized as advanced-stage disease. Approximately 80–90% of patients with FL present with advanced-stage disease.

Accurate staging is important for prognosis and for planning treatment, especially for determining whether patients with apparent early-stage disease have any distant involved lymph nodes or extranodal sites of disease. Staging routinely involves bone marrow aspirate and trephine biopsy and imaging of the neck, chest, abdomen and pelvis. Since FL is uniformly FDG-avid, positron emission tomography (FDG-PET) scanning is increasingly used alongside cross-sectional imaging. FDG-PET has been demonstrated to be more sensitive than CT and leads to upstaging in 18–30% of patients. FDG-PET scan may be particularly important in apparent early-stage disease where one study showed that 62% of patients were upstaged by FDG-PET; this has therapeutic implications since early- and advanced-stage disease are typically managed differently. Bone marrow infiltration in FL is not reliably detected by FDG-PET, and bone marrow aspirate and trephine therefore remains an important component of staging, even when FDG-PET is performed.

Additional investigations performed at the time of diagnosis include full blood count, biochemistry, including serum lactate dehydrogenase (LDH), β-2-microglobulin (β2MCG), immunoglobulins and protein electrophoresis, and virology (hepatitis B, hepatitis C and HIV testing are considered essential prior to starting immunochemotherapy). Baseline physiological investigations such as measurement of creatinine clearance or cardiac function are required when intensive treatment is being considered.

Predicting prognosis in FL

Determining prognosis in patients with newly diagnosed FL is based on a number of clinical and laboratory parameters. The Follicular Lymphoma International Prognostic Index (FLIPI) incorporates age (>60 versus ≤60 years), stage (III–IV versus I–II), number of involved nodal groups (>4 versus ≤ 4), anaemia (Hb <120 g/L versus ≥120 g/L) and serum LDH (>upper limit of normal (ULN) versus ≤ULN). The FLIPI score was developed from retrospective analysis and uses overall survival (OS) as its end point. Patients are divided into three risk groups, low risk (FLIPI 0–1) with 10-year OS of 71%, intermediate risk (FLIPI 2) with 10-year OS of 51% and high risk (FLIPI 3–5) with 10-year OS of 36% (see Tables 33.2 and 33.3).

The FLIPI was developed in the prerituximab era and, although it has been demonstrated to retain predictive power in patients treated with immunochemotherapy, patients included in the initial analysis are not representative of patients currently treated. Another drawback to the FLIPI score is its use of OS as the endpoint. Whilst OS is a valuable end point in aggressive lymphoma, in low-grade NHL such as FL, surrogate end points such as progression-free survival (PFS) or time to treatment failure (TTF) are often used, as OS is affected by subsequent lines

Table 33.3 Outcomes according to risk group defined by the FLIPI and FLIPI2 scores.

FLPI				FLIPI2			
Risk group	No. of risk factors	5-year OS (%)	10-year OS (%)	Risk group	No. of risk factors	3-year PFS (%)	3-year OS (%)
Low	0–1	91	71	Low	0–1	91	99
Intermediate	2	78	51	Intermediate	2	69	96
High	3–5	53	36	High	3–5	51	84

of therapy and it takes a long time to generate statistically meaningful OS data. These factors have led to the development of a revised prognostic system, the FLIPI2 score.

The FLIPI2 score determines risk according to age (>60 versus ≤60 years), haemoglobin (<120 g/L versus ≥120 g/L), bone marrow infiltration (present versus absent), β2MCG (>ULN versus ≤ULN), and greatest diameter of largest involved lymph node (>6 cm versus ≤6 cm). This score was generated prospectively, with many patients treated with rituximab-containing regimens, and uses PFS as the primary outcome measure. Patients with a FLIPI2 score of 0 are low risk and have a 3-year PFS of 91%, score 1–2 is intermediate risk with a 3-year PFS of 69% and score 3–5 is high risk with a 3-year PFS of 51% (see Tables 33.2 and 33.3).

The FLIPI and FLIPI2 scores are widely used for estimating the prognosis of newly diagnosed patients and they are helpful in clinical trials for describing the distribution of high- and low-risk patients in the population. However, there is no prospective evidence to support their use in determining treatment decisions.

Depth of response to treatment has also been used to provide prognostic information with evidence to show that patients achieving a complete remission (CR) after treatment have significantly longer duration of response than patients achieving only partial remission (PR). Remission status determined by FDG-PET has also been shown to be more predictive of outcome than by conventional imaging. In three studies of patients with advanced-stage disease treated with rituximab-containing induction treatment, 22–26% of patients were PET-negative at the end of induction treatment. Patients who were PET-negative after treatment had a 36–44% improvement in PFS compared to patients who were PET-positive. Significantly worse OS for patients who were PET-positive at end of treatment was also identified in one study (HR 7.0, 95% CI 1.8 – 2.7). Determining the level of molecular response using PCR for the quantification of *BCL2-IGH* rearrangement in the bone marrow after rituximab-based treatment has also been demonstrated to be predictive of duration of response and this is being assessed in ongoing prospective studies.

Management of early-stage FL

Early-stage FL represents 10–20% of all new FL diagnoses. It is important to identify this group because the prognosis, treatment intent and the treatment modalities employed are different from advanced-stage disease; whereas it is anticipated that most treatments in FL will not be curative, its high sensitivity to radiotherapy means that true early-stage disease can sometimes be cured with radiotherapy alone; 24 Gray (Gy) involved field radiotherapy (IFRT) is now recommended with evidence from a randomized trial that this dose is not inferior to higher doses of 40–45 Gy. Very low-dose radiotherapy (4 Gy) has also been assessed and, whilst this gives good disease control and symptomatic relief in the palliative setting, it is inadequate for long-term disease control.

Localized radiotherapy may cure just under 50% of patients. Relapses are typically outside of the original radiotherapy field, suggesting that the disease was in fact disseminated at time of treatment, but was not detectable; for this reason it is likely that patients staged with FDG-PET may have better response to localized radiotherapy. Late relapses beyond 10 years are extremely rare in early-stage FL treated with local radiotherapy.

Despite radiotherapy being recommended for early-stage FL by many national and international guidelines, it is not universally adopted. Studies from the USA showed that only 27% and 34% (respectively) of early-stage patients were treated with radiotherapy alone; other patients were treated with rituximab plus chemotherapy (with or without subsequent radiotherapy), rituximab alone, or observation only.

There is a need for the formal assessment of the treatment options in early-stage disease in randomized trials and the role of FDG-PET should be assessed in parallel; however, trials are slow to recruit due to the low incidence of early-stage disease and require prolonged follow-up due to the excellent outcomes with standard therapy.

In cases where the entire tumour has been removed by excision biopsy, one series has shown that up to 80% of patients have long-term remissions without additional therapy. It may therefore be reasonable for such patients to be observed without additional treatment. The role of FDG-PET scanning in this setting has not been formally assessed, but it would seem logical to consider FDG-PET scan and bone marrow biopsy to confirm the absence of disease elsewhere.

Management of advanced-stage asymptomatic FL

For asymptomatic patients with advanced-stage disease, there is no proven benefit of early chemotherapy over observation. Three separate trials demonstrated that, in this cohort of patients, delaying treatment did not impair disease-specific or overall survival. Many physicians therefore employ a period of observation or 'watchful waiting' in asymptomatic patients, withholding treatment until such a time as it is required due to progressive disease, increasing symptoms or falling peripheral blood counts. When this approach is followed, systemic therapy is delayed by 2–3 years, with approximately 20% of patients never requiring therapy, and, in the largest reported trial, 40% of patients over 70 years had still not required treatment after 10 years median follow-up.

This approach of delaying treatment in asymptomatic patients was originally adopted due to the failure of any treatments to prolong survival and in order to delay exposure to potentially

harmful chemotherapy, but, in the modern treatment era, where survival is improving and treatments are less toxic (leading to fewer treatment-related deaths and potentially fewer secondary malignancies) this approach is being challenged.

In a recent trial performed by the UK study group, patients with advanced-stage, asymptomatic disease were randomized to receive no treatment or rituximab alone with no cytotoxic chemotherapy; this was initially a three-arm study with two rituximab schedules (rituximab induction only or rituximab induction followed by maintenance), but the rituximab induction arm was closed early. Analysis of the observation versus rituximab induction followed by maintenance arms (n = 192 and 187, respectively) demonstrated that patients receiving rituximab had a significantly longer time to initiation of next treatment than the observed cohort (31.1 months versus not reached), and at 3 years median follow-up, the proportion of patients not requiring further therapy was significantly higher in the rituximab arm (88% versus 46%, HR 0.21, 95% CI 0.14–0.31). Another important finding of this study was that there was no deterioration in quality of life in patients treated with rituximab. Longer follow-up is required to determine the effect of single agent rituximab on response to subsequent lines of therapy, survival and transformation.

Another recently reported trial in the treatment of low-tumour-burden FL is the RESORT trial conducted by ECOG (Eastern Cooperative Oncology Group) in which continuous rituximab maintenance was compared with re-treatment with rituximab at relapse in patients who had initially responded to rituximab monotherapy induction (four doses at 1-week intervals). They found no significant difference in the time to treatment failure (TTF) between the two arms, and although a significantly higher proportion of patients in the re-treatment arm required subsequent cytotoxic therapy, this came at the price of a large difference in the number of doses of rituximab received, (median number of doses, 18 in the maintenance arm versus 4 in the re-treatment arm). Another interesting finding of this trial is that 50% of patients randomized to receive only induction rituximab remained in remission at 3 years, with relapsing patients demonstrating a very high response rate to subsequent single agent rituximab in first and second relapses (61% and 67%, respectively), which suggests that the development of rituximab resistance was not a frequent occurrence in this cohort.

In summary, an approach of watchful waiting remains appropriate for most patients with advanced-stage asymptomatic disease, and there is a strong evidence base to support this decision. However, this approach was developed in an era when no treatments had demonstrated improvements in OS and the available treatments were all cytotoxic. There is now a growing body of evidence that single-agent rituximab is an effective, well-tolerated therapeutic option that may be a suitable alternative to watchful waiting, although long-term follow-up of the UK trial is needed before recommending a change in practice.

First-line management of advanced-stage symptomatic FL

To determine whether a patient with advanced-stage disease needs to start treatment, the GELF (Groupe d'Etude des Lymphomes Folliculaires) criteria are widely used. According to these criteria, patients with any site of disease >7 cm in maximal diameter, at least three nodal sites each ≥3 cm in diameter, B symptoms, splenic enlargement, local symptoms, organ compromise or cytopenias should be considered for active treatment.

Current therapy for advanced-stage symptomatic FL is rarely curative, but aims to reduce symptoms, prolong the time until next therapy and improve survival. Contemporary regimens consist of induction typically followed by maintenance therapy. The treatment of grade 3b FL should follow the recommendations for high-grade NHL.

Induction therapy

First-line therapy of advanced-stage symptomatic FL is now rarely with chemotherapy alone, but is usually with rituximab in combination with chemotherapy, or sometimes rituximab alone.

Rituximab was first demonstrated to have activity in FL when used as a single agent in the relapsed or refractory setting, and was subsequently demonstrated to be highly active as a single agent in the frontline setting, where 4 × weekly doses of 375 mg/m^2 gave an overall response rate of 73%. Subsequent trials assessed the addition of rituximab to existing chemotherapy regimens. The addition of rituximab to CVP (cyclophosphamide, vincristine and prednisolone) chemotherapy in the first-line treatment of advanced-stage FL was demonstrated to lead to significant improvements in all outcome measures assessed, including longer TTF, CR rate, time to progression, response duration and, crucially, OS, with the benefits seen in both high- and low-risk patients, as assessed by FLIPI score. The 4-year OS in the R-CVP arm was 83% compared to 77% in the CVP arm (P = 0.029) and this improvement in OS through the addition of rituximab to combination chemotherapy represents a landmark in the treatment of FL.

Three further studies have demonstrated the benefit of adding rituximab to chemotherapy in first-line treatment (see Table 33.4). Hiddemann and colleagues compared R-CHOP (rituximab, cyclophosphamide, doxorubicin, vincristine and prednisolone) to CHOP, and demonstrated a 60% reduction in risk of treatment failure, significantly longer TTF and improved OS at 2 years (95% versus 90%, P = 0.016) in the R-CHOP arm. Herold *et al.* compared R-MCP (rituximab, mitoxantrone, cyclophosphamide and prednisolone) with MCP alone with similar results, and a large French trial demonstrated that rituximab led to improved outcome when added to CHVP+I (cyclophosphamide, doxorubicin, etoposide, prednisolone and interferon), although the survival benefit in this trial was limited to patients with high FLIPI scores. The survival benefit of rituximab has been confirmed in a meta-analysis, which

Table 33.4 Summary of results of key trials of rituximab in combination with chemotherapy in the first line treatment of follicular lymphoma.

Trial experimental regimen (first author, year of publication)	Patients (n)	Median follow-up (months)	ORR (%)	Response outcome measure and duration	OS (%)
R-CVP (Marcus, 2008)	321	53	81	Median TTF 27 months	83 at 4 years
R-CHOP (Hiddemann, 2005)	428	58	96	Median TTF not reached, 2-year TTF 85%	90 at 2 years
R-bendamustine (Rummel, 2013)	139	34	93	Median PFS not reached	84 at 4 years
R-MCP + interferon (Herold, 2007)	201	47	92	Median PFS not reached	87 at 4years
R-CHVP + interferon (Bachy, 2013)	358	96	81	Median EFS 66 months	79 at 8 years

EFS, event-free survival; ORR, overall response rate; OS, overall survival; PFS, progression free survival; R-CHOP, rituximab, cyclophosphamide, doxorubicin, vincristine, prednisolone; R-CHVP, rituximab, cyclophosphamide, doxorubicin, etoposide, prednisolone; R-CVP, rituximab, cyclophosphamide, vincristine, prednisolone; R-MCP, rituximab, mitoxantrone, chlorambucil, prednisolone; TTF, time to treatment failure.

concluded that there is a significant improvement in OS in patients treated with rituximab plus chemotherapy compared to various chemotherapy combinations (HR 0.63, 95% CI 0.51–0.79).

Whilst the benefit from the addition of rituximab to chemotherapy has been conclusively demonstrated, the optimal regimen to use in conjunction with it has not yet been established. A trial conducted by the Fondazione Italiana Linfomi (FIL) study group has shown that R-CVP is associated with inferior TTF compared to R-CHOP or R-FM (rituximab, fludarabine and mitoxantrone), and significantly higher toxicity with R-FM than R-CHOP. Recently, bendamustine has been shown to be highly effective in combination with rituximab (BR), with better outcome and fewer toxicities than R-CHOP (PFS 69.5 months versus 31.2 months, HR 0.58, $P < 0.001$); no difference in OS has been demonstrated to date. The improved outcome with BR is perhaps due in part to a lower toxicity profile, which may facilitate the timely administration of full-dose treatment to a higher proportion of patients.

The evidence to recommend any particular chemotherapy regimen to which rituximab should be added is limited, but six to eight doses of rituximab in combination with six to eight cycles of chemotherapy should be considered the standard of care at present, and this can be anticipated to give an overall response rate of 81–96% with CR in 20–63%. Many clinicians are choosing to use bendamustine as it is well tolerated and appears to be highly effective, although the evidence for this comes from only one randomized trial. Another potential advantage of using bendamustine as first-line rather than CHOP is that an anthracycline can be used at subsequent relapses or transformation if the patient has not received first-line R-CHOP. Fludarabine-based regimens are not usually recommended as first-line due to their increased toxicity, the risk of myelodysplasia and the impairment of stem cell mobilization.

For patients unable to tolerate a full course of immunochemotherapy, options include attenuated chemotherapy with full-dose rituximab or single-agent rituximab given weekly for 4 weeks, as reported in the RESORT trial, which gives durable remissions of 3 years in approximately 50% of patients. Since this was performed in patients with low tumour burden, the results may not be directly applicable to all patients with advanced-stage disease.

Maintenance or consolidation therapy

Rituximab maintenance

Previously, maintenance therapy with interferon was found to be partially effective, but poorly tolerated and difficult to administer. It has now been demonstrated that rituximab maintenance is effective at prolonging remissions after initial single-agent rituximab, after CVP chemotherapy, after rituximab-containing immunochemotherapy (R-CHOP) in the relapsed setting and more recently after rituximab-containing induction immunochemotherapy. This was the main result of the PRIMA trial conducted by the French GELA group in which 1217 patients with previously untreated, advanced-stage FL received one of three induction regimens (R-CHOP, R-CVP, or R-FCM) with responding patients (1019, 84%) randomized to maintenance rituximab (375 mg/m^2 every 2 months for 2 years) or observation. At a median of 3 years' follow-up, the rate of progression was significantly lower in the maintenance arm, PFS 74.9% versus 57.6% (HR 0.55, 95% CI 0.44–0.68, $P < 0.0001$), with the benefit observed in all FLIPI risk groups and patients both in PR and CR at the end of induction. No difference in OS has been identified to date. This important trial has led to the widespread use of rituximab maintenance following first-line treatment with immunochemotherapy.

A number of questions about the use of maintenance rituximab after induction therapy remain unanswered, these include

the optimal duration of maintenance, maintenance versus re-treatment at relapse and whether the benefit of maintenance is restricted to certain subsets of patients.

Whilst rituximab administered every 8–12 weeks for 2 years is considered safe and has proven efficacy, the optimal duration of maintenance has not been determined. Evidence from the RESORT trial suggests that a longer schedule is not harmful, but no benefit of prolonged maintenance over re-treatment at relapse was demonstrated. Provisional data from a trial conducted by the Swiss study group (SAKK) in which patients were treated with single-agent rituximab followed by either 5 years of maintenance or short-term maintenance of only four doses also suggests that prolonged maintenance for 5 years is safe and, in contrast to the RESORT trial, identified improvements in event-free survival (EFS) and PFS with the prolonged schedule over the short regimen; full publication of this trial is awaited.

It is has also been suggested that patients achieving a CR by FDG-PET may not require rituximab maintenance, this is yet to be confirmed in prospective studies. It may also be possible in the future to use the level of molecular response, or other biomarkers to predict patients who are most likely to derive benefit from maintenance therapy after induction therapy.

Radioimmunotherapy consolidation

The conjugation of radioactive isotopes to monoclonal antibodies targeting B cells allows the directed delivery of radiation to the tumour. Given the high radiosensitivity of FL to radiotherapy, radioimmunotherapy (RIT) is an appealing technique for consolidation therapy to eradicate residual sites of disease. Compared to no further treatment, consolidation with ^{90}yttrium-ibritumomab tiuxetan or ^{131}iodine-tositumomab in patients achieving a response after induction therapy has been demonstrated to lead to increases in PFS and conversion of PR to CR is seen in some patients. Although this treatment option is well tolerated in most patients, it has been associated with an increased risk of myelodysplasia, and there have been no direct comparisons of rituximab maintenance versus RIT consolidation. Rituximab should therefore be considered the standard maintenance approach at present, and RIT in this setting should probably only be used within clinical trials.

Autologous stem cell transplantation consolidation in first remission

Another approach to consolidating induction therapy is to use high-dose therapy and autologous stem cell transplant (ASCT) in first remission. Four randomized trials have been conducted comparing ASCT to no further treatment; most were performed before rituximab was routinely used in the frontline setting. Significant improvements in PFS were identified in all but one of these trials, but this has not been shown to lead to an improvement in OS, partly due to excess toxicity (including secondary malignancies) in the ASCT arms and also due to the success of salvage therapy in patients not receiving ASCT. In the current era of highly efficacious induction therapy and maintenance with rituximab, which has a low toxicity profile, there is no role for consolidating first remission with ASCT; it does however remain an important therapeutic option at relapse (see below).

Management of relapsed FL

There are numerous options but no standard treatment for relapsed FL. Relapsing patients should undergo re-staging, repeat biopsy, given the risk of transformation, assessment of symptoms and a review of previous treatment(s).

Localized relapses can be treated effectively with radiotherapy as in the first-line setting; 24 Gy is highly effective but lower doses (4 Gy) may be appropriate, especially for symptom control in the palliative setting.

If relapse is advanced stage and asymptomatic, there is no proven benefit of immediate treatment over an active observation approach and it is therefore common practice for treatment to be deferred until such a time as it is required.

In patients with symptomatic advanced-stage, non-transformed relapse there are a number of options and determining the most appropriate intervention depends on patient characteristics (e.g. age and comorbidities), prior treatment and duration of previous remission.

Immunochemotherapy

For patients who did not receive rituximab at first-line therapy, there is very strong evidence that rituximab-containing immunochemotherapy confers significant improvements in PFS and OS over chemotherapy alone at relapse, and this should therefore be considered in all cases of rituximab-naïve patients who relapse with symptomatic disease. Most patients now receive rituximab as part of their first-line treatment and the benefit of further rituximab has not been formally demonstrated, but most clinicians chose to add it to therapy, especially if durable remissions have been achieved with prior rituximab. The same chemotherapy should be considered if a prolonged remission (>2 years) has been achieved, but if the patient received CHOP or other anthracycline-containing regimens at first-line then there is little scope to receive further anthracyclines. Where remissions have been shorter than 2 years, consideration should be given to switching to a non-cross-resistant chemotherapy regimen. Fludarabine is effective in patients previously treated with alkylating-agent-based therapy, but it has significant toxicities, including profound immunosuppression, is associated with secondary malignancies, especially myelodysplasia, and can impair subsequent attempts at stem cell mobilization and so is not a suitable option in all cases. Bendamustine-rituximab has been demonstrated to be superior to fludarabine-rituximab in the relapsed setting, with significant improvement in PFS, and is therefore preferred particularly if it has not been used as first-line. In fit patients who do not respond to standard regimens, an intensive salvage regimen

should be considered e.g. IVE (ifosphamide, epirubicin and etoposide), ESHAP (etoposide, methylprednisolone, cytarabine and cisplatin) or ICE (ifosfamide, carboplatin and etoposide). Rituximab should be added to these treatment protocols unless patients have received this agent in the previous 6–12 months. When frailty or comorbidities prevent the administration of full-dose immunochemotherapy, consideration can be given to attenuated regimens or single-agent rituximab, which gives response rates of up to 50% in rituximab-naïve patients and evidence shows that it can be successful after prior exposure too.

A number of randomized trials and a meta-analysis have demonstrated that rituximab maintenance leads to significant improvements in both PFS and OS over observation alone after second-line therapy, even in patients who have previously received rituximab-containing immunochemotherapy. It is not clear whether patients who have previously received maintenance rituximab benefit from a further course.

Radioimmunotherapy (RIT) is an effective alternative to immunochemotherapy for relapsed FL, but there are no direct comparisons of RIT in the relapsed setting against other modalities. ^{90}Yttrium-ibritumomab tiuxetan has been demonstrated to give high response rates with durable remissions in heavily pretreated (albeit rituximab-naïve) patients. It is especially useful in patients with non-bulky tumours (maximum diameter of largest involved node <5 cm) and, since it is well tolerated, it is a suitable treatment option for patients who are refractory to other modalities or unable to receive further chemotherapy. The main toxicity is myelosuppression and close monitoring of blood counts after administration is necessary. There is also a concern about increased risk of myelodysplasia, especially in patients heavily pretreated with purine analogues.

For some patients with relapsed FL, the application of newer therapies should be considered within the context of a clinical trial. Emerging novel therapies for the treatment of FL are discussed below.

Transplantation in relapsed FL

Both autologous and allogeneic HSCT have a role in a subset of patients with relapsed or transformed FL.

Autologous stem cell transplantation

A number of non-randomized studies performed before the widespread use of rituximab suggested that ASCT prolongs PFS in relapsed FL compared to historical controls treated with standard chemotherapy. The CUP trial performed in Europe randomized patients who responded to three cycles of CHOP at relapse to either three further cycles of chemotherapy or ASCT. A significant improvement in OS for patients receiving ASCT was demonstrated (2-year OS, 71% versus 46%, HR 0.4 (95% CI 0.18–0.89), P = 0.026). Long-term follow-up of a number of series have indicated that there is a plateau in the survival curves approximately 10 years after ASCT with 12-year OS of 48% in one study suggesting possible cure in some patients. Whilst these data confirm the benefit of ASCT in the treatment of relapsed FL, the studies were performed in the prerituximab treatment era, which makes the application of these findings to today's patients difficult, especially given the relative toxicity of ASCT.

Second malignancies are the most common cause of late non-relapse mortality after ASCT. Although patient age, number of prior chemotherapeutic regimens and use of fludarabine, etoposide and alkylating agents may predispose patients to myelodysplastic syndrome (MDS) or acute myeloid leukaemia (AML), it is now clear that ASCT independently adds to the overall risk. The use of total body irradiation (TBI) in the conditioning regimen is logical, due to the radiosensitivity of FL, but also further increases the risk of subsequent MDS/AML and so is now rarely used.

Despite the plateau in survival after ASCT reported in some series, most patients will eventually relapse due to contamination of the harvest product with lymphoma cells or proliferation of malignant cells that survived conditioning. Molecular remission before transplantation appears to be linked to improved PFS. Therefore, purging the stem cells *in vivo* with rituximab before harvest, and further rituximab post transplant to eradicate residual cells that survived high-dose chemotherapy have both been assessed. The European Group for Blood and Bone Marrow Transplantation (EBMT) LYM1 study assessed *in vivo* purging with rituximab prior to PBSC harvesting and maintenance rituximab for 2 years post transplant in a randomized trial of rituximab-naïve patients in second or subsequent remissions. Although no benefit from prior *in vivo* purging was identified, the second randomization to maintenance or observation showed a significant improvement in PFS with maintenance. Again, it is difficult to interpret how this finding fits with patients who are now almost universally not rituximab naïve at the time of relapse.

In summary, ASCT is a powerful therapeutic option in the management of relapsed FL. It is recommended for the consolidation of remissions in relapsed disease, especially in patients with poor-risk features, such as short duration of response or high FLIPI score, but consideration should also be given to allogeneic HSCT in this setting (see below).

Allogeneic stem cell transplantation

Despite the advances in treatment of FL over recent years, it remains incurable in most cases. In a subset of younger patients with high-risk relapsed FL who have a suitable donor, allogeneic HSCT offers the possibility of cure. Reduced-intensity conditioning (RIC) HSCT is increasingly used over myeolablative conditioning for FL, due to its lower transplant-related mortality (TRM) and enhanced graft-versus-tumour effect. The results from a number of retrospective and registry studies of RIC allogeneic HSCT in FL suggest a relapse rate of 10–44% and TRM of 11–40% at 3 years, thus the 3-year OS is approximately 50–65%. The best published results are from a UK series of 82 patients treated with T-deplete RIC allogeneic HSCTs in whom the TRM

was very low (15%), with an excellent 4-year PFS of 76%; the median age of these patients was 45 years, which highlights that this intensive intervention is not applicable to most patients with FL, the median age of whom is 64 years at presentation.

At present RIC allogeneic HSCT is recommended for patients who relapse after prior ASCT, and should also be considered in younger patients with early relapse and high-risk disease characteristics, although there have been no prospective trials assessing autologous versus allogeneic transplantation.

Novel therapies for FL

There are a large number of new therapies that are currently being assessed in the treatment of B-cell lymphoproliferative disorders, including FL, and it is likely that they will find a role in both the relapsed and first-line settings. Novel approaches include the new generation of anti-CD20 antibodies, restoring the immune response against the tumour, inhibiting the B-cell receptor or its downstream pathways, inhibiting BCL2 or immunoconjugates. These are discussed briefly below.

There are several new anti-CD20 monoclonal antibodies currently being assessed in clinical trials, including ofatumumab and GA101 (obinutuzumab), which have a higher affinity for CD20 than rituximab and have been shown to lead to increased antibody-dependant cellular cytotoxicity *in vitro*. Studies in relapsed disease show efficacy as single agents and that they can be safely combined with chemotherapy to good effect without apparent increased toxicity over rituximab. A large international trial is underway, comparing rituximab chemotherapy with GA101 chemotherapy, followed by rituximab or GA101 maintenance in previously untreated FL and other low-grade NHLs.

Lenalidomide has been shown to be effective in FL and other NHLs. The mechanism of action is not completely understood, but probably includes direct cytotoxicity and restoration of the immune response against the tumour. As a single agent, lenalidomide gave response rates of only 23% in relapsed disease, but much higher response rates have been achieved when it has been used with rituximab in either the relapsed (ORR 74%) and more recently in the first-line setting (ORR 93%).

Other emerging therapies that directly influence the microenvironment include the anti-PD1 antibody, pidilizumab, which has considerable clinical efficacy when used in combination with rituximab in patients with relapsed FL, perhaps through restoration of the anti-tumour immune response.

Exciting results are being achieved in various B-cell malignancies by targeting the downstream signalling of the B-cell receptor with agents such as the PI3Kδ inhibitor idelalisib or the Bruton's tyrosine kinase (BTK) inhibitor ibrutinib, and results from early-stage clinical trials of these agents in low-grade NHL are very promising, with high response rates, even in patients who have received many previous lines of treatment.

The proteasome inhibitor bortezomib has been tested in relapsed FL with low response rates as a single agent and an ORR of approximately 60% in combination with rituximab. It has also been added to R-CHOP, giving an ORR of 90%, but significant neurotoxicity was a problem. The combination of bortezomib, bendamustine and rituximab gave a high response rate, but relatively short duration of response in a Phase II trial. Bortezomib may not be as active in FL as in other NHLs.

Inhibition of *BCL2* with compounds such as the small molecule inhibitor ABT-199 restores the tumour cell's ability to undergo apoptosis. Early-phase clinical trials show a very high level of activity, even in heavily pretreated patients.

Other novel treatment approaches include antibodies conjugated to chemotherapeutic agents such as inotizumab ozogamicin (anti-CD22 conjugated to calicheamicin) and histone deacetylase inhibitors such as vorinostat.

It is likely that the treatment for FL is going to change radically over the coming years and it is possible that treatment comprising of immunotherapy (rituximab or another anti-CD20 antibody) and novel, non-chemotherapy agents will replace current treatments. Numerous clinical trials with long follow-up will be required to determine the optimal combinations, doses and duration of these agents in induction and maintenance therapy. To facilitate the introduction of new treatments, parallel investigation into the use of surrogate end points and the development of biomarkers to predict which patients will benefit most from which treatments will be needed.

Management of transformed FL

FL has approximately a 3% per year risk of transformation to high-grade lymphoma; this often occurs relatively early in the course of the disease and may be accompanied by the development of B symptoms, high LDH and nodal or extranodal masses that may grow rapidly and often discordantly from other involved areas. Areas of transformation have high SUV on FDG-PET scans and this has been used to guide biopsies. In the pre-rituximab era, the outcome after chemotherapy was poor, with reported median survival of 1.2–1.8 years, although better outcomes are reported for patients reaching CR after treatment.

There is limited evidence to direct the optimal therapy of transformed FL. For patients who have not previously received anthracyclines, CHOP should be used, but if CHOP was used previously then alternative salvage regimens should be considered, such as ESHAP, ICE or IVE. In most cases, rituximab is added to chemotherapy regardless of prior rituximab, as long as the tumour continues to express CD20, but this has not been assessed in randomized trial and it is unlikely that it will be. Radiotherapy can be added to chemotherapy if the disease at time of transformation is of limited stage. If remission is achieved, consolidation with high-dose therapy followed by ASCT is often performed in patients able to tolerate this approach. There is some evidence that previously

untreated patients with transformed FL have a good outcome following R-CHOP, and ASCT may be omitted in this cohort.

Suggested algorithm for management of FL

- If grade 3b, treat as high-grade lymphoma.
- If early-stage disease is confirmed by comprehensive staging investigations (bone marrow aspirate and trephine, and FDG-PET scan, if available), treatment is with 24 Gy localized radiotherapy; other options should routinely only be considered within clinical trials.
- If advanced-stage asymptomatic disease, standard of care is watch and wait; if treatment is considered there is some evidence for single-agent rituximab, but the long-term implications of this are not known.
- In symptomatic advanced-stage disease, standard of care is six to eight cycles of rituximab plus chemotherapy. The chemotherapy of choice is not clear, but CHOP may be superior to CVP in patients who can tolerate it, and bendamustine superior to CHOP with lower toxicity. Fludarabine-based treatments are not recommended in the first-line setting
- For patients who are unable to receive full-dose immunochemotherapy, consider attenuated chemotherapy with full-dose rituximab or rituximab alone.
- Consolidation of first remission with rituximab maintenance for 2 years is recommended in most instances, with evidence that this prolongs remissions and time to next treatment, but no evidence that it prolongs survival.
- Relapsing patients should be re-staged and repeat biopsy should be performed wherever possible. Treatment depends on stage at relapse, prior treatments and duration of previous remission.
- ASCT is a useful treatment option for patients achieving second or subsequent remissions. There is a lack of evidence to guide when it is best used in the rituximab treatment era.
- Allogeneic haematopoietic stem cell transplantation after reduced-intensity conditioning can be considered in younger patients, with a suitable donor, who have aggressive early relapsed disease.
- Transformation has a poor prognosis and optimal treatment is not known. Options include CHOP, salvage regimens (e.g. ESHAP) with ASCT in patients achieving remission.
- Novel therapies are highly promising, but good-quality randomized trials are required to understand how these treatment options should best be used.

Marginal-zone lymphomas

There are three types of marginal-zone lymphoma (MZL): nodal, splenic and extranodal lymphomas of mucosa-associated lymphoid tissue (MALT). MALT lymphomas are the commonest of these, representing about 8% of all B-cell NHLs, of which at least one-third are primary gastric lymphoma. MZLs are characterized by the presence of mature B cells expressing surface immunoglobulin, CD20 and pan-B-cell markers and are typically negative for CD5 and CD10. Their pathology is covered in detail in Chapter 31. There is strong evidence that the pathogenesis of the MZLs is linked to inflammation, chronic antigenic stimulation and autoimmunity.

MALT lymphoma

The mean age of patients with MALT lymphomas at presentation is 60 years. Presentation depends on the site(s) of involvement; B symptoms are rare. These lymphomas are often localized to the site of origin, but disseminated disease is reported in 25% and 46% of cases of gastric and non-gastric MALT lymphoma, respectively. The commonest sites of presentation are stomach (33%), intestine (3–9%), salivary glands (16%), orbit (10–12%) and lung (6–10%). Dissemination can be either to other sites of mucosal tissue or to lymph nodes, bone marrow, spleen or liver. MALT lymphomas typically follow an indolent course, but transformation to high-grade disease is reported in about 10% of cases. The relatively frequent finding of disseminated disease at presentation indicates the need for comprehensive staging investigations, which, in addition to cross-sectional imaging and bone marrow biopsy, should include investigations appropriate to the site of presentation. The use of FDG-PET scanning is not routine, as MZL has a low proliferation fraction and is not uniformly FDG-avid. There is a lack of consensus regarding optimal staging since the Ann Arbor system does not adequately cover the breadth of presentations and therefore does not accurately guide treatment decisions or prognosis.

Gastric MALT lymphoma

The stomach is the most common site of MALT lymphomas. It is often diagnosed at endoscopy performed for the investigation of upper gastrointestinal tract symptoms (dyspepsia, pain, bleeding) or anaemia. Investigation should include endoscopic biopsy from each region of the stomach and duodenum, with samples sent for histopathology, cytogenetics to detect the t(11;18) translocation, and for determination of *Helicobacter pylori* status, which can be performed histologically as well as by urea breath test or serological testing. The depth of lymphoma penetration has prognostic importance and can be assessed by endoscopic ultrasound.

It has been demonstrated that chronic inflammation secondary to *H. pylori* infection is strongly linked to the development of gastric MALT lymphoma. *H. pylori*-positive, localized gastric MALT lymphoma can be effectively treated by clearance of *H. pylori* infection with triple therapy, typically consisting of a proton pump inhibitor in combination with amoxicillin and clarithromycin. This approach is most effective in cases where

the disease is limited to the mucosa without nodal involvement and when t(11;18) is not detected. There is a wide range in the reported CR rate with this approach, which is partly due to differences in response criteria used, but CR rates of 60–90% have been reported, with low relapse rates of between 5–17%. The time to histological remission is also variable: the median time for optimal response is 6 months, but can be up to 24 months after *H. pylori* eradication.

The addition of chemotherapy after treatment for *H. pylori* was assessed in the LY03 trial (chlorambucil versus observation after *H. pylori* eradication). Although the power of this study was low, it demonstrated that there was no benefit in PFS or OS through the addition of chlorambucil, and therefore consolidation with chemotherapy after successful *H. pylori* eradication is not routinely used.

Based on the available data, it is reasonable to treat localized disease with *H. pylori* eradication therapy, followed by strict endoscopic follow-up, including multiple biopsies two months after therapy to confirm elimination of the bacteria, with repeated endoscopies and biopsies at least twice per year for 2 years to assess histological regression of the tumour. Even when histological CR has been achieved, a population of monoclonal B cells remain detectable by PCR in approximately 50% of patients; the clinical implication of this is not yet clear.

Cases with the t(11;18) translocation are frequently resistant to *H. pylori* eradication, but have a low rate of high-grade transformation; although eradication therapy may be attempted it is unlikely to lead to remission and additional treatment options should be considered.

There is limited evidence to guide the optimal treatment of patients with localized disease that is *H. pylori* negative, or fails to respond to eradication therapy. In cases of unsuccessful eradication, a second-line high-dose antibiotic regimen may be administered. If the tumour fails to regress, despite second-line therapy, or in cases of *H. pylori*-negative localized disease, chemotherapy (e.g. oral cyclophosphamide/chlorambucil) in combination with rituximab, radiotherapy, surgery or combinations of these may be tried. Data from small series suggest that 30 Gy involved field radiotherapy to the stomach is highly effective for patients with stage I/II disease, providing a 5-year disease free survival of 93%, and this is a suitable treatment option for some patients with limited-stage disease that is *H. pylori* negative, fails to respond to eradication therapy or carries the t(11;18) translocation. A randomized trial comparing rituximab in combination with chlorambucil versus chlorambucil alone in MALT lymphoma (including patients with localized gastric MALT lymphoma that was resistant to antibiotic therapy) has demonstrated improvements in CR rate and EFS, but no difference in OS through the addition of rituximab. Given the indolent nature of MALT lymphoma, advanced stage disease can be managed with close observation in asymptomatic cases or with chemotherapy in combination with anti-CD20 monoclonal antibody therapy in those who are symptomatic. There is no evidence at present that maintenance anti-CD20 immunotherapy is beneficial.

Non-gastric MALT lymphoma

Non-gastric MALT lymphomas can arise at many sites, often in response to chronic infection, inflammation or autoimmune disease. Ocular adnexal MALT lymphoma is linked to infection with *Chlamydia psittaci*, MALT lymphoma of the thyroid is associated with Hashimoto's thyroiditis and salivary gland MALT lymphoma is more common in patients with Sjogren's syndrome. Disseminated disease is more frequently identified in non-gastric than gastric MALT lymphomas.

The management of non-gastric MALT lymphoma depends on site, stage and the presence or absence of symptoms. Radiotherapy is often effective in early-stage disease, doses of 25–35 Gy are typically used and this leads to local control and long-term freedom from disease in a high proportion of cases. In general, advanced-stage disease can be managed with observation alone until such a time as symptoms develop. Advanced-stage symptomatic disease may be treated with chemotherapy in combination with rituximab; radiotherapy retains a role for the management of local symptoms.

OAML may be effectively managed with antibiotics alone; analysis of seven studies (four retrospective) comprising 131 patients, mostly treated with doxycycline, found an overall response of 45%, with 25-month median follow-up. Interestingly, responses are seen, even in patients in whom *Chlamydia psittaci* cannot be detected, but best responses were attained in patients with evidence of infection. Similar responses to antibiotics have not been observed in MALT lymphoma occurring at other sites.

Splenic marginal-zone lymphoma

Splenic marginal-zone lymphoma (SMZL) is a rare disorder, comprising less than 2% of lymphoid neoplasms. SMZL often presents with splenomegaly, moderate lymphocytosis with bone marrow involvement and associated cytopenias. The circulating lymphocytes are characterized by the presence of cytoplasmic projections, and are termed villous lymphocytes. Splenic hilar lymph nodes may be infiltrated, but peripheral lymph nodes and extranodal sites are usually spared. Bone marrow infiltration is typically intrasinusoidal. A paraprotein, usually IgM, can be detected in up to 25% of cases, and autoimmune phenomena occur in 20% of patients. The immunophenotype of SMZL is not unique and the diagnosis therefore requires the exclusion of other B-cell NHLs (see Chapter 31). There is no pathognomonic chromosomal aberration associated with SMZL; allelic deletion of 7q32 appears in around one-third of patients and there is conflicting evidence as to whether this deletion defines an unfavourable prognostic subgroup. There is a possible association with hepatitis C virus (HCV) infection in some cases; a Japanese group identified HCV in 36% of SMZL cases and,

in areas where HCV is endemic, higher rates of SMZL are recorded.

SMZL has a generally indolent course, even where there is marrow involvement; approximately 10% of cases will eventually transform to high-grade lymphoma. The median OS is approximately 10 years. A prognostic index comprising haemoglobin <120 g/L, LDH greater than upper limit of normal, and serum albumin <35 g/L has been proposed with 5-year survivals of 88%, 73%, and 50% in the low- (no risk factors), intermediate- (one risk factor) and high-risk (more than one risk factor) groups, respectively. No definitive curative therapy has been established and therefore therapeutic intervention is generally reserved for patients with symptomatic disease; up to 50% of patients never require treatment.

Splenectomy has been the principal treatment for symptomatic patients or those with cytopenias, providing they are fit enough to undergo this procedure. Splenectomy results in resolution of cytopenias and improvement in quality of life in up to 90% of patients, but does not lead to resolution of marrow infiltration or lymphocytosis. The median time to next treatment after splenectomy is 8 years. Retrospective studies have reported better OS in splenectomized patients than those treated with chemotherapy, although this may reflect selection bias. Splenic irradiation may be considered in patients with poor performance status or as a palliative measure. Single-agent rituximab leads to durable remissions in up to 90% of patients in retrospective studies, and rituximab in combination with chemotherapy is increasingly used as an alternative to splenectomy.

There are reports of patients with SMZL and concurrent HCV infection being successfully treated with interferon-α, with or without anti-viral therapy, which has been associated with durable partial and occasional complete remissions. New drug treatments for HCV that show high response rates need further investigation as a treatment modality in this subgroup of patients.

Given that there is upregulation of NFκB and evidence of dependency on signalling through the B-cell receptor in SMZL, new treatments being used in other B-cell lymphomas, which require further investigation, include the proteasome inhibitor bortezomib and BTK inhibitors such as ibrutinib.

Nodal marginal-zone lymphoma

Nodal marginal zone-lymphoma is a rare entity comprising approximately 10% of MZL. There is little evidence to guide the optimal management and therefore recommendations tend to follow those for other low-grade NHLs, i.e. in localized disease, radiotherapy is effective and may be curative, in advanced-stage asymptomatic disease, a watch and wait approach is acceptable, whilst in symptomatic advanced-stage disease, chemotherapy in combination with rituximab is frequently used.

Waldenström macroglobulinaemia

Waldenström macroglobulinaemia (WM) is a clinical syndrome that occurs in patients with a pathological diagnosis of lymphoplasmacytic lymphoma. WM represents a clonal expansion of post-germinal-centre lymphoid cells expressing IgM, CD19 and CD20, with an associated IgM paraprotein. A point mutation in the *MYD88* gene leading to activation of NFκB and other pathways has recently been identified in 90–100% of cases of WM and this may help to distinguish it from other lymphomas with a similar phenotype. It is mainly a disease of the elderly, with a median age at presentation of over 70 years and a median survival from diagnosis of 60 months. WM is rare, with an incidence in the UK of 0.55 per 100,000 per year.

High IgM levels of greater than 50 g/L may cause hyperviscosity syndrome, with symptoms including fatigue, headache, blurred vision, mucosal bleeding, impaired cognition or reduced level of consciousness. Other clinical manifestations of WM include haemostatic abnormalities, polyneuropathies (which may be related to anti-myelin-associated glycoprotein (MAG) antibodies), cryoglobulinaemia, cold agglutinin haemolytic anaemia (CHAD), acquired von Willebrand's disease and, rarely, the consequences of light-chain tissue deposition such as amyloidosis.

Depending on the degree of marrow infiltration, patients may present with mild anaemia or with severe pancytopenia. Anaemia is the most common indication for initiation of treatment. Approximately one-third of patients presents with lymphadenopathy, splenomegaly or hepatomegaly.

Factors predictive of poor prognosis include age over 65 years, haemoglobin less than 115 g/L, thrombocytopenia, serum monoclonal protein concentration above 70 g/L and raised β2MCG. These factors form the basis of the International Prognostic Staging System for WM (IPSSWM) which defines three risk groups (see Table 33.5).

Response assessment in WM is according to the International Workshop on WM (IWWM) guidelines and includes

Table 33.5 Risk stratification in Waldenström macroglobulinaemia according to the International Prognostic Scoring system for WM.

Risk group	Number of risk factors	Median survival (months)
Low	0–1 (except age)	143
Intermediate	Age (alone) or any two risk factors	99
High	3–5	44

Risk factors: Age >65 years, haemoglobin ≤115 g/L, platelet count ≤100×10^9/L, β2MCG >3 mg/L, monoclonal IgM >70 g/L.

assessment of serum paraprotein, bone marrow infiltration, lymphadenopathy/organomegaly (if present at baseline), as well as assessment of signs and symptoms attributed to the disease. Assessment of bone marrow is particularly important and serum IgM should not be used alone to assess response, especially after rituximab therapy. Responses can be slow to achieve in WM and the timing of response assessment needs to be carefully considered.

Management of Waldenström macroglobulinaemia

A 'watch and wait' approach is appropriate for asymptomatic patients, but most patients will eventually require treatment. One study showed that 59% of patients progress within 5 years and 75% of patients eventually require treatment with a median follow-up of 15 years. Indications for treatment have been defined by the IWWM and include constitutional symptoms, symptomatic lymphadenopathy or splenomegaly, hyperviscosity syndrome, haematological suppression due to marrow infiltration and IgM-related syndromes (including CHAD and neuropathy). Patients presenting with hyperviscosity symptoms or asymptomatic high plasma viscosity in the presence of comorbidities, should receive adjuvant plasmapheresis.

Alkylating agents such as chlorambucil with or without prednisolone have been the mainstay of treatment for patients with symptomatic WM, producing remissions in 50–75% of patients. Several studies have shown rituximab to be active in WM, leading to responses in up to 50% of previously untreated patients when used as a single agent, and a rituximab-containing regimen should now be used in the treatment of all patients with symptomatic WM. Rituximab can cause a sudden rise in serum IgM and viscosity in some patients, the so called 'IgM flare' phenomenon, and can occur in up to a quarter of rituximab-treated patients, therefore, close monitoring for evidence of hyperviscosity is recommended. A number of groups have recommended omitting rituximab until the second or third cycle of chemotherapy, which seems a sensible precaution; alternatively, pre-emptive plasmapheresis may be used alongside systemic therapy in patients with a high serum IgM.

Frontline therapy of WM should now include rituximab in combination with one of: dexamethasone and cyclophosphamide, with a 2-year PFS 67%, a suitable regimen for frail elderly patients; bendamustine (ORR 86%); or purine analogues, which are active in both previously untreated patients and those who fail to respond to alkylating agents or in relapsed disease, with response rates of 38–85% and 30–50%, respectively. Disadvantages of purine analogue therapy include prolonged lymphopenia with consequent risk of infections; prophylaxis against opportunistic infections with aciclovir and cotrimoxazole or pentamidine during treatment and for approximately 6 months after treatment is therefore essential. Another concern is the well-recognised risk of Richter transformation that occurs in in 5–10% of WM patients and the development of secondary MDS/AML in patients treated with purine-analogue-containing regimens. These data favour limiting exposure of WM patients to nucleoside analogues, particularly younger patients. The high response rates and low toxicity of bendamustine make this a suitable first-line treatment option in WM, especially if there is disease bulk, it does not appear to prevent successful stem cell harvest. There is some evidence from retrospective series that maintenance rituximab improves PFS and time to next treatment in patients responding to rituximab-containing regimens and this is currently being assessed prospectively following bendamustine-rituximab induction.

Bortezomib has been shown to be effective as a single agent or in combination with dexamethasone and rituximab. Response rates of up to 83% are reported and the time to response is short, making this useful when rapid disease control is needed, for example if there is a high IgM paraprotein. The current UK 'R2W' trial is randomizing patients with previously untreated WM to either FCR or BCR (bortezomib, cyclophosphamide, rituximab). Neurotoxicity is the principal concern with bortezomib-based therapies and close monitoring is required.

The immunomodulatory drugs thalidomide and lenalidomide have been demonstrated to be effective either as single agents or in combination with steroids or rituximab, although neuropathy limits the use of thalidomide and profound anaemia has been reported with lenalidomide.

Everolimus, an inhibitor of the mammalian target of rapamycin (mTOR) pathway shows activity and is currently being assessed prospectively, the selective BTK inhibitor ibrutinib has also been shown to be active in a number of prospective studies and in a recently published trial inducing high levels of durable responses specifically in patients whose cells express the MYD88 mutation. This has led to ibrutinib gaining a US license in this setting.

It remains to be seen whether the new generation of anti-CD20 monoclonal antibodies, such as obinutuzumab (GA101) or ofatumumab are superior to rituximab in WM. A combination of novel agents with chemotherapy or immunotherapy may increase treatment efficacy by inhibiting diverse pathways that are important for tumour growth and survival.

Stem cell transplantation for Waldenström macroglobulinaemia

Several small studies have reported encouraging results with autologous SCT in heavily treated patients, giving high response rates (95%), with improved PFS in some patients. However, at least 50% of patients subsequently relapse, and the impact of this approach on long-term survival is not known. Autologous transplantation is most appropriate in younger, fitter patients with chemosensitive disease and short duration of response to first therapy. There is evidence to suggest a graft-versus-disease

effect in WM and, therefore, allogeneic HSCT may be appropriate in some young patients, with a 5-year PFS of 49–56% and OS of 62–64% reported. The 3-year non-relapse mortality was 33% after myeloablative conditioning and 23% after reduced-intensity conditioning in the largest registry series reported to date. The emergence of many new treatment options in WM makes it difficult to define the place of allogeneic HSCT in WM.

Guidelines for management of Waldenström macroglobulinaemia

The IWWM published recommendations for the treatment WM in 2009 and these have been updated in 2014 to incorporate bendamustine and bortezomib in first-line treatment options, whilst CHOP and purine analogues are no longer considered first-line options. Suggestions for the management of WM, taking into account the 2009 and 2014 IWWM recommendations are presented below:

- First-line treatment for symptomatic patients should be with DRC, bendamustine-rituximab, or bortezomib-rituximab. CHOP and purine analogue based treatments are not recommended in most cases. Frailer elderly patients can be treated with single-agent rituximab or oral fludarabine.
- Patients with relapsed disease can be treated with an alternative first-line agent or with the same agent again. In the salvage setting, re-use of a first-line regimen or use of a different regimen should be considered alongside novel agents, preferably within a clinical trial.
- Autologous SCT may be considered for patients with refractory or relapsing disease.
- Allogeneic transplantation should usually only be undertaken in the context of a clinical trial.
- Plasmapheresis should be considered as interim therapy until definitive therapy can be initiated.
- Patients with WM should be enrolled into clinical trials whenever possible.
- Patients should be stratified according to the IPSSWM.
- Response assessment should be according to the IWWM criteria.

Mantle-cell lymphoma

Mantle-cell lymphoma (MCL) comprises approximately 5% of all NHL cases. It typically occurs in older patients with a median age at diagnosis of 65 years and is more common in males. Most cases are characterized by the translocation t(11;14)(q13;q32) leading to over-expression of cyclin D1 and subsequent cell-cycle dysregulation. Upregulation of the neural transcription factor SOX11 has been identified in 90% of cases. MCL is staged according to the Ann Arbor system, presentation is usually with advanced-stage disease and extranodal involvement is common, often involving the gastrointestinal tract. CNS involvement is rare at presentation, but may occur at relapse. MCL is typically FDG-avid and therefore FDG-PET scanning is increasingly used for staging and response assessment.

Prognosis in MCL

MCL has a worse prognosis than the other 'low-grade' NHLs; the median overall survival is only 3–5 years, but in approximately 15% of patients the disease has a more indolent course. Patients with indolent disease usually have lymphocytosis and splenomegaly, with infrequent nodal involvement. These cases are also associated with hypermutated immunoglobulin heavy-chain variable gene and the absence of SOX11.

Patients with MCL can be stratified into three main prognostic groups according to the MCL International Prognostic Index (MIPI), which incorporates age, performance status, LDH and leucocytosis. The low-risk MIPI group has a median OS in excess of 5 years, whilst the OS in the intermediate- and high-risk MIPI groups is only 51 months and 29 months, respectively. Additional factors that may affect prognosis include Ki67 score, β2MCG and CNS involvement. Ki67 has now been incorporated into the MIPI score to give a 'biological MIPI' (MIPIb), although significant interobserver variability in quantification of Ki67 limits its general applicability.

First-line management of MCL

The management of MCL is now undergoing rapid change, as increased understanding of the pathways implicated in the pathogenesis of the disease is leading to the use of novel, molecularly targeted therapies.

The majority of patients require therapy at presentation, but approximately 25% of patients may be observed initially, and retrospective data show that this delays initiation of therapy for a median of 12 months. In the rare cases of early stage-I disease, radiotherapy, with or without adjuvant chemotherapy, can lead to durable remissions.

The addition of rituximab to first-line chemotherapy increases overall response and CR rates, PFS and, in at least two studies, improvements in OS. Rituximab plus one of the 'standard' chemotherapy regimens such as CHOP, FC (fludarabine and cyclophosphamide), MCP (mitoxantrone, cyclophosphamide, prednisolone), FCM (fludarabine, cyclophosphamide, mitoxantrone) or bendamustine leads to an ORR of 71–96% with a CR rate of 32–60%, but the risk of relapse is high, the median PFS in these studies ranging between 17–37 months. The optimal chemotherapy regimen has not been established, but the European MCL network trial found that R-CHOP was significantly better than R-FC, partly due to lower toxicity, and the R-bendamustine trial published by Rummel and colleagues demonstrated a PFS advantage for R-bendamustine over R-CHOP in the 94 patients who had MCL with a median

PFS of 35 versus 22 months. Single-agent rituximab has low efficacy in MCL and should only be considered if chemotherapy is contraindicated. The European MCL network trial for elderly MCL patients included a second randomization to rituximab or interferon maintenance and reported a PFS and OS benefit for rituximab maintenance; this should therefore be considered in patients not fit for more intensive therapies (see below).

Despite high remission rates with 'standard' rituximab-chemotherapy combinations, remissions are typically short lived and therefore, in patients who can tolerate it, a more intensive treatment strategy is appropriate. There is evidence from large, albeit non-randomized Phase II studies, that intensive chemotherapy followed by ASCT in first remission may be superior to standard treatment; this observation is independent of rituximab use. In the Nordic 2 study, 160 patients were treated with dose-intense R-CHOP (R-maxi-CHOP), alternating with cycles of rituximab plus high-dose cytarabine followed by ASCT. The CR rate after ASCT was 90% and an impressive 10-year OS of 58% has been reported. The success of this regimen may well be due to the intensive use of cytarabine and *in vivo* purging with rituximab prior to harvest.

A randomized trial by the European MCL network in patients <65 years of age demonstrated improved survival, with three cycles of R-CHOP and three cycles of R-DHAP (dexamethasone, cytarabine, cisplatin), followed by high-dose cytarabine-based ASCT, than with six cycles of R-CHOP followed by standard myeloablative ASCT, confirming the importance of cytarabine in the context of intensive therapy in the treatment of MCL. Despite this, the role of ASCT has not been definitively confirmed in any prospective RCT.

A dose-intensified approach with rituximab, hyperfractionated cyclophosphamide, vincristine, doxorubicin and dexamethasone, alternating with high-dose methotrexate and cytarabine (R-Hyper-CVAD) without ASCT also achieved very high response and survival rates in Phase II studies, but it is associated with significant dose-limiting toxicity.

For patients fit enough to tolerate an intensive treatment approach, standard of care would be a first-line chemotherapy regimen, incorporating rituximab and high-dose cytarabine, with ASCT consolidation. The achievement of molecular remission by flow cytometry or PCR has been independently associated with improved outcome after a variety of treatments, but it is not yet routinely used to inform treatment decisions outside of clinical trials. For less fit patients, the optimal treatment at present is either rituximab-bendamustine or R-CHOP, followed by rituximab maintenance. Novel agents are likely to change this approach in the near future.

Allogeneic transplantation is usually reserved for patients with relapsed disease and best outcomes have been achieved with reduced-intensity conditioning in patients with chemosensitive relapse. A retrospective analysis of reduced-intensity allogeneic HSCT for MCL reported to BSBMT comprising 70 patients showed a 5-year relapse risk of 67%, PFS of 14% and OS of 37% with a TRM of over 20% at 1 year.

Relapsed disease and novel treatments

There is no standard treatment for relapsed MCL. Treatment decisions should be based on duration of response to prior treatments and performance status. Therapeutic options in relapsed disease are now changing rapidly, with the introduction of numerous novel therapies, and some of these options are now also being investigated in the first-line setting.

Bortezomib is highly efficacious in MCL and has been used either alone or in combination with chemotherapy regimens such as CHOP or R-CHOP in the relapsed setting, and it is now being assessed in prospective trials as part of first-line treatment.

Targeting B-cell receptor signalling pathways with drugs such as the BTK inhibitor ibrutinib or the PI3Kδ inhibitor idelalisib is a promising approach, with high response rates in multiply relapsed patients. The addition of ibrutinib to bendamustine-rituximab is now being assessed as first-line therapy in patients >65 years.

Lenalidomide is also a highly effective single agent in MCL and is also currently being evaluated in conjunction with bendamustine-rituximab in newly diagnosed elderly patients, with ibrutinib in relapsed disease and as a maintenance strategy with rituximab.

Temsorilimus and everolimus are mTOR inhibitors that have been demonstrated to be active in relapsed MCL and are now also being investigated in the frontline setting.

As in FL, BCL2 inhibitors such as ABT-199 may prove to be effective, although there is little data specifically in MCL at present. Similarly, the new generation of anti-CD20 monoclonal antibodies will likely be of benefit in patients who are resistant to rituximab and may prove more efficacious than rituximab in first-line treatment too.

Given the numerous new treatment options in MCL it is clear that wherever possible, patients should be enrolled in clinical trials; this is especially true for relapsed disease.

First remissions of between 5 and 6 years are now being achieved with intensive therapies in younger patients and with somewhat shorter durations of first response in older patients; MCL remains an incurable disease that significantly shortens life expectancy in the majority of patients. It is to be hoped that novel agents, including pathway inhibitors, will increase efficacy, with no corresponding increase in toxicity.

Selected bibliography

Ardeshna KM, Smith P, Norton A *et al.* (2003) Long-term effect of a watch and wait policy versus immediate systemic treatment for asymptomatic advanced-stage non-Hodgkin lymphoma: a randomised controlled trial. *Lancet* **362**(9383): 516–22.

Dimopoulos MA, Kastritis E, Owen RG *et al*. (2014) Treatment recommendations for patients with Waldenstrom macroglobulinemia (WM) and related disorders: IWWM-7 consensus. *Blood* **124**(9): 1404–11.

Dreyling M, Ghielmini M, Marcus R *et al*. (2014) Newly diagnosed and relapsed follicular lymphoma: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. *Annals of Oncology* **35**(3): iii76–82.

Geisler CH, Kolstad A, Laurell A *et al*. (2008) Long-term progression-free survival of mantle cell lymphoma after intensive front-line immunochemotherapy with in vivo-purged stem cell rescue: a non-randomized phase 2 multicenter study by the Nordic Lymphoma Group. *Blood* **112**(7): 2687–93.

Marcus R, Imrie K, Solal-Celigny P *et al*. (2008) Phase III study of R-CVP compared with cyclophosphamide, vincristine, and prednisone alone in patients with previously untreated advanced follicular lymphoma. *Journal of Clinical Oncology* **26**(28): 4579–86.

McKay P, Leach M, Jackson R *et al*. (2012) Guidelines for the investigation and management of mantle cell lymphoma. *British Joural of Haematology* **159**(4): 405–26.

Rummel MJ, Niederle N, Maschmeyer G *et al*. (2013) Bendamustine plus rituximab versus CHOP plus rituximab as first-line treatment for patients with indolent and mantle-cell lymphomas: an open-label, multicentre, randomised, phase 3 non-inferiority trial. *Lancet* **381**(9873): 1203–10.

Salles G, Seymour JF, Offner F *et al*. (2011) Rituximab maintenance for 2 years in patients with high tumour burden follicular lymphoma responding to rituximab plus chemotherapy (PRIMA): a phase 3, randomised controlled trial. *Lancet* **377**(9759): 42–51.

Solal-Celigny P, Roy P, Colombat P *et al*. (2004) Follicular lymphoma international prognostic index. *Blood* **104**(5): 1258–65.

Treon SP, Tripsas CK, Meid K *et al*. (2015) Ibrutinib in previously treated Waldenström's macroglobulinemia. *New England Journal Medical* **272**: 1430–40.

Zucca E, Conconi A, Laszlo D *et al*. (2013) Addition of rituximab to chlorambucil produces superior event-free survival in the treatment of patients with extranodal marginal-zone B-cell lymphoma: 5-year analysis of the IELSG-19 Randomized Study. *Journal of Clinical Oncology* **31**(5): 565–72.

CHAPTER 34

Non-Hodgkin lymphoma: high grade

34

Jessica Okosun[1] and Kate Cwynarski[2]

[1]Barts Cancer Institute, London, UK
[2]Royal Free Hospital, London, UK

Introduction

The term high-grade non-Hodgkin lymphomas (NHL) originally referred to one of three categories defined by the International Working Formulation (IWF) classification system introduced in 1981 and included just three histological subtypes: large-cell immunoblastic lymphoma, lymphoblastic lymphoma and Burkitt lymphoma. Today a number of clonal heterogeneous pathological entities fall under this broad umbrella term, which in essence describes lymphomas that are fast-growing or highly proliferative and considered as aggressive lymphomas (see also Chapter 31). They exhibit distinct epidemiology, aetiology, pathogenesis, morphologic, immunophenotypic and clinical features. Paradoxically, despite the aggressive nature of the lymphoma, a significant number of patients can now be cured with the advent of more intensive combination chemotherapy regimens. Recently, significant strides have been made in our understanding of the molecular pathogenesis of lymphomas and how this can be potentially translated into novel therapeutic strategies.

This chapter will focus on the pathogenesis, clinical diagnoses and management of the common high-grade lymphomas. In addition, some of the rarer high-grade lymphomas (B- and T-cell) will also be discussed. Common T-cell lymphomas are covered in Chapter 28.

Epidemiology

High-grade lymphomas represent approximately 60% of all NHLs, with diffuse large B-cell lymphoma (DLBCL) accounting for the bulk. The annual incidence rate varies markedly, as expected, with lymphoma subtype and also with age, with particular types being more prevalent in children and younger adults, whilst others are increased in elderly individuals. The overall incidence of NHL has increased in the Western World over recent decades, rising by 3–5% per annum, in part explained by improved diagnostic techniques, increased life expectancy and a somewhat higher incidence of NHL in immune-suppressed patients. There remains a slight male predominance.

Classification of high-grade lymphomas

The classification of high-grade lymphomas is dependent on morphological, immunophenotypic, immunohistochemical, genetic and molecular features, which form the basis of the current World Health Organization (WHO) classification system (Table 34.1). Precise subtype classification is discussed in detail in Chapter 31. Although most high-grade lymphomas fall into well-defined categories based on their characteristic profiles, there remain subgroups of patients with intermediate or overlap features that present a diagnostic and therapeutic challenge. New developments in the understanding of the pathogenesis and improved molecular diagnostic tools may further refine the current classification system.

Aetiology

With the exception of a handful of high-grade NHL subtypes, the causative factors that underpin lymphomagenesis are unknown. Certain infectious agents are known to play a role in the pathogenesis of specific high-grade lymphomas, such as

Postgraduate Haematology, Seventh Edition. Edited by A Victor Hoffbrand, Douglas R Higgs, David M Keeling and Atul B Mehta.

Table 34.1 Classification of high-grade lymphomas in accordance with WHO classification.

- Diffuse large B-cell lymphoma (DLBCL), NOS
- Primary mediastinal (thymic) large B-cell lymphoma
- Burkitt lymphoma
- B-cell lymphoma, unclassifiable, with features intermediate between DLBCL and Burkitt lymphoma
- Primary central nervous system lymphoma
- Lymphomas associated with HIV infection
- Post-transplant lymphoproliferative disorders (PTLD)
- Peripheral T-cell lymphoma (covered in Chapter 28)

Rarer variants:

- Intravascular large B-cell lymphoma
- Plasmablastic lymphoma
- Primary effusion lymphoma
- Adult T-cell leukaemia/lymphoma
- Extranodal NK/T-cell lymphoma, nasal type
- Enteropathy-associated T-NHL
- Hepatosplenic T-cell lymphoma
- Anaplastic large-cell lymphoma (cutaneous and systemic)
- Aggressive NK-cell lymphoma

Epstein–Barr virus (EBV) in Burkitt lymphoma, post-transplant lymphoproliferative disorders, plasmablastic lymphoma and extranodal NK/T-cell lymphoma; human herpesvirus-8 (HHV8) in primary effusion lymphoma and multicentric Castleman's disease; and human T-cell lymphoma/leukaemia virus 1 (HTLV1) in adult T-cell lymphoma/leukaemia. Immunosuppression as a result of human immunodeficiency virus (HIV) or after allogeneic stem cell or orthotopic transplantation is a well-recognized risk factor, with an increased incidence of lymphomas in this group of patients. Familial aggregation of high-grade lymphomas has rarely been reported, suggesting there is no clear genetic susceptibility. Scant epidemiological studies have implicated occupational exposure to insecticides and pesticides as a risk factor in the development of specific lymphomas. However, the exact impact of these factors and their putative role in contributing to lymphomagenesis is not apparent.

Molecular basis of lymphomas (see also Chapter 18)

In the majority of lymphomas, chromosomal translocations were considered the primary genetic event and have been the most extensively studied genetic alteration. A common feature is balanced reciprocal chromosomal translocation, mostly involving the immunoglobulin (IG) gene loci, usually the IGH locus (14q32) or less often, the IGL loci (κ; 2p11 or λ; 22q11), and a variety of partner oncogenes that subsequently become deregulated leading to apoptosis inhibition and/or uncontrolled cell proliferation (Table 34.2).

Since the last edition of this chapter, the advent of newer technologies, such as microarray chips and next-generation sequencing (NGS) and reduced costs for its application, has allowed the genetic landscape and molecular mechanisms that underlie almost all of the major lymphomas to be enumerated at a much finer resolution. The details will be discussed within each subtype.

Diagnosis

The diagnosis of high-grade lymphoma is made based on a combination of clinical features, histological and, more recently, molecular confirmation (Chapter 31).

Clinical features

High-grade lymphomas present in a variety of ways. Localized or widespread painless lymphadenopathy is usually evident and may be accompanied by B symptoms (unexplained fever of 38 °C or higher, night sweats and loss of more than 10% of body

Table 34.2 Chromosomal translocations found in high-grade non-Hodgkin lymphoma.

Type of lymphoma	Translocation	Genes involved	Frequency	Effect of alteration
Burkitt lymphoma	t(8;14)(q24;q32) t(8;2)(p11/2;24) t(8;22)(q24;q11)	c-*MYC* and IgH c-*MYC* and Igκ c-*MYC* and Igκ	100%	Cell proliferation, differentiation and apoptosis
Diffuse large B cell lymphoma	t(3;4)(q27;q32) and variants t(14;18)(q32;q21) 8q24 and various partners	*BCL-6* and various partners *BCL-2* and IgH c-*MYC* and IgL or other non-IG genes	35% 15–30% 5-14%	Deregulated BCL6 activity Apoptosis inhibition
Anaplastic large B cell lymphoma	t(2;5)(p23;q35)	*ALK* and *NPM*	80%	Signal transduction

weight in 6 months). Given the rate of tumour proliferation, patients can present with compression symptoms and/or organ dysfunction due to rapidly enlarging lymphadenopathy or extranodal masses. Superior vena cava syndrome caused by a bulky mediastinal mass is also more frequent in certain subtypes of NHL, such as primary mediastinal B-cell lymphoma. Neurological presentation due to CNS infiltration is rare, although not an uncommon feature of Burkitt lymphoma. Unlike indolent NHL, bone marrow involvement is less common, occurring in less than 20% of cases, and has been associated with poorer outcomes despite intensive treatment. Patients can present with symptoms secondary to deranged haematological or biochemical parameters, such as anaemia and hypercalcaemia.

Laboratory investigations

Histological diagnosis from either excision or core biopsy of a lymph node, bone marrow or extranodal mass is essential. Fine-needle aspiration (FNA) has been used to exclude an infection (i.e. mycobacterial) or suggest lymphoma, but can lead to delay in establishing a diagnosis. FNA should not be performed, except in suspected cases of T-lymphoblastic lymphoma or Burkitt lymphoma, when a diagnosis can be made on the morphology and immunophenotype, and therapy can be urgent. Immunohistochemistry is routinely performed on tissue sections of lymph nodes and, in specialized units, cytogenetic and immunoglobulin or TCR rearrangement analysis, as well as, more recently, DNA microarray and genetic mutation testing.

Anaemia at diagnosis is usually indicative of widespread disease and may reflect a non-specific manifestation of malignancy. However, it can also be due to hypersplenism or bone marrow infiltration. The white count is variable. Overspill of lymphoma cells into the blood can occur in aggressive lymphoma, representing the so-called 'leukaemic phase', but this is infrequent and usually seen in later stages of disease, and can infrequently be seen by light microscopy at diagnosis. Hypoalbuminaemia is another non-specific feature associated with a systemic disturbance and is indicative of a poor prognosis. A raised level of lactate dehydrogenase (LDH) is usually associated with advanced disease and was identified as an important independent prognostic factor in 1993 (within the International Prognostic Index) and remains so today (Table 34.3). Rarely, a paraprotein is detectable in less than 5% of histologically aggressive lymphomas, but the significance of its presence is unclear.

Table 34.3 Prognostic factors for aggressive lymphoma: International Prognostic Index (IPI).

Age >60 years	
Ann Arbor stage III or IV	
Elevated serum lactate dehydrogenase	
ECOG Performance Score ≥2	
Number of extranodal areas involved >1	
Risk groups (Score)	**5-year OS (%)**
Low (0–1)	73
Low-intermediate (2)	51
High-intermediate (3)	43
High (4 or 5)	26

Staging

The Ann Arbor staging system, initially developed for Hodgkin lymphoma, is used in adults with NHL (see Chapters 32 and 33). Inspection of Waldeyer's ring is particularly important, and a bone marrow biopsy is routinely performed. The presence of high-grade disease in the bone marrow is prognostically important and may direct CNS prophylaxis and concomitant low-grade disease may also be identified. However, whether assessment of the bone marrow is mandatory in all patients at diagnosis is debated in the positron emission tomography (PET) era. Radiography with computerized tomography (CT), magnetic resonance imaging (MRI) or PET scanning is used for initial staging of the disease and is of value in monitoring response to therapy and detection of residual disease or relapse.

As treatment may differ for patients with stage I disease and more advanced stages, especially in the context of DLBCL, a PET scan at diagnosis can have significant therapeutic implications. At the end of therapy PET has been incorporated into the revised response criteria.

Cerebrospinal fluid (CSF) examination should be considered if there are clinical signs of CNS disease, or for patients presenting with high-grade NHL at certain anatomical sites, such as paranasal sinus, breast, epidural area or testis, where CNS prophylaxis is generally recommended. Cytological assessment by cytospin and immunophenotyping should be considered. The prognostic and therapeutic implications of identifying disease in the CSF solely by flow cytometry are much debated. Indications for CNS prophylaxis, using either intrathecal or high-dose intravenous chemotherapy, is described later (see section on CNS prophylaxis in high-grade non-Hodgkin lymphoma). Intrathecal prophylaxis should be administered at the time of the first CSF examination in these patients. Serology for HIV and hepatitis B (HBsAg, HBcAb and HBsAb) and hepatitis C should be performed in all patients as appropriate concomitant antiviral therapy may be indicated during and after therapy.

Treatment

Treatment of NHL is mainly based on the histological findings, although therapeutic decisions are influenced by the patient's age, comorbidities, performance status and disease extension.

There are also a number of subtypes of lymphoma with distinct clinical and pathological features and therapeutic strategies that will be discussed separately. Some of the NHLs are potentially curable and thus access to timely sophisticated diagnostic services and expert clinical opinion is necessary to ensure accurate diagnosis and tailored therapy.

Particular considerations prior to therapy

A formal assessment of cardiac function should be considered for older patients and those with a previous history of diabetes or cardiac disease if anthracycline therapy is considered. Accurate analysis of renal function is necessary before the administration of high-dose methotrexate, platinum-containing regimens and high doses of other renally-excreted agents.

Diffuse large B-cell lymphoma

DLBCL is by far the most common NHL, representing 35–40% of all newly diagnosed NHL and more than 80% of aggressive lymphomas. According to the latest Haematological Malignancy Research Network's registry data, the overall incidence rate for DLBCL in the UK is increasing at 8.4 per 100,000, with a marginal preponderance in males compared to females at a ratio of 1.2:1. The median age of patients at diagnosis is approximately 70 years, but the age range is broad, although it is rarely diagnosed in children and young adults. Although DLBCL is potentially curable, it remains a challenging lymphoma to manage because of the biological and clinical heterogeneity of the disease, and this is reflected by the 2008 WHO classification of a broad category, designated DLBCL, not otherwise specified, to encompass several morphological, molecular and immunohistochemical variations.

Aetiology

The majority of DLBCLs have no underlying causative factor and arise *de novo*. However, a small proportion of cases occur as a result of transformation from a previously indolent disorder such as follicular lymphoma, marginal-zone lymphoma or chronic lymphocytic leukaemia. A small proportion of cases are associated with risk factors occurring in the setting of an underlying immunocompromised state, including concomitant infections with EBV and HIV, post solid-organ transplantation and in elderly patients.

Pathogenesis and molecular basis of DLBCL

DLBCL is a clonal neoplasm of large B-lymphoid cells with a diffuse growth pattern. Importantly, pivotal studies over the last decade have provided substantial insights to explain the genetic and biological heterogeneity of DLBCL. Firstly, large-scale gene expression profiling (GEP) studies were instrumental in dividing DLBCL cases that were morphologically indistinguishable at the histological level into at least three distinct molecular subtypes: germinal-centre B cell (GCB), activated B cell (ABC) and primary mediastinal B-cell (PMBL). The distinct genes expressed by each of these subtypes reflected the different stages of B-cell development from which the malignant B cell arose, the putative cell of origin (COO). The derivation of each subtype from differing cells of origin in addition explains their distinctive molecular features, clinical characteristics and outcomes.

Cell of origin subtypes

The GCB subtype largely expresses genes that are reminiscent of normal germinal-centre B cells such as *CD10* and *LMO2*. Additionally, GCB DLBCLs are characterized by high expression of the transcriptional repressor gene, *BCL6*, which can also be translocated or mutated in DLBCL and exhibit highly mutated immunoglobulin genes with ongoing somatic hypermutation (SHM), further evidence that supports its germinal-centre origin. In contrast, ABC DLBCL characteristically expresses several genes that are normally upregulated in plasma cells with activated chronic B-cell receptor (BCR) signalling, such as nuclear factor (NF)-κB and IRF4 (MUM1), but appear to be blocked from full differentiation into plasma cells by additional genetic alterations that interfere with *Blimp1*, the master regulator of plasmacytic differentiation.

A block in differentiation seems to be an important early step in the pathogenesis of the ABC subtype, but the nature of the exact COO is unknown. Nevertheless, these results suggest that the GCB and ABC subtypes are derived from B cells at different stages of differentiation: GCB DLBCL appears to originate from GC centroblasts, whereas ABC DLBCL is postulated to arise from post-GCB cells blocked from plasmacytic differentiation. The third molecular subtype, PMBL, will be discussed later, as this now represents a distinct entity in the WHO classification. Importantly, stratification of DLBCL patients according to the cell of origin has clinical prognostic relevance and potential therapeutic implications, with GCB DLBCL demonstrating significantly more favourable response and overall survival rates compared with ABC DLBCL, independent of the IPI, following treatment with anthracycline-based chemotherapy. Moreover, even in the setting of combination immunochemotherapy, the worse outcome for ABC-DLBCL was not abrogated (3-year progression-free survival (PFS): 74% GCB; 40% ABC; 3-year overall survival (OS): approximately 80% GCB; 45% ABC, $P < 0.001$) (Figure 34.1).

Use of immunohistochemistry for subtype classification

Although GEP provides an in-depth landscape of the relative gene expression and clearly subdivides DLBCL into its molecular subtypes, its translation as a diagnostic tool into routine clinical practice has been hampered by: (i) the cost, (ii) turnaround

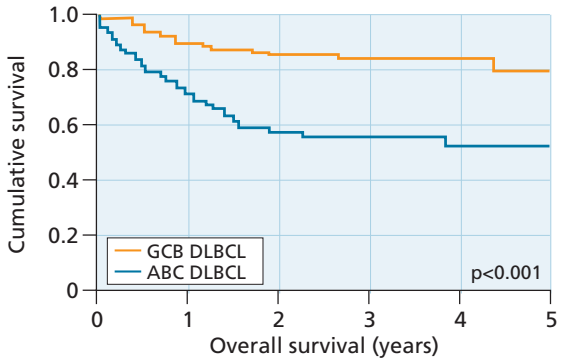


Figure 34.1 Kaplan Meier survival curves demonstrating significantly different outcomes according to molecular subtype. (Source: Frick *et al*., 2012 [*Best Pract Res Clin Haematol*. 2012 **25**(1): 3–12.]. Reproduced with permission of Elsevier.)

time required for analysis and interpretation, (iii) the inability to simultaneously assess the histology of the tumour and (iv) the need for RNA extracted from good-quality fresh lymphoma biopsies. Given the prognostic nature of subtype categorization, surrogate approaches were evaluated using protein-based tests, in the form of immunohistochemistry (IHC) antibody panels for specific germinal-centre and post-germinal-centre markers, in order to distinguish GCB from ABC lymphomas, as this could be performed cheaply and quickly on routine formalin-fixed paraffin-embedded (FFPE) tissues. One of the earliest IHC algorithms used, the Hans algorithm, relied on a three-antibody-staining approach, using CD10, BCL6 and MUM1/IRF4 to discriminate GCB from ABC DLBCLs, with 83% concordance compared with GEP. A number of other algorithms were developed, such as Choi and Tally, which built on the Hans algorithm by incorporating additional IHC markers such as GCET1, FOXP1 and LMO2. However, there remain discrepancies in the standardization and utility of IHC algorithms in approximating GEP, and this was confirmed in a recent study that compared a number of the IHC classifier systems, finding a poor concordance across the systems for reliably determining COO subtypes and conflicting prognostic relevance, especially for patients treated in the rituximab era, and although these algorithms are routinely used in clinical laboratories, they have yet to be incorporated into the standard of care.

More recently, two different GEP assays (DASL and Nanostring) have been reported, demonstrating feasibility with FFPE lymphoma tissue and both could accurately discriminate the COO subtypes with high accuracy, concordance and rapid turnaround time, allowing for real-time decision-making regarding treatment based on COO. There remains a need for international standardization of the methodologies used to determine molecular DLBCL sub-typing if this is to be integrated into routine clinical diagnostics and indeed inform therapy stratification.

Genetic abnormalities in DLBCL

The advent of high-resolution next-generation sequencing technologies has allowed a comprehensive and unbiased enumeration of cancer genomes at an unprecedented rate that could only be imagined a decade ago. In DLBCL, these types of studies have highlighted, in particular, that the ABC and GCB DLBCL are driven by their dependence on different oncogenic pathways.

In GCB DLBCL, in addition to the expression of germinal-centre markers, they harbour a number of genetic alterations restricted to its subtype. A significant proportion of GCB DLBCLs have a t(14;18) (q32; q21) that leads to constitutive over-expression of BCL2 and over 50% of GCB DLBCLs harbour mutations in the BCL2 gene. Previous studies have suggested that BCL2 over-expression in DLBCL to be associated with poor prognostic significance, but this appears to have been overcome by the inclusion of rituximab into therapy. One of the most relevant findings from recent DLBCL genomic studies was the unexpected high incidence of alterations in genes encoding proteins involved in epigenetic regulation, particularly histone and chromatin-modifying enzymes. Mutations in a histone methyltransferase, *EZH2*, that encodes a protein that serves as a critical enzymatic component of a complex involved in gene silencing, was identified in approximately 20% of GCB DLBCL patients. Notably, all the mutations were centred within the catalytic SET domain of the protein. Its exact role in lymphomagenesis is unclear, but it appears to be essential for germinal-centre formation. Concurrently, recurrent aberrations affecting *CREBBP* and *EP300*, two related histone acetyltransferase (HAT) enzymes, have been identified in about 15–30% of DLBCLs. Functionally, the mutant proteins were deficient in their ability to acetylate BCL6 and TP53, supporting a tumour-suppressor role. More recently, mutations in another histone methyltransferase, *MLL2*, encoding a protein involved in gene transcription, were identified in about 30% of DLBCL patients. Together these findings implicate the importance of epigenetic deregulation in lymphomagenesis.

However, in contrast, constitutive NF-κB activation and signalling is the pathogenic hallmark of ABC DLBCLs, occurring in virtually all cases and promoting cell proliferation and survival. This constitutive activation is mediated by a number of genetic alterations. Components of the B-cell receptor (BCR), CD79A and CD79B are mutated in 20% of ABC-DLBCLs with these aberrations leads to chronically active BCR signalling and downstream NF-κB activation. Likewise, about 30% of ABC DLBCLs harbour mutations in *MYD88*, specifically in the intracellular Toll/IL-1 receptor domain of this adaptor molecule, which has the potential to activate NF-κB, as well as the JAK/STAT signalling pathway. Additionally, both positive and negative regulators of NF-κB signalling are aberrant in ABC-DLBCLs. Roughly 10% of ABC DLBCL patients harbour activating mutations in the coiled-coil domain of *CARD11*, a positive regulator of NF-κB signalling, with the gain-of-function mutations further promoting NF-κB signalling. The negative

regulator of NF-κB signalling, A20 (TNFAIP3), is subject to disruptive mutations and deletions in up to 30% of ABC DLBCLs. Loss of A20 could thus promote lymphomagenesis, at least in part, by inducing inappropriately prolonged NF-κB signalling. Interestingly, in most cases, there is an overlap of these genetic alterations suggesting a complementary and non-redundant role of these alterations.

These unique oncogenic pathways for both DLBCL subtypes present many potential targets for biological agents that are currently being exploited and tested in preclinical and early-phase clinical studies.

Role of the microenvironment in DLBCL

In addition to the biological properties of the lymphoma cells, the composition of the tumour microenvironment milieu and genes expressed by these non-tumour cells have been shown to have relevance in the pathogenesis and outcomes of DLBCL. About 10% of DLBCLs aberrantly express CD5, a surface antigen normally expressed by T- cells, with these patients commonly presenting with aggressive clinical behaviour, including advanced stage, elevated LDH and extranodal site involvement. Moreover, CD5 positivity has independent prognostic importance and is additive to the IPI in the rituximab era. Additionally, by separating the tumour from the non-tumour population, GEP-derived unique stromal gene signatures were identified, that also predict survival in patients with DLBCL. The stromal-1 signature reflects genes expressed by the extracellular matrix and histiocyte infiltration, such as secreted protein, acidic and rich in cysteine (SPARC) and connective tissue growth factor (CTGF), are associated with a good prognosis. The stromal-2 signature, which reflects blood vessel density, is prognostically unfavourable. A more recent study used a simplified two-gene model to predict overall survival in patients with DLBCL, the first gene *LMO2*, expressed by the tumour cells and the second gene, *TNFRSF9*, expressed by the tumour microenvironment. The precise role of the expression of a number of these tumour microenvironment markers is unknown.

Prognostic factors

Despite advancements in the treatment of DLBCL, the outcome of patients with DLBCL remains variable. Although potentially curable, about 40% of patients will die with relapsed or refractory disease. There is therefore a pressing need to identify patients for whom standard treatment measures are inadequate, so that alternate therapeutic strategies may be employed. A number of clinical and biological prognostic or predictive markers have been described.

The International Prognostic Index (IPI) was devised in 1993 (Table 34.3) and remains the most widely used prognostic tool. It takes into account five clinical parameters determined at diagnosis: age >60, elevated serum LDH, ECOG performance status (PS) >1, Ann Arbor stage III or IV and number of involved extranodal sites ≥2. Based on the number of features scored by each patient, they are categorized into one of four risk groups with predicted 5-year OS ranging from 26% to 73%. An age-adjusted IPI was also defined for patients <60 years old (Table 34.4) and a revised IPI (R-IPI) outlined in the rituximab era with three risk groups identified. More recently, clinical data from the National Comprehensive Cancer Network (NCCN) database (n = 1650) collected during the rituximab era (2000–2010) has allowed an 'enhanced IPI' tool to be devised. Comparable to the IPI, five similar predictors (age, LDH, sites of involvement, stage, PS) were identified, with a potential maximum of 8 points. However, this prognostic tool accounted for the continuous nature of specific variables such as LDH (normalized LDH of >1 to ≤3 scored 1 point and 2 points if ≥3) rather than the '+' or '–' nature of the original IPI. Additionally, the specific site of extranodal involvement was found to be more prognostic in this model compared to one based solely on the number of extranodal sites, irrespective of location. Four risk groups were formed and compared favourably with the IPI, with the NCCN-IPI being additionally better at discriminating low- and high-risk subgroups (5-year OS: 96% versus 33%) compared to the IPI (5-year OS: 90% versus 54%), respectively. Despite the validity of the various IPI risk models, they remain limited in their ability to predict which patients will follow a very aggressive course. A number of additional clinical predictive factors have been reported. Sex-dependent variability in outcomes has been demonstrated in a number of studies following rituximab-based treatment, with males having an inferior outcome due to their more rapid rate of rituximab clearance compared to females, therefore raising the question of differential dosing. Bone marrow involvement is implicated in poorer prognosis after adjusting for the IPI. Other factors associated with adverse outcomes include: absolute lymphocyte:monocyte ratio, maximum tumour diameter and elevated serum free light chains.

Clinical prognostic features remain surrogates for the underlying biological differences in patients. DLBCL COO subtype (GCB versus ABC) is one of the key molecular prognostic markers, as already described. Recent studies have also implicated MYC or BCL2 alone or in combination to have a negative impact

Table 34.4 Age-adjusted International Prognostic Index (IPI).

Ann Arbor stage III or IV Increased lactate dehydrogenase concentration Performance score of >2	
Risk groups (Score)	**5-year OS (%)**
Low (0)	83
Low-intermediate (1)	69
High-intermediate (2)	46
High (3)	32

Table 34.5 Chemotherapy protocols used in high-grade NHL.

ACVBP	Doxorubicin, cyclophosphamide, vindesine, bleomycin, prednisolone
BEAM/mini-BEAM	BCNU, etoposide, cytarabine, melphalan
CHOP	Cyclophosphamide, doxorubicin, vincristine, prednisolone
CODOX–M/IVAC	Cyclophosphamide, doxorubicin, high-dose methotrexate/ifosfamide, etoposide and high-dose cytarabine
DHAP	Dexamethasone, high-dose cytarabine and cisplatin
DVIP	Dexamethasone, etoposide, ifosfamide and cisplatin
EPOCH/DA-EPOCH	Etoposide, prednisolone, vincristine, cyclophosphamide, doxorubicin
ESHAP	Etoposide, cytarabine, methylprednisolone, cisplatin
GCVP	Gemcitabine, cyclophosphamide, vincristine, prednisolone
Hyper CVAD	Hyperfractionated cyclophosphamide, vincristine, doxorubicin and dexamethasone
ICE	Ifosfamide, carboplatin, and etoposide
MACE–CYTABOM	Doxorubicin, etoposide, prednisolone, cytarabine, bleomycin, vincristine, methotrexate
MACOP B	Methotrexate, doxorubicin, cyclophosphamide, vincristine, prednisolone, bleomycin
VACOP B	Etoposide, doxorubicin, cyclophosphamide, vincristine, prednisolone, bleomycin

on DLBCL outcomes. *MYC* rearrangement occurs in 5–10% of DLBCL patients and demonstrated significantly inferior 5-year PFS and OS in retrospective series when compared to non-MYC rearranged DLBCLs (PFS: 31% versus 66%; OS: 33% versus 72%). BCL2 over-expression occurs in approximately 40–60% of DLBCLs and although, in the prerituximab era, its presence was associated with worse outcomes, this does not appear to be the case using rituximab-based therapies. The presence of both *BCL2* and *MYC* rearrangements, termed 'double hit', occurs in about 5–10% of DLBCLs and has dismal outcomes of less than a year with standard R-CHOP (rituximab, cyclophosphamide, doxorubicin, vincristine, prednisolone; see Table 34.5) therapies. Strategies using more intensive chemotherapies or concomitant use of biological agents in this subgroup of patients are being explored.

Somewhat surprisingly, despite the wealth of information on prognostic markers, none are currently in routine use for stratifying patients and directing therapy.

Treatment

First line: early-stage disease

Patients with stage 1A and 2A DLBCL are considered to have early-stage disease. The British National Lymphoma Investigation (BNLI) reported a 10-year DFS and OS of 35% and 67%, respectively, in 243 patients who were treated with local radiotherapy (RT) for stage I/IE disease. Although DLBCL is a highly radiosensitive tumour, it is clear that local radiotherapy is insufficient and systemic treatment with chemotherapy is necessary.

A significantly improved DFS and OS in early-stage DLBCL patients who received combined modality treatment with a brief chemotherapy course (three cycles of R-CHOP) followed by consolidative radiotherapy has been reported in a prospective randomized trial. The five-year OS and PFS for stage I and II disease treated with R-CHOP and consolidative radiotherapy were 92% and 82%, respectively; whereas without radiotherapy, OS and PFS were 73% ($P = 0.0007$) and 68% ($P = 0.0003$), respectively. At present, abbreviated chemo-immunotherapy (three to four cycles of R-CHOP) followed by involved field radiotherapy (IFRT) is a preferable approach for early stage DLBCL with non-bulky disease, especially at sites of presenting disease where radiation is well tolerated, with lower risks of late toxicities. An alternative approach in patients whose site of presenting disease are at potential risk of long-term late toxicities from radiotherapy is following a full course of chemo-immunotherapy (six cycles of R-CHOP) without radiotherapy.

One randomized study has shown that the outcome for patients older than 60 years with limited stage disease and no risk factors is equivalent after four courses of R-CHOP chemotherapy to that after four courses of R-CHOP chemotherapy and IFRT. Consideration of this approach may be appropriate if radiotherapy is not thought feasible.

First line: advanced-stage disease

Patients with more advanced stages of disease (stage 2B to stage 4) should receive combination chemo-immunotherapy. Several studies have shown the benefit of adding rituximab to chemotherapy making R-CHOP or R-CHOP-like regimens the gold standard in advanced-stage DLBCL. Furthermore, the age of the patient (<60 or >60 years) and IPI risk group (score <3 or ≥3) provides certain considerations for choice of therapy. A summary of selected completed clinical trials is presented in Table 34.6.

Based on results from the MiNT study conducted prior to the use of PET scans in response assessment, R-CHOP with the addition of radiotherapy to sites of primary bulky disease has been recommended for young patients with low- or intermediate-risk IPI. A recent study demonstrated that dose-intensive R-ACVBP (see Table 34.5) followed by sequential consolidation (rituximab, ifosfamide, etoposide and cytarabine) represents an efficacious alternative showing significantly

Table 34.6 Results from selected Phase III clinical trials for first-line DLBCL treatment in patients <60 years and >60 years.

Study	Regimen	No. of patients	Criteria	Age (years)	Outcome	Reference
MInT	**A** R-CHOP-like ×6 vs. **B** CHOP-like ×6	824	aaIPI: 0 or 1	18–60	3-year EFS: 79% (A) vs. 59% (B)* 3-year OS: 93% (A) vs. 84% (B)*	Pfreundschuh *et al*. 2006
LNH03-2B (GELA study group)	**A** AR-ACVBP ×4 and consolidation vs. **B** R-CHOP ×8	380	aaIPI: 1	18–59	3-year EFS: 81% (A) vs. 67% (B)* 3-year PFS: 87% (A) vs. 73% (B)* 3-year OS: 92% (A) vs. 84% (B)*	Recher *et al*. 2011
NCRI	**A** R-CHOP-14 ×6 vs. **B** R-CHOP-21 ×8	1080	All IPIs	>18	2-year PFS: 75.4% (A) vs. 74.8% (B)# 2-year OS: 82.7% (A) vs. 80.8% (B)#	Cunningham *et al*. 2013
NHL-B2 (German High-grade Non-Hodgkin Lymphoma Study Group)	**A** CHOP-14 ×6 vs. **B** CHOP-21 ×6	831 (70% DLBCL)	IPI: ≥1	61–75	5-year EFS: 43.8% (A) vs. 32.5% (B)* 5-year OS: 53.3% (A) vs. 40.6% (B)*	Pfreundschuh *et al*. 2004
LNH98-5 (GELA study group)	**A** R-CHOP-21 ×6 vs. **B** CHOP-21 ×6	399	All IPIs	60–80	5-year EFS: 47% (A) vs. 29% (B)* 5-year PFS: 54% (A) vs. 30% (B)* 5-year OS: 58% (A) vs. 45% (B)*	Feugier *et al*. 2005
RICOVER-60 (German High-grade Non-Hodgkin Lymphoma Study Group)	**A** CHOP-14 ×6 vs. **B** CHOP-14 ×8 vs. **C** R-CHOP-14 ×6 vs. **D** R-CHOP-14 ×8	1222		61–80	3-year EFS: 47.2% (A) vs. 53% (B) vs. 66.5%(C) vs. 63.1% (D) 3-year OS: 67.7% (A) vs. 66% (B) vs. 78.1% (C) vs. 72.5% (D)	Pfreundschuh *et al*. 2008
LNH03-6B (GELA study group)	**A** R-CHOP-14 ×8 vs. **B** R-CHOP-21 ×8	602	aaIPI ≥ 1	60–80	3-year EFS: 56% (A) vs. 60% (B)#	Delarue *et al*. 2013

aaIPI, age-adjusted IPI; GELA, Groupe d'Etudes des Lymphomes de l'Adulte; NCRI, National Cancer Research Institute.

improved survival of patients aged <60 years with low-intermediate risk compared to standard R-CHOP (3-year PFS 87% versus 73%, HR 0.48, p = 0.0015; 3-year OS 92% versus 84%, p = 0.0071), although associated with a higher risk of haematological toxicity, thereby limiting its use in elderly DLBCL patients.

To intensify treatment, the German High-Grade NHL Study Group (Deutsche Studiengruppe Hochmaligne Lymphome; DSHNHL) shortened intervals between CHOP from 3 weeks (CHOP-21) to 2 weeks (CHOP-14). A follow-up study, RICOVER-60, performed by the DSHNHL for patients over 60 years with DLBCL, demonstrated that the addition of rituximab to six cycles of CHOP-14 significantly improved event-free survival (EFS), PFS and OS without relevant additional toxicity. No further benefit was observed after eight cycles of chemoimmunotherapy. However, the outcome of this dose-intensification strategy was not reproduced by the Japanese Clinical Oncology study group (CHOP-14 versus CHOP-21), further confirmed by both the NCRI-led study (R-CHOP-21 versus R-CHOP-14 in all age groups) and the GELA-led study (R-CHOP-21 vs R-CHOP-14 in elderly patients) with no improvement in efficacy with a shorter interval between chemoimmunotherapy.

For patients with a high-risk IPI score of 3–5, for which the PFS of <50% at 3 years following R-CHOP is unsatisfactory, there remains no consensus on the appropriate treatment strategy. A number of approaches have been or are currently being examined in clinical trials to improve the outcome in this subset of patients, including increasing dose intensity, number of drugs and/or including high-dose therapy (HDT) and autologous stem cell transplant (ASCT). In the UK, the NCRI Phase II study has explored the feasibility of using the intensified R-CODOX-M - R-IVAC regimen (see Table 34.5) in high-risk DLBCL (IPI 3–5).

Rituximab maintenance following standard chemoimmunotherapy has demonstrated improved outcomes in indolent NHLs such as follicular lymphoma and mantle-cell lymphoma; however, the Eastern Cooperative Oncology Group (ECOG) 4494/Cancer and Leukemia Group B (CALGB) 9793 study showed no benefit for rituximab maintenance following R-CHOP chemotherapy and it is therefore not routinely recommended.

Special considerations: DLBCL in the elderly

The incidence of DLBCL rises with age, with 50% of DLBCL diagnoses occurring in patients older than 60 years, and this number is expected to rise with the increasing age of the population. Patients older than 60 years also represent a heterogeneous population with varying levels of general health, coexisting comorbidities and an increased susceptibility to treatment toxicities, a higher incidence of ABC-DLBCL subtype and therefore the outcome in this population is worse than that observed in younger adults. Attempts to improve patients' outcome using intensified chemotherapy regimens (compared with R-CHOP) commonly leads to a trade-off with increased toxicity without necessarily an improvement in survival, whereas less toxic regimens are often less effective. Results of the GELA study (LNH98-5) demonstrated that the addition of rituximab to CHOP-21 improves PFS and OS. Dose-dense regimens with R-CHOP-14 have failed to demonstrate superiority over R-CHOP-21. A comprehensive geriatric assessment, and routine G-CSF prophylaxis and supportive therapies should be considered in this patient population. R-CHOP-21 remains the standard therapy for fit patients without cardiac dysfunction, whilst omission of anthracycline is commonly considered in those that are unfit or have proven cardiac dysfunction. For this latter cohort of older patients with cardiac comorbidities, a recent Phase II study has demonstrated efficacy using a regimen where anthracycline is substituted with gemcitabine, R-GCVP, with an ORR of 61% and 2-year PFS and OS of 49.8% and 55.8%, respectively. For patients older than 80 years without significant cardiac dysfunction, the attenuated regimen R-miniCHOP was shown to be a reasonable option with good results. For patients with bulky disease, additive radiotherapy following chemotherapy has been shown to improve PFS and OS compared to those not receiving this treatment modality, although it is unclear if this approach is recommended for those patients already achieving complete metabolic responses.

Up-front autologous stem cell transplantation

The merit of up-front (performed in CR1) ASCT is unproven. A meta-analysis of 15 randomized control trials including 3079 patients was conducted in the prerituximab era. Overall treatment-related mortality was 6.0% in the ASCT group and not significantly different compared to conventional chemotherapy (OR 1.33, 95% CI 0.91 to 1.93, P = 0.14). Thirteen studies, including 2018 patients, showed significantly higher CR rates in the group receiving ASCT (OR 1.32, 95% CI 1.09–1.59, P = 0.004). Despite higher CR rates after ASCT, there was no OS or EFS benefit when compared to conventional chemotherapy. The pooled HRs were 1.04 (95% CI 0.91–1.18, P = 0.58) and 0.93 (95% CI 0.81–1.07, P = 0.31), respectively. In the rituximab era, a number of prospective randomized trials have been conducted to assess the merit of HDT and ASCT in high-risk patients, summarized in Table 34.7. The conclusion from these studies was that the addition of HDT and ASCT failed to show benefit on survival and is therefore not recommended outside of a clinical trial. Whether the response of histological subtypes (ABC versus GCB) is different is as yet unclear and may warrant future evaluation.

Prophylactic therapy to the central nervous system

Patients with DLBCL who present with a high IPI (particularly those with high LDH and/or extranodal disease) are at higher risk for CNS relapse and should therefore be considered for CNS prophylactic therapy (see CNS prophylaxis section below).

Relapsed or refractory DLBCL

Up to 40% of patients fail first-line therapy and fall into one of three categories: those relapsing after achieving CR, those achieving PR with residual disease and refractory patients. Once a patient has failed initial chemotherapy, the prognosis is poor and long-term overall survival is less than 10%. The majority of the failures occur within the first 18 months, with relapses occurring more than 5 years after diagnosis being rare.

Salvage HDT followed by ASCT is recommended for patients that are fit enough, as established from the results of the landmark prospective PARMA trial, which demonstrated the superiority of intensified treatment in patients with sensitive relapse. The efficacy of HDT followed by ASCT is only well-proven in those who have shown chemosensitivity to conventional dose second-line (salvage) chemotherapy prior to ASCT (PR or chemoresponsive relapse). A retrospective UK study of 57 lymphoma patients (26 with histologically aggressive NHL) identified that of 16 patients with progressive disease after first-line salvage, only one patient (4%), subsequently responded to second-line salvage regimens. The PFS and OS at 3 years in this subgroup was only 4%. Since patients who progress on first-line salvage therapy have a very low chance of responding to a second-line salvage regimen, they should be considered instead for experimental therapies.

The salvage chemotherapy of choice, however, remains a subject of debate. The majority of favoured first-line salvage regimens include rituximab with either one or both of a platinum compound or ifosfamide. A recent study showed there was no difference between salvage regimens R-ICE (rituximab, Ifosfamide, carboplatin, and etoposide) and R-DHAP (rituximab, dexamethasone, high-dose cytarabine and cisplatin) in terms of 3-year EFS and 3-year OS. However, a subgroup analysis demonstrated GCB-subtype DLBCLs to have a better PFS when treated with R-DHAP, suggesting that COO subtype remains an independent risk factor in relapsed/refractory DLBCL. Rituximab is often used as a component of salvage regimens, having previously been shown to significantly improve the outcome of patients who were not previously treated with rituximab. However, the impact of rituximab as part of the salvage therapy in patients previously treated with rituximab appears to be inferior,

Table 34.7 Results from randomized Phase III trials conducted for upfront autologous stem cell transplantation in aggressive NHLs.

Study	Regimen	No. of patients	Criteria	Age (years)	Outcome	Reference
SWOG-9704 (US and Canadian Intergroup)	**A** CHOP or R-CHOP ×6 followed by ASCT vs. **B** CHOP or R-CHOP ×8	370 (248DL BCL)	aaIPI: 2 (high-intermediate) or 3 (high risk)	15–65	2yr PFS: 69% (A) vs. 55% (B)* 2yr OS: 74% (A) vs. 71% (B)#	Stiff *et al*. 2013
DSHNHL 2002-1 (German High-Grade Non-Hodgkin Lymphoma Study Group)	**A** R-MegaCHOEP vs. **B** R-CHOEP-14	275	aaIPI: 2 or 3	18–60	3yr EFS: 61.4% (A) vs. 69.5% (B)# 100% grade 4 leucopenia and thrombocytopenia in R-MegaCHOEP arm	Schmitz *et al*. 2012
DLCL04 (Fondazione Italiana Linfomi (FIL)	**A** R-dose-dense chemotherapy followed by R-HDC and ASCT vs. **B** R-dose-dense chemotherapy	399	aaIPI: 2 or 3	18–65	3yr PFS: 70% (A) vs. 59% (B)* 3yr OS: 81% (A) vs. 78% (B)#	Vitolo *et al*. 2012 abstract #688 ASH
GOELAMS 075	**A** R-HDT (high-dose CHOP ×2, followed by high-dose methotrexate and cytarabine then ASCT vs. **B** R-CHOP-14	286	All IPI, aaIPI: 2 or 3 (58%)	18–60	No significant difference in OS and EFS between the two arms	Milpied *et al*. 2010 abstract #685 ASH

*Statistically significant difference ($P < 0.05$); #Not statistically significant.
ASH, American Society of Hematology; EFS, event-free survival; GOELAMS, Groupe Ouest-Est des Leucémies et Autres Maladies du Sang; HDC, high-dose cytarabine, mitoxantrone, dexamethasone; OS, overall survival; PFS, progression-free survival; SWOG, Southwest Oncology Group.

potentially as a result of rituximab resistance and results from prospective studies addressing this issue are awaited.

The question of the benefit of maintenance or consolidation with rituximab following ASCT has been evaluated as part of the CORAL study; no improvement in EFS was demonstrated between post-transplant rituximab or observation arms.

The role of PET scanning following second-line chemotherapy prior to ASCT has been evaluated in a number of studies. A positive FDG-PET scan pre-HSCT has been correlated with inferior PFS and OS and an increased risk for relapse after HSCT.

For relapsed or refractory patients who are ineligible or unfit for HDT and ASCT, treatment options include clinical trial participation using novel biological agents or experimental therapies, or consideration of palliation with radiotherapy as these patients have little chance of prolonged disease control with conventional therapies.

Allogeneic stem cell transplantation in high-grade non-Hodgkin lymphoma

Some patients are likely to have a very poor prognosis with conventional therapy, including those who fail ASCT, those presenting with a very high IPI and those with special subtypes of aggressive NHL (e.g. γδ T-cell NHL). Such patients are potential candidates for allogeneic SCT, as several reports have suggested allogeneic SCT to be potentially curative in high-risk patients with chemosensitive disease. However, there is no certainty regarding the existence of a strong graft-versus-lymphoma (GvL) effect in aggressive NHLs, whereas mortality following transplantation remains high, indicating that this treatment should be reserved for selected patients only.

The choice between myeloablative and non-myeloablative (reduced-intensity) conditioning regimens is challenging, although myeloablative approaches are only recommended in a subset of younger and fitter patients given the high transplant-associated mortality. Recent studies have reported the feasibility of adopting a reduced-intensity conditioning approach in relapsed DLBCL patients. A UK collaborative study reported OS and PFS of 47% and 48%, respectively, in a cohort of relapsed DLBCL (including transformed follicular lymphoma) using reduced intensity conditioning allografts. The French Society of Marrow Transplantation and Cellular Therapy reported an estimated 2-year OS and PFS of 49% and 44%, respectively, with a 1-year cumulative incidence of non-relapse mortality (NRM) of 23%. More recently, the European Bone Marrow Transplantation (EBMT) registry data supported these findings with a 3-year OS and PFS after transplantation of 52% and

42%, respectively, in DLBCL patients who relapsed after an ASCT. In this study, although a lower NRM was observed in the group of patients treated with a reduced-intensity approach (3 year NRM: 20% versus 40%), this did not translate to a significant difference in PFS or OS compared to myeloablative conditioning approaches, because of the higher relapse rate in this cohort. In all of these studies, the quality of response prior to transplantation was a predictor of outcome and therefore patients who are chemotherapy refractory at the time of transplant are unlikely to benefit from this strategy.

Role of interim PET/CT in response assessment

Unlike in Hodgkin lymphoma, where the prognostic utility of interim PET-CT is established, its role in DLBCL remains debatable. Initial studies using PET-CT for response assessment after two, three or four cycles of immunochemotherapy suggested an association between the result of the interim scan and patient outcome. However, some studies have demonstrated interim PET-CT assessment in DLBCL to have low prognostic and positive predictive value with a high number of false positive interim scans confirmed by a biopsy of the FDG-positive site. What is evident on comparing these studies is a current lack of standardization for the appropriate time point for interim response assessment (2, 3 or 4) and the method of interpretation and response criteria used, as different grading systems from visual assessment scales (Deauville five-point scale; see Chapter 32) to quantitative measures (such as standardized uptake value; SUV) are employed and subject to wide inter- and intraobserver variability. Greater harmonization of interim PET-CT interpretation is required before this response modality can be transitioned to the standard of care, and there are a number of ongoing large prospective trials that aim to address these concerns.

Emerging therapeutic approaches

Despite our standard therapies in DLBCL offering the majority of patients a favourable outcome, there remain several areas of unmet clinical need, especially patients with prognostically poorer outcomes, including elderly patients, patients with relapsed or refractory disease, patients with inferior molecular risk factors, such as ABC-DLBCL and MYC-rearranged DLBCLs. There is now a plethora of novel targeted therapies based on our improved understanding of the biological mechanisms that determine DLBCL diversity and prognosis. A number of agents are particularly focused on targeting biological pathways of relevance in ABC-DLBCL to improve outcomes. Recent proof of principle studies using bortezomib, an inhibitor of NF-κB signalling, or ibrutinib, an inhibitor of BCR signalling, have shown preferentially higher response rates in ABC-DLBCL compared to GCB-DLBCL, supporting the benefit of employing therapies targeted to the COO. Here, the addition of bortezomib, a proteasome inhibitor that inhibits NF-κB signalling, demonstrated a higher response in ABC-DLBCL compared to GCB-DLBCL. As a consequence, two separate Phase III clinical trials are currently being conducted to examine the benefit of the addition of drugs directed at NF-κB signalling to improving response and outcomes in patients in accordance with their molecular COO subtype. In the UK, the REMoDL-B trial is currently investigating if first-line treatment with R-CHOP plus bortezomib will benefit ABC-DLBCL or GCB-DLBCL compared to standard R-CHOP. Similarly, an international Phase III study is evaluating the benefit of the addition of ibrutinib to R-CHOP in newly diagnosed non-germinal centre DLBCL compared to R-CHOP alone.

Primary mediastinal (thymic) large B-cell lymphoma

Primary mediastinal large B cell lymphoma (PMBL) is an aggressive lymphoma that accounts for 2–4% of all NHLs. PMBL is now recognized as a separate entity in the WHO classification of lymphoid neoplasms, with distinct clinicopathological features to DLBCL (Table 34.6). It characteristically presents with a large anterior mediastinal mass, frequently accompanied by clinical complications, such as pericardial or pleural effusions and superior vena cava obstruction, as a result of local intrathoracic tumour invasion. Extrathoracic involvement and bone marrow infiltration is rare, although at relapse, there is a significantly high incidence of extranodal site involvement. The median age at presentation is in the third to fourth decade, with a female predominance. Interestingly, PMBL shares not only clinical and pathological features with classic Hodgkin lymphoma (cHL), but also molecular similarities with the GEP, demonstrating greater resemblance to cHL than to diffuse large B-cell lymphoma (DLBCL).

Genetic and molecular features

The cell of origin of PMBL is unclear, although the pattern of somatic hypermutation of typical gene targets, such as BCL6 and the immunoglobulin heavy-chain gene (IgVH), and GEP supports derivation from thymic B-cells with a germinal-centre or post-germinal-centre origin. Earlier and more recent genetic profiling studies have identified key molecular pathways and biological mechanisms involved in PMBL, including deregulated NF-κB, Janus kinase signal transducers and activators of transcription (JAK-STAT) signalling and immune surveillance. A number of these genetic and molecular aberrations are shared by cHL.

Chromosomal gains and amplification of chromosome 2p16.1, containing the gene locus for *REL*, encoding a member of the NF-κB complex, occurs in over 50% of patients, with increased NF-κB activity seen in these cases. Additionally, truncating mutations and deletions of *TNFAIP3*, encoding A20 protein, a negative regulator of NF-κB signalling, occurs in 36% of PMBL, promoting constitutive NF-κB activation. Moreover, directly inhibiting the NF-κB complex was toxic in a PMBL

cell line, demonstrating the critical dependence on NF-κB signalling.

The *JAK2* gene loci within chromosome 9p24.1 is amplified in up to 75% of PMBL cases and is a major molecular hallmark of the disease. More recently, other components of the JAK-STAT signalling pathway have been found to be altered in PMBL, including recurrent mutations in *PTPN1* (22% of cases), encoding a protein tyrosine phosphatase, the tumour-suppressor gene, *SOCS1* (45% of cases) and *STAT6 (*36% of cases), with functional studies in cell lines demonstrating that these alterations promoted constitutive pathway activation.

MHC Class II antigens such as HLA-DR are downregulated on the surface of PMBL cells and this correlates with a diminished number of infiltrating cytotoxic T cells and has been linked to inferior outcome. Translocations in the transcriptional regulator of MHC II class expression, *CIITA*, occur in 38% of PMBLs and have been shown to suppress HLA class II expression *in vitro*. Highly recurrent genomic rearrangements involving T-cell anergy-inducing programmed death ligands (PDL1 and PDL2) lend further support to the role of escape from immune surveillance as a feature of PMBL.

Treatment and prognosis

Initial therapy is critical in treating PMBL; however there has been limited prospective trial activity in PMBL, making it difficult to establish a standard of care. This is further compounded by the dilemma of how to deliver a therapy with the highest cure rate, yet minimize long-term morbidity and sequelae in this young population. Currently, therapies for PMBL consist of a multiagent anthracycline- and rituximab-containing backbone. Retrospective datasets suggest dose-intensive third-generation chemotherapy regimens like MACOP-B (methotrexate, doxorubicin, cyclophosphamide, vincristine, prednisolone and bleomycin) are superior to the more commonly used conventional CHOP regimen. A large multicentre retrospective analysis from the IELSG demonstrated that complete remission rates were comparable in patients treated with CHOP or CHOP-B and dose-intensive regimens like MACOP-B, VACOP-B (modification of MACOP-B, with methotrexate substituted with etoposide) or ProMACE CytaBOM, and improved with the addition of involved field radiotherapy (IFRT) in patients in partial remission, although the relapse rate at 3 years was significantly lower in the dose-intensive chemotherapy group compared to those treated with CHOP (12% versus 23%; $P = 0.02$) and the projected 10-year OS and PFS were superior, at 71% and 67%, compared with 44% and 33% ($P = 0.0001$ and $P = 0.0003$, respectively). Subsequent studies have reproduced these observations with OS and EFS superior in MACOP-B or VACOP-B groups compared to CHOP. The integration of rituximab with CHOP or third-generation regimens has resulted in improved outcomes, with cure rates generally in the range of 80% compared with approximately 50–60% in the prerituximab era. Current treatment approaches are focused on limiting or removing the need for consolidative mediastinal radiotherapy to reduce the long-term risks of breast cancer, cardiac toxicities, hypothyroidism, pulmonary fibrosis and cancer. The recent NCI study using dose-adjusted EPOCH-R (DA-EPOCH-R; see Table 34.5) reported 5-year EFS and OS rates of 93% and 97%, respectively, in a cohort of 51 patients, demonstrating that regimens can obviate the need for radiotherapy.

There are conflicting data regarding the need for consolidation radiotherapy in patients achieving a first CR, and a major challenge is the management of patients with residual masses post chemotherapy. In patients with confirmed PR, there is clear evidence from the IELSG series that many patients completing chemotherapy in PR may be converted to CR following radiotherapy. The prognostic utility of PET-CT scanning following chemo-immunotherapy for PMBL has been examined in a recent prospective study performed by the IELSG (IELSG-26). In 115 evaluable patients, 47% achieved a complete metabolic response (CMR) following treatment using rituximab combined with either MACOP-B, VACOP-B or CHOP, the majority of which went on to receive consolidative mediastinal radiotherapy, irrespective of end-of-treatment PET-CT results. Excellent clinical outcomes were attained with a 5-year PFS and OS of 92% and 86% for the entire cohort. Importantly, although a post-treatment PET-CT scan demonstrated a good negative predictive value, the positive predictive value was poor, ranging from 18% to 32%, depending on the mode of assessment employed (mediastinal blood pool versus liver pool cut-off). Based on these data, the role of post-treatment PET scans in informing our decisions for receiving or omitting consolidative radiotherapy, remains unclear. The IELSG37 study is one trial that is addressing this issue.

Relapse of PMBL is infrequent compared to DLBCL and the majority of treatment failures occurs within the first 6–12 months after therapy completion, and are rare beyond 2 years. A repeat biopsy is advocated to confirm histology at the time of recurrence. The approach for PMBL relapse, similar to DLBCL, is salvage therapy followed by consolidation with ASCT in those deemed fit enough. Chemosensitivity to the salvage regimen is predictive of a more favourable outcome with ASCT.

Intravascular large B-cell lymphoma

Intravascular large B-cell lymphoma is an extremely rare form of NHL and is characterized by an accumulation of large neoplastic B cells within the lumens of blood vessels. These lymphomas preferentially affect the elderly and involve the CNS, reported in approximately 30% of patients, and lungs and skin. Prognosis is generally poor. Patients presenting with only skin lesions appear to have a significantly better survival than patients with other clinical presentations (3-year overall survival: 56% versus

22%). Multiagent chemotherapy using agents with higher CNS bioavailability is the preferred mode of treatment. The role of ASCT in younger patients with unfavourable features is being explored.

Burkitt lymphoma

Burkitt lymphoma (BL) was first described in Ugandan children by Denis Burkitt in 1958. BL is a highly aggressive mature B-cell lymphoma that constitutes less than 2% of all NHL diagnoses in the UK occurring in children and adults. It represents the most frequent B-cell lymphoma in children, accounting for 30–50% of all paediatric lymphomas. There are three clinical variants recognized: endemic, sporadic and immunodeficiency-associated BL. The latter is discussed in detail in a later section of this chapter.

The endemic BL variant occurs predominantly in equatorial Africa and Papua New Guinea, mainly affecting children with an estimated incidence of 3–6 cases per 100,000 children per year, and a peak incidence of 4 to 7 years. The risk factors for endemic BL are incompletely understood. However, EBV is widely accepted to be a contributing factor, with the EBV genome detected in about 95% of endemic BL cases. Infection with *Plasmodium falciparum* malaria has also been linked as a risk factor for endemic BL. The ubiquity of EBV and malaria in these regions suggest that these factors alone cannot explain all endemic BL cases. Sporadic BL is distributed worldwide, occurring mainly in children and young adults. EBV can be detected in about 30% of sporadic BL cases.

Presentation varies according to the clinical subtype, but extranodal site involvement, especially CNS involvement, is a common feature. In endemic BL, the jaw and facial bones are characteristically affected. For sporadic BL, presentation with intra-abdominal masses is more common, with jaw tumours rarely seen. Lymph node presentation is seen more commonly in adults than children.

Genetic and molecular features

The molecular hallmark of BL is the translocation t(8;14)(q24;q32) or its variants t(2;8) and t(8;22), occurring in nearly all affected individuals**,** irrespective of the clinical variant. These chromosomal translocations juxtapose the *MYC* oncogene at 8q24 with one of the three immunoglobulin loci partners: Ig heavy chain (IGH) at 14q32, κ light chain (IGK) at 2p12 or λ light chain (IGL) at 22q11, leading to deregulation of MYC expression. While *MYC* rearrangement is a diagnostic feature of BL, it is not specific and occurs less commonly in DLBCL and other more aggressive high-grade lymphomas. BL is derived from a germinal-centre cell of origin yet has a unique gene expression profile that distinguishes it from DLBCL. Recent characteri**z**ation has shown that BL most closely resembles cells from germinal-centre dark zones, whereas the GCB subtype of DLBCL is more akin to cells from germinal-centre light zones.

Current models for BL assume that IG-MYC translocation is an early genetic event and additional cooperating oncogenic events are required to drive lymphomagenesis. *MYC* and *TP53* mutations occur in approximately 60% and 40% of BL cases, respectively. Recent genomic sequencing studies have revealed novel activating mutations in the transcription factor, *TCF3* and inactivating mutations in its negative regulator, *ID3*, affecting up to 70% of cases. Additionally, activating mutations in *CCND3* (encoding cyclin D3) have been identified in 38% of sporadic BL. Together, these recurrent oncogenic mutations cooperate with MYC to promote proliferation, survival and growth of BL cells.

Treatment and prognosis

Therapies in BL are typically more intensive than DLBCL. Due to the high proliferation rates of the tumour, a dose-intense, short-duration treatment strategy is employed to avoid rapid growth and development of resistance between chemotherapy cycles. Effective treatment of patients with BL is therefore based on a multidrug combination chemotherapy of antimetabolites and alkylating agents (e.g CODOX–M/IVAC regimen, hyper CVAD; Table 34.5) and includes drugs which highly penetrate the blood–brain barrier, such as methotrexate and cytarabine, this being a necessity. Given the high expression of CD20 on BL cells, rituximab is commonly added to combination regimens, showing improvement in outcome parameters in some studies, although this requires validation in a prospective randomized study. With these rituximab-containing regimens, CR and 3-year PFS rates are in the region of 80–94% and 70–90%, respectively. Addition of granulocyte colony-stimulating factors (G-CSF) is required in these regimens to maintain dose intensity and reduce the duration of severe neutropenia. A number of less dose-intense regimens such as DA-EPOCH-R, have been reported, with good clinical outcomes. Notably, high-dose cytarabine and methotrexate do not feature in this regimen. However, this regimen, like the others described above, does include intrathecal therapy.

The risk of tumour lysis syndrome (TLS), especially after initiating chemotherapy, is relatively high, reflecting the high proliferation rate of tumour cells. Vigorous hydration with alkalinization, accompanied by administration of recombinant urate oxidase (rasburicase) (or a uric acid synthesis inhibitor such as allopurinol) are very important in preventing/reducing the severity of this complication (see also Chapter 24). Hypercalcaemia is also frequent at presentation and needs appropriate therapy with hydration and biphosphonates.

Overall, approximately 50–80% of adult patients can be cured with these intensive chemotherapy regimens, and in paediatric populations the cure rate is even higher. Relapses typically occur within the first 12 months following treatment. Although BL is a highly FDG-avid tumour, there is currently no consensus on

its usefulness for response assessment, and this requires further exploration. There remains a lack of improvement in the outcome of patients older than 60 years, who make up a quarter of BL diagnoses and experience a higher rate of adverse events, and for whom the standard dose-intense regimens may be unsuitable and alternative approaches are warranted. Newer molecularly targeted therapies may be beneficial in this cohort of patients.

B cell lymphoma, unclassifiable, with features intermediate between diffuse large B-cell lymphoma and Burkitt lymphoma

B-cell lymphoma unclassifiable (BCLU) with features intermediate between DLBCL and BL is a provisional category in the WHO 2008 lymphoma classification to encompass a 'grey zone' of tumours displaying transitional gene expression profiling and morphological features between DLBCL and BL (Table 34.8). From a clinical perspective, a clear distinction between BL and DLBCL is of major importance for the outcome of patients, as the different lymphoma types require different therapies. In contrast to BL, BCLU tumours normally harbour complex karyotypes. In addition to *MYC* translocations, translocations in BCL2, BCL6 or both occur in a significant proportion of patients, referred to as double-hit (DH) or triple hit (TH) lymphomas (further indicated as BCL2+/MYC+, BCL6+/MYC+ or BCL2+/BCL6+/MYC+). BCLU occurs in older patients, most of whom present with adverse clinical features: advanced stage, elevated LDH, high IPI, bone marrow and CNS involvement. The disease is aggressive and outcome is poor with conventional chemotherapy.

Transformed lymphomas (see also Chapter 33)

Progression to an aggressive lymphoma, usually resembling DLBCL histologically, is part of the natural history of indolent lymphomas. The risk of high-grade transformation is highest

Table 34.8 Differential clinical and molecular features between DLBCL, PMBL, BL and BCLU.

	DLBCL	PMBL	BL	BCLU
Clinical features				
Age at presentation	All age groups, median 70 years	30–40 years	Children and young adults	Mainly older adults
Gender	Male > female (1.2:1)	Female predominance	No real predominance	Unknown
Site of disease	Nodal and extranodal	Mediastinal mass and extension to supraclavicular nodes	Often extranodal (jaw and GI tract)	Widespread disease Often extranodal
Bone marrow involvement	<20%	Rare	Common	Common
CNS involvement	Rare	Rare	Common	Common
Pathologic features				
Cell of origin	Germinal-centre (GCB) or activated B-cell (ABC) derived	Thymic	Germinal-centre derived	Usually germinal-centre derived
Genetic and molecular features	**GCB DLBCL**: *BCL2* translocation/mutations; mutations in epigenetic regulators (*CREBBP*, *EP300*, *EZH2*, *MEF2B* and *MLL2*); *PTEN* deletions **ABC DLBCL**: BCL2 amplification; constitutive NF-κB activation by mutations in components (*TNFAIP3, MYD88, CARD11, CD79A/B*)	*REL* and *JAK2* gains *CIITA* translocation Alterations in components of NF-κB and JAK-STAT signalling	*MYC* rearrangement in nearly all patients *MYC*, *TP53*, *ID3*, *TCF3* and *CCND3* mutations	Complex karyotypes. Often combination of *MYC* translocations with *BCL2* translocations +/– *BCL6* translocations

in follicular lymphoma, but has also been described in other B-cell lymphoproliferative disorders, including CLL/SLL, Hodgkin lymphoma, marginal-zone lymphoma and lymphoplasmacytic lymphoma. The true incidence of transformed lymphomas is unknown as there remains a lack of consensus about how one defines transformation or indeed the modality of diagnoses required for confirmation (clinical versus histological transformation). Transformed lymphomas tend to be associated with adverse clinical features and poorer clinical outcomes, although the majority of these observations have been extrapolated from data obtained in the setting of follicular lymphoma transformation and have yet to be examined prospectively. There remains no standard therapeutic approach, with the majority of clinicians adopting a DLBCL-like regimen after consideration of the patient's prior therapy. Whether consolidation with autologous SCT is indicated is contentious.

Double-hit or triple-hit lymphomas

Double-hit (DH) and triple-hit (TH) lymphomas are rare subtypes of lymphoma defined by two or three recurrent chromosome translocations, as described above. The term is not limited to DLBCL and BCLU, as this genetic composition has also been observed in follicular lymphoma and lymphoblastic lymphomas. In general, this entity is usually associated with an aggressive clinical course and dismal outcomes with R-CHOP therapy. Intensive approaches employing dose-dense regimens like R-CHOEP-14 and R-CODOX-M/R-IVAC, and consideration of ASCT or allogeneic SCT consolidation in first remission are usually considered.

DLBCL in HIV-positive patients

People living with HIV have an increased risk of developing DLBCL, an AIDS-defining illness, and the second most common tumour in individuals with HIV. The development of DLBCL is related to older age, low CD4 cell count and no prior treatment with combination antiretroviral therapy (cART). Patients tend to present with advanced clinical stage, B symptoms and extranodal involvement, including bone marrow. Factors that are associated with survival in the post-cART era are the IPI score and in some studies, a low CD4 cell count, but median survival is beginning to approach that observed in the HIV-negative population.

The introduction of cART has led to better control of HIV viral replication and improved immune function. Opinions differ as to whether cART should be continued during chemotherapy or not. Treatment centres in the US that use the infusional regimens, have previously suspended cART because of concern regarding potential adverse pharmacokinetic and pharmacodynamic interactions with chemotherapy and the potential for increased toxicity. In these studies, despite a high response rate, CD4 cell counts fell dramatically during chemotherapy and took months to recover to baseline levels, despite the re-introduction of cART on completion of chemotherapy. Although this strategy did not appear to adversely affect lymphoma outcome or increase infectious complications, the treatment groups have not been large. There is concern that the interruption of cART in patients on therapy prior to lymphoma diagnosis might lead to the development of viral resistance. In Europe, a combined approach to care, involving HIV physicians and haemoncologists, is adopted and it is usual to continue cART during chemotherapy, avoiding boosted protease inhibitors wherever possible as they are associated with greater toxicity and drug interactions

There is much debate on whether R-CHOP or infusional regimens, such as R-EPOCH, are most efficacious as first-line treatment in HIV-positive patients. However, there has been no randomized study comparing these regimens. A pooled analysis of US studies suggested that R-EPOCH resulted in superior response rates and survival compared to R-CHOP, but the former regimen was used during a later time period (2002–2006) than the R-CHOP study (1998–2002), suggesting other variables, including supportive care and antiretroviral drug options, may have differed. Consistent with this, the patients treated with R-EPOCH routinely received concurrent antifungal and antibacterial prophylaxis, which was omitted from those treated earlier with R-CHOP. More recently a multicenter retrospective study from the UK reported an excellent outcome after R-CHOP for 97 HIV-positive patients with DLBCL. The reported 5-year DFS and OS were 94% and 78%, respectively, and significantly longer than HIV negative patients (n = 208).

The role of rituximab in HIV-associated DLBCL has been controversial ever since a randomized Phase III study of CHOP (n = 50) versus R-CHOP (n = 99), conducted in the US was published. Although there was a trend to improved response rate with rituximab (58% versus 47%, P = 0.15), a significant reduction in progression of lymphoma on treatment, and in death due to lymphoma, an increased death rate from infectious complications, particularly (9/15) in those with a CD4 count below 50 cells/mm^3, was observed. However, maintenance rituximab was included every 3 months in those who responded to R-CHOP. Although, this remains the only Phase III study addressing the role of rituximab in HIV-positive patients with DLBCL, subsequent Phase II studies of R-CHOP (without maintenance rituximab) from Europe did not show an increased risk of infectious deaths, instead showing that rituximab was beneficial. A recent meta-analysis of prospective studies has confirmed the benefit in response rate and OS of the addition of rituximab to chemotherapy. Thus, the addition of rituximab to chemotherapy is now recommended for all patients with DLBCL, with appropriate antimicrobial prophylaxis (cotrimoxazole, fluconazole, aciclovir), pre-emptive G-CSF and prompt treatment of opportunistic infection.

Treatment of relapsed/refractory AIDS-related lymphoma

Treatment of refractory or relapsed DLBCL in the pre-cART era was disappointing, with few clinically useful responses. Two large, retrospective, multicentre studies performed by the EBMT have confirmed the feasibility and efficacy of HDT and ASCT for HIV-positive patients that respond to second-line chemotherapy. In one of these studies a comparative analysis was performed between 53 HIV-positive lymphoma patients and a matched cohort (66% non-Hodgkin and 34% Hodgkin lymphoma) of 53 HIV negative patients. The incidence of relapse, OS and PFS were similar in both cohorts. Thus, in the cART era, HIV patients with chemosensitive relapsed ARL should be considered for ASCT according to the same criteria adopted for HIV-lymphoma patients.

Burkitt lymphoma in HIV-positive patients

A number of studies have demonstrated the feasibility of administering intensive chemotherapy regimens, such as CODOX-M/IVAC, with or without rituximab, to HIV-positive patients with BL. These studies report a 2-year EFS of approximately 60%, similar to that observed in HIV-negative patients treated with the same regimen, with no increase in toxicity and prompt immune recovery, suggesting that a uniform approach to treatment of BL should be used, regardless of HIV status. The DA-EPOCH regimen with rituximab has also been shown to be effective in this cohort.

Primary central nervous system lymphoma

Primary central nervous system lymphoma (PCNSL) is a rare form of NHL comprising 2.2% of all central nervous system (CNS) tumours. It encompasses lymphoma exclusively involving the brain, spinal cord, eyes, meninges and cranial nerves, with 90–95% classified histologically as DLBCL. The majority of PCNSL are sporadic and the incidence increases with age. A minority are attributable to immunosuppressed states, including HIV infection or iatrogenic immunosuppression following organ transplantation. In the era of effective cART the frequency of HIV-associated PCNSL has diminished. The involvement of critical sites within the CNS presents both diagnostic and therapeutic challenges, with outcomes consistently inferior to systemic DLBCL. Neurocognitive dysfunction and impaired performance status are frequent at clinical presentation, whilst histological confirmation is inherently risky and often yields small tissue biopsies. Moreover, choice of cytotoxic therapy is limited by the inability of many drugs employed for systemic NHL treatment to penetrate the blood–brain barrier (BBB) efficiently. Since the initial description of PCNSL in 1975, treatment algorithms have evolved from whole-brain radiotherapy (WBRT) as a single-modality treatment towards a multiagent, high-dose, methotrexate (MTX)-based chemotherapy approach, where WBRT is reserved for consolidation or for relapsed disease. Given the rarity of PCNSL, together with challenges conducting clinical trials in this patient group, data from randomized studies are scarce and the level of evidence to guide therapeutic decisions is often low.

Diagnosis

The diagnosis of CNS lymphoma can be a particular challenge because of lesional response to corticosteroids and MRI features that are shared with other pathologies. The majority of PCNSL are diagnosed via stereotactic biopsy or, less commonly, by flow cytometric analysis of cerebrospinal fluid (CSF) lymphocytes. Whole body ^{18}F-FDG PET-CT has an increased sensitivity for detection of systemic DLBCL over conventional CT staging and has an important role in the exclusion of systemic lymphoma at presentation. PCNSL lesions characteristically exhibit homogeneous, high-avidity ^{18}F-FDG uptake, and one small study has suggested that this may assist in differentiating PCNSL from other intracranial malignancies where MRI findings are equivocal. Preimaging corticosteroid therapy is a potential confounding factor, however, and the additional diagnostic value of ^{18}F-FDG PET over modern MRI brain imaging remains poorly defined.

Genetic and molecular features

Improved characterization of PCNSL genotype and phenotype, albeit from small studies with restricted availability of diagnostic material, has the potential to provide prognostic information and identify key molecular pathways that may serve as potential targets for novel therapeutics. An ABC-like phenotype is typical, but evidence of ongoing somatic hypermutation suggests ongoing germinal-centre exposure. Therefore PCNSL does not neatly conform to either of the principal molecular profiles identified in systemic DLBCL. The most frequent genomic aberration identified in PCNSL tissue is deletion of 6p21, involving the HLA locus (56–79%), a lesion found commonly in DLBCL arising in immune-privileged sites and represents a potential mechanism for immune escape. Activating *CARD11* and *MYD88* mutations (especially the L265P mutation) are common molecular aberrations and the resultant activation of the NF-κB pathway highlights a key tumour survival pathway and potential therapeutic target.

Treatment

Remission induction

Chemotherapy regimens incorporating HD-MTX are considered the standard of care as induction therapy for newly diagnosed PCNSL, achieving high rates of initial response when

combined with other agents. There is general consensus that MTX should be administered as a rapid infusion (2–4 hours) at a dose of at least 3 mg/m^2 to maximize therapeutic CSF concentrations, at an interval of 14–21 days. To improve on outcomes with single-agent HD-MTX, the IELSG20 trial assessed the role of combined antimetabolite therapy, with HD-MTX and cytarabine. This study demonstrated a superior CR rate of 46% compared to HD-MTX alone (18%, P = 0.006), with significant improvements in PFS, but not OS (3-year OS 46% versus 32%, P = 0.07). Several studies have evaluated the additional value of BBB-penetrating alkylating agents, such as temozolomide, procarbazine and thiotepa, providing non-cross-resistant agents that, unlike antimetabolite chemotherapy, are also cytotoxic to cells in G0 of the cell cycle. Recent data from IELSG32, an international randomized phase-II trial, has demonstrated improved outcome with the addition of rituximab and thiotepa to a MTX/Ara-C-containing chemotherapy (MATRIX) regimen. Although associated with higher hematological toxicity, MATRIX was not associated with higher rates of severe complications.

Consolidation treatment: WBRT

Following introduction of HD-MTX-based chemotherapy, WBRT (23–45Gy) has continued to be employed to consolidate responses and provide more durable disease control. Delayed neurotoxicity, however, is a major limitation, clinically evident in approximately one-third of patients, particularly with increasing age and increased doses of WBRT and this is associated with significant morbidity and mortality. The only Phase III trial thus far to complete accrual in PCNSL aimed to demonstrate that omission of consolidation WBRT after MTX-based chemotherapy resulted in non-inferior OS rates. Although outcomes suggested that WBRT can be safely omitted in selected patients achieving CR with induction chemotherapy, this large study failed to meet its primary statistical endpoint and issues regarding the design and amendments to the study and protocol adherence limit interpretation of these data. A recent systematic review, which assessed outcomes of chemotherapy versus combined modality treatment using a decision-analytic model, has suggested improvements in both survival and quality-adjusted life years with consolidation WBRT for those <60 years only. The emerging outcome data for HDT-ASCT consolidation for younger, sufficiently fit patients, however, challenges this approach. Deferring WBRT in patients achieving CR with chemotherapy is a particularly attractive concept for those >60 years.

Relapsed PCNSL

Currently there exists no standard treatment approach for patients with relapsed and refractory PCNSL. Published data are largely restricted to small, uncontrolled, retrospective studies, in which outcomes are influenced by a number of parameters, including prior therapies and response durations. In studies where eligible patients have proceeded to HDT-ASCT, available data support the use of thiotepa-based HDT-ASCT in relapsed PCNSL, resulting in 5-year EFS and OS rates of 37.8% and 51.4%, respectively, in the largest study to date. Studies addressing the efficacy of agents such as temozolomide, pemetrexed, bendamustine, topotecan, lenalidomide, pomalidomide, temsirolimus, ibrutinib and other agents targeting B-cell signalling are ongoing.

Post-transplant lymphoproliferative disorder

Post-transplant lymphoproliferative disorder (PTLD) ranges from polyclonal plasmacytic hyperplasia to monoclonal B-cell proliferation, polymorphic hyperplasia/lymphoma, immunoblastic lymphoma or myeloma. EBV causes the majority of these disorders: impaired T-cell function leads to proliferation of latently EBV-infected polyclonal B cells, whereas subsequent genomic mutations result in monoclonal PTLD. Prevention of EBV infection/re-activation may therefore be essential. Antiviral agents (e.g. aciclovir, ganciclovir), passive immunotherapy and pre-emptive immunotherapy, using cytotoxic T lymphocytes targeted against EBV, are all applied to prevent PTLDs, with variable success.

The optimal treatment of PTLD is still not clearly defined. Most published data are in the form of case series. Therapeutic options in PTLD include immunosuppression reduction, chemotherapy and immunotherapy. Patients may respond to reduction in immunosuppressive therapy, especially those who develop PTLD within 1 year from transplant (ORR 80% versus 10% in patients who developed PTLD beyond 1 year). Case series and Phase II studies of rituximab monotherapy following failure of reduction of immunosuppression have confirmed efficacy in inducing remission in a proportion of patients (44–65%) with PTLD. Toxicity appears to be low, but significant numbers of patients progressed on therapy or relapsed after rituximab, some died of lymphoma and many required intensification to chemotherapy. A risk score for identifying patients with PTLD most likely to respond to rituximab monotherapy has been proposed, using age over 60 years, performance status ECOG 2–4 and raised LDH as poor-risk markers.

The largest prospective Phase II trial, the PTLD-1 trial, has demonstrated the efficacy and safety of sequential therapy (rituximab followed by CHOP chemotherapy) for solid-organ transplant recipients with CD20-positive PTLD unresponsive to reduction of immunosuppression. Patients received 4 weekly courses of 375 mg/m^2 rituximab followed by 4 weeks without treatment and four cycles of CHOP-21 chemotherapy at 3-week intervals starting at day 50. Supportive treatment included mandatory G-CSF support and antibiotic prophylaxis (cotrimoxazole and ciprofloxacin were recommended). Of 70 patients

assigned to sequential treatment, 76% had late, 96% monomorphic and 44% EBV-associated PTLD. Main adverse events were grade 3/4 leucopenia in 68% and grade 3/4 infections in 41% of patients. The overall response rate was 90%, with 67% complete responses; 74% of responders were progression-free at 3 and 5 years and median overall survival was 6.6 years.

Anti-EBV cytotoxic T lymphocytes have been used for chemorefractory disease, but their availability is restricted, and their infusion is associated with regulatory challenges.

CNS prophylaxis in high-grade non-Hodgkin lymphoma

There is a 5% incidence of CNS relapse in most large studies of aggressive NHL. The outcome of those who develop CNS relapse is extremely poor, with a median survival of less than 6 months.

The use of rituximab has recently been shown to be associated with a significantly lower incidence of CNS disease in elderly patients with aggressive CD20-positive lymphoma treated in the RICOVER-60 trial. However, this was not a primary endpoint of the trial and there were a high proportion of CNS prophylaxis treatment violations. Patients treated with R-CHOP-14 had a relative risk for CNS disease of 0.58 (95% CI 0.3–1.0, P = 0.046) compared to patients treated with CHOP-14. CNS prophylaxis with intrathecal methotrexate was offered to patients with involvement of bone marrow, testes, upper neck or head. The estimated two-year incidence of CNS disease was 6.9% (CI 4.5; 9.3) after CHOP-14 and 4.1% (CI 2.3; 5.9) after R-CHOP-14. This association of rituximab use with a lower incidence of CNS disease has not been observed in other studies.

As a number of studies have shown that CNS relapse occurs soon after diagnosis, during treatment or just after treatment, administration of preventative CNS-directed therapy during first-line treatment is now routinely undertaken in patients considered at 'high risk'. Patients considered 'high risk', are those with high-grade NHL and either high IPI score, elevated LDH, involvement of more than one extranodal site and specific anatomical sites of involvement such as breast, epidural space, paranasal sinuses, adrenals and testes. Intrathecal methotrexate (12–15 mg flat dose) is usually the CNS-directed therapy of choice, commenced as early as practically possible, with the patients receiving between three and six doses, and administered alongside standard chemotherapy. Alternative prophylaxis includes intrathecal cytarabine or systemic chemotherapy known to cross the blood–brain barrier (such as methotrexate, cytarabine or ifosfamide). Systemic high-dose methotrexate with folinic acid rescue can also be considered in special circumstances although there are insufficient data regarding efficacy and toxicity to suggest this modality can replace or improve the results using intrathecal methotrexate. Whether both IT and IV methotrexate should be incorporated into regimens for patients with a high risk of CNS disease such as testicular DLBCL is unclear. The use of other systemic chemotherapy has not been shown to be beneficial, but there have been few prospective studies that have addressed this issue directly.

All patients with BL or lymphoblastic lymphoma should receive CNS prophylaxis, as the observed incidence of CNS relapse without is about 20%. Standard treatment incorporates both intrathecal therapy and high-dose systemic chemotherapy with methotrexate and cytarabine, although DA-EPOCH does not include the latter and has been reported to be highly effective in a small group of patients.

Suggested algorithm for therapy of aggressive non-Hodgkin lymphoma (summary)

At diagnosis

- Anthracycline-containing combination chemotherapy, with rituximab (CHOP–rituximab). Alternative dose-intense regimens should be considered for particular lymphoma subtypes (e.g. R-CODOX-M/IVAC for BL).
- Patients with intermediate-high/high IPI (particularly those who present with raised LDH and extranodal involvement) should receive CNS prophylactic therapy.
- Adjuvant radiotherapy may be considered in patients with bulky disease.
- No clear role for rituximab maintenance following first-line chemotherapy.
- No clear role for up-front ASCT (performed in CR1).

At relapse

- Second-line regimens (e.g. ESHAP, DVIP, DHAP, ICE, mini-BEAM +/−R; Table 34.5), followed by ASCT.
- Allogeneic SCT may be considered in patients who fail ASCT/special subtypes of NHL (e.g γδ lymphoma) and demonstrate chemosensitivity.

Rare aggressive T-cell and NK-cell lymphomas (see also Chapter 31)

Adult T-cell leukaemia/lymphoma

Adult T-cell leukaemia/lymphoma (ATLL) is a rare mature T-cell neoplasm caused by HTLV-1 with most affected individuals originating from endemic regions, including Japan, Africa, Caribbean and Latin America. It has four different clinical subtypes with the acute (leukaemic) and lymphomatous variants being more aggressive and common than the chronic and smouldering forms. Leucocytosis, lymphadenopathy, hepatosplenomegaly, skin lesions and hypercalcaemia are

observed in the majority of patients. Patients are immunocompromised and opportunistic infections are common, including *Pneumocystis jirovecii* pneumonia, aspergillosis or candidiasis, *Strongyloidies stercoralis* and cytomegalovirus.

Although the initial role of HTLV-1 in ATLL via integration into host DNA is well elucidated, little is known about the subsequent molecular mechanisms involved in the multistep process towards leukaemogenesis. Treatment of ATLL is dependent on the subtype. Responses to conventional chemotherapy are poor. A worldwide meta-analysis showed a combination of the antiretroviral drug zidovudine (AZT) and interferon-α (IFN-α) had significant activity in patients with ATLL, especially with the leukaemic subtype, confirmed in a UK retrospective analysis. Patients with lymphoma subtype benefit from the addition of chemotherapy regimens such as CHOP to AZT and IFN-α. CNS prophylaxis should be considered and special consideration of anti-infective prophylaxis. For patients with the chronic and smouldering forms, long-term survival has been shown to be dismal when managed with a watchful waiting approach and expert recommendations are for these patients to be treated with AZT and IFN-α. Conjugated and unconjugated monoclonal antibodies (anti-CD25, -CD4, -CD52) have all been tested in small numbers of patients. Other agents being evaluated include arsenic trioxide and bortezomib.

There appears to be minimal long-term benefit in autografting patients with ATLL, with the majority of patients relapsing within 1 year of transplant. Prolonged overall survival has been described after allogeneic SCT with the largest retrospective study from Japan reporting 3-year overall survival at 33% and this should be considered for a selected subset of patients as consolidation therapy. Age, remission status and source of the graft (cord versus HLA-matched versus HLA-mismatched) have been identified as significant predictors of survival and since the median age at presentation with ATLL is approximately 60 years, reduced-intensity conditioning is generally favoured.

The prognosis for acute and lymphoma subtypes remains poor, with a median survival of only 6.2 and 10.2 months, respectively, and approximately 2 years for the chronic and smouldering forms, because of chemoresistance and severe immunosuppression.

Extranodal NK/T-cell lymphoma, nasal type (see also Chapter 31)

This is an aggressive, largely extranodal lymphoma, usually of NK-cell type, but with recognized T-cell phenotypic variants. These are very rare in the Western world, but commoner in Asia and South America. It is usually EBV-virus associated and often presents as localized disease in and around the nasal structures. Local extension and dissemination is frequent, usually to regional nodes and distant extranodal structures such as skin, testis and gut. An association with the haemophagocytic syndrome has been described.

The cell-surface expression of P-glycoprotein mediates multidrug resistance (MDR) and renders CHOP or other anthracycline-based regimens ineffective in this lymphoma. Since the introduction of chemotherapeutic agents not affected by P-glycoprotein, such as methotrexate and L-asparaginase, the clinical outcomes have improved. Sequential involved field radiation (IFRT) followed by etoposide- and cisplatin-containing chemotherapy is suggested for patients with stage I/II disease with a durable efficacy of 70% 5-year OS. For advanced-stage disease, novel therapeutic approaches using L-asparaginase-containing regimens such as SMILE (consisting of steroid, methotrexate, ifosfamide, L-asparaginase and etoposide) showed excellent responses in a Phase II clinical trial, with an overall response rate (ORR) of 79% and 3-year overall survival of 50%. No superiority has been demonstrated using either autologous or allogeneic stem cell transplantation.

Future directions

Continual progress in biological insights into high-grade lymphomas has significant implications for the future. We can expect refinement in the way we classify and subclassify certain high-grade lymphomas based on this stream of information. The face of diagnostics will change dramatically in the years to come, with the integration of standardized routine genetic and molecular analyses of tumours, an approach that is currently being piloted in the UK and internationally. Not only will this approach yield important diagnostic information, but it will ultimately identify patients that can be stratified to several different treatment algorithms and novel targeted therapies, rather than continuing with our single treatment combination 'one drug fits all' approach.

Selected bibliography

Cunningham D, Hawkes EA, Jack A *et al*. (2013) Rituximab plus cyclophosphamide, doxorubicin, vincristine, and prednisolone in patients with newly diagnosed diffuse large B-cell non-Hodgkin lymphoma: a phase 3 comparison of dose intensification with 14-day versus 21-day cycles. *Lancet* **381**: 1817–26.

Gisselbrecht C, Glass B, Mounier N *et al*. (2010) Salvage regimens with autologous transplantation for relapsed large B-cell lymphoma in the rituximab era. *Journal of Clinical Oncology* **28**: 4184–90.

Récher C, Coiffier B, Haioun C *et al*. (2011) Intensified chemotherapy with ACVBP plus rituximab versus standard CHOP plus rituximab for the treatment of diffuse large B-cell lymphoma (LNH03-2B): an open-label randomised phase 3 trial. *Lancet* **378**: 1858–67.

Roschewski M, Staudt LM, Wilson WH (2014) Diffuse large B-cell lymphoma-treatment approaches in the molecular era. *Nature Reviews Clinincal Oncology* **11**: 12–23.

Rosenwald A, Wright G, Chan WC *et al*. (2002) The use of molecular profiling to predict survival after chemotherapy for diffuse large-B-cell lymphoma. *New England Journal of Medicine* **346**: 1937–47.

Wilson WH, Young RM, Schmitz R *et al*. (2015) Targeting B cell receptor signaling with ibrutinib in diffuse large B cell lymphoma. *Nature Medicine* [Epub ahead of print]. Accessed 31 July 2015. Available at: http://www.ncbi.nlm.nih.gov/pubmed/26193343.

Shaffer AL, Young RM, Staudt LM (2012) Pathogenesis of human B cell lymphomas. *Annual Review of Immunology* **30**: 565–610.

Swerdlow SH, Campo E, Harris NL *et al*. (eds.) (2008) *WHO Classification of Tumours of Haematopoietic and Lymphoid Tissues*. Lyon, France: IARC Press.

Trappe R, Oertel S, Leblond V *et al*. (2012) Sequential treatment with rituximab followed by CHOP chemotherapy in adult B-cell post-transplant lymphoproliferative disorder (PTLD): the prospective international multicentre phase 2 PTLD-1 trial. *Lancet Oncology* **13**: 196–206.

Zhou Z, Sehn LH, Rademaker AW *et al*. (2014) An enhanced International Prognostic Index (NCCN-IPI) for patients with diffuse large B-cell lymphoma treated in the rituximab era. *Blood* **123**: 837–42.

CHAPTER 35

Stem cell transplantation

35

Charles Craddock[1] and Ronjon Chakraverty[2]

[1]Centre for Clinical Haematology, Queen Elizabeth Hospital, Birmingham, UK
[2]University College London, London, UK

Immunological basis of stem cell transplantation

The major complications of allogeneic SCT are caused by the immunological responses triggered by the infusion of donor haemopoietic progenitors and lymphocytes into an immunosuppressed host. These can take the form of either a host-versus-graft (HVG) or a donor-derived graft-versus-host (GVH) response. Clinically, the HVG response can result in graft rejection, while a GVH response may manifest itself as either graft-versus-host disease (GVHD) or a graft-versus-leukaemia (GVL) reaction. It is now possible to blunt the HVG reaction by optimizing the immunosuppressive properties of the conditioning regimen and consequently graft rejection is rare in most clinical settings. In contrast, GVHD and disease relapse remain the major complications of allogeneic transplantation and novel approaches that permit the induction of GVL without inducing host injury are required. Conversely, autologous SCT is a relatively unremarkable immunological event in which these allogeneic responses are absent.

The human leucocyte antigen system

The human leucocyte antigen (HLA) genes are located within the major histocompatibility complex (MHC) on the short arm of chromosome 6 (6p21.3) (Figure 35.1). HLA genes found in the class I region of the MHC differ in structure from those found in the class II region. HLA class I genes encode polypeptides expressed as cell-surface transmembrane glycoproteins in complex with a soluble protein, β_2-microglobulin. The three classical class I proteins, HLA-A, HLA-B and HLA-C are expressed on virtually all nucleated cells and also on platelets. HLA class II molecules are similar in structure to HLA class I molecules and are composed of two MHC-encoded polypeptide chains (α and β). The three HLA class II molecules, HLA-DR, HLA-DQ and HLA-DP, have a restricted tissue distribution; they are mainly expressed on antigen-presenting cells such as B cells, dendritic cells and macrophages, but can be expressed upon other cells in the context of inflammation.

Both class I and II molecules share a similar structure of four extramembranous domains. The two most membrane-distal domains of HLA class I (α1 and α2) and HLA class II (α1 and β1) contain a cleft available for binding of small peptides. Thus, both class I and class II proteins can be considered trimolecular proteins consisting of three subunits: HLA heavy chain, β_2-microglobulin and peptide form class I molecules, and HLA α- and β-chains and peptide form class II molecules. The binding of peptides is key to the functions of the classical HLA molecules. The source of peptides and the binding procedure differs for the two classes of molecules. HLA class I molecules are assembled within the endoplasmic reticulum (ER), where chaperone proteins mediate the association with β_2-microglobulin and peptide. Peptides are derived from endogenous molecules (e.g. self or viral proteins) by cytosolic degradation within the proteasome, a multicatalytic protein complex, before being actively transported into the ER. Once the HLA class I trimolecular complex is formed, the molecule leaves the ER and traffics to the cell surface via the Golgi apparatus. The formation of a peptide-bound HLA class II molecule is distinct in several ways. Both α- and β-chains are directed to the ER, where they first associate with the transmembrane invariant chain. The invariant chain blocks binding of ER-resident peptides and instead directs the

Postgraduate Haematology, Seventh Edition. Edited by A Victor Hoffbrand, Douglas R Higgs, David M Keeling and Atul B Mehta.

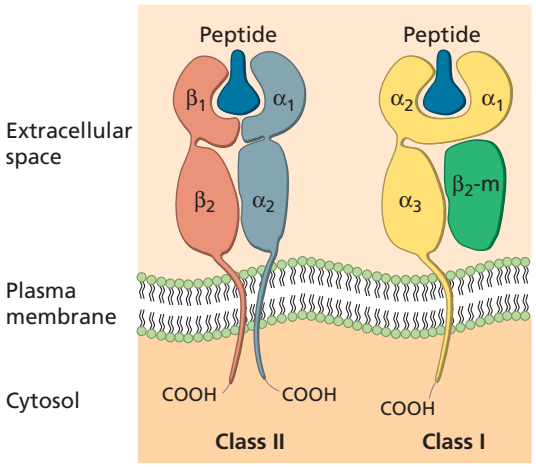


Figure 35.1 Diagrammatic representation of HLA class I and II molecules showing protein domains and bound peptide.

class II molecule via the Golgi apparatus to endosomal vesicles. Peptides contained in the endosome are derived from the processing of exogenous proteins (e.g. derived from microbes or dead cells). Within the endosome, the invariant chain is cleaved in a series of steps and specific peptide binding is catalysed by HLA-DM before the HLA class II–peptide complex is transported to the cell surface.

T cells and natural killer (NK) cells both possess receptors that recognize and interact with HLA molecules (see also Chapter 16). T cells expressing the CD4 molecule recognize HLA class II molecules and T cells expressing CD8 recognize HLA class I molecules. The specificity of interaction is determined during development of T cells within the thymus, such that circulating T cells should not interact with HLA molecules presenting peptides derived from normally expressed 'self' proteins. However, the presentation of 'non-self' peptides, for example peptides derived from viral or bacterial proteins initiate activation signals mediated by the T-cell receptor and coreceptors, resulting in the generation of an immune response against the cells expressing the target HLA protein–peptide complex. NK cells express a highly diverse repertoire of receptors, including killer-cell immunoglobulin-like receptors (KIRs) that elicit either activation or inhibitory signals. Interaction between an NK cell and a target cell is influenced by the interaction between certain HLA class I molecules on the target cell (e.g. HLA-B and C) and KIRs on the NK cell. Interactions between an HLA class I molecule and a corresponding KIR may result in a negative signal that outweighs any activation signals; thus the NK cell will not attack the target cell. However, if a cell has lost expression of HLA molecules as a result of malignancy or viral infection, the absence of the inhibitory signal may result in activation signals being dominant and allow NK cell-mediated attack on the target cell.

The outstanding feature of the HLA genes is their extensive polymorphism (www.ebi.ac.uk/imgt/hla/). At each of the genes, there are multiple possible variant alleles and the majority of this variation occurs within regions encoding the peptide-binding domains. Because different HLA proteins bind peptides with different sequences, one of the functions of HLA polymorphism is to allow the presentation of numerous different peptides to the immune system. As there are six antigen-presenting 'classical' HLA molecules (HLA-A, -B, -C, -DR, -DQ, -DP) and most individuals are heterozygous for these loci, potentially each individual has 12 different HLA molecules, each of which can bind thousands of different peptides. Evolution of HLA polymorphism occurred as a result of selective pressure imposed by geography, climate and infectious pathogens so that the repertoire of HLA alleles varies between different populations.

Detection of HLA polymorphisms is now routinely performed using DNA methods rather than serotyping. Current methods include PCR with sequence-specific primers or sequence-specific oligonucleotide probes, although these approaches will be rapidly superseded by high-throughput, low-cost automatic next generation sequencing technologies. Each HLA allele is given a unique numerical designation. For example, for the allele *HLA-A*02:101:01:02N,* the gene name *HLA-A* is followed by an asterisk (*) and then a series of four fields separated by colons (:). The first field indicates the allele group (02), which often corresponds to the broad serological antigen encoded by the allele. The second field (101) indicates the subtype, a number assigned in the order in which the DNA sequences was determined. Alleles whose numbers differ in the first two fields must differ in one or more nucleotide substitutions that change the amino acid sequence of the encoded protein. The third field is used to name alleles that differ only by synonymous nucleotide substitutions (also called silent or non-coding). The fourth field is used to name alleles that differ in either intron, or 3′ or 5′ regions of the gene. Lastly, an allele may have a suffix indicating aberrant expression; for example *N* indicates that it is a null allele with no protein being expressed, *L* indicates low cell-surface expression and *S* indicates that the molecule is expressed only in a soluble form.

HLA matching for transplantation

The frequency of T cells capable of reacting against a foreign antigen is of the order of 1 in 10–100,000. However, if cells from two HLA disparate individuals are mixed, the frequency of responding (or 'alloreactive') cells can be as high as 1–10%. To overcome alloreactivity following allogeneic SCT, it is necessary to define the HLA type of donor and recipient and to select an HLA-matched donor wherever possible. Excessive alloreactivity occurring after allogeneic SCT can result in graft failure or GVHD.

In situations where the donor is genetically identical to the recipient (syngeneic transplantation), graft rejection or significant GVHD are not observed. For patients undergoing allogeneic transplantation using an HLA-identical sibling, there is still risk of GVHD as a result of polymorphisms outside the HLA system that generate so-called 'minor' histocompatibility antigens that are presented as processed peptides in the context of 'self' HLA molecules. For example, male patients will possess antigens encoded by the Y chromosome that are absent in females (HY antigens). Female donors, specifically multiparous individuals who have had male infants, may have been primed against HY antigens and as a consequence have circulating T cells that recognize male patient cells expressing these proteins. This is likely to underlie the increased risk of GVHD in male recipients of female grafts.

The chances of a sibling being HLA-identical to a patient is theoretically one in four and, taking into account the size of most families, about a third of patients will have an HLA-identical sibling donor. Another source of HLA-matched donors are the various volunteer donor registries (such as the Antony Nolan Registry and the US National Marrow Donor Program) that exist throughout the world and that have accrued over 20 million donors. The likelihood of finding an optimal donor matched for *HLA-A, HLA-B*, *HLA-C* and *HLA-DRB1* alleles is about 75% amongst whites of European descent, but significantly lower for other ethnic groups. Although single HLA mismatches can be reasonably well tolerated clinically, greater degrees of mismatch are associated with increased rates of graft rejection or GVHD. Certain HLA mismatches may be 'permissive' so that they do not result in a worse transplant outcome. For example, it has been proposed recently that mismatches at HLA-DPB1 (a locus not usually considered in matching with unrelated donors) may actually be associated with an improved outcome because of a lower risk of disease relapse.

In the absence of an adult donor, there is the option of using umbilical cord blood cells because a reduced degree of HLA matching is acceptable, most likely reflecting the reduced alloreactivity of fetal T cells. Cord blood units mismatched for one or two loci are available for the majority (>80%) of adult patients. More extensive HLA disparity between donors and recipients following related haploidentical transplantation, is also becoming more feasible as a result of novel transplantation strategies involving transfer of unmanipulated grafts followed by *in vivo* purging of alloreactive T cells using cyclophosphamide, or alternatively using methods to completely deplete T cells from the graft. For both cord and haploidentical SCT, there is growing interest in other factors that influence outcomes. For example, the risks of transplantation may also be influenced by the complex interaction between mismatching at HLA class I and the repertoire of activating or inhibitory KIR genes inherited by the donor. Furthermore, there is also evidence that recipients of cord or haploidentical grafts may be less prone to develop GVHD if they possess a haplotype that is absent in the donor, but is shared with the donor's mother; in this scenario, *in utero* exposure to non-inherited maternal antigens (NIMAs) is postulated to induce tolerance of donor T cells.

Acute graft-versus-host disease

GVHD is a complex immunological disorder in which donor T cells initiate tissue damage consequent upon recognition of recipient antigens not expressed in the donor. The recipient is unable to reject the donor T cells because of the immunosuppressive effects of conditioning therapy. Chemotherapy- or radiation-induced tissue damage leads to the release of proinflammatory cytokines and altered chemokine or adhesion molecule expression. These changes impact on the developing alloreactive response, leading to further cytokine dysregulation and enhanced trafficking of cellular effectors to the organs targeted in GVHD.

Initiation of GVHD (Figure 35.2) requires the interaction of donor T cells with recipient antigen-presenting cells (APCs) within secondary lymphoid organs, such as lymph nodes and gut-associated lymphoid tissue, although there appears to be considerable redundancy in the type of cells involved. Conditioning-related tissue injury leads to the release of damage-associated molecular patterns (DAMPs, e.g. uric acid and ATP) or exposure to pathogen-associated molecular patterns (PAMPs, resulting from entry of microbes across damaged epithelia). Recognition of DAMPs or PAMPs by cell-surface or intracellular pattern-recognition receptors induces the release of proinflammatory cytokines, such as interleukin (IL)-1β. Enhanced expression of HLA and costimulatory (e.g. B-7 family) molecules upon antigen-presenting cells promotes their interaction with donor CD4+ and CD8+ T cells. Donor T cells with specificity for host antigens are then activated by the APCs, rapidly undergoing proliferation and developing a multitude of effector functions, including the generation of proinflammatory cytokines (e.g. TNF-α, interferon-γ and IL-17) or the upregulation of perforin/granzyme B or Fas ligand pathways. The primed effector T cells then traffic in large numbers to inflamed peripheral tissues, where they induce apoptosis of epithelial target cells and recruit other cellular effectors such as neutrophils, macrophages and NK cells, which in concert cause further damage. In the ensuing inflammatory response, T cells recognizing other antigens may also become activated (so called 'epitope spreading') leading to even worse injury.

Chronic graft-versus-host disease

Chronic GVHD is a disorder associated with continuing host injury and profound immunodeficiency. The pathogenesis of this disorder is less well defined, partly as a result of the lack of preclinical models that fully recapitulate its onset and clinical features. The thymus is a target of GVHD and this may lead to impaired negative selection of autoreactive T cells that can

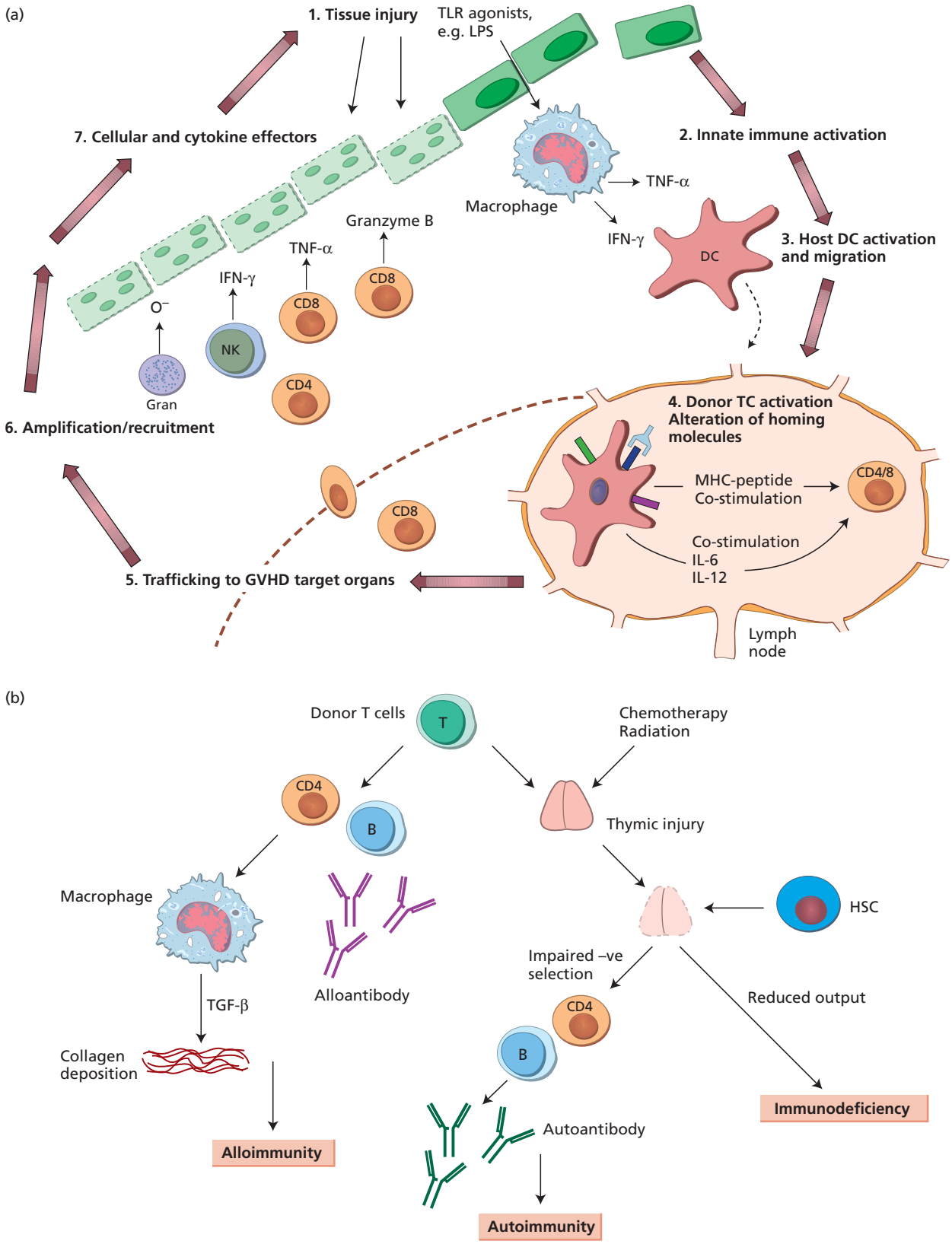


Figure 35.2 Pathophysiology of acute GVHD (a) and chronic GVHD (b). See text for for explanation of abbreviations.

then cause further injury to peripheral tissues. This may help to explain why individuals with chronic GVHD develop some clinical features normally associated with autoimmunity (e.g. immune cytopenias). Host-reactive CD4+ T cells might provide helper functions for B cells leading to the production of allo- or autoantibodies and indeed, some chronic GVHD in some patients responds to anti-B-cell monoclonal antibody therapy. Skin infiltration by donor macrophages may drive excess collagen deposition in the skin, a hallmark feature of the scleroderma-tous form of the disease.

Graft-versus-leukaemia effect

The importance of an immunologically mediated GVL effect in contributing to the curative effect of allogeneic SCT is supported by the observations that T-cell depletion (TCD) increases the risk of relapse and that patients who develop GVHD have a lower risk of disease relapse. Conclusive evidence was provided by the demonstration that infusion of donor lymphocytes can produce durable remissions in patients who have relapsed after allogeneic SCT. It has subsequently been shown that donor lymphocyte infusion (DLI) is a remarkably effective salvage therapy in chronic myeloid leukaemia (CML), with more than 80% of patients achieving a sustained molecular remission. Durable responses can also be achieved in patients transplanted for indolent lymphoma and Hodgkin lymphoma, with responses being observed less commonly in patients with aggressive lymphoma, myeloma or acute leukaemias. As would be predicted, DLI can be complicated by the development of GVHD, although the risk of this life-threatening complication is reduced if the lymphocytes are infused following a significant delay (>12 months) after transplantation or by using an escalating dose schedule rather than a single 'bulk' infusion. For the most part, the antigens recognized in a GVL reaction overlap with those recognized during GVHD. Clinical separation of GVL and GVHD may reflect increased sensitivity of normal or malignant haemopoietic tissue to an emerging GVH reaction compared with epithelial cells. Selective expression of antigenic targets of a GVH reaction on haemopoietic tissues (e.g. the minor H antigen HA-1, proteinase 3 or WT1) or leukaemic blasts (e.g. the product of the *BCR–ABL1* fusion gene) may also underlie the development of a GVL reaction in the absence of GVHD. This concept forms the basis of novel strategies to deliver a GVL effect without a concomitant risk of GVHD, such as peptide vaccination or gene transfer of T-cell receptors (TCRs) or chimeric antigen receptors (CARs) specific for leukaemic antigens.

Immune reconstitution

Allogeneic SCT is followed by a prolonged period of cellular and humoral immunodeficiency, while donor-derived immune recovery occurs. Reconstitution of an immune response after transplantation can be evaluated in the clinic by monitoring the absolute numbers of T (CD4, CD8), B and NK cells. NK cell numbers recover most rapidly while other subsets (especially CD4+ T cells and B cells) recover more slowly. In patients who have received a T-cell-depleted graft, T-cell recovery is substantially delayed. Analysis of thymic function after transplantation can be evaluated by measuring the episomal DNA excision circles of the TCR δ locus deleted during recombination of the TCR in functional αβ T cells (known as TREC, for T-cell receptor excision DNA circles). TREC levels are low for the first 6 months following transplantation and recover thereafter. As a consequence, the T-cell repertoire is limited and mostly dependent on expansion of donor memory T cells. Thymic function is reduced in adults and may be further compromised by the effects of chemoradiotherapy and GVHD. Significant HLA mismatching between donor and recipient may also lead to 'holes' within the T-cell repertoire due to perturbations in thymic selection. These defects reflect a failure of donor-derived thymic emigrants to interact with peptides presented in the context of 'foreign' host HLA molecules.

Quantitative B-cell deficiency is present in virtually all patients in the first months after transplantation and may persist for a number of years post transplantation as a consequence of reductions in the number of marrow B-cell precursors, particularly in patients with chronic GVHD. This defect in B-cell production has multifactorial causes, including damage to the bone marrow stroma, the deleterious effect of inflammatory cytokines and the lympholytic effects of glucocorticoid therapy.

The source of stem cells also influences the kinetics of immune reconstitution. Thus, peripheral blood stem cell (PBSC) grafts, which contain higher numbers of haemopoietic progenitors and mature T cells, are associated with more rapid immune reconstitution than bone marrow grafts. Cord blood transplantation may be affected by poor immune reconstitution in adults since the number of haemopoietic progenitors is often limited and the transferred T cells are naive. Thus, patients are at heightened risk of re-activation of cytomegalovirus (CMV) or Epstein–Barr virus (EBV).

Stem cell engraftment

Biology of stem cell engraftment

The establishment of durable donor haemopoiesis after SCT depends on the engraftment of long-term reconstituting haemopoietic stem cells (HSCs) (see Chapter 1). These cells, defined by their capacity for self-renewal as well as their ability to differentiate into all haemopoietic lineages, are normally resident in the bone marrow at low frequency, but can be mobilized into the peripheral blood by cytokines or chemotherapy. A high frequency of haemopoietic stem and progenitor cells is also present in umbilical cord blood (UCB). The cell-surface glycoprotein CD34 is expressed on haemopoietic progenitors and

Table 35.1 Factors determining stem cell engraftment.

Autologous transplantation
Stem cell dose
Allogeneic transplantation
Stem cell dose
Intensity of host immunosuppression delivered by the conditioning regimen
Numbers of donor T cells in the stem cell inoculum
Degree of genetic disparity between donor and host

HSCs and is currently widely used as a stem cell marker in clinical transplantation. However, it is important to remember that xenograft studies have demonstrated that HSCs reside within a $CD34^+CD38^-$ subpopulation and that many of the cells within the $CD34^+$ population lack the properties of a long-term reconstituting stem cell. Murine transplant studies have established that in a syngeneic setting, where there is no HVG reaction, transplantation of very small numbers of HSCs can result in durable engraftment of lethally irradiated recipients. In contrast, host immunosuppression is required to blunt an HVG reaction capable of rejecting the transplanted stem cell inoculum when there is donor–host HLA disparity, whether the donor is an HLA-identical sibling or an alternative donor.

Clinical factors determining stem cell engraftment

The sole determinant of durable engraftment after an autologous or syngeneic transplant is stem cell number and graft failure is exceedingly rare, providing at least 2×10^6 $CD34^+$ cells per kilogram are transplanted. Clinically durable engraftment in allograft recipients is determined by the degree of HLA disparity, the number of transplanted T cells and the size of the stem cell inoculum (Table 35.1). The immunosuppressive properties of the conditioning regimen play a critical role in blunting the HVG response and incorporation of total body irradiation (TBI) and highly immunosuppressive drugs such as fludarabine, coupled with optimal post-transplant immunosuppression, are critical in ensuring durable stem cell engraftment. Historically, high rates of graft failure were observed when T-cell depletion (TCD) was introduced as a form of GVHD prophylaxis in the late 1980s. However, the recognition that donor T cells play a critical role in facilitating stem cell engraftment led to re-design of conditioning regimens in those patients receiving T-cell-depleted grafts such that their immunosuppressive properties were optimized. The widespread availability of PBSCs from sibling or unrelated donors (see below) allows transplantation of five to ten times more CD34-positive cells than if harvested bone marrow were used. Consequently, the use of PBSCs has played an important role in optimizing engraftment in settings such as TCD or transplantation of a mismatched donor, where a significant risk of graft failure rate would be expected were bone marrow to be used as the stem cell source. Incorporating these principles into clinical practice has markedly reduced the risk of graft failure such that it occurs in fewer than 1% of patients undergoing an HLA-identical sibling allograft and fewer than 5% of those transplanted from an unrelated donor. For a long time graft failure was a major complication of cord blood transplantation (CBT), particularly in adult recipients, and this likely reflected the low stem cell dose. The recent demonstration that the simultaneous transplantation of two cord blood units delivers durable rates of engraftment in the region of 95% has dramatically improved outcome of UCB transplants in adults.

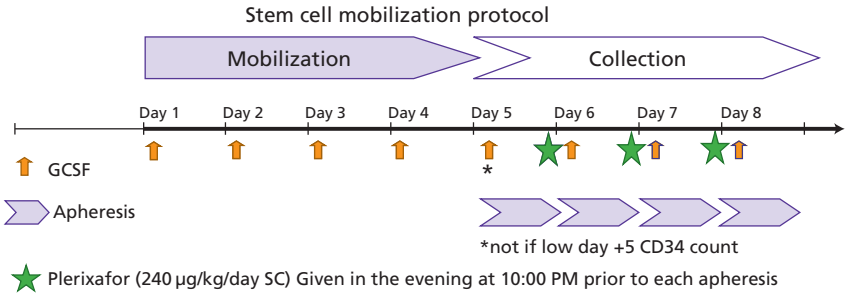
Stem cell mobilization

Biology of stem cell trafficking

In steady-state haemopoiesis only very small numbers of haemopoietic stem and progenitor cells are present in the peripheral blood. The localization of haemopoietic stem and progenitor cells within the bone marrow cavity is mediated through the binding of a range of adhesion molecules, including CXCR4 and VLA4, with their cognate ligands on bone marrow stroma. Recent studies have demonstrated that administration of haemopoietic growth factors such as granulocyte colony-stimulating factor (G-CSF) disrupts the adhesion of progenitors to the bone marrow stroma, resulting in their mobilization into the peripheral blood in large numbers. This has had a dramatic impact on transplant practice, resulting in peripheral blood replacing bone marrow as the commonest source of haemopoietic progenitors in both autologous and, more recently, allogeneic transplants. Antagonists to CXCR4, such as plerixafor, significantly augment G-CSF-mediated stem cell mobilization and this has transformed mobilization options in patients considered for an autologous transplant. Adhesive interactions between haemopoietic progenitors and the bone marrow stroma also play a critical role in the homing of transplanted HSCs and stem cell engraftment can be augmented in animal models by manipulation of the levels of adhesion molecule expression on haemopoietic progenitors and bone marrow stroma. As yet these observations have not been translated into clinical practice.

Stem cell mobilization in clinical practice

In patients undergoing autologous SCT, G-CSF administered alone or after myelosuppressive chemotherapy represents the commonest method of stem cell mobilization. G-CSF administered in conjunction with salvage chemotherapy is an effective mobilization strategy in most patients with relapsed lymphoma. It is now clear that administration of G-CSF alone, albeit at somewhat higher doses, can be equally effective and is substantially less toxic, and this option is now widely used as the



Day	1	2	3	4	5	6	7	8
G-CSF	✓	✓	✓	✓	✓	✓[a]	✓[a]	✓[a]
Plerixafor injection [0.24 mg/kg] (only if low circulating CD34 count on day +5)					✓[a]	✓[a]	✓[a]	
Apheresis*					✓[a]	✓[a]	✓[a]	✓[a]

[a] If necessary

Figure 35.3 'Pre-emptive' mobilization schedule.

mobilization option of choice in patients with myeloma who are eligible for an autograft. This option, consisting of four daily subcutaneous injections of G-CSF, with stem cell harvesting on days 5 and 6, can be easily delivered as an out-patient and has modest but distinct toxicity, which typically takes the form of myalgia, bone pain and headache. Of note there are rare reports of splenic rupture after G-CSF administration. Taken together both options allow a minimum target cell dose of 2×10^6 $CD34^+$ cells/kg to be harvested in more than 80% of patients. In patients who fail to mobilize the target number of $CD34^+$ cells the CXCR4 antagonist plerixafor can be successfully combined with G-CSF and this advance now permits mobilization of the required numbers of PBSC in more than 95% of autograft candidates. Initially used in patients who had previously failed to mobilize in response to G-CSF alone or G-CSF/chemotherapy, plerixafor can now be pre-emptively administered in patients undergoing G-CSF mobilization who can be predicted to be poor mobilizers on the basis of low numbers of circulating $CD34^+$ cells in the peripheral blood on the first day of leucapharesis (Figure 35.3). In the small minority of eligible patients with relapsed lymphoma who fail to mobilize adequate numbers of PBSCs using combined G-CSF/plerixafor administration, allogeneic transplantation should be considered.

G-CSF-mobilized PBSCs now represent the commonest stem cell source in patients undergoing allogeneic SCT. The apparent safety of G-CSF doses in the range routinely used for autologous stem cell mobilization (10–15 µg/kg for 4–6 days) has resulted in its adoption as the standard approach for stem cell mobilization in sibling and, increasingly, adult unrelated donors. The options of both PBSC donation and bone marrow harvest, as well as any expressed preference of the transplant centre, are discussed with volunteer unrelated donors. Donors who elect to receive G-CSF also need to be aware of the 2–3% chance that an additional bone marrow harvest will be required if G-CSF fails to mobilize sufficient PBSC. Combined G-CSF/plerixafor administration may become an alternative option in this population of normal donors in the future. Initial concerns that G-CSF may predispose to the subsequent development of secondary haematological malignancies in normal donors have not been substantiated. Although the effects of G-CSF on the fetus are not known, all female donors of childbearing age should have a negative pregnancy test prior to its administration.

Choice of stem cell source and dose

Autologous transplantation

The use of PBSCs as a stem cell source is associated with more rapid engraftment than is observed with harvested bone marrow and this, coupled with their ease of procurement, has secured their role as the standard stem cell source in autologous transplants. The generally accepted minimum number of haemopoietic progenitors required for engraftment after autologous SCT is 2×10^6 $CD34^+$ cells/kg. While increasing the number of $CD34^+$ cells transplanted hastens neutrophil and platelet engraftment, there is little evidence that transplantation of more than 5×10^6 $CD34^+$ cells/kg is beneficial and there remain concerns that higher stem cell doses may be associated with an increased risk of tumour contamination. In patients with myeloma, many centres aim to harvest a minimum of 4×10^6 CD34-positive cells/kg, using G-CSF alone or in combination with plerixafor, in order that sufficient cells are available so that a tandem transplant, or second autograft at the time of relapse, is possible.

Sibling and unrelated donor transplantation

PBSCs are the most frequently used stem cell source for patients undergoing a sibling allograft and are increasingly used in unrelated donor transplant recipients. While PBSCs are increasingly utilised as the preferred stem cell source in reduced-intensity allografts in order to optimise engraftment, it is less clear

whether they are preferable to harvested bone marrow in patients transplanted using a myeloablative conditioning regimen. Transplantation of PBSCs results in earlier neutrophil and platelet engraftment, which may reduce transplant-related mortality (TRM), particularly in patients with advanced leukaemia, there are emerging data that their use may be associated with an increased incidence of chronic GVHD compared with bone marrow in patients transplanted without the use of TCD. This is likely to be consequent on the five- to tenfold greater dose of T cells in a PBSC harvest compared with bone marrow. Thus while PBSCs are commonly used in patients being allografted for advanced leukaemia, where TRM is a major cause of treatment failure, there are considerable centre differences in what is deemed the optimal stem cell source in standard risk allografts. In paediatric transplantation, bone marrow remains the preferred stem cell source and the use of PBSCs is less common.

Cell dose is an important factor determining outcome after both matched sibling and unrelated donor transplants an observation is noted in recipients of both bone marrow and PBSCs. A component of this effect is mediated through a reduction in TRM likely consequent on accelerated immune reconstitution in recipients of a higher stem cell dose. The lowest acceptable stem cell dose to secure engraftment in recipients of a PBSC allograft is considered to be 2×10^6 CD34-positive cells/kg (2×10^8 mononuclear cells/kg if bone marrow is being used), although in practice the great majority of patients transplanted with a PBSC graft containing in excess of 1×10^6 CD34-positive cells/kg will still engraft. Since outcome is improved with higher doses of CD34-positive cells, most centres aim for a dose in the region of 4×10^6 CD34-positive cells/kg. In patients transplanted from an HLA-identical sibling using a T-replete myeloablative regimen, there is some evidence that the incidence of chronic GVHD is increased if the stem cell dose exceeds 8×10^6 CD34-positive cells/kg. In this setting it may therefore be reasonable to have a target dose of $4–8 \times 10^6$ CD34-positive cells/kg. However, there is no evidence, as yet, that there should be an upper limit on cell dose in patients transplanted using an RIC regimen or if the graft is T cell depleted.

Cord blood transplantation

UCB contains a high proportion of haemopoietic progenitors and HSCs and is an increasingly important stem cell source in paediatric and adult transplantation. Importantly, HLA disparity appears better tolerated in recipients of UCB and as a consequence the incidence of severe GVHD is lower with mismatched UCB than would be expected using a comparably mismatched unrelated donor. This has important implications in terms of donor identification in patients with uncommon HLA types for whom a suitably matched unrelated donor cannot be readily identified. A major factor limiting the uptake of CBT has been delayed or failed engraftment, which used to be a common problem, particularly in adult recipients. The two most important factors determining the likelihood of neutrophil and platelet engraftment after CBT are nucleated cell dose and HLA disparity. More recently, it has been shown that transplantation of two cord blood units, harvested from separate donors, increases the speed of engraftment and is associated with a decreased risk of primary graft failure. Consequently results approaching those achievable using a well-matched adult unrelated donor are now routinely observed in patients with high-risk leukaemia after a double cord transplantation performed using cord blood units that meet well defined cell dose and matching criteria. Immune reconstitution is markedly delayed after CBT and viral infections, particularly CMV, HHV 6 and adenovirus, contribute to the significant infectious-related TRM. A number of studies identify the potential of *in vitro* manipulation of one or both cord blood units, utilizing cytokines or other stem cell expansion technologies, as a strategy to hasten engraftment and immune reconstitution and are the subject of ongoing prospective studies. For all the reasons outlined above there is a compelling rationale that cord blood unit selection and transplantation should be performed in transplant centres with expertise in this area.

Haploidentical stem cell transplantation

Transplantation from a haploidentical donor has historically been associated with excessive rates of graft failure and severe GVHD consequent upon the major degree of HLA disparity. Initial approaches utilized stringent *ex vivo* T-cell depletion coupled with an intensely immunosuppressive conditioning regimen and transplantation of high doses of CD34-positive cells to overcome these barriers. Whilst encouraging results were reported from a few centres, an unacceptably high incidence of viral infections and consequent complex post-transplant course prevented this approach being widely adopted. A refinement of this approach in which CD3-positive and CD19-positive cells are depleted from the graft, thereby preserving dendritic cells, monocytes and natural killer cells in the graft, has yielded promising results, particularly in paediatric recipients. A radically different approach, in which cyclophosphamide is administered in the first few days post transplant, resulting in depletion of alloreactive host and donor T cells, has recently been shown to both result in high rates of engraftment and substantially reduce the risk of severe GVHD. Although considerable uncertainty remains concerning the degree to which such a manoeuvre abrogates the GVL effect and increases the relapse risk, there are now sufficient encouraging data to support the prospective examination of haploidentical grafts, delivered using the described novel technologies, compared with other alternative stem cell sources, such as cord blood.

Conditioning regimens: basic principles

The combination of drugs and radiotherapy administered prior to stem cell infusion is termed the conditioning or preparative regimen. In autologous SCT, where there is no alloreactive

response, the sole purpose of the conditioning regimen is disease eradication. The most common conditioning regimens used in autologous transplantation utilize alkylating agents, alone or in combination, and other chemotherapeutic drugs.

In the setting of allogeneic SCT the conditioning regimen serves two purposes: immunosuppression designed to abrogate an HVG reaction, thereby preventing graft rejection, and dose escalation of chemoradiotherapy in order to de-bulk and potentially eradicate tumour. Historically, only myeloablative (MA) conditioning regimens, which lead to irreversible cytopenias, were utilized in patients undergoing allogeneic SCT. However, increasing recognition of the potency of the GVL effect, coupled with awareness that the toxicity of the myeloablative regimen precludes their deployment in patients older than 50–55 years has led to the development of less intensive conditioning regimens, which are designed to be tolerable in older patients and rely on a GVL effect as the dominant antitumour activity. Associated with a substantially reduced TRM, non-myeloablative (NMA) (which results in only transient cytopenias) and RIC regimens (which delivers significant cytopenias and requires stem cell support, but employs substantially reduced doses of chemoradiotherapy compared with myeloablative regimens) allow allogeneic transplantation to be safely performed in patients in whom it would previously have been contraindicated on the grounds of age or comorbidity. The remarkable success of RIC and NMA regimens in reducing early transplant toxicity has resulted in a dramatic increase in the number of allografts performed worldwide and has transformed treatment options for many patients with AML, MDS, ALL and relapsed lymphoma in whom conventional chemotherapy would be unlikely to deliver long-term disease-free survival. Currently there is a proliferation of different RIC and NMA regimens that require evaluation in prospective clinical trials. In addition, although the reduction in conditioning intensity appears to be associated with some increase in the risk of disease relapse, compared with MA regimens, there may be an important role for RIC allografts in younger patients because of their dramatically reduced early mortality and possible longer-term benefits, such as preserved fertility.

Conditioning regimens in autologous SCT

Conditioning regimens in autologous SCT are designed with dose intensification in mind and are limited mainly by considerations of extramedullary toxicity. High-dose melphalan (200 mg/m^2) is the standard conditioning regimen in myeloma autografts. BEAM (carmustine, etoposide, cytarabine, melphalan) is widely used in patients with lymphoma. In patients with relapsed CNS lymphoma, thiotepa plays an important role. Autologous transplants are indicated in certain patients with AML, principally those with APML in second CR, but sometimes also in patients with AML associated with a CBF abnormality who lack an allogeneic donor, and both busulfan/cyclophosphamide and cyclophosphamide/TBI are effective preparative regimens in this setting. A number of other drug combinations incorporating melphalan, busulfan and thiotepa are used in solid tumours. The major extramedullary toxicities of these regimens are mucositis and gastrointestinal toxicity. Disappointingly, there are few prospective randomized data on which to base the choice of conditioning regimen in autologous SCT. In myeloma a randomized comparison between a TBI-containing regimen and high-dose melphalan alone demonstrated that no benefit was associated with the use of TBI. Encouraging results have been reported in patients with myeloma who receive a combination of busulfan and melphalan, but the results of prospective randomized trials are awaited.

Myeloablative conditioning regimens in allogeneic SCT

The two commonest MA conditioning regimens used in allogeneic SCT employ combinations of cyclophosphamide and either TBI or busulfan.

TBI/cyclophosphamide

Cyclophosphamide is an alkylating agent that, when administered in the doses routinely used in myeloablative conditioning regimens (120–200 mg/kg), has both immunosuppressive and antileukaemic properties. It is a prodrug that must be metabolized by the cytochrome P450 system in the liver to produce metabolically active derivatives, principally phosphoramide mustard, which exert their cytotoxic activity through the production of interstrand DNA links. The two major complications of cyclophosphamide at the doses employed in allogeneic transplantation are haemorrhagic cystitis and cardiac toxicity. Haemorrhagic cystitis results from the toxic effects of a cyclophosphamide metabolite, acrolein, on the uroepithelium and can be reduced by use of sodium 2-mercaptoethanesulfonate (MESNA), while cardiac toxicity is very rare at doses of cyclophosphamide below 150 mg/kg.

TBI has dual immunosuppressive and antileukaemic properties when administered in myeloablative doses (typically 12–14.4 Gy). Radiobiological principles predict that the toxicity of TBI can be reduced by either decreasing the overall dose of radiation or, as is now common, by giving it in fractionated form over a number of days (e.g 14.4 Gy divided into eight fractions over 4 days). The degree of immunosuppression produced by TBI-containing regimens is related to the total dose of irradiation delivered. The use of higher TBI doses is therefore an effective method of optimizing engraftment in myeloablative allografts where there is use of TCD or an alternative donor and consequently an increased risk of graft failure. Haematological malignancies are highly radiosensitive and the risk of disease relapse is reduced if a higher dose of TBI is used, although this benefit is blunted by increased transplant toxicity (see below). Early complications associated with the use of TBI include nausea,

vomiting, diarrhoea and parotitis, which can usually be managed symptomatically. Increased doses of TBI are also associated with pneumonitis and veno-occlusive disease (VOD) of the liver, both of which may be life-threatening. Long-term complications include cataract formation, hypothyroidism, infertility and, in children, growth retardation.

Busulfan/cyclophosphamide

Busulfan/cyclophosphamide (Bu/Cy) is a well-established MA conditioning regimen which has the practical advantage of not requiring the presence of irradiation facilities on site. Busulfan is an alkylating agent with potent activity against leukaemic progenitors and is a core component of both allogeneic and autologous transplant regimens. Historically, busulfan has only been available as an oral preparation, used at a dose of 14–16 mg/kg delivered 6-hourly over a period of 4 days. In this formulation, VOD and pulmonary and central nervous system (CNS) toxicity represent major complications. The pharmacokinetics of oral busulfan are highly variable because it undergoes first-pass metabolism in the liver and there is therefore substantial interpatient variability in plasma drug levels using a standard dosing schedule. Because the incidence of VOD is closely correlated with higher plasma busulfan levels and patients with low busulfan levels also have an increased risk of relapse, oral preparations of busulfan are far from ideal. Two approaches have been used to overcome this problem. A number of groups use frequent (6-hourly) measurement of plasma busulfan levels during the first 24 hours of administration of an oral preparation, followed by dose adjustment over the following 3 days in order to achieve a therapeutic busulfan level. Using such a targeted approach, the Seattle group have achieved excellent results using busulfan/cyclophosphamide with a very low risk of VOD and reduced relapse rates compared with those achieved using oral busulfan. The recent development of an intravenous formulation of busulfan which is associated with more predictable pharmacokinetics is much better tolerated than oral preparations and is rapidly becoming the preferred formulation of this important constituent of both MA and RIC regimens. It should be remembered that both oral and intravenous preparations require the use of prophylactic phenytoin or clonazepam to prevent seizures, a complication associated with the administration of high doses of busulfan.

Alternative myeloablative conditioning regimens

Fludarabine augments alkylator-induced cell killing *in vitro* and regimens combining fludarabine with intravenous busulfan are active and well tolerated. The incidence of VOD using this combination appears to be low, and prospective studies of this potentially important new MA conditioning regimen in patients with high-risk myeloid malignancies are ongoing. Treosulfan is a novel alkylating agent that shows promise as a component of both MA and RIC regimens in patients allografted for myeloid disorders – specifically MDS and MPN. In patients undergoing a sibling allograft for ALL, etoposide is often substituted for cyclophosphamide and a TBI/etoposide regimen have been reported to be associated with improved outcome.

In non-malignant disorders, such as aplastic anaemia, cyclophosphamide alone can be used as a conditioning regimen and is sufficiently immunosuppressive to permit engraftment of allogeneic stem cells, provided an adequate stem cell inoculum is transplanted. Addition of fludarabine is increasingly used in conjunction with busulfan coupled with alemtuzumab or antithymocyte globulin (ATG) in patients with sickle cell disease and thalassaemia.

Comparison of myeloablative conditioning regimens

Sibling allografts

There are two central questions to be considered in the design of myeloablative conditioning regimens: is there any survival benefit to be gained from intensifying the conditioning regimen, and are Cy/TBI and Bu/Cy equally effective preparative regimens? Prospective randomized trials in patients undergoing a sibling allograft for AML have failed to show any improvement in survival using an increased dose of TBI. Studies performed over two decades ago showed that while increasing the TBI dose reduces the risk of leukaemic relapse, this benefit is offset by a concomitant increase in TRM. Similarly, there is no evidence that addition of busulfan to a cyclophosphamide/TBI regimen has any impact on disease-free survival, because of the increased toxicity associated with this regimen.

The decision to use Cy/TBI or Bu/Cy as a conditioning regimen in sibling allografts has been studied in a number of randomized studies, utilizing oral busulfan. Given the substantial toxicity associated with the use of oral busulfan, the results of these studies, which demonstrated broad equivalence of the two regimens in patients with myeloid malignancies, need to be interpreted with caution since they may well underestimate the potential benefit of a busulfan/cyclophosphamide regimen utilizing targeted or intravenous busulfan. Indeed, two recent large retrospective analyses, including patients who received intravenous busulfan, suggest a survival advantage with Bu/Cy in patients with AML. The results of a prospective randomized comparison of BU/CY and Flu/Bu are awaited.

Unrelated donor transplants

The optimal conditioning regimen in patients undergoing an unrelated donor transplant has not been determined and there are few randomized trials in this setting. The greater degree of HLA disparity associated with the use of an unrelated donor results in a higher risk of graft failure than that observed using an HLA-identical sibling. For this reason many groups elect to use a Cy/TBI-based regimen, with its greater immunosuppressive properties, although it should be noted that equivalent results

have been reported in a number of large series using a Bu/Cy regimen. The other important factor determining the choice of conditioning regimen in unrelated donor transplant is whether TCD is employed. Regimens utilizing rigorous *ex vivo* TCD are associated with a higher risk of graft failure. In all patients receiving a TCD unrelated donor transplant, it is important to ensure that the conditioning regimen is sufficiently immunosuppressive if primary graft failure is to be avoided.

Cord blood transplants

The increased HLA disparity and high rates of primary graft failure in CBT makes it critical to optimize the immunosuppressive properties of the preparative regimen. In adults undergoing transplantation using a myeloablative regimen, engraftment rates in the region of 90–95% can be achieved by the addition of fludarabine to a cyclophosphamide/TBI regimen, providing an adequate stem cell inoculum is used. ATG was initially used as additional GVHD prophylaxis in myeloablative CBT, but is less commonly used now because of delayed immune reconstitution and an increased risk of post-transplant lymphoproliferative disorders.

Strategies for GVHD prophylaxis in myeloablative regimens

Post-transplant immunosuppression using various combinations of ciclosporin, methotrexate, prednisolone and mycophenolate mofetil represent the commonest forms of GVHD prophylaxis in patients after a myeloablative sibling or unrelated donor transplant. Randomized trials from the Seattle group established the use of intravenous ciclosporin (2.5–5 mg/kg daily) and short-course methotrexate (administered on days 2, 4, 8 and 12 post transplant) as the most effective form of GVHD prophylaxis in patients transplanted undergoing a T-replete allograft using either an HLA-identical sibling or volunteer unrelated donor. Rates of chronic extensive GVHD are in the region of 30% and 66% for recipients of T-replete sibling and unrelated donor transplants, respectively, despite the use of ciclosporin/methotrexate GVHD prophylaxis. TCD is an additional, and highly effective method of reducing the risk of both acute and chronic GVHD. TCD can be achieved either by manipulating the stem cell inoculum *ex vivo* or by the *in vivo* administration of T-cell-depleting antibodies such as ATG or alemtuzumab (a humanized monoclonal antibody that recognizes CD52). Although a highly effective method of GVHD prophylaxis, TCD is associated with an increased risk of relapse and graft failure and delayed immune reconstitution, increasing the risk of post-transplant infections such as CMV and EBV. An important recent randomized trial confirmed the ability of *in vivo* ATG to reduce the risk of both acute and chronic GVHD without increasing relapse risk or impacting overall survival in recipients of a matched unrelated donor transplant. Clearly, the form of GVHD prophylaxis used for any particular patient should be selected with their individual risk of both GVHD and relapse in mind. Thus it may be desirable to avoid the use of TCD in patients with advanced leukaemia in whom the risk of relapse is high. In contrast, patients with a low risk of disease recurrence may benefit from more intensive GVHD prophylaxis. However, there remains no consensus concerning the use of TCD in allogeneic transplantation. Many UK and European groups, while performing T-replete sibling allografts, would choose to use *in vivo* TCD in unrelated donor transplant recipients. Compromise strategies in which TCD is used at the same time as further intensifying the conditioning regimen show promise.

Reduced-intensity conditioning regimens in malignant and non-malignant disease

Reduced-intensity regimens which incorporate either fludarabine or low-dose (200 cGy) TBI as a core immunosuppressive component reliably deliver durable donor stem cell engraftment and are associated with a markedly reduced toxicity, particularly in older patients or those with significant comorbidities. The recent demonstration that the risk of relapse after a RIC allograft is both reduced in the presence of chronic GVHD and correlates with the intensity of post-transplant immunosuppression confirms that these transplant platforms also have the capacity to exert a potent GVL effect. However, disease relapse and GVHD remain major causes of morbidity and mortality after RIC allografts and despite a proliferation of differing transplant regimens there is little randomized data to inform the choice of the optimal conditioning regimen. One of the only randomized comparison demonstrates contrasting activities of two of the most widely used conditioning regimens in patients transplanted for AML in CR1: the Seattle low-dose TBI regimen and the Flu/Bu/ATG regimen. Whilst the low-dose TBI regimen is associated with a markedly reduced 100-day transplant-related mortality this benefit is blunted by an increased risk of disease relapse observed using this regimen. On the basis of these data it is not possible currently to recommend an optimal conditioning regimen in older patients with AML and the results of ongoing randomized comparisons examining other reduced intensity conditioning regimens is a priority if patient outcomes are to be improved.

Low-dose (200 cGy) TBI-based regimens

Pioneering work in a canine model by Storb's group demonstrated that durable engraftment could be achieved with an NMA conditioning regimen that used immunosuppressive doses of TBI (200–450 cGy) in conjunction with post-transplant immunosuppression in the form of ciclosporin and mycophenolate mofetil. Combined administration of fludarabine and low-dose TBI (200 cGY) reliably achieves high rates of engraftment and full donor chimerism with a substantially reduced TRM, compared with myeloablative preparative regimens in patients with a range of haematologic malignancies up to (and

on occasions beyond) the age of 70. The main causes of treatment failure using this regimen include a substantial risk of chronic GVHD and disease relapse, particularly in patients with advanced haematologic malignancies Up to 66% of patients experience chronic extensive GVHD, the management of which is often very challenging in elderly and more frail patients.

Fludarabine-based RIC regimens

The most commonly utilized RIC regimens all incorporate an immuosuppressive fludarabine backbone administered in combination with a range of myelosuppressive drugs including melphalan, busulfan, treosulfan and cyclophosphamide. In common with NMA regimens, RIC regimens deliver long-term disease-free survival in patients with both myeloid and lymphoid malignancies and are associated with markedly reduced early transplant toxicity. Initial RIC regimens utilized T-replete stem cell incocula and were associated with a significant risk of severe acute and extensive chronic GVHD. Given the significant risk of GVHD associated with the use of both regimens, a major area of recent clinical research has focused on defining the optimal form of GVHD prophylaxis.

Strategies for GVHD prophylaxis in RIC regimens

Currently, two contrasting approaches are used for GVHD prophylaxis in patients undergoing an RIC allograft. The first uses combinations of ciclosporin or tacrolimus, methotrexate and mycophenolate mofetil as the sole form of GVHD prophylaxis and is used in patients transplanted using a T-replete stem cell inoculum. Methotrexate can be omitted by using either sirolimus (an MTOR inhibitor) or mycophenolate mofetil. The alternative approach uses ciclosporin alone in combination with either ATG or alemtuzumab as *in vivo* TCD. Supporters of such a T-depleted approach argue that the incidence of acute, and particularly chronic, GVHD is otherwise unacceptably high, particularly in older patients transplanted using an unrelated donor. It can also be argued that this strategy generates an effective platform for the subsequent delivery of DLI with less GVHD-related toxicity. They point to the effectiveness of ATG and alemtuzumab in reducing the risk of GVHD while at the same time securing high rates of durable donor engraftment. Those favouring a T-replete strategy highlight potential abrogation of a GVL effect by TCD. They also cite the delay in immune reconstitution associated with the use of TCD and consequent increased risk of viral, particularly CMV, infection. As with many debates in transplantation, there are no randomized data to support either approach. Substantial numbers of patients achieve durable remissions with both T-replete and TCD regimens. While there may be an increased risk of disease relapse with a T-cell-depleted regimen, much would appear to depend on the intensity of TCD, the form of post-transplant immunosuppression employed and whether adjunctive donor lymphocyte administration (DLI) in patients with mixed T-cell chimerism is mandated. Randomized clinical trials in this important area remain a priority.

Clinical management of patients undergoing stem cell transplantation

The use of PBSCs coupled with improved supportive care has decreased the 100-day TRM of autologous transplantation to 1–3% in most centres. The morbidity and mortality of allogeneic SCT has also continued to fall over the last two decades, but the 1-year TRM remains in the region of 15–25%, depending on patient age, donor type, disease status and stem cell source. This progress reflects advances in supportive care and, in patients undergoing an unrelated donor transplant, more accurate tissue typing. Despite this organ toxicity, acute GVHD and infections consequent on delayed immune reconstitution still represent major causes of morbidity and mortality in the first few months after allogeneic transplantation (Figure 35.4). In addition, chronic GVHD remains an increasing burden for patients and transplant services alike, with the increasing use of mismatched unrelated and cord blood donors. The long-term complications of allogeneic transplantation, which can significantly compromise a patient's quality of life, are also increasingly recognized and have led to the introduction of specific late-effects clinics where multidisciplinary input from transplant physicians, endocrinologists, gynaecologists, ophthalmologists and psychologists can be provided.

Practicalities of stem cell infusion and blood product support

Stem cell products are infused in the same way as other blood products except for the fact that online blood filters should not be used. Stem cell products must not be irradiated. They may either be infused once collected or, in the case of cryopreserved products, be infused immediately after being unfrozen, at the patient's bedside. The most common side-effect of the cryopreservative (dimethyl sulfoxide, DMSO) is nausea, but since it is excreted by the lungs, a garlic-like odour is also observed for 2–3 days after stem cell infusion. Damage to red blood cells releases free haemoglobin, which can precipitate acute renal failure, and thus patients should be adequately hydrated prior to stem cell infusion and for the following hours. In all patients urine output must be monitored closely.

Patients will require red cell and platelet support during the immediate post-transplant period. In allograft recipients, particularly cord blood transplants and patients receiving myelotoxic drugs such as ganciclovir, platelet transfusion may be required for a number of months after transplantation. All cellular blood products should be irradiated (25 Gy) prior to administration in order to prevent transfusion-related GVHD (see Chapter 13). This should be commenced 6 weeks prior to transplant and is

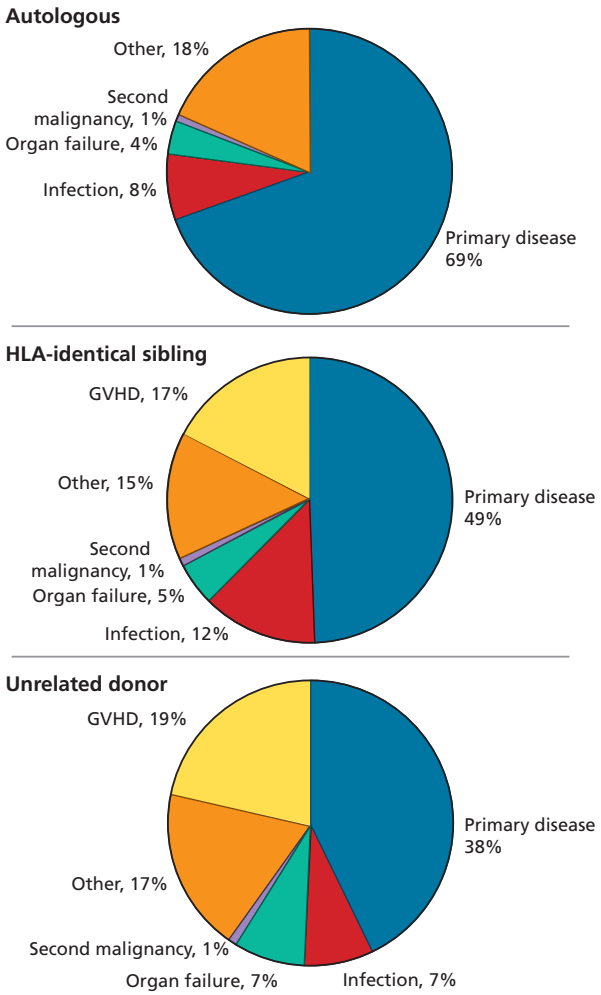


Figure 35.4 Causes of death after stem cell transplantation 2001–2006. (Courtesy of the International Bone Marrow Transplant Registry.)

continued for 6 months after an autologous transplant or indefinitely for allogeneic transplants. In patients undergoing allogeneic SCT where there is a major ABO incompatibility between donor and recipient (e.g. a group O recipient receiving group A bone marrow), the graft must be depleted of red cells prior to administration, unless PBSCs are being transplanted, in which case red blood cells are usually effectively depleted during leucapheresis, although on occasions the stem cell product may need to be administered after two days. Measurement of anti-A or anti-B titres, as appropriate, is mandated in all patients undergoing an allogeneic transplant associated with a major ABO incompatibility. The subsequent choice of blood group for platelets or red cells depends on the precise nature of the ABO incompatibility, the time from SCT and the results of blood grouping. For example, if the donor is group A and the recipient is group B, red cells with blood group O would be used until the patient began to type as blood group A, whereas group AB platelets (lacking anti-A and anti-B) would be infused. Delayed erythroid engraftment or haemolysis caused by continuing synthesis of isohaemagglutinins by host lymphocytes may occur several weeks after stem cell infusion and is associated with the presence of a positive direct antiglobulin test and antidonor red cell antibodies in serum or red cell eluates.

Complications of allogeneic SCT

Early complications (days 0–90)

Graft failure

Primary graft failure is defined as failure to achieve a neutrophil count above 0.5×10^9/L for three consecutive days within 28 days of stem cell infusion. Risk factors include transplantation of a low CD34 or mononuclear cell dose, increased donor–host HLA disparity and the deployment of TCD, particularly using *ex vivo* depletion technologies. The mortality of primary graft failure is in excess of 50% and it should be suspected in any who fail to demonstrate evidence of neutrophil engraftment by day +21. Mandatory investigations include an urgent bone marrow aspirate and trephine and chimerism studies. Patients with no morphological evidence of engraftment (as demonstrated by a good-quality trephine biopsy) and an absence of donor chimerism require a second transplant, preferably utilizing G-CSF mobilized PBSC, after administration of a second conditioning regimen incorporating fludarabine with or without low-dose TBI designed to maximize host immunosuppression. Alternatively, on the rare occasions when the liklihood of securing an additional incoculum of stem cells from the original donor is low, an urgent search for a new, usually unrelated, donor or infusion of cryopreserved autologous stem cells, can be considered.

Secondary graft failure is defined as the occurrence of sustained neutropenia and thrombocytopenia (usually for more than 7 days) after initial donor engraftment in the absence of disease relapse. It is rare in sibling allografts and occurs predominantly in recipients of mismatched or unrelated donor transplants. The aetiology of secondary graft failure is often complex, but causes that need to be considered include late graft rejection, drugs (particularly ganciclovir and on rare occasions cotrimoxazole), viral infection (CMV, HHV-6 and parvovirus B19) and, rarely, hypersplenism. Secondary graft failure is associated with significant mortality related to fungal infections and must be investigated urgently with a bone marrow aspirate and trephine, urgent chimerism studies and virological assessment to exclude parvovirus B19 or human herpesvirus (HHV)-6 infection. In patients with a hypocellular marrow and evidence of donor T-cell engraftment, a second infusion of CD34-positive selected PBSC is required.

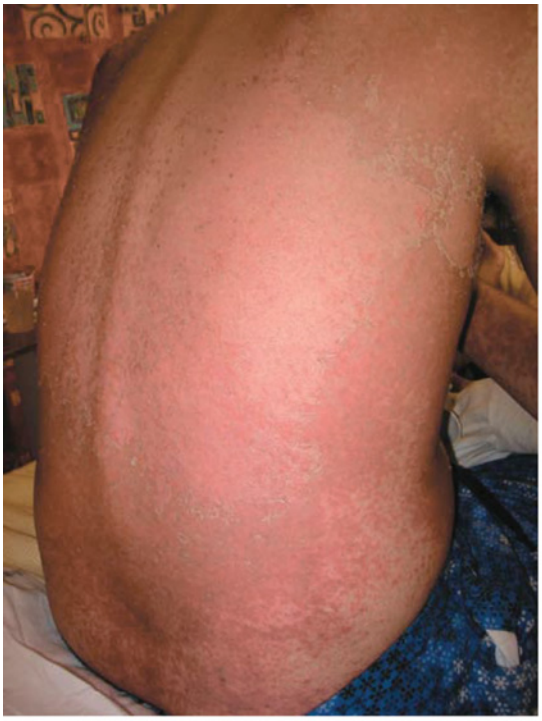
Acute graft-versus-host disease

Acute pattern GVHD usually occurs within the first 3 months post transplant, at or near the time of engraftment, and is characterized by the presence of rash, diarrhoea or abnormal

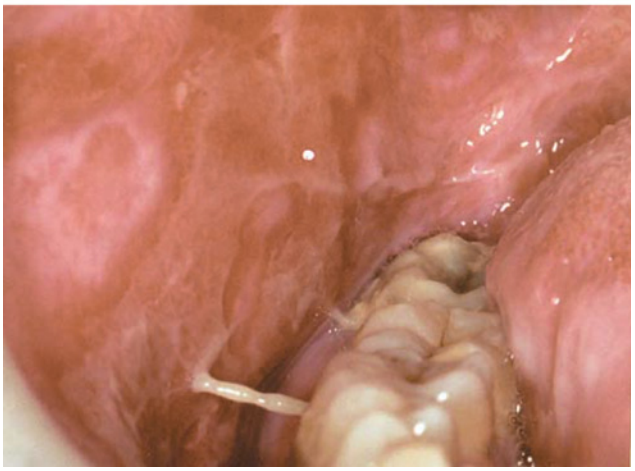
liver function tests. Depending on the conditioning regimen and GVHD prophylaxis employed, between 20% and 70% of patients will develop this disorder. Acute pattern GVHD occurring after 3 months, with no features of chronic GVHD, is also increasingly recognized in patients receiving RIC transplants, often folllowing withdrawal of immunosuppression. Risk factors for the development of acute GVHD include increased recipient age, unrelated or HLA-mismatched donor transplantation and the use of a female donor. Children have a lower risk of acute GVHD and recipients of CBT may have a lower incidence of GVHD compared with adult bone marrow transplantation with similar degrees of HLA disparity.

Skin GVHD typically presents as a maculopapular rash involving the face, neck, palms and soles, but may extend to involve the whole body (Figure 35.5). In the worst cases it progresses to erythroderma, with bullae formation and painful blistering. Histology shows apoptosis at the base of dermal crypts, dyskeratosis and evidence of lymphocytes in a perivascular distribution or adjacent to dyskeratotic keratinocytes. Gastrointestinal involvement presents with nausea, vomiting, secretory diarrhoea and/or abdominal pain. In more advanced forms, severe abdominal pain and distension associated with voluminous, occasionally bloody, diarrhoea may occur. Gastric, antral and rectal biopsies have a high diagnostic yield, with diagnostic features including the presence of apoptotic cells in the base of crypts and a lymphocytic infiltrate. A well-defined manifestation of upper gut GVHD is the development of anorexia and nausea, both of which usually resolve rapidly if treated with low-dose methylprednisolone (1 mg/kg). A rising bilirubin and raised alkaline phosphatase are the initial features of liver GVHD, which typically develops later than skin or gut GVHD. Liver histology is diagnostic and demonstrates a portal tract lymphocytic infiltration, pericholangitis and bile duct loss.

Accurate and early diagnosis of acute GVHD is essential for effective management of this potentially life-threatening disorder. Where possible, diagnostic biopsies should be taken to both confirm the presence of GVHD and assist in the exclusion of other aetiologies. It is also important to stage GVHD accurately and the criteria devised by Glucksberg (and recently updated by the International Bone Marrow Transplant Registry) are widely used (Table 35.2). This staging system is a reliable indicator of prognosis and guides the intensity of treatment required. Grade 2–4 acute GVHD should be treated with high-dose methylprednisolone (typically 2 mg/kg daily), which is tapered according to response. In the setting of limited skin GVHD and upper gut GVHD, topical or oral steroids (1 mg/kg) coupled with optimization of ciclosporin levels may be sufficient to control symptoms. Approximately 70% of patients will improve significantly with oral or intravenous corticosteroid therapy, but a number will either fail to respond or relapse when immunosuppression is tapered. Failure to respond to 5–7 days of intravenous corticosteroid therapy defines steroid-resistant acute GVHD, and these patients require the use of second-line therapies. Steroid-refractory acute GVHD has a poor prognosis, with a non-relapse mortality in excess of 70%, which is consequent predominantly on infectious complications – specifically fungal infections. Currently, robust evidence supporting efficacy of second-line treatments is lacking. Some patients with acute pattern GVHD affecting skin may respond to phototherapy. Patients with predominant gut involvement may respond to oral non-absorbable steroids or the monoclonal antibody infliximab. In other



(a)



(b)

Figure 35.5 Acute skin GVHD: (a) acute cutaneous GVHD; (b) chronic oral GVHD.

Table 35.2 Glucksberg staging of acute GVHD: (a) clinical staging; (b) clinical grading.

(a)

Stage	Skin	Liver bilirubin	Gut
+	Maculopapular rash <25% body surface	34–51 μmol/L	Diarrhoea 500–1000 mL/day or persistent nausea
++	Maculopapular rash 25–50% body surface	51–102 μmol/L	Diarrhoea 1000–1500 mL/day
+++	Generalized erythroderma	102–255 μmol/L	Diarrhoea >1500 mL/day
++++	Desquamation and bullae	>255 μmol/L	Pain ± ileus

(b)

Overall grade	Skin	Liver	Gut	Functional impairment
0 (none)	0	0	0	0
I (mild)	+ to ++	0	0	0
II (moderate)	+ to +++	+	+	+
III (severe)	++ to +++	++ to +++	++ to +++	++
IV (life-threatening)	+ to ++++	++ to ++++	++ to ++++	+++

(Source: Blume *et al*., 2004 [*Thomas' Hemopoietic Stem Cell Transplantation*, 3rd edn]. Reproduced with permission of Wiley.)

patients, treatment with other monoclonal antibodies such as daclizumab or pentostatin may be of benefit. More recently, encouraging results have been reported with the use of donor or third-party mesenchymal stromal cells, with a response rate of 60–70%. Application of this approach will require evaluation in prospective randomized studies. Patients with acute severe GVHD often require intensive supportive measures, including replacement of gastrointestinal losses, parenteral nutrition, pain control and infectious prophylaxis.

Infectious complications (see also Chapter 23)

Bacterial, fungal, protozoal and viral infections are a major cause of morbidity and mortality after allogeneic transplantation. Host factors include neutropenia, post-transplant immunosuppression and acute or chronic GVHD requiring steroid therapy. Numerous additional factors contribute to delayed immune reconstitution after an allograft and include thymic atrophy in adults, the use of TCD and a suboptimal stem cell inoculum. Infectious complications are a particular challenge after CBT. The temporal pattern of infectious complications after allogeneic SCT is shown in Figure 35.6.

Considerable progress has been made in the development of strategies to reduce the risk of infection after allogeneic SCT. All patients should be nursed in single rooms, preferably with laminar airflow or high-efficiency particulate air filtration. Evidence supports the use of triazole antifungals, such as fluconazole 400 mg daily, as an effective means of reducing *Candida* infection. Aciclovir (200–400 mg four times daily) is usually administered to prevent herpes simplex virus (HSV) re-activation. Quinolone antibiotics (e.g. ciprofloxacin 500 mg twice daily) are used by some units to reduce the risk of severe Gram-negative infections, although the evidence supporting this measure is inconclusive and practice should be guided by advice from local microbiologists concerning the prevalence and sensitivity of drug-resistant organisms. Patients should receive cotrimoxazole (480 mg twice daily three times per week) at the time of neutrophil engraftment (neutrophils $>1.0 \times 10^9$/L) to prevent *Pneumocystis jirovecii* infection. If allergic to cotrimoxazole, nebulized pentamidine (300 mg monthly) can be substituted, although it should be remembered that this provides incomplete protection from *Pneumocystis* pneumonia and for this reason some units prefer to use dapsone.

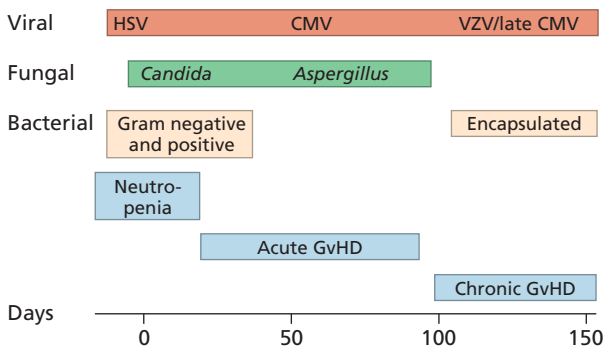


Figure 35.6 Temporal pattern of infectious complications after allogeneic stem cell transplantation.

Allogeneic transplants, particularly recipients of TBI-containing regimens, continue to be at long-term risk from infections caused by encapsulated bacteria such as *Streptococcus pneumoniae* and *Haemophilus influenzae* and require lifelong prophylaxis with penicillin (250 mg twice daily) or erythromycin (250 mg twice daily) if allergic to penicillin. Antibody titres to diseases for which childhood vaccination is performed decline after SCT. Re-vaccination to bacterial organisms which are targeted by childhood immunizations is therefore recommended, particularly in allograft recipients, and most centres commence such a programme 6–12 months after transplantation.

Herpesvirus infection: CMV, HSV, varicella zoster and HHV-6

Human CMV is a ubiquitous herpesvirus present in up to 60% of the general population, which re-activates after allogeneic SCT in seropositive recipients or those transplanted from a seropositive donor, giving rise to either asymptomatic infection or, less commonly, end-organ damage (CMV disease). Patients at the highest risk of CMV re-activation are seropositive recipients, especially those who receive TCD or unrelated donor grafts, and patients who develop GVHD requiring steroid therapy. Recent studies from the EBMT have demonstrated improved transplant outcome in CMV-positive recipients transplanted from a CMV-positive donor, and CMV status is now an important consideration in the choice of an optimal volunteer unrelated donor. CMV re-activation occurs in 40–80% of at-risk patients and until recently a substantial number of such patients developed CMV disease most commonly manifests as pneumonitis, but can rarely cause gastrointestinal ulceration, hepatitis and retinitis. Historically, CMV was the commonest cause of infectious death after allogeneic transplantation, but it is now possible to detect low levels of CMV infection after transplantation using polymerase chain reaction (PCR)-based detection of CMV, which allows the timely introduction of pre-emptive therapy with intravenous ganciclovir, a nucleoside analogue that inhibits viral thymidine kinase, or its oral prodrug valganciclovir. This has markedly reduced the incidence of early CMV disease, although late CMV infection (beyond 100 days post transplant) and on occasion disease is an increasingly recognized problem. Primary infection of seronegative patients may occur as a result of the infusion of stem cell or blood products from a CMV-positive donor, but is rare. For this reason seronegative transplant recipients should receive CMV-negative or leucodepleted blood products to limit the possibility of primary infection.

All patients at risk of CMV infection (all CMV-seropositive patients and any patient with a seropositive donor) should undergo weekly PCR or CMV antigenaemia testing from engraftment until 100 days after transplantation. Patients with CMV re-activation should be treated promptly with ganciclovir or valganciclovir. The major side-effect of ganciclovir is myelosuppression, which is especially problematic in patients transplanted using an unrelated or cord blood donor. Randomized studies have confirmed that this pre-emptive treatment strategy reduces the risk of CMV disease and death after sibling allogeneic transplantation. The use of prophylactic ganciclovir, which is administered regardless of whether there is evidence of CMV infection, does not improve outcome and is associated with significant bacterial and fungal infections consequent on high rates of myelotoxicity. For this reason a pre-emptive approach to prevention of CMV disease is generally preferred. Foscarnet, a DNA polymerase inhibitor, has less myelotoxicity than ganciclovir and is effective as part of a pre-emptive approach, although it is associated with significant nephrotoxicity.

The incidence of CMV pneumonitis after allogeneic transplantation has substantially reduced since the advent of effective screening and pre-emptive treatment strategies. It occurs in patients with evidence of CMV reactivation within the first 100 days after transplantation and typically presents with dyspnoea, hypoxaemia and pulmonary infiltrates. Ganciclovir and foscarnet are often ineffective in patients with established CMV pneumonitis. However, recent studies have demonstrated significant activity of cidofovir, which is considered in some units as first-line treatment in all patients with CMV pneumonitis. Cidofovir is nephrotoxic, but can usually be safely administered if attention is paid to adequate hydration and other nephrotoxic drugs, particularly foscarnet, are discontinued. The role of high-titre CMV immunoglobulin in the treatment of CMV pneumonitis remains unclear, although it is still widely used, if available. The effective treatment of CMV infection delays the development of an immune response to CMV and as a result late (beyond 100 days post-transplant) CMV re-activation and disease is increasingly observed. Risk factors for late CMV infection include previous CMV re-activation, lymphopenia and the presence of active GVHD.

Other members of the herpesvirus family have the potential to cause significant morbidity after allogeneic SCT. The incidence of HSV, which used to be very common in the first 30 days after SCT, has been sharply reduced by the use of prophylactic aciclovir. Re-activation of varicella zoster virus (VZV) occurs in up to 50% of at-risk patients after allogeneic SCT and typically presents as shingles with severe pain and a dermatomal vesicular eruption. Less commonly, VZV re-activation presents with atypical pain (headache or undiagnosed abdominal pain) in the absence of a rash. Prompt treatment of VZV infections with high-dose intravenous aciclovir is indicated after allogeneic SCT to prevent dissemination, but also to reduce the severity of post-herpetic neuralgia. HHV-6 is also increasingly being reported in association with a syndrome variously associated with delayed engraftment, encephalitis, classically with a psychiatric component including hallucinations, and hepatitis in patients receiving T-cell-depleted allografts or cord blood transplants.

Fungal infections

Fungal infections remain a major complication after SCT, reflecting the absence of accurate diagnostic tests and the inadequacy of current therapies. A high index of clinical suspicion is therefore required in transplant patients, particularly allografts, and most units administer systemic antifungal therapy early in the management of neutropenic fever. Risk factors for the development of fungal infection include prolonged neutropenia after SCT, the use of high-dose corticosteroids for treatment of GVHD and a history of prior fungal infection.

Effective strategies exist for the prophylaxis and treatment of infection with yeasts (*Candida* spp.), but are lacking for infection with moulds such as *Aspergillus* spp. *Candida* infections typical manifest as oral thrush and less commonly as oesophageal candidiasis. Hepatosplenic candidiasis is seen occasionally, presenting with high spiking fevers at the time of engraftment in association with abnormal liver function tests. Ultrasound or computed tomography (CT) of the liver and spleen will confirm the diagnosis. Prophylactic use of fluconazole (100–400 mg daily) has proved effective in reducing the incidence of both superficial and invasive candidiasis. Patients who develop either hepatosplenic candidiasis or candidaemia should be treated with systemic antifungals, usually liposomal amphotericin. All indwelling catheters must be removed. Emergence of fluconazole-resistant *Candida* species such as *Candida kruseii* or *Candida glabrata* is of concern, and sensitivity data must be requested in any patient with *Candida* infection.

Aspergillus infections usually present prior to or shortly after engraftment. The most common manifestation is as invasive pulmonary aspergillosis (IPA), which typically presents with an antibiotic-resistant fever, a significantly raised C-reactive protein, and abnormal chest radiography or high-resolution CT, often in the absence of respiratory symptoms. Rarely invasive *Aspergillus* infections can present with cerebral or hepatic disease. Accurate diagnosis of *Aspergillus* infections remains problematic since spores are only rarely cultured from lavage fluid or infected tissues and the sensitivity and specificity of other currently available diagnostic techniques is low. Contradictory results have been obtained using galactomannan detection assays and the initially encouraging results with PCR technology have not been confirmed by all groups. Operationally, the most helpful test in deciding whether IPA is a clinical possibility is high-resolution CT of the chest, which should be obtained in all patients with a neutropenic fever that has persisted for more than 72 hours. While the characteristic radiographic features of peripheral nodular shadows, with or without evidence of cavitation or a 'halo' sign, may take weeks to develop, the presence of any significant pulmonary infiltrate substantially increases the likelihood of *Aspergillus* infection and is an indication for the consideration of treatment doses of liposomal amphotericin or voriconazole. Importantly, the liklihood of IPA in a patient with a normal chest on high-resolution CT is low. Improved antifungal prophylaxis and treatment now make it possible to contemplate allogeneic transplantation in patients with a previous proven or suspected invasive fungal infection. Risk factors for recrudescence of fungal disease in this setting include a short period (<6 weeks) of antifungal treatment prior to transplant, the use of bone marrow or cord blood cells as opposed to PBSCs, occurrence of grade 2 or greater acute GVHD, and CMV re-activation.

Organ toxicity

Gastrointestinal toxicity

Mucositis is frequently observed in patients transplanted using an MA conditioning regimen, but is less common in patients transplanted using an RIC regimen. Symptoms of oral pain and pain on swallowing typically develop in the first few days after stem cell infusion and peak approximately 8 days post transplant. The severity of mucositis is closely correlated with the intensity of the conditioning regimen, the use of higher doses of fractionated TBI and the inclusion of methotrexate as a component of the GVHD prophylaxis regimen. Patients with severe mucositis should receive adequate (often opiate) analgesia and be monitored for evidence of airway obstruction. In patients who develop symptoms of severe oesophagitis, the possibility of superadded infection with *Candida albicans* or HSV must be considered, the latter typically being associated with intractable vomiting.

Diarrhoea and/or abdominal pain are common manifestations of conditioning toxicity. Alternative causes of diarrhoea in the first 4 weeks after transplantation include enteritis due to *Clostridium difficile*, rotavirus or CMV, acute GVHD and pancreatitis. VZV re-activation can occasionally present with severe abdominal pain, often in association with markedly deranged liver function tests sometimes days in advance of, or even in the absence of, the development of classic vesicles.

Liver toxicity

Abnormalities in liver function tests are commonly observed after allogeneic SCT and causes include VOD, drugs, infectious complications, cholelithiasis, acute GVHD and transfusional iron overload (Table 35.3). VOD is a clinical syndrome characterized by a triad of hyperbilirubinaemia (bilirubin >34 μmol/L), weight gain (>5% over baseline) and painful hepatomegaly. It is associated with evidence of damage to sinusoidal endothelial cells and hepatocytes and subsequent damage to the central veins in zone 3 of the hepatic acinus. In severe cases hepatic venular occlusion and widespread zonal disruption may lead to portal hypertension, hepatorenal syndrome, multiorgan failure and death. Risk factors for the development of VOD include the use of conditioning regimens containing busulfan or higher doses of TBI, pretransplant abnormalities of liver function tests, previous abdominal irradiation and recent exposure to the anti-CD33 antibody gemtuzumab ozogamicin (Mylotarg). A

Table 35.3 Causes of abnormal liver function tests after allogeneic SCT.

Precipitant	Clinical presentation
Veno-occlusive disease	Hyperbilirubinaemia associated with weight gain, ascites and painful hepatomegaly
Drugs	Ciclosporin (hyperbilirubinaemia), azole antifungals (raised AST)
Haemolysis	Fall in Hb associated with unconjugated hyperbilirubinaemia and positive Coombs test
Biliary obstruction	Dilated biliary tree associated with gallbladder 'sludge' or cholelithiasis
Infection	Viral hepatitis, fungal infection (disseminated aspergillosis or candidiasis), cholangitis lenta (progressive development of obstructive picture occurring days after an episode of sepsis)

diagnosis of VOD is most commonly made on clinical criteria. Hepatic Doppler studies demonstrating evidence of reversal of portal flow support the diagnosis. Definitive diagnosis requires transjugular venous liver biopsy, which has significant morbidity in the early post-transplant period and is therefore often avoided. It is important to exclude other causes of hyperbilirubinaemia, particularly ciclosporin toxicity (which can present with a very similar clinical picture), haemolysis (typically consequent on donor–recipient ABO mismatch or transplantation-associated microangiopathic haemolytic anaemia) and the hepatitis of sepsis (cholangitis lenta). Management is supportive, consisting of careful fluid balance, the judicious use of diuretics and, where necessary, haemofiltration. There is now convincing evidence that defibrotide can effectively treat severe VOD. Because of the clinical similarities between VOD and ciclosporin toxicity, it is wise to discontinue ciclosporin for at least 48 hours in any patient in whom a diagnosis of VOD is suspected.

Renal toxicity

The importance of the daily monitoring of weight, fluid balance and renal function in the effective management of patients after SCT cannot be over-estimated. Impairment of renal function is frequently observed after allogeneic SCT unless careful attention is paid to fluid balance and the nephrotoxic potential of drugs commonly used during SCT, especially ciclosporin, amphotericin, aminoglycosides and loop diuretics. Ciclosporin-related renal toxicity is usually easily reversible by temporary omission and dose reduction. Occasionally, ciclosporin toxicity manifests itself as a microangiopathic haemolytic anaemia (MAHA), with features of thrombotic thrombocytopenic purpura/haemolytic–uraemic syndrome. Withdrawal of the drug is mandatory. There is no convincing evidence that plasmapheresis is beneficial. Introduction of an alternative immunosuppressant (e.g. mycophenolate mofetil) in conjunction with oral corticosteroid is required in patients who develop ciclosporin-induced MAHA. It is important to monitor for late renal toxicity in patients receiving long-term ciclosporin.

Pulmonary infections and non-infectious complications

A range of pulmonary infections occur after allogeneic SCT. In the first month after transplantation bacterial and fungal pneumonias are common. CMV, respiratory syncytial virus, influenza and parainfluenza are important causes of pneumonitis, and typically occur in the first 90 days after transplantation. *Pneumocystis* pneumonia and *Toxoplasma* infection are still seen occasionally in the first few months after transplantation in patients who do not receive, or who are not compliant with, cotrimoxazole prophylaxis.

Non-infectious pulmonary complications occurring after allogeneic transplantation can be reversible, but demand prompt diagnosis if treatment is to be effective. Pulmonary oedema consequent on either increased capillary hydrostatic pressure caused by fluid overload or increased capillary permeability due to irradiation or sepsis is frequently seen in the post-transplant period. Idiopathic pneumonia syndrome, defined as diffuse lung injury occurring after SCT for which no infectious or non-infectious aetiology can be identified, typically occurs 30–50 days after transplantation. The classic presentation includes dyspnoea, non-productive cough, hypoxaemia and non-lobar infiltrates on chest radiography and can progress rapidly to acute respiratory distress syndrome. Treatment is supportive, but frequently unsatisfactory, and steroids have little effect on outcome. Diffuse alveolar haemorrhage is seen predominantly in patients undergoing autologous SCT, but can also be seen in allogeneic recipients. This complication usually occurs within the first 2–3 weeks of transplantation and presents with dyspnoea, non-productive cough and hypoxaemia. Radiographic changes include interstitial or alveolar shadowing. Definitive diagnosis requires bronchoscopy, which shows fresh blood on repeated lavage. Early recognition of this disorder is essential since early intervention with high-dose steroids may significantly improve survival.

Post-transplant lymphoproliferative disease

Post-transplant lymphoproliferative disease (PTLD) includes a spectrum of EBV-driven B-cell hyperproliferative states that range from polyclonal benign proliferations to life-threatening neoplastic disease. Most cases of PTLD involve EBV-seropositive donors and present with lymphadenopathy and fever. In

contrast to the PTLDs that develop following solid organ transplantation, the majority of cases after SCT are of donor origin and arise as a result of inadequate T-cell control of proliferation of EBV-infected B cells. Risk factors for the development of PTLD include TCD, particularly the use of ATG and alemtuzumab, and increased HLA disparity. Weekly PCR quantitation of EBV viraemia is indicated in patients undergoing allogeneic transplantation that incorporates ATG or alemtuzumab and in all CBT procedures until 100 days post-transplant. Patients with a rising EBV load may be managed by a reduction in immunosuppression or treated with pre-emptive anti-CD20 monoclonal antibody (rituximab), although the triggers for intervention differ between centres. In patients with a documented PTLD, early treatment with rituximab and, where possible, cessation of immunosuppressive therapy are the key to treatment. DLI using EBV-specific cytotoxic T lymphocytes has also been used with effect. Chemotherapy in patients with rituximab-resistant PTLD is rarely of value.

Intensive care support

The likelihood of developing organ failure requiring admission to the intensive therapy unit (ITU) after SCT varies widely according to stem cell source (autologous versus allogeneic), the intensity of the conditioning regimen and the presence of pre-transplant comorbidities. While recent studies suggest a reduction in the proportion of transplant recipients requiring ITU admission, it remains the case that at least 10% of allografted patients, and in some institutions significantly more, will require ITU admission at some stage during their treatment.

Analysis of outcome in patients admitted to ITU has identified factors determining survival at varying stages after transplantation, allowing rational decisions about the likelihood of benefit of sustained intensive care in patients with multiorgan failure. Further sophistication in defining who will benefit from admission to ITU has been provided by the application of scoring systems. These include the Acute Physiology and Chronic Health Evaluation (APACHE) II and III mortality prediction models and Simplified Acute Physiology Score (SAPS) prognostic systems. Application of these scores prior to admission to ITU and, where appropriate, after a 72-hour period on ITU is helpful in identifying patients who will benefit from institution or continuation of intensive support. A number of studies have indicated that the development of progressive organ failure or the presence of multiorgan failure 72 hours after admission to ITU predicts an extremely poor outcome. The careful integration of these data into discussion with relatives and, where possible, patients can spare many the indignity of prolonged and futile intervention.

Late complications

As the results of allogeneic SCT have improved, its long-term complications have come to be better recognized. Chronic GVHD, secondary malignancies, and growth and fertility disorders represent an increasingly important cause of morbidity in patients who have been cured of their underlying disease.

Chronic GVHD

Chronic GVHD refers to a complex syndrome occurring more than 3 months following allogeneic transplantation and is its commonest long-term complication. There is increasing recognition that the onset of chronic GVHD onset can be variable and also that patients may present features of both acute and chronic GVHD (often referred to as overlap syndrome). Registry data show that chronic GVHD occurs in approximately one-third of patients undergoing T-replete transplants from HLA-identical siblings, rising to two-thirds of patients receiving T-replete grafts from unrelated donors. Risk factors for the development of chronic GVHD include increased recipient age, the use of PBSCs as opposed to bone marrow as the stem cell source, transplantation from an unrelated or HLA-mismatched donor, the use of a T-replete stem cell inoculum and the presence of prior acute GVHD. Chronic GVHD may develop directly from acute GVHD (progressive), after the resolution of an episode of acute GVHD, or *de novo* in patients with no history of chronic GVHD. The clinical manifestations are characterized by features of both immunodeficiency and 'autoimmunity' (Figure 35.7). Previous grading systems for chronic GVHD have classified its severity as limited or extensive (Table 35.4), although the utility of this staging system has been questioned. Patients with progressive-type onset, extensive skin involvement, thrombocytopenia or bronchiolitis obliterans have a particularly poor prognosis. A new scoring system based on the number of organs involved and the severity within each affected organ is likely to be adopted, particularly in trials of novel therapies.

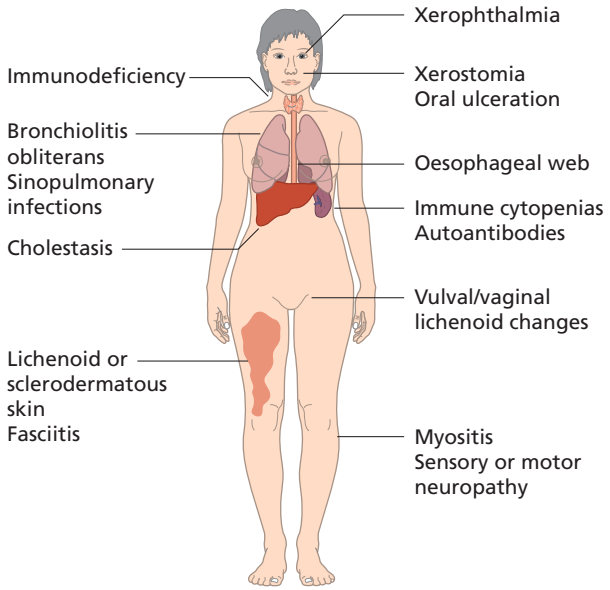


Figure 35.7 Clinical manifestations of chronic GVHD.

Table 35.4 Classification of chronic GVHD.

Limited chronic GVHD
Either or both
Localized skin involvement
Hepatic dysfunction as a result of chronic GVHD
Extensive chronic GVHD
Either
Generalized skin involvement
or
Localized skin involvement and/or hepatic dysfunction as a result of chronic GVHD
plus
Liver histology showing chronic aggressive hepatitis bridging necrosis or cirrhosis
or
Involvement of eye (Schirmer test with <5 mm wetting)
or
Involvement of minor salivary glands or oral mucosa demonstrated on labial biopsy
or
Involvement of any other target organ

(Source: Blume *et al.*, 2004 [*Thomas' Hemopoietic Stem Cell Transplantation*, 3rd edn]. Reproduced with permission of Wiley.)

Current treatment options are unsatisfactory and in the first instance involve immunosuppression with ciclosporin or prednisolone prior to the institution of a gradual reduction in responding patients. In patients, who fail to respond or who show steroid dependency, second-line therapies including extracorporeal photopheresis, sirolimus and rituximab can be effective. Extracorporeal photopheresis is particularly beneficial in patients with sclerodermatous involvement but appears to be less effective in patients with visceral disease. Rituximab has been employed successfully in some patients, with responses observed particularly in patients with musculoskeletal involvement. Other treatments suggested include non-absorbable steroids for those with gut involvement, tacrolimus (FK506), low-dose interleukin-2, thalidomide and imatinib. Definitive prospective randomized data supporting any of these strategies are lacking.

Secondary malignancies

The incidence of solid tumours in recipients of allogeneic SCT is increased compared with control populations and includes a higher rate of skin or buccal cavity squamous cell carcinoma and melanoma. In the largest study to date, the cumulative incidence of all solid tumours was 2.2% at 10 years and 6.7% at 15 years, although the latter figure may represent an overestimate. Risk factors for the development of secondary malignancies include the use of TBI in the conditioning regimen and the presence of chronic GVHD.

Growth, puberty and fertility

Retarded growth is seen predominantly in children who receive transplants at a younger age (<10 years) and in those who have received irradiation, particularly cranial irradiation. Growth retardation is multifactorial and although growth hormone levels may be diminished as a consequence of hypothalamic or pituitary irradiation, this does not entirely account for the observed reduction in height. Gonadal failure (both testicular and ovarian) is a common consequence of myeloablative conditioning regimens, particularly those that contain TBI or busulfan. While prepubertal females receiving cyclophosphamide alone (as for transplantation in severe aplastic anaemia) have a high probability of experiencing a normal puberty, most receiving other preparative regimens will fail to regain normal ovarian function and will require sex hormone replacement therapy for the induction of puberty. Thereafter sex hormone replacement is indicated to maintain the menstrual cycle and normal bone turnover/mineralization, but the optimal duration of therapy and the potential long-term risks of hormone replacement therapy in this setting are unknown. Although about 10% of women who receive TBI-containing regimens may have some recovery of gonadal function, the overall incidence of pregnancy is very low. The advent of RIC raises the prospect that some women may remain fertile after transplantation and in this case prior therapies and age will be additional factors determining fertility. In males, cryopreservation of sperm should be performed at diagnosis since chemotherapy and transplantation both result in a high rate of male infertility.

Complications of autologous SCT

The morbidity and mortality of autologous SCT is much lower than that observed in allogeneic transplantation. The major toxicities relate to immediate or delayed organ failure caused by the conditioning regimen or complications consequent on neutropenia or thrombocytopenia (Table 35.5). The commonest infectious complications are bacterial and fungal and occur in the main in the first 28 days after transplantation. Late complications are rare, but include pulmonary fibrosis in patients who have received busulfan and carmustine, and gonadal failure in recipients of TBI or busulfan-containing regimens. Female recipients of BEAM chemotherapy may recover some gonadal function. The development of myelodysplasia is a well-recognized late complication, particularly in patients transplanted using a TBI-based regimen. It appears that prior chemotherapy is an important factor in the development of MDS or AML since the leukaemia-associated cytogenetic abnormality, most commonly involving chromosome 5 or 7, is frequently detected in the infused stem cell inoculum.

Table 35.5 Immediate complications of autologous SCT.

Organ	Complications
Lung	Diffuse alveolar haemorrhage, interstitial pneumonitis especially in recipients of bleomycin/carmustine- and TBI-containing regimens
Gastrointestinal	Mucositis, nausea, diarrhoea
Hepatic	Asymptomatic elevation of bilirubin or AST, veno-occlusive disease (especially recipients of busulfan-containing regimens)
Renal	Prerenal acute renal failure, renal toxicity (nephrotoxic antibiotics), interstitial nephritis
Cardiac	Arrhythmias, cardiac failure (cyclophosphamide)

Patient factors determining outcome after allogeneic stem cell transplantation

The major causes of treatment failure after allogeneic transplantation consist of transplant-related mortality, principally organ toxicity, infectious complications GVHD and disease relapse (see Figure 35.4). The risk of disease relapse is determined by tumour biology and the intensity of the conditioning regimen. However, although relapse risk is increased in patients transplanted using an RIC rather than an MA conditioning regimen, this is compensated for by an attendant reduction in TRM. The last decade has seen a considerable advance in the development of scoring systems that permit the relatively accurate prediction of TRM after allogeneic transplantation. Whilst the degree of patient–donor HLA mismatch and stem cell source are important determinants of TRM, it is becoming increasingly clear that patient comorbidities are a critical determinant of outcome. The haemopoietic cell transplant comorbidity index (HCT-CI) incorporates a range of patient-specific variables and predicts outcome after both MA and RIC allografts. The advent of RIC regimens and alternative stem cell sources has increased the importance of developing a robust estimate of TRM. These important studies require validation in the context of specific conditioning regimens, particularly TIC regimens utilizing alternative donors or TCD, but offer the prospect of developing a more rational basis of selecting candidates for allogeneic transplantation. Thus, decisions concerning whether to proceed to an allogeneic transplant are increasingly viewed as a patient-specific appraisal of the predicted reduction in relapse risk if patients are allografted compared with the predicted TRM on the basis of the calculated HCT-CI, stem cell source and donor–patient HLA disparity.

Indications for transplantation

SCT is now an essential constituent of the treatment strategy for many haematological malignancies. In the case of autologous transplantation, current practice has been informed by the results of randomized controlled trials and registry studies comparing the outcome of transplanted patients with those who have received standard treatment. Improvements in the efficacy of first- and second-line treatments (for example, the use of rituximab, brentuximab and bendamustine in lymphoma and the immune-modifying drugs in myeloma) mean that the role of autologous SCT will continue to be re-defined. Small molecular inhibitor drugs are also showing great promise in the treatment of indolent lymphoproliferative disorders and will likely change current treatment paradigms. It remains very challenging to perform randomized trials in the setting of allogeneic transplantation. Obstacles not only include randomization biases introduced because of the fixed perceptions of either physicians or patients, but also the very real problem that the allograft arm may include a smaller proportion of patients with high-risk disease, given the propensity of such patients to relapse before they reach transplantation. The only effective approach that allows these biases to be removed has been the use of a 'donor versus no-donor' analysis, which exploits the availability of an HLA-identical sibling donor as a form of biological randomization. However, such an approach is hard to apply in rarer diseases and when most patients potentially have access to an alternative stem cell donor. Consequently, despite the development of novel statistical methodologes such as the Mantel–Byer analysis, many of the recommendations concerning the role of allogeneic transplantation in specific diseases will continue to be based on registry data or institutional Phase II studies.

In order to provide a robust framework for SCT, the EBMT publishes a regularly updated summary of indications for transplantation (Table 35.6), allowing a consistent approach to be adopted. This categorizes specific diseases according to whether autologous or allogeneic. SCT is accepted as standard therapy, is a promising treatment modality requiring further evaluation, or is not indicated.

Factors determining the choice of an allogeneic stem cell donor: the donor algorithm

An HLA-identical sibling donor can be identified in about one-third of patients eligible for an allogeneic transplant. In older patients, where potential sibling matches may have medical conditions precluding stem cell donation, the likelihood of locating a suitable matched donor is significantly lower. Three potential donor options now exist for adults lacking a suitable HLA-identical sibling donor: (i) an unrelated adult donor, (ii) a cord blood donor and (iii) a haploidentical related donor.

Table 35.6 Indications for SCT.

		Allogeneic SCT	
	Autologous SCT	**Sibling transplant**	**VUD transplant**
AML first CR			
Good-risk cytogenetics	NR	NR	NR
Standard-risk cytogenetics	R	R	NR
Poor-risk cytogenetics	R	R	R
AML second CR	R	R	R
ALL (normal cytogenetics) first CR	D	R	NR
ALL t(9;22) first CR	R	R	R
ALL second CR	R	R	R
CML first CP	NR	R (resistant to TKIs)	R (resistant to TKIs)
MDS	NR	R	R
Myeloma	R	R	D
Hodgkin lymphoma first CR	NR	NR	NR
Hodgkin lymphoma relapsed	R	R	D
NHL DLBCL first CR	D	D	D
NHL DLBCL relapse	R	R	D
NHL follicular	D	R	D
Aplastic anaemia	NR	R	D
Haemoglobinopathies	NR	R	D

ALL, acute lymphoblastic leukaemia; AML, acute myeloid leukaemia; CML, chronic myeloid leukaemia; CR, complete remission;D, developmental; DLBCL, diffuse large B-cell lymphoma; MDS, myelodysplastic syndrome; NHL, non-Hodgkin lymphoma; NR, not recommended; R, recommended; VUD, volunteer unrelated donor.
(Source: Ljungman *et al*., 2010 [11]. Reproduced with permission of Nature Publishing.)

Thanks to the rapid expansion of unrelated donor registries, over 20 million unrelated donors are now registered worldwide. As a result it is possible to identify a suitable donor in more than 75% of whites of European descent, although this figure is much lower in other ethnic groups. Preferably the donor should be matched for HLA A, B, C; DRB1 and DQB1 (10/10). Initial results with unrelated donor transplantation were associated with a high TRM and an increased incidence of severe GVHD. In the last decade there has been a marked improvement in outcome of unrelated donor SCT because of the availability of molecular typing at class I and II loci of the HLA complex. As yet we lack sufficient information to allow us to identify whether specific allelic mismatches are of particular clinical significance. Although the recent growth in the size of unrelated donor registries internationally has increased the availability of volunteer unrelated donors, their unrepresentative ethnic composition limits the availability of donors for ethnic minorities. In addition, lack of availability of a potential donor (for example through being lost to contact) remains a significant problem. Consequently, the timely identification of a suitable alternative donor remains a challenge for a significant number of patients, particularly those with high-risk acute leukaemia who require urgent transplantation.

UCB has long been known to be a rich source of haemopoietic progenitors and recent results have demonstrated that high rates of durable engraftment can be achieved in adults, comparable to those achieved in children, providing the cell dose is optimized. Importantly, it appears that greater degrees of HLA mismatch can be tolerated using UCB than with a volunteer unrelated donor. As a result it is possible to transplant a cord blood unit mismatched at two or fewer HLA-A, -B or -DRB1 antigens, as long as an adequate cell dose is available. This, coupled with the speed of procurement of UCB units, has resulted in the emergence of cord blood as an increasingly important alternative stem cell source in both adults and children. Although initial results using UCB in adults were compromised by delayed engraftment and a high rate of primary graft failure, these problems have largely been overcome by ensuring the use of a high total nucleated cell dose or two cord blood units. As a result, UCB is increasingly recognized as an important stem cell source in adults with high-risk leukaemia lacking a suitable unrelated donor, providing strict matching and cell dose requirements are adhered to, although its precise role needs to be defined in prospective clinical trials.

Until recently, haploidentical transplants were associated with unacceptable rates of graft failure and GVHD because of the

presence of HLA disparity at all major HLA antigens. However, the ability to maximize the stem cell dose, using G-CSF-mobilized progenitors, coupled with the development of efficient methods of TCD, has significantly improved outcome, and it is now an important treatment option in children lacking an unrelated donor. In adults, recent studies from groups in Baltimore and Pennsylvania have shown the feasibility of T-cell-replete haploidentical SCT followed by alloreactive T-cell purging *in vivo* using a strategy of post-transplant cyclophosphamide; this approach removes alloreactive T cells that are dividing rapidly in the first few days following transplant, but leaves intact the non-alloreactive repertoire containing cells that can respond to microbial and viral pathogens. Again, the role of this, currently experimental, approach will need to be addressed in clinical trials.

Management of disease relapse

Relapse after autologous SCT is caused by either resistance of the underlying haematological malignancy to before it becomes adopted as standard practice. The conditioning regimen or infusion of contaminating tumour cells at the time of transplantation. Recent data demonstrates that patients with myeloma in whom the duration of response to the first transplant was greater than 18 months benefit from a second autograft, although whether this is superior to the outcome of treatment with one of the newer agents such as bortezomib or lenalidomide remains the subject of ongoing studies. Allogeneic transplantation using an RIC regimen is an important treatment option in patients with Hodgkin or non-Hodgkin lymphoma who have relapsed after an autologous transplant.

Disease relapse remains the commonest cause of treatment failure after allogeneic transplantation and is caused by either resistance of tumour cells to the conditioning regimen or failure of the donor immune system to eradicate residual malignant haemopoiesis. Possible mechanisms for tumour cells evading a GVL effect include downregulation of HLA class I or II expression absence of costimulatory molecules on leukaemic blasts or the location of residual cells in privileged sites such as the CNS or testes. While it is possible to reduce the risk of disease relapse after an MA allograft by increasing the intensity of the preparative regimen, this approach has not yet translated into improved overall survival because of the increased TRM associated with a more intense conditioning regimen. Consequently, optimization of the GVL effect represents an important mechanism by which the antileukaemic activity of an MA allograft can be augmented. In routine clinical practice the options for manipulating a GVL effect are, however, limited and largely are restricted to modifying the intensity of post-transplant immunosuppression, for example through the institution of a rapid ciclosporin taper in patients at a high risk of disease relapse. While the prophylactic use of DLI in patients deemed at high risk of relapse is an attractive option, such an approach is difficult to effect in practice because many patients will relapse early while still on immunosuppression and before DLI can be administered. In those patients in whom a course of prophylactic DLI can be commenced before disease relapse, the risk of severe GVHD associated with its use in the first 12 months after transplantation represents a significant problem.

In patients transplanted using an RIC regimen, there is limited room for increasing the intensity of the conditioning regimen without increasing transplant toxicity, although the sequential FLAMSA regimens represent an important model of how this might be achieved in future. In common with MA allografts, a GVL effect can be optimised by limiting post-transplant immunosuppression and both the dose and duration of CsA or tacrolimus post transplant represent important and manipulable variables that require prospective evaluation in clinical trials. There remains considerable debate as to whether the presence of mixed T-cell chimerism after an RIC allograft, which is commonly observed in recipients of alemtuzumab-based allografts, is associated with an increased risk of disease relapse. On the basis that it is an indicator of bidirectional tolerance and therefore permissive for disease relapse, a number of units administer prophylactic DLI in patients who have been transplanted using an RIC regimen in whom mixed T-cell chimerism persists after the withdrawal of post-transplant immunosuppression. Definitive data concerning the effectiveness of this approach in improving outcome are not yet available. Alternative strategies with the potential to reduce the risk of disease relapse include the adminstration post-transplant of disease-modifying agents such as tyrosine kinase inhibitors in the setting of AML associated with an FLT3 ITD mutation or Ph+ leukaemias or lenalidomide in patients allografted for multiple myeloma. Post-transplant azacitidine has also been investigated in patients allografted for high-risk AML, given its ability *in vitro* to upregulate tumour antigens and potentially augment a GVL response.

The management of patients who relapse after allogeneic transplantation is complex and in many patients should probably be palliative. This particularly applies to patients transplanted for acute leukaemia who relapse within the first 12 months post transplant, or patients with a history of severe acute or chronic extensive GVHD. However, there is now a significant population of patients, who relapse more than 12–18 months post transplantation and have no active GVHD, in whom DLI or a second transplant should be considered. Two factors determine the use of DLI in clinical practice. Firstly, the ability of DLI to deliver durable remissions is highly dependent on the underlying haematological malignancy. Thus, up to 80% of patients who have relapsed after an allograft for CML will achieve a durable molecular remission with DLI, in contrast to patients transplanted for acute leukaemia or myeloma, where response rates are much lower. Secondly, prior cytoreduction, and acquisition of a CR, appears to increase the likelihood of patients with relapsed acute leukaemia achieving

a durable response to DLI. The major complications of DLI are myelosuppression and GVHD. Myelosuppression is most commonly observed in patients with overt haematological relapse and may require treatment with G-CSF and occasionally transfusion of donor CD34-positive selected PBSCs. In early studies, DLI was administered as a single bulk infusion and was associated with a significant risk of severe GVHD. However, the recognition that use of an escalating schedule of administration substantially reduces the risk of severe GVHD without compromising its efficacy has markedly reduced the toxicity of DLI. The explanation of why early DLI is associated with such a high risk of severe GVHD is unclear, but may relate to more rapid expansion of donor alloreactive T cells in the profoundly lymphopenic environment present in the first few months after transplantation. An alternative approach in patients who relapse late (more than 12–18 months post transplant) and in whom there is no evidence of active GVHD is a second allograft using an RIC regimen. Although associated with an increased TRM, a number of studies now show a significant chance of long-term disease-free survival after a second allograft in a significant proportion of such patients.

Future developments in stem cell transplantation

Autologous SCT

Relapse remains the commonest cause of treatment failure in patients after autografting. Most of the drugs contained in preparative regimens are already administered at close to the limit of extramedullary toxicity and there is therefore little room for further dose escalation. Consequently, efforts to reduce relapse rate will depend on either altering the mode of delivery of currently used agents or the development of new forms of chemotherapy or radiotherapy. The recent advent of an intravenous preparation of busulfan is an example of how improved drug delivery can improve the pharmacokinetic profile of an antitumour agent, thereby reducing toxicity and increasing activity. Many haematological malignancies are highly radiosensitive and there is interest in the possibility of increasing the effective dose of irradiation delivered to the marrow without increasing organ toxicity using radiolabelled immunoconjugates. Early-phase clinical studies using α (^{131}I) or β (^{90}Y) emitters conjugated to haemopoietic antigens, such as CD45 or CD66, demonstrate that such an approach is feasible, and Phase II studies in myeloma and lymphoma are ongoing. Alternatively, drugs that disrupt tumour cell binding to the bone marrow microenvironment, such as the CXCR4 antagonist plerixafor, may overcome cytoadhesion-mediated drug resistance and increase the activity of drugs such as melphalan. Although contamination of PBSC grafts by malignant cells can undoubtedly contribute to disease relapse after an autologous transplant, there is no evidence that purging of stem cell collections reduces relapse risk.

Allogeneic SCT

The major advances in allogeneic transplantation in the past decade have been the introduction of RIC regimens and the increased availability of alternative donors. As a result, allogeneic transplantation is increasingly employed as a treatment option in patients with haematological malignancies. However, the challenges of allografting remain substantially unchanged: GVHD, infection and disease relapse. Since all three, to a greater or lesser extent, are consequent on a failure to regulate or harness the donor immune response appropriately, advances in their management are rarely made alone, but are interlinked.

The introduction of novel immunosuppressive agents remains challenging, as few agents identified in preclinical studies are successfully translated into clinical practice. However, newer agents, including vorinostat (a histone deacetylase inhibitor), bortezomib (a proteasome inhibitor) and an antagonist of CCR5 (a chemokine receptor implicated in GVHD) have also shown promising results in early-phase trials in the setting of both GVHD prophylaxis and therapy. Refractory chronic GVHD may respond to imatinib (a tyrosine kinase inhibitor) or low-dose interleukin-2, although further prospective studies will be required to determine their role. In the setting of TCD, where the risk of GVHD is low, the emphasis is on minimizing the impairment in immune reconstitution and increased infectious complications caused by depletion of donor T cells. One promising approach by which this might be achieved is through the use of 'add backs' of antigen-specific T cells recognizing common post-transplant viral pathogens such as CMV, EBV or adenovirus. Such cells can be generated either by coculturing donor lymphocytes with APCs primed with candidate antigens, followed by expansion, or by early selection on the basis of interferon-γ secretion (a technique termed 'IFN-γ capture'). Alternatively, antigen-specific cells might be selected directly from peripheral blood using immunomagnetic techniques based on labelling of antigen-specific cells using HLA class I–peptide multimers. In situations where a donor has no relevant immunity (e.g. where a CMV-seropositive patient is transplanted from a seronegative donor), gene transfer of a TCR specific for the relevant viral antigen is under investigation. In order to improve coverage against multiple pathogens, infusion of donor lymphocytes from which alloreactive T cells have been depleted may hasten immune reconstitution without incurring the risk of GVHD normally associated with DLI. Such an 'allodepletion' strategy can be achieved by either directly removing donor CD8$^+$ T cells from the infusion, given that CD8 cells are postulated to be the major effectors of GVHD, or alloreactive T cells can be depleted from the graft by incubating cells *ex vivo* with recipient APCs. Activated T cells capable of mediating GVHD can be identified by their expression of CD25 or other activation markers and

removed using magnetic selection techniques or an immunotoxin. Preclinical models have suggested that transfer of memory phenotype T cells may also permit more rapid immune recovery without inducing GVHD, although whether such cells are truly representative of human memory T cells remains to be determined. There is also substantial interest in the feasibility of transferring naturally occurring or *ex vivo* expanded $CD4^+FoxP3^+$ or IL-10-producing regulatory T cells as a means of preventing conventional T cells from inducing GVHD. The best approach for isolating and expanding these regulatory populations remains unknown, but the results from early trials have shown promise.

From first principles, reducing the risk of relapse can be achieved by increasing the intensity of the conditioning regimen (ideally without increasing transplant toxicity), administering antileukaemic therapies in the post-transplant period or by optimizing the GVL effect. Incorporating radioimmunotherapy (e.g. with antibodies directed at antigens specific for haemopoietic cells or malignant cells) into the preparative regimen shows promise in the setting of allogeneic as well as autologous SCT and may prove of value in patients with advanced leukaemia, where relapse is the most important cause of treatment failure. Alternatively, patients judged to be at high risk of relapse may benefit from administration of adjunctive therapies after transplantation (e.g. tyrosine kinase inhibitors or demethylating agents), reducing the residual disease burden while a GVL response is established. Donor T cells that have been expanded selectively to recognize tumour-associated antigens or lineage-resticted minor H antigens (e.g. by expanding donor cells in the presence of peptide mixes) may prove effective. Perhaps the most exciting approach under investigation is the gene engineering of T cells to express TCRs or chimeric antigen receptors specific for tumour-associated antigens. TCR gene therapy has the advantage of targeting tumour antigens derived from intracellular proteins, but is limited by the need to use receptors that are restricted to particular HLA types, and also by the potential for cross-reactivity with other antigens. Chimeric antigen receptors (CARs) have the advantage that they are not HLA-restricted, but the disadvantage that they only recognize surface antigens. Initial data from Phase I clinical trials are very encouraging, even in patients with advanced, refractory disease, and it is likely that this strategy will emerge as useful adjunct to allo-SCT.

Selected bibliography

Appelbaum FR (2007) Hematopoietic-cell transplantation at 50. *New England Journal of Medicine* **357**(15): 1472–5.

Appelbaum FR (2009) Optimising the conditioning regimen for acute myeloid leukaemia. *Best Practice and Research in Clinical Haematology* **22**(4): 543–50.

Bacigalupo A, Ballen K, Rizzo D *et al*. (2009) Defining the intensity of conditioning regimens: working definitions. *Biology of Blood and Marrow Transplantation* **15**(12): 1628–33.

Blaise D, Tabrizi R, Boher JM *et al*. (2013) Randomized study of 2 reduced-intensity conditioning strategies for human leukocyte antigen-matched, related allogeneic peripheral blood stem cell transplantation: prospective clinical and socioeconomic evaluation. *Cancer* **119**(3): 602–11.

Boeckh M. (2011) Complications, diagnosis, management, and prevention of CMV infections: current and future. *Hematology American Society Hematology Education Program* **2011**:305–9.

Cornelissen JJ, Gratwohl A, Schlenk RF *et al*. (2012) The European LeukemiaNet AML Working Party consensus statement on allogeneic HSCT for patients with AML in remission: an integrated-risk adapted approach. *Nature Reviews Clinical Oncology* **9**(10): 579–90.

Ferrara JL, Levine JE, Reddy P, Holler E (2009) Graft-versus-host disease. *Lancet* **373**(9674): 1550–61.

Finke J, Bethge WA, Schmoor C *et al*. (2009) Standard graft-versus-host disease prophylaxis with or without anti-T-cell globulin in haematopoietic cell transplantation from matched unrelated donors: a randomised, open-label, multicentre phase 3 trial. *Lancet Oncology* **10**(9): 855–64.

Ljungman P, Bregni M, Brune M *et al*. (2010) Allogeneic and autologous transplantation for haematological diseases, solid tumours and immune disorders: current practice in Europe 2009. *Bone Marrow Transplantation* **45**(2): 219–34.

Majhail NS, Rizzo JD, Lee SJ *et al*. (2012) Recommended screening and preventive practices for long-term survivors after hematopoietic cell transplantation. *Biology of Blood and Marrow Transplantation* **18**(3): 348–71.

Markey KA, MacDonald KP, Hill GR (2014) The biology of graft-versus-host disease: experimental systems instructing clinical practice. *Blood* **124**(3): 354–62.

Schmid C, Schleuning M, Schwerdtfeger R *et al*. (2006) Long-term survival in refractory acute myeloid leukemia after sequential treatment with chemotherapy and reduced-intensity conditioning for allogeneic stem cell transplantation. *Blood* **108**(3): 1092–9.

Sorror ML, Giralt S, Sandmaier BM *et al*. (2007) Hematopoietic cell transplantation specific comorbidity index as an outcome predictor for patients with acute myeloid leukemia in first remission: combined FHCRC and MDACC experiences. *Blood* **110**(13): 4606–13.

CHAPTER 36

Normal haemostasis

36

Keith Gomez[1] and John H McVey[2]

[1]Haemophilia Centre and Thrombosis Unit, Royal Free London NHS Foundation Trust, London, UK
[2]School of Biosciences and Medicine, University of Surrey, Guildford, UK

Introduction

Haemostasis is one of a number of protective processes that have evolved in order to maintain a stable physiology. It interacts with other body defence mechanisms, such as the immune system and the inflammatory response. These links are most clearly seen in ancient species such as the horseshoe crab (*Limulus polyphemus*), where a primitive 'coagulation' pathway is initiated by entry of endotoxin into the haemolymph. Vestiges of this process still exist in humans and may give rise to serious clinical consequences. For example, disseminated intravascular coagulation (DIC) can be initiated by Gram-negative septicaemia. However, consequent upon the development of a high-pressure blood circulatory system, extra components have evolved and have resulted in a complex, highly integrated process in all vertebrates (Figure 36.1). Indeed, analysis of the haemostatic network in bony fish suggests that the network in its entirety evolved over 430 million years ago, prior to the divergence of bony fish from tetrapods.

The high blood pressure generated on the arterial side of the vertebrate circulation requires a powerful, almost instantaneous, but strictly localized procoagulant response in order to minimize blood loss from sites of vascular injury without compromising blood flow generally. Systemic anticoagulant and clot-dissolving components have also evolved to prevent extension of the procoagulant response beyond the vicinity of vascular injury resulting in unwanted thrombus formation. The resultant haemostatic system is thus a complex mosaic of activating or inhibitory pathways that integrates its five major components (blood vessels, platelets, coagulation factors, coagulation inhibitors and fibrinolytic elements).

This chapter reviews current concepts of haemostasis in humans to provide a background for the haemostatic disorders described in succeeding chapters. It starts with an overview of the response to blood vessel injury. Following this, each of the five main areas are discussed in more detail.

Overview of haemostasis

In the most simplistic terms, blood coagulation occurs when the enzyme thrombin is generated and proteolyses soluble plasma fibrinogen, forming the insoluble fibrin polymer, or clot. 'Haemostasis' refers more widely to the process whereby blood coagulation is initiated and terminated in a tightly regulated fashion, together with the removal (or fibrinolysis) of the clot as part of vascular remodelling. It is essentially the global process by which vascular integrity and patency are maintained over the whole organism, for its lifetime.

Although it is pedagogically convenient to present the subsystems of haemostasis as if they operate independently, the whole haemostatic mechanism is integrated *in vivo* so that thrombin generation is localized, limited and followed by fibrinolysis and tissue remodelling, which are also localized and limited. Furthermore, thrombin generation is not a simple exponential cascade, as was originally envisaged by Davie and Ratnoff when the 'waterfall' hypothesis was described in the 1950s (Figure 36.2). Rather, it is a complex network of interactions with

Postgraduate Haematology, Seventh Edition. Edited by A Victor Hoffbrand, Douglas R Higgs, David M Keeling and Atul B Mehta.

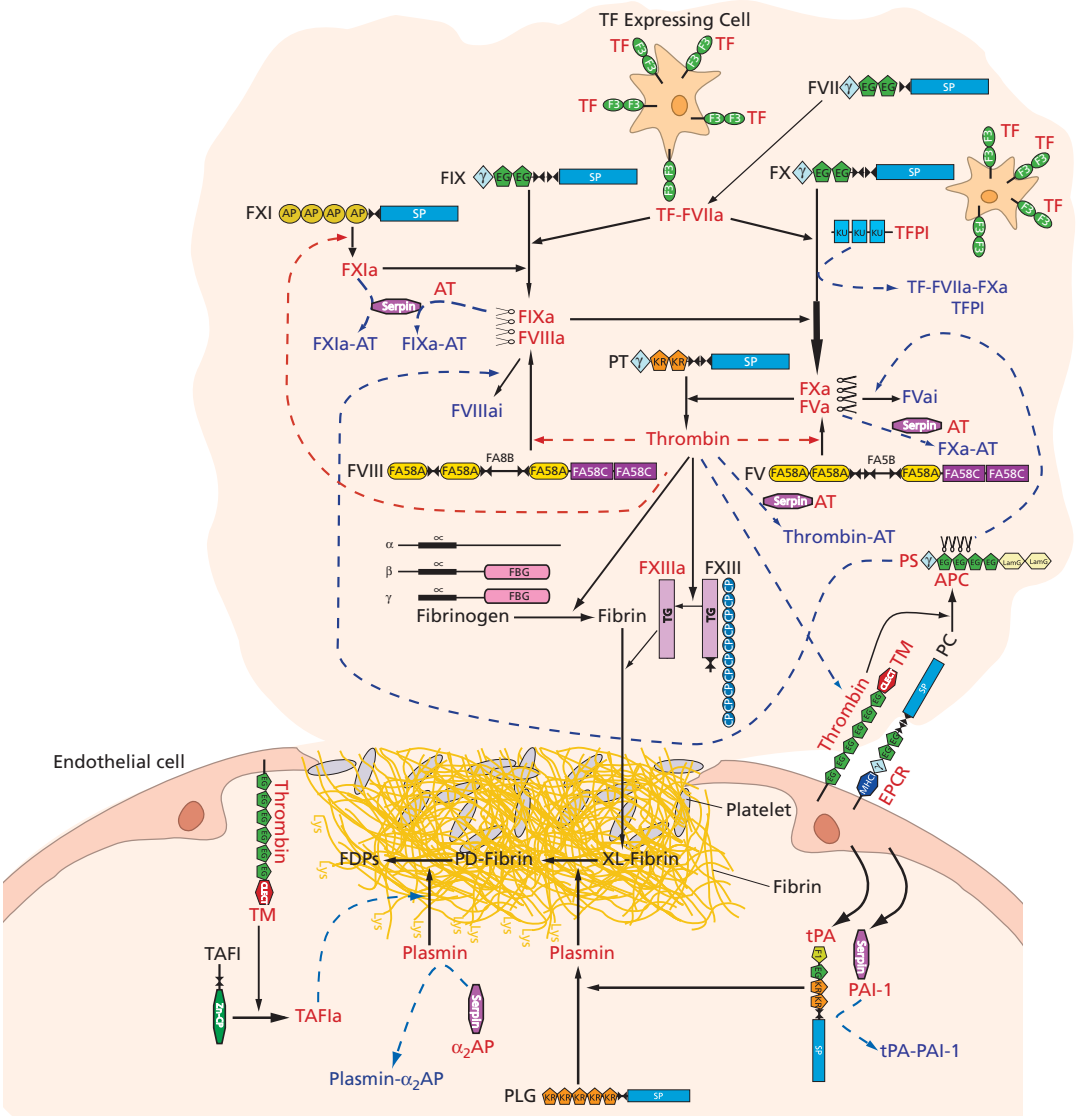


Figure 36.1 The haemostasis network. The network includes elements of coagulation initiation via TF–FVIIa interaction, amplification via the tenase (FVIIIa–FIXa) and prothrombinase (FVa–FXa) complexes, feedback inhibition through specific inhibitors, and clot dissolution by fibrinolysis. Module organization of proteins is indicated (see Figure 36.8) and all abbreviations are given in the text. Protein names are coloured as follows: functionally active proteins, red; inactivated or inhibited proteins, blue; precursors and fibrinogen-derived components, black. Lines and arrows are coloured as follows: dashed red lines, positive procoagulant feedback loops; dashed blue lines, negative feedback and inhibitory loops; solid black lines, interactions or processes. Cream boxes indicate assembly of macromolecular complexes on a cellular phospholipid surface. The process of blood coagulation is initiated by the exposure of cells expressing TF to flowing blood. Thrombin generation is propagated by a series of positive feedback loops, leading to fibrin deposition. This process is controlled by a series of negative feedback steps: the initiation complex is inhibited by the formation of the quaternary complex TF–FVIIa–FXa–TFPI and the active proteases FIXa, FXa and thrombin are inactivated by the serpin AT. In addition, thrombin initiates a negative feedback pathway by activating PC, leading to the inhibition of FVa and FVIIIa. The generation of fibrin leads to activation of the fibrinolytic pathway through binding of tPA, which leads to plasminogen activation and fibrin degradation. This pathway is also subject to inhibition through the binding of the serpins PAI-1 and α_2-antiplasmin (α_2AP) to tPA and plasmin, respectively. All these processes, plus platelet activation, occur simultaneously or with short lag times. (Source: McVey and Tuddenham, 2007. Genomics of haemostasis. In: *Genomics and Clinical Medicine* (Kumar, ed.). Reproduced with permission of Oxford University Press.)

positive and negative feedback loops. Thus, the effects of varying the concentration of any component in such a system are not intuitively obvious.

Tissue factor initiates blood coagulation

Exposure of blood to cells expressing tissue factor (TF) on their surface is both necessary and sufficient to initiate blood coagulation *in vivo*, both in normal haemostasis and in pathological situations such as thrombosis. Figure 36.2 presents a version of the earlier cascade concept for intrinsic and extrinsic pathways of thrombin generation. Although the contact system does not appear to have a physiological role in haemostasis it remains a valuable tool for understanding the common coagulation tests that are used in clinical practice. TF is constitutively expressed at biological boundaries such as skin, organ surfaces, vascular adventitia and epithelial–mesenchymal surfaces, where it functions as a 'haemostatic envelope'. This ensures that, following disruption of vascular integrity, blood is immediately exposed to cells expressing TF, leading to the initiation of blood coagulation. The primary control of haemostasis is therefore the anatomical segregation of cells expressing functional TF from other components of the coagulation network present in blood.

Amplification of the initial stimulus

An updated concept of TF-initiated thrombin generation, including the important feedback reactions of thrombin, is shown in Figure 36.3. As stated above, it is now clear that the physiological initiator of blood coagulation is exposure of the circulating zymogen FVII to membrane-bound TF. Activation of FVII to the protease FVIIa and formation of a very high-affinity complex of TF with FVIIa results in the activation of FIX and FX by the TF–FVIIa complex. In the absence of its activated cofactor FVa, FXa generates only trace amounts of thrombin from prothrombin. Although insufficient to initiate significant fibrin polymerization, thrombin formed in this initiation stage of coagulation is able to back-activate FV, FVIII and FXI by limited proteolysis. In the amplification phase of coagulation, FXIa activates FIXa, which forms a complex with FVIIIa. This is the intrinsic tenase complex (FVIIIa–FIXa) that activates sufficient FXa to form a complex with FVa, producing the prothrombinase

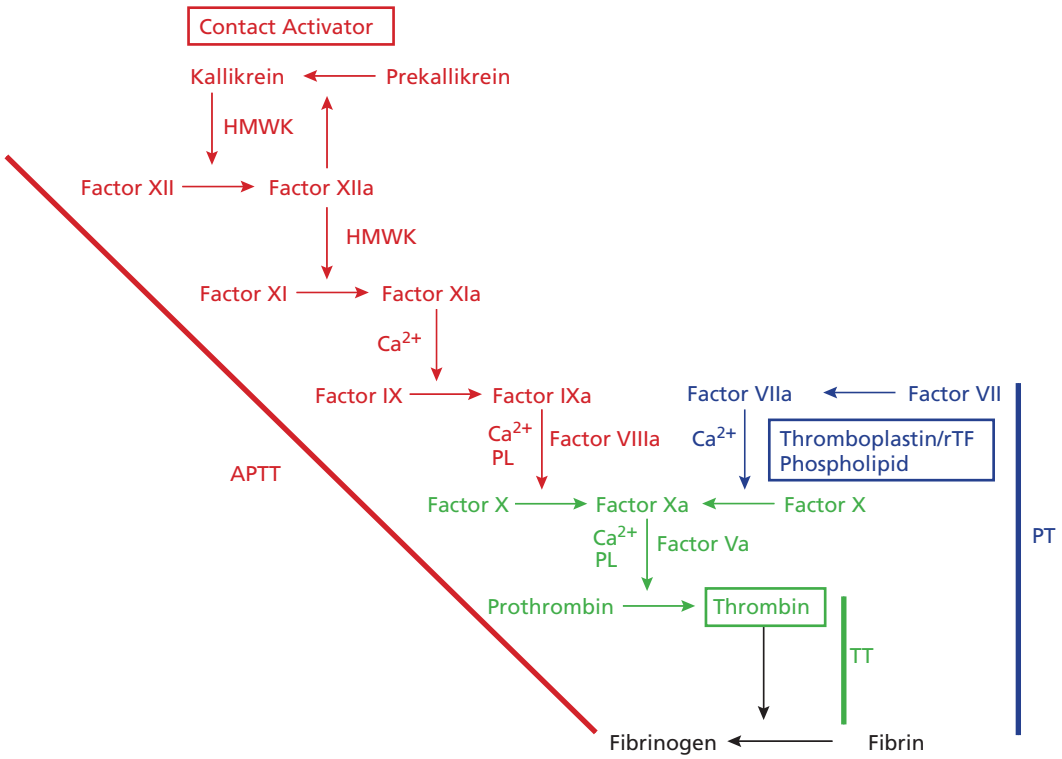


Figure 36.2 The coagulation cascade. The traditional concept of blood coagulation with separate intrinsic (red) and extrinsic (blue) pathways converging on the common pathway (green) with the generation of FXa. The activated partial thromboplastin time (APTT) is initiated by the addition of a contact activator (e.g. kaolin or silica) and tests for deficiencies in the intrinsic pathway. The prothrombin time (PT) is initiated by addition of thromboplastin and phospholipids and tests for deficiencies in the intrinsic pathway. The thrombin time (TT) is initiated by addition of thrombin and tests for an inhibitor of thrombin (most commonly heparin) or a problem with fibrin cleavage. HMWK, high-molecular-weight kininogen; PL, phospholipid.

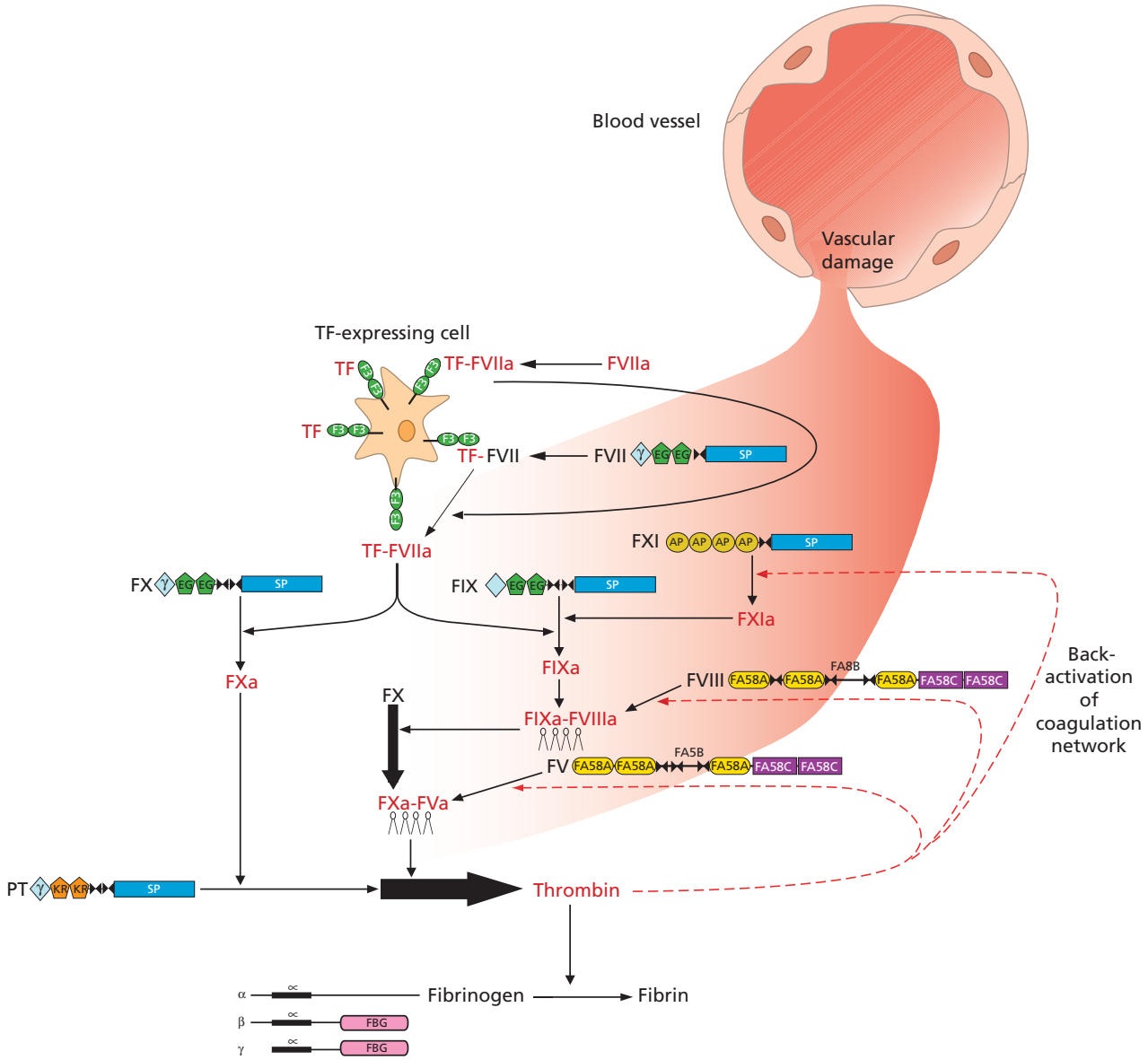


Figure 36.3 The haemostasis network: fibrin generation. Following vascular damage, blood coagulation is initiated by exposure of FVII to cells expressing the integral membrane protein TF. Activation of FVII to the protease FVIIa results in the activation of FIX and FX by the TF–FVIIa complex. In the absence of its activated cofactor FVa, FXa generates only trace amounts of thrombin. Although insufficient to initiate significant fibrin polymerization, trace amounts of thrombin formed in this initiation stage of coagulation are able to back-activate FV and FVIII by limited proteolysis. In the amplification phase of coagulation, FVIIIa forms a complex with FIXa, activating sufficient FXa that (in complex with FVa) leads to the explosive generation of thrombin, ultimately leading to generation of a fibrin clot. The tenase (FVIIIa–FIXa) and prothrombinase (FVa–FXa) complexes assemble on phospholipid surfaces. Colours and symbols as in Figure 36.1.

complex (FVa–FXa). This results in the explosive generation of thrombin that ultimately leads to generation of a fibrin clot.

A key feature of these processes is the assembly of multimolecular complexes on a phospholipid surface provided *in vivo* by cell membranes. For procoagulant complexes such as tenase and prothrombinase, this surface is provided by activated platelets. Each of these complexes consists of a cofactor (TF, FVa, FVIIIa), an enzyme (FVIIa, FIXa, FXa) and a substrate that is a zymogen (FIX, FX and prothrombin) of a serine protease. The product of one reaction becomes the enzyme in the next complex. The cofactors FV and FVIII are themselves activated by limited proteolytic digestion.

Feedback inhibition of the procoagulant response

The TF–FVIIa complex is rapidly inactivated by tissue factor pathway inhibitor (TFPI), released from thrombin-activated platelets or found on the surface of endothelial cells, by the formation of a quaternary inhibited complex (Figure 36.4). This explains the requirement for further activation of FX via FIXa–FVIIIa, and back-activation of FXI by thrombin, to permit further activation of FIX. Figure 36.4 also indicates the role of one of the other important inhibitors of the network, antithrombin (AT), which damps down thrombin generation by forming

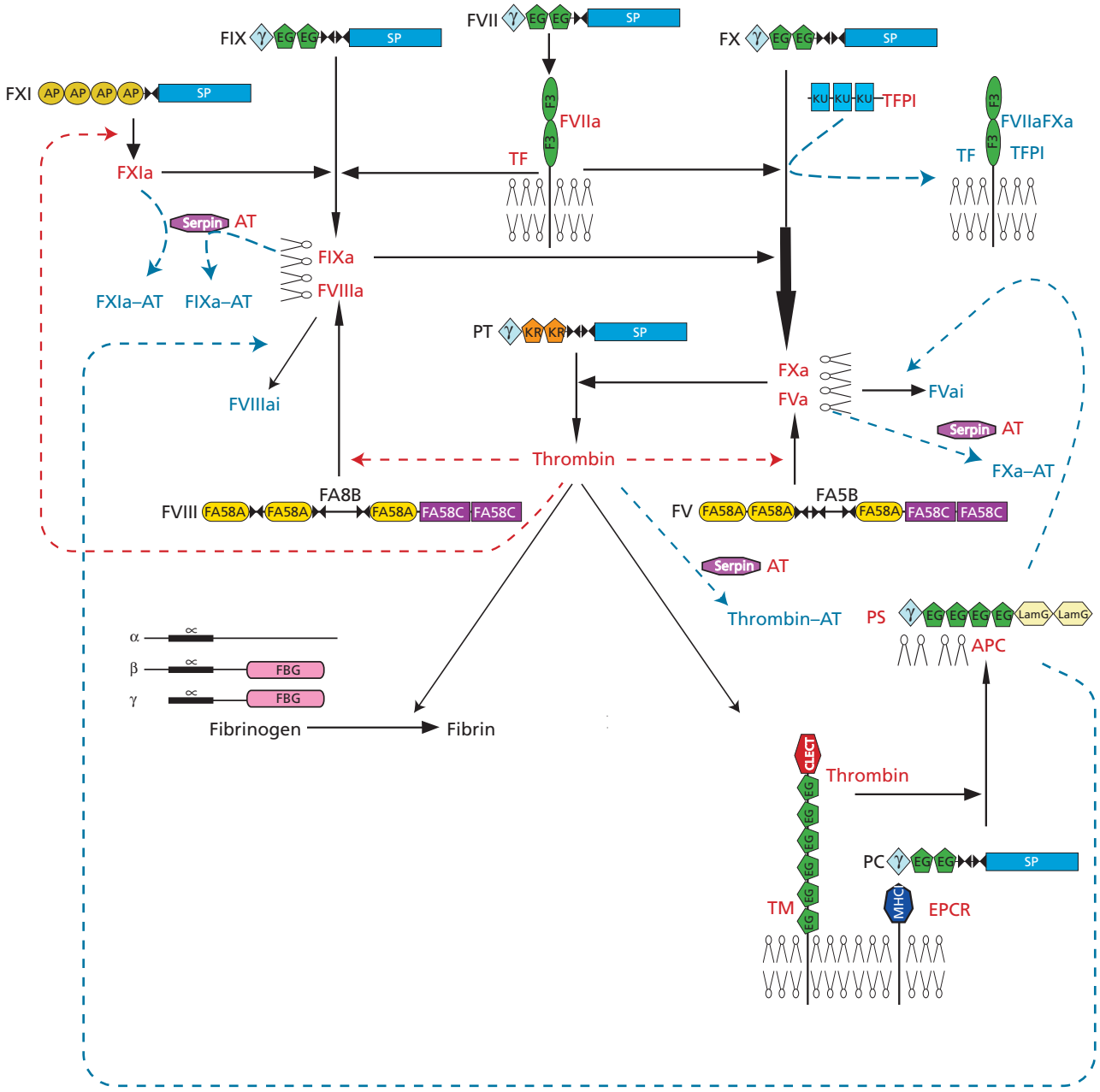


Figure 36.4 The haemostasis network: inhibition of thrombin generation. The initiation of blood coagulation via generation of FIXa and FXa by TF–FVIIa is shut down by the action of TFPI, which forms a quaternary complex with TF–FVIIa–FXa. Further generation of FIXa occurs following back-activation of FXI by thrombin. The serine proteases FIXa, FXa, FXI and thrombin are all inhibited by AT. At the endothelial cell surface, thrombin bound to TM activates PC bound to its receptor EPCR. APC in complex with its cofactor PS inactivates FVa and FVIIIa by further proteolytic cleavages. Colours and symbols as in Figure 36.1.

inactive complexes with FIX, FX, FXI and thrombin. The rate of inhibition by AT is substantially increased by binding a variety of polysulfated mucopolysaccharides known as glycosaminoglycans (GAGs) on the surface of endothelial cells.

Thrombin also proteolytically activates an anticoagulant pathway, the protein C (PC) pathway (Figure 36.4). Thrombomodulin (TM), an endothelial cell surface receptor, plays a pivotal role in the activation of this anticoagulant pathway. TM binds thrombin and, by an allosteric mechanism, alters its substrate specificity. The procoagulant substrates of thrombin, including FV, FVIII and fibrinogen, are no longer efficiently proteolysed. The preferred substrate of the thrombin–TM complex is PC, a zymogen that is proteolysed to activated PC (APC). The activation of PC can be augmented by another cellular receptor, endothelial cell protein C receptor (EPCR), which provides a direct endothelial cell binding site for PC and thus increases the apparent affinity of the thrombin–TM complex for PC; however, not all endothelial cells coexpress TM and EPCR. In complex with its cofactor, protein S (PS), APC rapidly inactivates the procoagulant cofactors FVa and FVIIIa by specific proteolysis, forming a negative feedback loop. The PC pathway is most active in the microvasculature, where the relative concentration of EPCR is highest.

Fibrinolysis

The final step in the regulation of fibrin deposition is the prevention and/or rapid removal of insoluble fibrin by the fibrinolytic system (Figure 36.5) described more fully later in the chapter. Briefly, once sufficient fibrin is generated, it binds tissue plasminogen activator (tPA), leading to the increased activation of plasminogen (PLG). This results in the formation of plasmin at

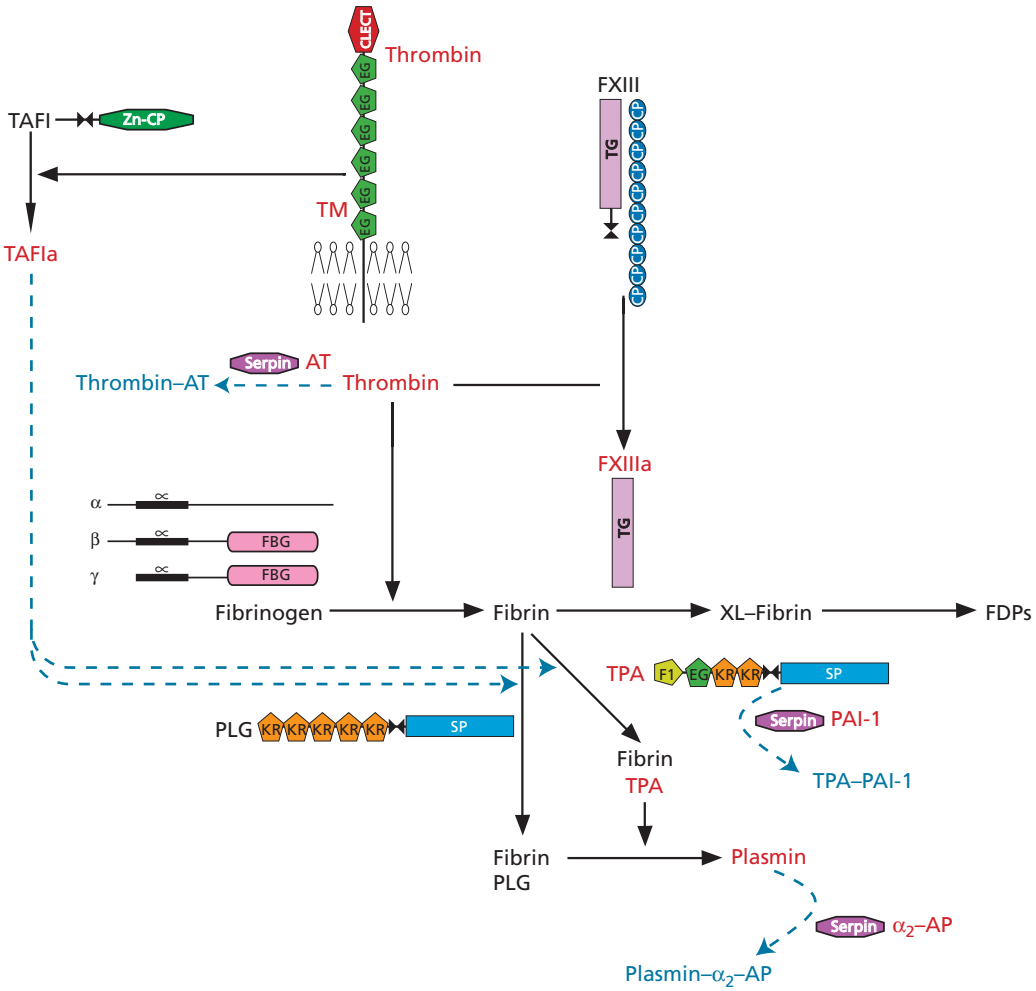


Figure 36.5 The haemostasis network: fibrinolysis. Both PLG and tPA bind to polymerized fibrin, promoting the cleavage of PLG by tPA to active plasmin, which then degrades cross-linked fibrin to soluble FDPs. Modulation of fibrinolysis is achieved by serpin inhibition: PAI-1 inactivates tPA, while α_2-antiplasmin (α_2AP) inhibits plasmin. In addition, TAFIa removes Lys residues from fibrin, removing PLG and tPA binding sites. Colours and symbols as in Figure 36.1.

the site of the fibrin clot, which breaks down fibrin into soluble fibrin degradation products. The fibrinolytic system is also subject to inhibition through the action of inhibitors of tPA and plasmin, namely plasminogen activator inhibitor (PAI)-1 and α_2-antiplasmin, respectively.

Blood vessels

The endothelium

The endothelium functions in a multitude of physiological processes, including intracellular transport, the regulation of vasomotor tone and maintenance of blood flow. Endothelial cells possess surface receptors for a variety of physiological substances, for example thrombin and angiotensin II, which may influence vascular tone directly or indirectly through various haemostasis-related events. Once activated, endothelial cells express a variety of intracellular adhesion molecules, some of which are released into the plasma. These include vascular cell adhesion molecule (VCAM), E-selectin, P-selectin and von Willebrand factor (VWF), which modulate leucocyte and platelet adhesion, inflammation, phagocytosis and vascular permeability.

However, the endothelium should not be regarded as a simple homogeneous cell type. It would appear that endothelial cell phenotypes are differentially regulated. At any given point in time, structural and functional phenotypes may vary between segments of the vascular tree, and at any given location the endothelial phenotypes may change from one moment to the next. Endothelial cell heterogeneity occurs between different organs, within the vasculature of a given organ, and even between neighbouring endothelial cells of a single vessel.

The platelet–vessel wall interaction

Healthy unperturbed endothelial cells exert a powerful inhibitory influence on haemostasis by virtue of the factors they release or express on their surface (see Figure 36.6a). Two of these, prostaglandin (PG)I_2 (or prostacyclin) and nitric oxide (NO), also known as endothelium-derived relaxing factor, have powerful vasodilator activity, acting on smooth-muscle cells in the vessel wall (basal-directed secretion) and hence modulating blood flow. Both substances inhibit aggregation of platelets and leucocytes (luminal-directed secretion) by raising intraplatelet levels of cyclic adenosine monophosphate (cAMP) and cyclic guanosine monophosphate (cGMP), respectively (see below).

PGI_2 is the major prostaglandin synthesized by endothelial cells, a small amount also being produced by fibroblasts and smooth-muscle cells. Its action on platelets involves binding to a specific G-protein-coupled receptor (PTGIR) that activates adenylate cyclase, thus increasing the intraplatelet cAMP concentration. This promotes Ca^{2+} uptake into the dense tubular system and inhibits phosphatidylinositol metabolism, both of which prevent platelet aggregation and the consequent release of storage granules containing procoagulant molecules such as VWF and FV. Thrombin, and other agents that are generated at the site of injury, stimulate the synthesis of PGI_2 by adjacent endothelial cells, which counteracts the platelet-aggregating activity of the protease and thereby helps to localize platelet plug formation. PGI_2 is rather unstable, with a half-life of only a few minutes in whole blood, and degrades spontaneously to 6-keto-$PGF_{1\alpha}$. This degradation product is biologically inert, but in some tissues a variable proportion of PGI_2 may be enzymatically converted to 6-keto-$PGF_{1\alpha}$ or 6,15-diketo-13,14-dihydro-$PGF_{1\alpha}$, which retain some platelet-inhibitory activity.

The once-held belief that PGI_2 is a circulating hormone that limits platelet–platelet interactions within the bloodstream is probably false, as plasma levels of PGI_2 are at least two orders of magnitude below that needed to inhibit platelet aggregation. Although in some settings, such as following severe trauma, it is possible that markedly raised systemic levels could occur transiently, it is more likely that PGI_2 serves mainly as a local hormone, principally concerned with vascular tone, but possibly also inhibiting the extension of the platelet plug beyond the immediate vicinity of any endothelial damage.

NO is synthesized in smooth-muscle cells, macrophages and activated platelets, as well as by endothelial cells. Its synthesis and secretion may be constitutive (when it serves as a local hormone to 'fine tune' blood flow). Stimulated (inducible) synthesis of NO also occurs in endothelial cells exposed to cytokines such as interleukin (IL)-1, tumour necrosis factor (TNF) or endotoxin. This produces a slower but longer-lasting rise in NO and can have undesirable side-effects, such as inflammation, cytotoxicity or prolonged hypotension. Like PGI_2, it has a very short half-life (3–5 s) and is rapidly oxidized to the inactive nitrite (NO_2^-) or nitrate (NO_3^-) forms.

In addition to PGI_2 and NO, endothelial cells also express an ecto-ADPase (CD39) on their cell surface, which rapidly metabolizes ATP and ADP, released from activated platelets, to AMP, thereby drastically reducing or abolishing platelet recruitment and aggregation in response to thrombin.

The 21-amino-acid peptide endothelin (ET)-1 is the predominant isoform of the endothelin peptide family, which includes ET-2, ET-3 and ET-4. It exerts various biological effects, including vasoconstriction and the stimulation of cell proliferation in tissues both within and outside the cardiovascular system. In the endothelium, ET-1 is predominantly released abluminally towards the vascular smooth muscle, where it acts as a potent vasoconstrictor and mitogen for smooth-muscle cells. In addition, ET-1 stimulates the production of cytokines and growth factors. Furthermore, it stimulates neutrophil adhesion and platelet aggregation and is chemotactic for macrophages.

Thrombospondin (TSP)-1 is an abundant constituent of the α-granules of platelets and Weibel–Palade bodies of endothelial cells. Upon release from platelets, TSP-1 binds to the platelet surface in a Ca^{2+}-dependent manner and interacts with

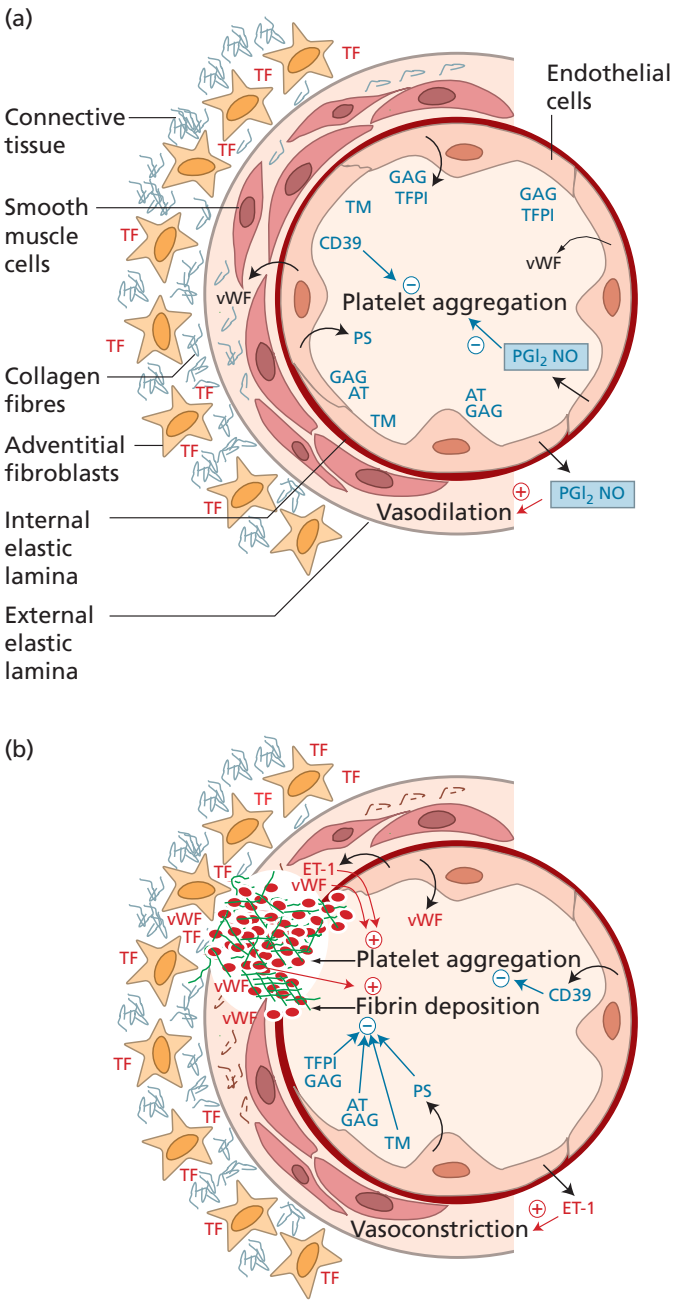


Figure 36.6 Blood vessel architecture and functions in haemostasis. (a) The anatomy of a generic vessel in the resting state. The layer thicknesses vary widely between arterial and venous circulation and between large and small vessels. In the absence of vessel damage, cells expressing TF are anatomically separated from components of the blood coagulation network, preventing initiation of coagulation. In addition, coagulation inhibitors present on the endothelial surface and inhibitors of platelet aggregation maintain vessel patency. (b) The localized haemostatic response to vessel injury: initiation of coagulation by TF exposure and platelet activation/aggregation by exposure of extravascular collagen to plasma VWF. Procoagulant responses away from the lesion site are damped down by anticoagulant and antiplatelet factors derived from the endothelium.

integrins $\alpha_{IIb}\beta_3$ and $\alpha_v\beta_3$, CD36 and the integrin-associated protein (IAP), integrin-bound fibrinogen and fibronectin. TSP-1 activates $\alpha_{IIb}\beta_3$ through its binding to IAP. This results in spreading of platelets and aggregation.

von Willebrand factor

VWF is a multimeric glycoprotein that plays an important role in primary haemostasis by promoting platelet adhesion to the subendothelium at sites of vascular injury under high shear-rate conditions. It is also a carrier of FVIII and this association protects FVIII from rapid proteolysis. Failure of FVIII to bind VWF, as occurs in type 2N von Willebrand disease (see Chapter 38), leads to its rapid turnover and resulting FVIII deficiency. VWF is synthesized by endothelial cells, megakaryocytes and platelets. In endothelial cells VWF may be secreted directly into the circulation or stored in Weibel–Palade bodies. The VWF produced in megakaryocytes and platelets is not secreted, but stored in α-granules. Release of VWF from these stores occurs following activation of endothelial cells or platelets. Some therapeutic products, such as desmopressin, may work by stimulating release of stored VWF.

Mature VWF is composed of disulfide-linked subunits, each comprising 2050 amino acid residues and up to 22 carbohydrate side-chains. Each subunit has a molecular mass of approximately 278 kDa, of which 10–19% is carbohydrate. Two subunits joined at their C-termini by a disulfide bridge form a dimer. Additional disulfide bridges between dimers lead to the formation of larger multimers that range in size from 500 kDa to in excess of 10,000 kDa (Figure 36.7). The ultra-large VWF multimers are predominantly stored, but when secreted by endothelial cells may appear as globular molecules or as thin filaments that are several microns long. These ultra-large VWF multimers can only be detected transiently in normal plasma because they are cleaved by a plasma metalloproteinase (ADAMTS-13). Deficiency or inhibition of ADAMTS-13 results in accumulation of ultra-large VWF multimers in plasma, leading to enhanced shear-induced platelet aggregation and microvascular thrombosis, as seen in thrombotic thrombocytopenic purpura (TTP) (see Chapter 43).

VWF contains binding sites for collagen, fibrinogen and the platelet receptor, glycoprotein (GP)Ib. When vascular injury exposes subendothelial collagen, VWF binds and becomes unwound from its globular form (Figure 36.7c). This exposes more binding sites and allows the capture of platelets from the circulation. Thus VWF, and to some extent fibrinogen, act as bridges between the platelet and the injured vascular wall. Mutations in the genes encoding the GPIbα complex that lead to loss of function cause the bleeding disorder Bernard–Soulier syndrome. Conversely, gain-of-function mutations in either VWF or GPIb that increase the interaction lead to type 2B von Willebrand disease or pseudo (platelet-type) von Willebrand disease (see Chapter 38).

Endothelial cell anticoagulant activities

Endothelial cells regulate coagulation by a number of mechanisms. The activity of the PC pathway is promoted by expression of two cellular receptors. TM is constitutively expressed on all endothelial cells with the exception of the brain. EPCR is expressed strongly in the endothelial cells of arteries and veins in the heart and lung, less intensely in capillaries in the lung and skin, and not at all in the endothelium of small vessels of the liver and kidney. In addition, endothelial cells synthesize and secrete PS, the cofactor for APC in the inactivation of FVa and FVIIIa.

Endothelial cells regulate the initiation of coagulation by synthesizing the main physiological inhibitor of the FVIIa–TF initiating complex, namely TFPI. This is described in greater detail below. Endothelial cell-surface GAGs are important in binding TFPI and also modulate the activity of AT, which inhibits thrombin and FXa and to a lesser extent FIXa and FXIa.

Endothelial cell-derived fibrinolytic factors

Three important fibrinolytic factors are detectable in the vessel wall: tPA and PAI-1 are synthesized primarily by endothelial cells, whereas urinary plasminogen activator (uPA, urokinase) is mainly derived from fibroblast-like cells in the kidney and gut. Unlike VWF and P-selectin, tPA and PAI-1 are not stored in Weibel–Palade bodies and, in the resting state, synthesis and secretion are slow, resulting in low circulating levels. However, stimulated synthesis and release occur in response to a variety of stimuli.

Coagulation factors

A summary of the characteristics of the proteins discussed, including gene location, number of exons, molecular mass, circulating concentration, plasma half-life and presumed function can be found in Table 36.1. The domain structure and modular nature of the proteins are shown in Figure 36.8. This demonstrates the similarity between coagulation factors that belong to the same protein family, such as the serine proteases (e.g. compare FVII, FIX, FX and PC).

Tissue factor

The formation of the FVIIa–TF complex is regarded as the sole initiator of coagulation in both normal and pathological coagulation. The processed mature TF protein is 263 amino acids in length after signal peptide cleavage. The 219-residue extracellular region consists of two fibronectin type III domains anchored into the cell membrane by a 23-residue transmembrane sequence. The extracellular region binds tightly to FVIIa to form a highly active procoagulant complex. The 21-residue cytoplasmic domain has been implicated in intracellular signalling; however, transgenic experiments in mice have shown that this domain is not required for either development or procoagulant function, so its role is still unclear.

Vascular adventitial cells, neuroglia, vascular smooth muscle and epidermal cells express TF constitutively. It therefore forms a protective envelope around blood vessels and organs, ready to initiate clotting as soon as blood leaks out of vessels. TF is also expressed by monocytes and endothelium after activation by inflammatory cytokines or by endotoxin, as occurs in sepsis, and on cancerous tissues. Intravascular exposure of TF by any route, but particularly after atherosclerotic plaque rupture, can result in pathological thrombosis. Recent work has also focused on the novel hypothesis that functionally active TF can circulate in flowing blood on small procoagulant particles (microparticles), which may also be transferred from one blood cell type to another via specific receptor-mediated interactions.

Factor VII

Plasma FVII binds to TF, for example after vessel trauma or plaque rupture, to form a complex that initiates coagulation by directly activating FX and to a lesser extent FIX. The FVII gene (*F7*) lies adjacent to the factor X gene (*F10*), suggesting gene

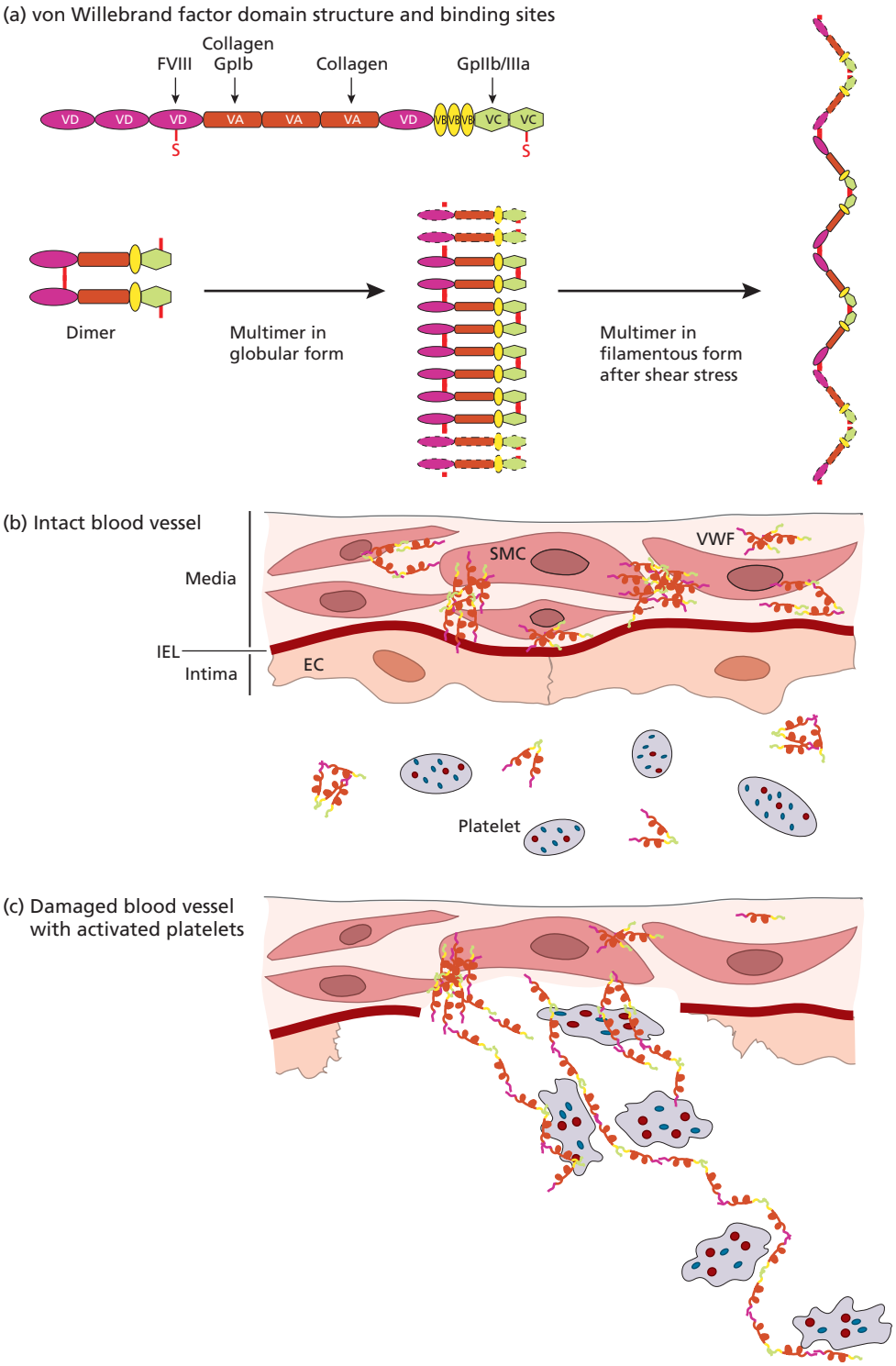


Figure 36.7 Role of von Willebrand factor (VWF) in haemostasis. (a) VWF domain structure with important binding sites. Dimers are formed when individual monomers become joined through disulfide bonds at the C-terminus. Multiple dimers are then joined at the N-terminus to form high-molecular-weight multimers that take on a globular shape. When exposed to flowing blood under high shear stress the multimers become filamentous, exposing their binding sites. (b) In normal vasculature the ultra-high-molecular-weight VWF multimers required for full VWF function are found only in the subendothelium. The circulation contains lower-molecular-weight multimers. After disruption of the endothelium (c), exposure of high-molecular-weight multimers to flowing blood applies shear stress, causing the multimers to unravel and capture activated platelets. The coagulation cascade forms on the surface of these platelets. Thus VWF acts to localize the process of clot formation to the site of vascular damage. EC, endothelial cell; IEL, internal elastic lamina; SMC, smooth-muscle cell.

Table 36.1 Key proteins involved in the haemostatic network.

Common name	Abbreviation	Subunit	Gene symbol	Gene location	No. of exons	Amino acids (mature)	M_r of monomer (kDa)	Plasma level (μg/mL)	Plasma level (nmol/L)	$t_{1/2}$ (hours)	Main action
Tissue factor	TF		*F3*	1p22–p21	6	263	44	NA	NA	NA	Cofactor for FVII/FVIIa
Prothrombin	FII		*F2*	11p11.1	14	579	72	90	1400	65	Clots FBG, activates PC, FXI, TAFI
Factor V	FV		*F5*	1q23	25	2196	330	10	30	15	Cofactor for FXa
Factor VII	FVII		*F7*	13q34	8	416	50	0.5	10	3	Activates FIX and FX
Factor VIII	FVIII		*F8*	Xq28	26	2332	330	0.1	0.3	10	Cofactor for FIXa
Factor IX	FIX		*F9*	Xq27	8	415	56	5	90	25	Activates FX
Factor X	FX		*F10*	13q34	8	445	59	8	135	40	Activates prothrombin
Factor XI	FXI		*F11*	4q35	15	607	80*	5	30	45	Activates FIX
Prekallikrein	PK		*KLKB1*	4q35	15	638	86	50	580	35	Anti-angiogenic, profibrinolytic
Factor XIII† (A chain)	FXIII	A	*F13A1*	6p25.3–p24.3	15	731	75†	10	30	200	Cross-links fibrin
Factor XIII† (B chain)	FXIII	B	*F13B*	1q31	12	641	80†	10	30	200	Cross-links fibrin
Fibrinogen (α-chain)‡	FGN	α	*FGA*	4q28	6	866	68‡	3000	9000	90	Mechanical stabilization of clot
Fibrinogen (β-chain)‡	FGN	β	*FGB*	4q28	8	491	52‡	3000	9000	90	Mechanical stabilization of clot
Fibrinogen (γ-chain)‡	FGN	γ	*FGG*	4q28	10	453	49‡	3000	9000	90	Mechanical stabilization of clot
von Willebrand factor	VWF		*VWF*	12p13.3	52	2050	255	10	40	12	Cell adhesion and FVIII carrier
Thrombomodulin	TM		*THBD*	20p11.2	1	557	60	NA	NA	NA	Cofactor in PC/TAFI activation
Endothelial protein C receptor	EPCR		*PROCR*	20q11.2	7	220	27	NA	NA	NA	Cofactor in PC activation

Table 36.1 (*Continued*)

Common name	Abbreviation	Subunit	Gene symbol	Gene location	No. of exons	Amino acids (mature)	M_r of monomer (kDa)	Plasma level (μg/mL)	Plasma level (nmol/L)	$t_{1/2}$ (hours)	Main action
Protein C	PC		*PROC*	2q13–14	9	419	62	4	65	6	Inactivation of FVa and FVIIIa
Protein S	PS		*PROS1*	3q11.2	15	676	69	10 (free)	145	?	Inactivation of FVa and FVIIIa
Tissue factor pathway inhibitor	TFPI		*TFPI*	2q32	12	304	42	0.08	2.5	?	Inhibition of coagulation initiation
Antithrombin	AT		*SERPINC1*	1q23–q25.1	9	464	58	140	2400	5	Inhibits thrombin, FIX, FX, FXI
Heparin cofactor II	HCII		*SERPIND1*	22q11.21	5	499	66	90	1200	60	Prevention of arterial thrombosis?
Plasminogen	PLG		*PLG*	6q26	14	791	92	200	2000	50	Dissolution of clot in wound repair
Tissue plasminogen activator	tPA		*PLAT*	8p12	14	562	69	0.005	0.07	0.03	Plasma activator of plasminogen
Prourokinase	UK		*PLAU*	10q24	11	431	54	0.0015	0.04	0.03	Tissue activator of plasminogen
Plasminogen activator inhibitor 1	PAI-1		*SERPINE1*	7q21.3–q22	9	379	52	10	200	0.1	Inhibition of tPA and uPA
α_2-Antiplasmin	α_2-AP		*SERPINF2*	17p13	9	452	67	70	1000	72	Inhibition of plasmin
Thrombin-activatable fibrinolysis inhibitor	TAFI		*CPB2*	13q14.11	11	401	60	5	75	0.2	Inhibition of fibrinolysis

*FXI circulates as a 160-kDa homodimer of two 80-kDa monomers.
†FXIII circulates as a 326-kDa tetramer of two A- and two B-chains.
‡Fibrinogen circulates as a 340-kDa complex of two each of A-, B- and C-chains.

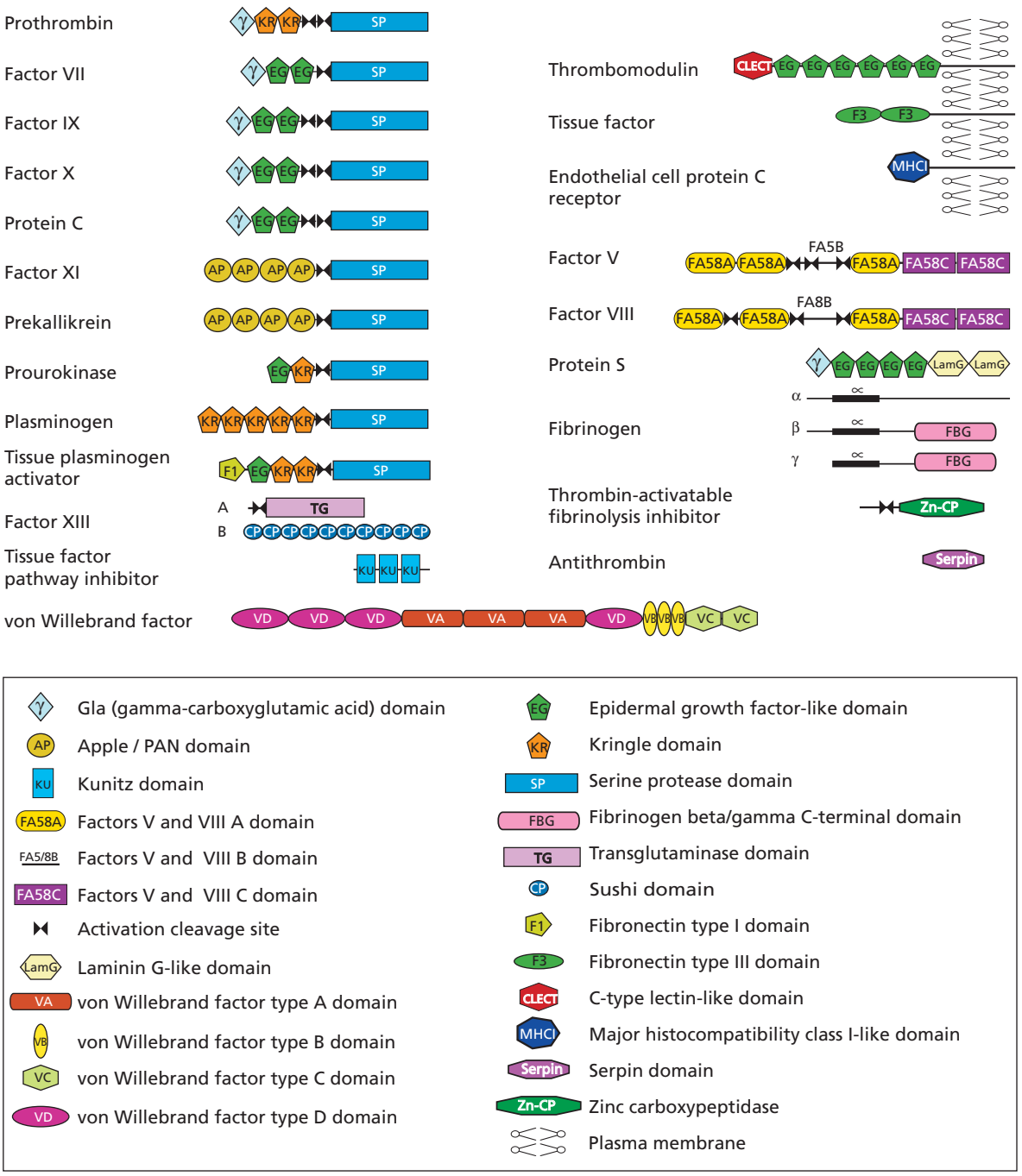


Figure 36.8 Modular organization of haemostasis proteins. The evolutionary relationship of many factors is suggested by their modular organization.

duplication during evolution, and close to that of protein Z (*PZ*) at 13q34. In common with the other serine proteases of the coagulation network (FIX, FX, prothrombin, PC), as well as PS and PZ, FVII has an N-terminal domain that contains a number (9–12) of glutamic acid residues that are post-translationally modified by the addition of a carboxyl group to the γ-carbon by a vitamin-K-dependent carboxylase. This γ-carboxyglutamic acid (GLA) domain confers affinity to negatively charged phospholipid membranes, such as those of activated platelets or endothelial cells, promoting the assembly of functional multiprotein complexes on these surfaces. Proteins containing GLA modules are commonly referred to as 'vitamin K-dependent factors' as

this vitamin is a cofactor in the carboxylation reaction. Blocking this post-translational modification with coumarin derivatives such as warfarin is one of the main treatments for the long-term prevention of thromboembolic events. FVII in common with other GLA-domain-containing proteins has a primary translation product that contains a propeptide sequence that directs the γ-carboxylation at 10 Glu residues within the GLA domain. The GLA domain is followed by two epidermal growth factor (EGF)-like domains, the connecting or activation peptide, and the serine protease domain. FVII is activated by cleavage between residues Arg152 and Ile153, yielding a two-chain disulfide-linked FVIIa molecule; however, this has little catalytic activity until it is bound to TF. The half-life of FVII zymogen in plasma is 3 hours and, remarkably, the half-life of the FVIIa enzyme is 2.5 hours, probably because there is no plasma inhibitor capable of effectively neutralizing free FVIIa. Recombinant FVIIa is in clinical use as a treatment for haemophiliacs with inhibitors and, for a wider range of bleeding problems and general surgical intervention.

Factor X

The gene and protein structures of FX closely resemble those of FVII. Unlike FVII, zymogen FX circulates as a two-chain disulfide-linked heterodimer due to excision of a tribasic peptide (Arg-Lys-Arg) at residues 139–141 during synthesis. The light chain of FX therefore consists of the GLA domain with 11 γ-carboxyglutamic acid residues, the amphipathic helix and two EGF-like domains. The half-life of FX in plasma is 36 hours. FX is activated by either FIXa–FVIIIa or TF–FVIIa on phospholipid surfaces in the presence of Ca^{2+} ions. FXa forms a phospholipid-bound complex with FVa, which efficiently activates prothrombin (prothrombinase complex).

Factor IX

The FIX gene (*F9*) is located at Xq26, about 15.2 Mb from the FVIII gene (*F8*); thus deficiencies of both factors are X-linked disorders. Deficiency of FIX results in clinical haemophilia B, as the main function of FIX is to participate in the tenase complex (FIXa–FVIIIa). The FIX promoter has been extensively studied and includes sites for liver-specific transcription factors and an androgen-responsive promoter element, which accounts for the haemophilia B Leiden phenotype (FIX deficiency, which corrects spontaneously at puberty). The plasma half-life of FIX is 18 hours.

The mature protein is a single-chain molecule consisting of 415 amino acids and the GLA domain contains 12 γ-carboxyglutamic acid residues. Either TF–FVIIa or FXIa cleaves FIX within the connecting peptide (between residues 145–146 and 180–181). Unlike FVII and FX, FIX requires two proteolytic cleavages, releasing an activation peptide, in order to achieve full activation.

Factor XI

FXI is a zymogen of a serine protease that has four Apple (PAN1–4) domains and a serine protease domain in each monomer. The protein circulates as a homodimer, with the two monomer subunits linked via a series of interactions between the two Apple 4 domains that includes a disulfide bridge. Activation of FXI is by a single cleavage performed by thrombin. It is clear that dimerization is required for full function of this factor, presumably due to obstruction of essential binding sites in the monomer form. FXIa then activates FIX directly in free solution. FXI activation by the contact protein FXIIa was originally thought to be a relevant step in intrinsic or contact-activated coagulation (see Figure 36.2), but it is now considered that the feedback activation of FXI by trace thrombin provides a physiologically relevant route for generation of increased amounts of FIXa to assemble tenase during the amplification of the initial TF stimulus (see Figure 36.3).

There is increasing evidence that FXI has other roles in haemostasis, such as regulation of fibrinolysis by the thrombin-activatable fibrinolysis inhibitor (TAFI). The fact that bleeding symptoms in FXI deficiency correlate poorly with FXI procoagulant level suggest that these other functions may be important in determining the phenotype in this mild bleeding disorder.

Factor XIII

FXIII circulates as a tetramer of two A-chains and two B-chains. The B-chains function as carriers for the A-chains which, after activation by thrombin, function as a transglutaminase to cross-link fibrin and other proteins in the clot, resulting in a stable structure. FXIIIa contains a free sulfydryl group at the active site. Platelets also contain FXIII A-chain dimers, which are fully functional after thrombin activation.

Factor VIII

The FVIII gene (*F8*) spans 187 kb of the X chromosome (Xq28) and encodes a protein of 2351 amino acids. After removal of the 19 amino acid signal sequence, a single chain of 2332 amino acids is transiently formed, but is subsequently processed by a single proteolytic cleavage prior to release as a two-chain molecule. Because of internal homology, the FVIII domain structure can be represented as A1–A2–B–A3–C1–C2. The B domain is not necessary for procoagulant function, as B-domain-deleted FVIII has full activity. The B-domain sequences of FV and FVIII share little sequence identity, although their lengths are similar and both are heavily glycosylated. It is likely that the glycosylated B domain functions in promoting endoplasmic reticulum/Golgi transport since deficiency of LMAN1 (also known as ERGIC 53), a mannose-binding protein, results in combined deficiency of FV and FVIII. FVIII also has three short acidic interdomain peptides (a1, a2 and a3) that are closely implicated in FVIII function.

FVIII is the essential cofactor for activation of FX by FIXa in the tenase complex. It has no function until proteolysed to FVIIIa by thrombin or FXa at Arg372-Ser373 and Arg1689-Ser1690. FVIIIa is directly inactivated by APC cleavage at Arg336-Ser337 and Arg562-Ser563; however, functional activity of FVIIIa also decays rapidly by dissociation of the A2 subunit from FVIIIa. The crystal structure of FVIIIa has recently been solved by two groups.

Factor V

The mature protein product is 2196 amino acids in length after cleavage of a 28 amino acid signal peptide. The structure of FV, like that of FVIII, can be represented as A1–A2–B–A3–C1–C2. In contrast with FVIII the B domain is required for full procoagulant function. FV also differs from FVIII in that it lacks the three short acidic interdomain peptides (a1, a2 and a3) implicated in FVIII function.

FV is the cofactor for the activation of prothrombin by FXa. It has no cofactor activity until proteolysed by thrombin or FXa at Arg709-Ser710, Arg1018-Thr1019 and Arg1545-Ser1547. FVa is inactivated by APC through cleavages at Arg506-Ser507 and Arg1765-Leu1766. The initial (and rate-limiting) cleavage is at Arg506. This is the site of the mutation in FV Leiden (FV Arg506Gln) that is resistant to APC, leading to the most common form of familial thrombophilia.

Fibrinogen

The fibrinogen gene cluster is located on chromosome 4q32 in the order β–α–γ, with β transcribed in the opposite direction to α and γ. The three chains of fibrinogen are disulfide cross-linked and folded together in an intricate manner. The overall structure of fibrinogen is a symmetrical dimer, $\alpha_2\beta_2\gamma_2$. Viewed by electron microscopy, the molecule is trinodular, with the outer two globular domains (fragments D) containing the C-termini of all three chains connected to the central globular domain (fragment E), which contains the N-termini of all six chains tethered together by disulfide bonds. Coiled regions, forming α-helical ropes, connect the lateral and central globular domains.

Polymerization of fibrinogen occurs when thrombin cleaves two short, negatively charged fibrinopeptides A and B from the N-termini of the α- and β-chains, respectively. This reveals new N-terminal sequences in the fragment E region (called knobs) that fit into holes in the fragment D regions. Polymerization then occurs spontaneously in a staggered half-overlap array, which can elongate indefinitely in either direction. Electron microscopy studies were and still are important in resolving the molecular architecture of fibrillar fibrin formation, but from 1997 elegant crystal structures of parts of the fibrin(ogen) molecule were successfully produced in a number of laboratories. These structures reveal the elegant mechanisms of fibrin polymerization.

Prothrombin

The mRNA encodes a preproleader sequence similar to that found in other vitamin-K-dependent proteins, followed by a GLA domain, two kringle domains, an activation peptide and a serine protease domain. FXa complexed with FVa activates prothrombin zymogen to thrombin on a phospholipid surface (prothrombinase) on cleavage of two peptide bonds. The first, between Arg271 and Thr272, releases the protease domain from the GLA and kringle domains; the second, between Arg320 and Ile321, generates the catalytic site of thrombin by a typical trypsin-like conformational rearrangement. Cleavage at Arg320 yields the protease (meizothrombin) that retains membrane-binding properties. Cleavage at Arg271 yields thrombin and severs covalent linkage with the N-terminal domain known as fragment 1.2. Fully cleaved thrombin is termed α-thrombin, and is rapidly released from its site of production to participate in numerous haemostatic functions free in solution: acting as a procoagulant against many substrates including fibrinogen, FV, FVIII and FXI; acting in complex with TM as an anticoagulant against PC; and acting to activate cellular transmembrane protease-activated receptors. Thrombin also activates thrombin-activatable fibrinolytic inhibitor (TAFI), in complex with TM, inhibiting fibrinolysis.

Naturally occurring inhibitors of blood coagulation

In common with other defence mechanisms, such as those resulting in kinin release and complement activation, the blood coagulation process can be activated very rapidly when the need arises. This involves the generation of proteolytic enzymes, such as thrombin, that are potentially lethal if their action is not limited. For example, 10 mL of plasma can generate, in theory, sufficient thrombin to clot all of the fibrinogen in the body in 30 s. That it does not normally do so is due partly to the fact that the procoagulant response is most pronounced in the vicinity of the platelet plug forming at the point of vascular injury, while elsewhere coagulation is inhibited by substances in plasma or on the vascular surface that exhibit anticoagulant activity.

Classification of physiological anticoagulants

Physiological anticoagulants fall into two main groups: those that inhibit the serine proteases of the coagulation cascade (both Kunitz-type and serpins) and those involved in destruction of the activated coagulation cofactors FVa and FVIIIa (components of the PC system). These inhibitors assume great physiological significance: relatively minor deficiencies (50–70% of average levels) of some of these inhibitors, such as might be found in individuals heterozygous for a genetic defect, are associated with an increased incidence of thrombosis. Homozygous deficiencies

are frequently either fatal in the first few years of life unless prophylactic replacement therapy is instituted, or are not found in nature, suggesting early fetal loss.

In addition to the specific inhibitors, there are some other inhibitory mechanisms that do not fit into either of the above categories, one of these being the detoxifying property of the liver, which plays an important role in removal of activated clotting factors, both directly and after their combination with natural inhibitors. Furthermore, the removal of free thrombin occurs as a result of its adsorption onto fibrin or onto fibrin(ogen) degradation products. The latter, which cannot themselves take part in a clot, may interfere with normal fibrin polymerization. Although less well defined than the others, this group appears to be physiologically important, as qualitative defects of fibrinogen (dysfibrinogenaemias) that result in reduced thrombin binding may be associated with a thrombotic disorder (see Chapter 44).

Tissue factor pathway inhibitor: a Kunitz-type inhibitor

The various serpins found in blood (see below) probably play no physiological role in the inhibition of FVIIa in the initiating TF–FVIIa complex. Instead, the action of this serine protease–cofactor complex is modulated by TFPI, which is located on the endothelial cell surface, within platelets, in plasma and on monocytes. Two alternatively spliced isoforms of TFPI (α and β) have been identified in humans that differ in their domain structure and mechanism of cell-surface binding. TFPIα has an acidic N-terminal region followed by three Kunitz-type protease inhibitory domains (K1–K3) and a basic C-terminal region. K2 is responsible for binding and direct inhibition of FXa and K1 inhibits FVIIa in an FXa-dependent manner. K3 and the C-terminal region do not directly inhibit proteolysis but they are necessary for rapid inhibition of FXa by K2. Recent data intriguingly suggest that this may be mediated through an interaction with PS acting as a cofactor in the formation of the TFPI–FXa complex. TFPIα indirectly associates with the endothelial cell surface through a non-covalent association with a glycosylphosphatidylinositol (GPI)-anchored coreceptor. It can also bind non-specifically to endothelial cell GAGs via the basic C-terminal region; it is thought that this interaction is responsible for the twofold to fourfold increase in plasma TFPI following infusion of heparin. TFPIβ has the N-terminal acidic region followed by K1 and K2, but lacks K3 and the C-terminal region of TFPIα. Instead it has a different C-terminal region encoding a GPI-anchor attachment sequence that allows it to directly associate with the cell surface.

Most TFPI in the vasculature (~85%) is associated with endothelial cells, particularly in the microcirculation, with a smaller amount (5–10%) being found in platelets and the rest in plasma. Virtually all circulating TFPI is associated with lipoproteins and has undergone variable amounts of proteolytic degradation within its C-terminal region. As a result plasma TFPI has reduced anticoagulant activity and is unlikely to play an important role as an *in vivo* inhibitor of blood coagulation.

Disruption of the murine *TFPI* gene is incompatible with normal development, suggesting that a deficiency of functional TFPI might lead to clinical thrombosis. However, as the bulk of intravascular TFPI is associated with the endothelial cell surface, simple measurement of plasma TFPI levels is not a clear indicator of a deficiency state. Moderately low plasma levels can occur in DIC, septicaemia and following major surgery, possibly due to increased utilization, but heparin infusion increases plasma TFPI in such patients. Whether congenital deficiency of TFPI is a risk factor for thrombosis remains uncertain: after several negative studies in this area, recently a case–control study found a relationship between low levels of free plasma TFPI (non-lipoprotein bound) and enhanced risk of deep vein thrombosis and myocardial infarction.

Serine protease inhibitors (serpins) and heparin

Human plasma contains at least seven inhibitors of serine protease coagulation factors. Apart from α_2-macroglobulin, all are single-chain serpins of molecular mass 40–65 kDa. Specificity is imparted by their tertiary structure, which engenders high affinity for a defined substrate or small range of substrates. Of this group, only AT and heparin cofactor II (HCII) assume haemostatic significance, acting predominantly on proteases generated late in the coagulation cascade (i.e. thrombin and FXa). A deficiency of any of the other serpins (although sometimes giving rise to a clinical disorder due to the failure of neutralization of a serine protease not involved in coagulation pathways) is asymptomatic in terms of haemostasis.

Several serpins contain GAG-binding sites that, particularly in the case of AT and HCII, but to a lesser extent for PAI-1 and APC inhibitor, greatly enhance the rate of interaction with (although not the affinity for) their specific protease(s). Since a major source of heparin-like material is heparan sulfate on the endothelial cell surface, another possible function of the heparin-binding site of these inhibitors (especially AT and HCII) might be to allow them to exert a general and constitutive anticoagulant effect on intact vessels; this may also act to prevent extension of a procoagulant response beyond an area of damaged endothelium.

Antithrombin

AT (formerly ATIII) is a 58-kDa serpin that is synthesized principally in the liver, with a high plasma concentration of 2.4 μmol/L. The turnover of AT *in vivo* is complex, with a rapid initial clearance ($t_{1/2}$ 10 min) due to equilibration with endothelial cell-bound AT, a slower clearance ($t_{1/2}$ 3 hours) due to

equilibration with the extravascular compartment, and a much slower linear phase with an overall $t_{1/2}$ of 90 hours.

Disruption of the murine AT gene results in death of all embryos shortly before birth due to severe cardiac and hepatic thrombosis. In humans, AT deficiency is associated with familial thrombosis.

AT forms a stable complex with several serine protease coagulation factors, predominantly thrombin and FXa, but also to some extent FIXa, FXIa, FXIIIa and kallikrein. Initially, the serine at the active centre of the protease cleaves a peptide bond (involving Arg393) in the 'reactive centre loop' near the C-terminus of AT: a conformational change in the serpin ensues, resulting in irreversible trapping of the protease against the inhibitor. The inactivated enzyme–serpin complex is then cleared rapidly. Complex formation is progressive and only with thrombin and FXa is it rapid enough to be of physiological significance. In purified *in vitro* systems, the $t_{1/2}$ is around 30 s for free thrombin, 90 s for FXa and 10–25 min for the other enzymes. In each case, the inhibitor–protease complex is rapidly cleared ($t_{1/2}$ 3 min) from the circulation by the liver.

Heparin, without altering the stoichiometry, induces a more than 2000-fold increase in the rate of thrombin inactivation by AT, such that its action becomes almost instantaneous ($t_{1/2} < 0.01$ s). The unique mechanism of this activation involves release of the reactive centre loop from partial insertion in the β-sheet core of AT following heparin binding. Heparin also strongly enhances AT neutralization of FXa and, to a lesser extent, FIXa, particularly in the presence of Ca^{2+}. The enhancement of AT inhibition of thrombin and FXa by heparin is the basis of its use as a therapeutic anticoagulant to prevent thromboembolism. At therapeutic concentrations there is no significant effect on AT inhibition of FXIa and FIXa.

Heparin cofactor II

This 65-kDa serpin is present in plasma at the high concentration of 90 mg/L (1.2 μmol/L). It appears to be a specific inhibitor of thrombin and to have little or no anti-FXa activity. The rate of thrombin neutralization by HCII is increased approximately 1000-fold by heparin, although because of its lower heparin affinity it requires five to ten times more heparin than does AT. Interaction seems to depend largely on the high anionic charge of heparin and related GAGs. Details of the molecular mechanism of HCII interaction with thrombin have been elucidated by X-ray crystallography.

That HCII has some physiological significance is suggested by the fact that it falls in parallel with AT in DIC. However, as AT is in twofold molar excess over HCII, the latter cannot altogether compensate for a deficiency of AT, which, as stated above, is a well-established cause of a thrombotic tendency. Whether HCII deficiency leads to a similar clinical picture remains to be established, as few cases have yet been described and only occasionally has concomitant thrombotic disease been present. Mice with HCII knockout develop normally and do not show spontaneous thrombosis; however, they show an enhanced propensity to carotid occlusion after deliberate injury to the endothelium, corrected by infusion of purified HCII, suggesting that HCII has a role in prevention of arterial thrombosis.

Heparin and heparin-like substances in plasma

Although the term 'heparin' implies a single compound, it in fact refers to a heterogeneous mixture of sulfated polysaccharides that are all members of the GAG group. The major pharmaceutical source is porcine intestinal mucosa. By virtue of its strong positive charge, heparin combines non-specifically with a number of cationic proteins such as albumin and reacts in a highly specific way with β-lipoproteins, fibrinogen, HCII and AT. It also mobilizes platelet factor (PF)4 and TFPI, which are bound to GAGs on the surface of endothelial cells, and releases lipoprotein lipase into plasma.

Unfractionated heparin (UFH) is an extremely heterogeneous polymer, being composed of between 10 and 100 saccharide units. Only about one-quarter of the polysaccharides has any anticoagulant activity *in vitro*. The pharmaceutical properties of heparin are unpredictable, because of its heterogeneity, with marked variation in pharmacokinetics and side-effects, the most serious of which is heparin-induced thrombocytopenia (see Chapter 45). The realization that the anticoagulant action of UFH was mostly delivered by short pentasaccharides has led to the use of more efficient fractionation procedures (including cleavage by nitrous acid, heparinase and oxidizing agents), yielding a number of low-molecular-weight heparin (LMWH) preparations. LMWHs contain 10–20 saccharide units (4–6 kDa) and because of their more uniform composition have much better pharmacokinetics and fewer side-effects. A synthetic pentasaccharide has come into therapeutic use and offers increased half-life compared with larger LMWHs.

UFH and the various LMWHs differ in their affinities for the plasma factors with which they interact and in the specific serine proteases that they inhibit. This, together with the timing of blood samples in relation to dosage interval, is critically important in laboratory monitoring of the therapeutic effect (see Chapter 46). The major anticoagulant activity of both UFH and LMWH is attributed to the pentasaccharide sequence containing at its centre a 2,3,6-(SO_3) trisulfated glucosamine group. It is this that binds to lysine or tryptophan residues in AT in the region of the D helix, inducing the conformational change in the inhibitor. A second anticoagulant action of UFH (at least 18 saccharide units are required) involves direct binding between thrombin (but not FXa or other proteases) and a heparin sequence adjacent to the pentasaccharide. This brings thrombin and AT into close proximity and is necessary for the antithrombin activity. As many chains in a LMWH have fewer than 18 saccharide units, they have less antithrombin action compared with their inhibition of FXa. There is no detectable heparin in normal plasma. However, endogenous heparin-like molecules

(e.g. dermatan sulfate and heparan sulfate) are present on the surface of endothelial cells and, by enhancing the action of AT and HCII, these would have antithrombotic effects. Such mechanisms seem to be of clinical importance, as recurrent thromboses have been reported in association with several dysfunctional AT molecules that inhibit thrombin normally in the absence of heparin, but which do not show enhanced reactivity in its presence, presumably due to defects at the heparin-binding sites. At points of vascular injury, local accumulation of activated platelets provides a source of heparin-neutralizing activity that could overcome the anticoagulant effects of endothelial cell-bound heparinoids and permit the procoagulant response to proceed.

Protein Z and protein-Z-dependent inhibitor

PZ is a 62-kDa vitamin K-dependent plasma protein that serves as a cofactor for the inhibition of FXa by PZ-dependent protease inhibitor (ZPI) a 72-kDa member of the serpin superfamily.

The organization of the PZ gene and the structure of the molecule are very similar to those of coagulation factors FVII, FIX, FX and PC; however, the PZ 'serine protease' domain lacks the canonical active-site His and Ser residues and therefore cannot function as a protease. PZ circulates in plasma in a complex with ZPI. Inhibition of FXa by ZPI in the presence of phospholipid and Ca^{2+} ions is enhanced 1000-fold by PZ, but ZPI also inhibits FXIa in a process that does not require PZ, phospholipid or Ca^{2+}. ZPI activity is consumed during coagulation through proteolysis mediated by FXa (with PZ) and FXIa. PZ may serve to dampen the procoagulant response *in vivo* as PZ deficiency dramatically increases the severity of the prothrombotic phenotype of FV Leiden mice. Studies to determine the potential roles of PZ and ZPI deficiency in human thrombosis are in progress.

α_1-Antitrypsin

This is a serpin whose primary targets are pancreatic and leucocyte elastases. In coagulation, the major inhibitory activity is directed against FXIa and FXa, although the chief physiological inhibitor of FXIa is probably protease nexin 2. It has little effect on overall thrombin inhibition and a straightforward deficiency of α_1-antitrypsin is not associated with hypercoagulability.

An abnormal molecular form of α_1-antitrypsin in which there is a Met→Ser substitution at the active centre (antitrypsin Pittsburgh) has been described, resulting in a higher affinity for thrombin. At a plasma concentration of 25 μmol/L, the variant circulates at a 10-fold higher level than AT and gives rise to a clinical bleeding tendency. However, it is unaffected by heparin.

Protease nexin 2

This is a Kunitz-type serine protease inhibitor that is found in the α-granules of platelets. Normally, there is very little in plasma, but it is released at sites of platelet activation, following which the single Kunitz domain enters into a tightly bound inhibitory complex with FXIa. It is one of several isoforms of Alzheimer β-amyloid protein precursor present in platelets. Deficiency of protease nexin 2 has not been described.

C1-esterase inhibitor

The primary target of this serpin is the activated form of the first component of complement, but it also contributes in a minor way to the neutralization of FXIa and plasmin. Deficiency of C1-esterase inhibitor, although of no haemostatic consequence, causes angioneurotic oedema, the characteristic lesions of which may sometimes be confused with haematomas.

α_2-Antiplasmin

This serpin is the principal inhibitor of the fibrinolytic enzyme plasmin. It also has weak activity against several coagulation proteases, especially the contact factors. However, any anticoagulant action against proteases late in the coagulation cascade (e.g. FXa) is only apparent at concentrations well in excess of those in normal plasma. Its mechanism of action and clinical importance are discussed further in the section on fibrinolysis.

α_2-Macroglobulin

α_2-Macroglobulin is composed of four identical chains and has a molecular mass of 740 kDa. It is not a member of the serpin superfamily and its effects are not restricted to serine proteases. It binds to coagulation factors at a site away from the active site, the interaction involving the formation of a bond between cysteine and glutamate residues in the inhibitor and a lysine residue in the protease. Inhibition is produced by stearic hindrance rather than by active site inactivation and indeed the proteases retain some esterolytic and amidolytic activity, particularly against small peptides, a fact that should be borne in mind when using chromogenic substrates to assay coagulation inhibitors. It is responsible for approximately 50%, 20% and 10% of the inhibition of kallikrein, thrombin and FXa, respectively.

Deficiency of α_2-macroglobulin is not associated with a thrombotic tendency. However, it is an acute-phase reactant and it is possible that, when elevated under conditions of stress or when the other major antithrombins or antiplasmins are overwhelmed, it might become a significant inhibitor of coagulation or fibrinolysis. Moreover, it has been suggested that the raised level of α_2-macroglobulin that exists in children (150–200% of adult values) may compensate for a low level of AT, and explain why thrombotic episodes do not usually occur before puberty in congenitally AT-deficient patients.

The protein C pathway: inhibition of cofactors FVa and FVIIIa

The activated forms of coagulation cofactors FV and FVIII (FVa and FVIIIa) are potent procoagulants that enormously enhance the activity of serine protease factors in the tenase and prothrombinase complexes (see above). It is not unexpected that

these cofactors should be subject to a negative feedback mechanism that limits their procoagulant activity. This is achieved by a complex series of reactions collectively referred to as the protein C pathway. Four key factors are now known to be involved, and their role in the haemostatic network is shown in Figure 36.4.

Protein C

This vitamin-K-dependent serine protease has an identical modular composition to the procoagulant factors FVII, FIX and FX. It circulates in blood as a two-chain disulfide-linked molecule. During biosynthesis, the primary translation product is proteolytically processed, releasing the dipeptide Lys156Arg157. In order to exert its anticoagulant effect, PC must first be activated to APC. This is achieved by the action of thrombin, which cleaves the heavy chain to release a 12-residue (Gly158–Arg169) activation peptide, revealing the active site by the usual chymotrypsin-like mechanism. Thrombin activation of PC is slow in free solution, but is markedly accelerated by specific endothelial cell receptors for both thrombin (TM) and PC (EPCR), which coordinate the assembly of a membrane complex for PC activation.

APC interacts with PS bound to the phospholipid surface of activated platelets, enhancing the anticoagulant activity of APC (see below) against FVa and FVIIIa. These procoagulant cofactors are inactivated by APC on the platelet surface by specific cleavage in their A domains, terminating the activity of the tenase and prothrombinase complex by disrupting their binding sites for FIXa and FXa respectively. In the absence of PS, this reaction is inefficient. The most common form of familial thrombosis is FV Leiden, which is caused by mutation of Arg506 to Gln, which makes the molecule resistant to cleavage by APC. It appears that APC inactivation of FVa is a prerequisite for efficient inactivation of FVIIIa *in vivo*, as FVIIIa inactivation is also impaired in these patients, although there is no abnormality in the FVIII molecule. Since the discovery of FV Leiden, other less common FV mutants have been found with a similar phenotype. None of the mutations reported in the FVIII molecule have been shown to cause a similar prothrombotic phenotype. PC is synthesized in the liver, and, being vitamin-K-dependent, is often low in the newborn. Even though its substrates (FV and FVIII) are normal at birth, this deficiency is compensated for by the reduction in plasma of the vitamin K-dependent procoagulant factors (prothrombin, FVII, FIX and FX). Because warfarin treatment initially decreases PC levels faster than FVII, FIX, FX and prothrombin levels, it can paradoxically increase the procoagulant tendency when anticoagulant treatment is first begun (many patients starting on warfarin are given heparin in parallel to combat this). Disruption of the murine PC gene results in lethal perinatal consumptive coagulopathy, whereas in humans homozygous PC deficiency is associated with lethal purpura fulminans (in the absence of PC replacement therapy) and heterozygous individuals have a high risk of venous thrombosis. There is increasing appreciation of the role of the PC pathway in regulation of inflammation and the concept of signalling mechanisms allowing 'cross-talk' between the haemostatic and inflammatory networks. This is well demonstrated by the protective effect of administration of APC concentrate to animals with induced sepsis. However the concentrate has not proved to be quite so efficacious in human studies. Recently, variant recombinant PC molecules have been produced that appear to have either anticoagulant or anti-inflammatory effects independent of each other.

Thrombomodulin

TM is an integral transmembrane receptor found on endothelial cells in virtually all body tissues. It appears to be absent in the brain vasculature and in hepatic sinusoids and lymph node venules. TM is essential for normal fetal development as shown by mouse knockout studies, although $TM^{+/-}$ heterozygous mice appear to be completely healthy.

The protein has an extracellular region composed of an N-terminal lectin-like or CLECT domain, six EGF-like domains, a transmembrane sequence and a small cytoplasmic region. TM forms a complex with thrombin (via the protease's anion-binding exosite I and TM's EGF-like domains 4–6), preventing binding of the protease to its various procoagulant substrates (fibrinogen, FV, FVIII, FXIII and protease-activated receptors involved in platelet aggregation). TM also plays a part in binding of PC zymogen and, after formation of the TM–thrombin complex on the cell surface, there is a 20,000-fold increase in the rate of activation of PC, so that thrombin effectively becomes an anticoagulant. Binding of PC to TM is also enhanced by EPCR.

The TM–thrombin complex is short-lived, being endocytosed by endothelial cells, where the thrombin is taken up and degraded by lysosomes while TM re-circulates to the cell membrane.

Endothelial protein C receptor

EPCR is a transmembrane receptor on endothelial cells that binds PC and promotes its activation by the thrombin–TM complex. The extracellular portion of EPCR is related to the MHC class I/CD1 proteins; however, it lacks the α3 domain found in the extracellular portion of most of the family: there is good evidence that the α1 and α2 extracellular domains of EPCR interact with the GLA domain of PC. EPCR may function by localizing the PC molecule to the endothelial cell surface, moving laterally on the cell surface to locate a thrombin–TM complex, then presenting the PC molecule optimally for thrombin cleavage to form APC.

The essential role of EPCR was demonstrated by targeted deletion of the murine EPCR gene, resulting in early fetal death, with evidence of thrombosis at the maternal–fetal interface. Abrogation of EPCR function *in vitro* reduces the rate of PC activation by thrombin–TM. In contrast, inhibition of EPCR action by a specific anti-EPCR monoclonal antibody in baboons given thrombin resulted in reduced APC levels amid evidence

of excessive coagulopathy. Blockade of EPCR on challenge with *Escherichia coli* also led to worse outcomes in an animal model of sepsis, whereas administration of APC improved survival in experimental sepsis.

Protein S

PS is a single-chain vitamin-K-dependent glycoprotein chiefly synthesized in the liver by endothelial cells. Unlike the other vitamin-K-dependent coagulation proteins, it is not a serine protease, having two C-terminal LamG domains rather than a serine protease domain, and the N-terminal GLA domain is followed by four EGF-like domains rather than the two usually seen in the serine proteases (see Figure 36.8). About 40% of the PS in plasma is in the free form, whereas the remaining 60% is associated in a complex with C4b-binding protein (C4bBP). Both forms bind strongly via the PS GLA domain to negatively charged phospholipids exposed on the surface of activated platelets. Free PS (although having no strong inhibitory effect on FVa or FVIIIa itself) forms a Ca^{2+}-dependent complex with APC, helping to orient the APC active site above the phospholipid surface and enhancing its anticoagulant activity against both proteins. However, C4bBP-bound PS does not enhance PC function against FVa. Free thrombin cleaves PS between the GLA domain and EGF1, removing its ability to bind to both phospholipid and PC and thereby abolishing its PC cofactor activity.

Protein C inhibitors

As with other serine proteases, APC is subject to inhibition by serpins (half-life of APC in plasma is 15–20 min), including APC inhibitor (PCI), PAI-1 and α_1-antitrypsin. The relative contributions of these *in vivo* are difficult to assess. PCI slowly but progressively blocks the action of APC (and, to a much lesser extent, thrombin and FXa), in each case forming a 1:1 complex; in addition, PCI action is enhanced 20-fold by heparin and more weakly by other GAGs, but the physiological significance of this is uncertain. Despite the name, disruption of the murine PCI gene leads not to coagulopathy, but to infertility due to lack of inhibition of another protease target of the serpin, the sperm protein acrosin.

Fibrinolysis

It is widely acknowledged that the principal functions of the fibrinolytic system are to ensure that excess fibrin deposition is either prevented or rapidly removed (i.e. that a localized procoagulant response is achieved without compromising blood circulation generally) and, following re-establishment of haemostasis, the fibrin mesh is removed during wound healing. The system of profibrinolytic and antifibrinolytic factors that has evolved to meet these requirements is closely coupled to that which results in fibrin clot formation. Fibrinolysis is essentially a localized, surface-bound phenomenon, with most events being catalysed by the presence of cross-linked fibrin itself, i.e. 'fibrin orchestrates its own destruction'. For this reason, the assays of fibrinolytic factors carried out in the soluble phase, in particular in systemic blood, may be misleading and should be interpreted with great caution.

Components of the fibrinolytic system

These include plasminogen (PLG) and plasmin, several endogenous (tissue or plasma derived) or exogenous (e.g. bacterial or venom derived) PLG activators, and a number of inhibitors of plasmin or of the PLG activators. Both endogenous and exogenous fibrinolytic factors have been used clinically to treat venous and arterial thrombosis, with varying degrees of success (see Chapter 47).

Plasminogen and plasmin

PLG is a single-chain glycoprotein zymogen of the serine protease plasmin, which carries out the enzymatic degradation of cross-linked fibrin. Besides its active site serine, plasmin contains five kringle domains, four of which have a lysine-binding site, through which the molecule interacts with lysine residues in its substrates (e.g. fibrin), its activators (e.g. tissue PLG activator and urinary PLG activator) and its inhibitors (principally PLG activation inhibitor type 1). In its native form, it has a glutamic acid residue at its N-terminus and is known as Glu-PLG. Conversion of PLG to plasmin can proceed via two routes. Most PLG activators (see below) cleave the Arg561–Val562 bond to form Glu-plasmin, a disulfide-linked two-chain molecule. The heavy chain is derived from the N-terminal region and bears the lysine-binding sites, whereas the C-terminal light chain contains the serine active centre. Glu-plasmin, despite being a serine protease, is functionally ineffective, as its lysine-binding sites remain masked. It is converted autocatalytically to Lys-plasmin by N-terminal cleavage, chiefly between Lys76 and Lys77, which exposes the lysine-binding sites and thus markedly enhances its interaction with fibrin.

Both Glu- and Lys-plasmin also attack the same Lys76–Lys77 bond in Glu-PLG to form the zymogen Lys-PLG. This binds to fibrin before activation to the protease and is thus brought into close proximity with the physiological PLG activators (which also bind to fibrin) that convert it to Lys-plasmin. As a consequence, the conversion of PLG to plasmin by tPA is enhanced by two to three orders of magnitude; this serves to localize the fibrinolytic response to the fibrin clot, where plasmin is to some extent protected from the effects of circulating antiplasmins, which (as indicated below) would otherwise neutralize plasmin extremely rapidly (<50 ms). The fact that Lys-PLG is potentially a much more effective agent in fibrinolysis than Glu-PLG is reflected in its half-life, which is around 20 hours compared with 50 hours for the latter.

Action of plasmin on fibrin and fibrinogen

Plasmin can hydrolyse a variety of substrates, including FV and FVIII, but its major physiological targets are fibrin and fibrinogen, which are split progressively into a heterogeneous mixture of small soluble peptides (plasmin attacks at least 50 cleavage sites in fibrinogen) known collectively as fibrin degradation products (FDPs). The first stage in the proteolysis of fibrinogen involves the removal of several small peptides (fragments A, B and C) from the C-terminus of the A α-chains, each involving cleavage after a lysine residue. This is rapidly followed by removal of the first 42 amino acids from the N-terminal end of the B β-chain (the Bβ1–42 fragment). The large residual portion, which is known as fragment X, and which still contains fibrinopeptide A, remains thrombin-clottable and will agglutinate some species of *staphylococcus* spp. Assay of the Bβ1–42 fragment released from fibrinogen by plasmin gives a sensitive index of fibrinogenolytic activity.

Asymmetrical digestion of all three pairs of chains of fibrin or fibrinogen then occurs with the release of the D fragment, in which the chains remained linked by disulfide bonds. The residue, known as fragment Y, is again attacked by plasmin, cleaving a second fragment D and leaving the disulfide-linked N-terminal ends of all six chains, which are referred to as fragment E. Fragments Y, D and E are not thrombin-clottable and do not agglutinate staphylococci. Their presence can be detected immunologically using an antibody-coated latex bead agglutination assay, which provides a simple test for most FDPs, although carefully prepared serum must be used to prevent cross-reactivity of the antibody with fibrinogen in plasma. These assays detect the degradation products of both fibrin and fibrinogen indiscriminately.

Following thrombin generation and consequent activation of FXIII, intermolecular or intramolecular transamidation of the α- or β-chains by FXIIIa occurs and then the action of plasmin yields characteristic D-dimer, D-dimer–E fragments and oligomers of fragments X and Y (collectively known as cross-linked FDP or XDP), in addition to X, Y, D and E. These XDPs can be detected very simply using monoclonal-antibody-coated latex beads. Because the monoclonal antibodies to XDPs do not cross-react with fibrinogen, they can be detected directly in citrated plasma. The presence of D-dimers in blood samples can be used in a clinical algorithm that predicts the likelihood of the presence of venous thrombosis (see Chapter 47).

Furthermore, plasmin-induced cleavage of the N-terminal end of the β-chain of fibrin (the Bβ1–14 fibrinopeptide B fragment having been removed by thrombin) produces a β15–42 fragment, the detection of which indicates fibrin (as opposed to fibrinogen) degradation. Consequently, assays for the Bβ1–42 and β15–42 fragments used in combination may be clinically useful by indicating whether fibrinogen or fibrin has been degraded, and thus whether fibrinolytic activity is primary or secondary to fibrin formation. However, clinically, FDP assays are used to detect DIC, when mixed fibrin/fibrinogen degradation products appear in the circulation.

Plasminogen activators

Tissue plasminogen activator

tPA is a serine protease secreted by endothelial cells. It is not synthesized by the liver or kidney, but is found in most extravascular body fluids, including saliva, milk, bile, cerebrospinal fluid and urine. Intravascular tPA is quickly cleared by the liver or inactivated by the fast-acting tPA inhibitor (see below), the half-life of tPA in plasma being approximately 2 min. The resting level of tPA in plasma is around 70 pmol/L, most of which is in an inactive complex with tPA inhibitors (see below).

Several physical and biochemical stimuli, including venous occlusion, strenuous exercise, thrombin, adrenaline and vasopressin or its analogues such as desmopressin (see Chapter 38), markedly increase the rate of tPA release, although its biological activity remains negligible until it becomes bound to fibrin, whereupon its affinity for and action upon PLG is greatly potentiated. The activity of tPA is further enhanced by plasmin itself, which cleaves tPA at Arg275-Ile276 into a two-chain molecule, whose binding sites are exposed, thus enabling it to form a complex with PLG and fibrinogen more readily. The ability of venous occlusion to stimulate tPA release from endothelial cells forms the basis of a test of fibrinolytic activity known as the 'cuff test'.

Both single-chain tPA and the two-chain form possess very little serine protease activity until they bind to fibrin, whereupon tPA affinity for, and activation of, fibrin-bound PLG is increased at least 100-fold. The principal interactions involve binding between the second kringle domain of tPA and lysine residues on the α- and β-chains of fibrin (in particular Lys157 on the α-chain of partly degraded fibrin). This association enhances cleavage of the Arg561–Val562 bond in adjacent PLG molecules, forming active plasmin.

Urinary plasminogen activator

So called because it was first extracted from urine, uPA is synthesized chiefly by the tubules and collecting ducts in the kidney and by fibroblast-like cells in the gastrointestinal tract. It is a serine protease secreted as an inactive single-chain zymogen (prourokinase) that is cleaved by activators in plasma (including kallikrein and plasmin) at Lys158–Ile159 to produce active two-chain uPA. The protease activity of uPA is associated with the heavy chain, which may dissociate from the light chain carrying the PLG-binding site. The isolated heavy chains, which are also known as low-molecular-weight urokinase, are therefore poorer PLG activators than the two-chain form.

uPA cleaves the same Arg561–Val562 bond in PLG as tPA. Although uPA contains a kringle domain, this does not impart high affinity for fibrin and it binds instead (via its EGF-like domain) to its cell-associated receptor. Thus it has been

proposed that uPA may be preferentially involved in cellular events (such as differentiation and mitogenesis) rather than with dissolution of fibrin clots.

Exogenous plasminogen activators

These are derived from non-human sources, including animals (e.g. vampire bat saliva and some snake venoms) and certain plants and microorganisms. The best known of these is streptokinase (SK), which is derived from some strains of β-haemolytic streptococci and which has for many years been used, with moderate success, as a fibrinolytic agent for the treatment of life-threatening thrombotic states. SK is a non-enzymatic polypeptide that forms a stable complex with PLG, as a result of which the latter undergoes a conformational change, unmasking its serine-active centre. The 'plasmin' that is formed remains associated with SK, but can convert free PLG to plasmin.

Inhibitors of fibrinolysis

The plasmin-generating potential of plasma is sufficient to completely degrade all the fibrinogen in the body in a very short period of time. It is prevented from doing so by the PLG activator inhibitors or PAIs, most of which belong to the serpin family, and by a number of circulating inhibitors of plasmin itself (the antiplasmins).

Inhibitors of plasminogen activation

Plasminogen activator inhibitor type 1

PAI-1 is an important fast-acting serpin inhibitor of tPA, uPA and, to a small extent, plasmin, and is secreted by endothelial cells. It is also found in platelet α-granules. In plasma, it occurs in two forms: a functionally active 'free' form (that is stabilized by association with vitronectin) and as an inactive complex with tPA. Basal PAI-1 concentration in plasma is low at 0.5 nmol/L, of which at least 80% is in complex with tPA or uPA. It follows a diurnal rhythm, with an early morning peak that is around twice that in the late afternoon, and its activity is also increased by heparin. Elevated levels of PAI-1 are associated with an increased incidence of venous and arterial thrombosis, and there is a suggested association between the early morning peak level of PAI-1 and a higher incidence of myocardial infarction at that time, the extent of this diurnal variation being associated with polymorphisms in the PAI-1 gene (*SERPINE1*).

The three main profibrinolytic serine proteases (tPA, uPA and plasmin) all cleave the same bond (Arg346–Met347) in the reactive centre loop of PAI-1, and are thus inhibited in the same way as described in the section on AT. PAI-1 binds non-covalently to fibrin, but although it can then complex with and inhibit fibrin-bound tPA (with most of the complexes remaining bound to the fibrin), it does so less effectively than with free tPA. Soluble tPA–PAI-1 complexes are rapidly removed by the liver ($t_{1/2}$ 4 min), as are uPA–PAI-1 complexes that have dissociated from fibrin.

Plasminogen activator inhibitor type 2

This serpin inhibitor of tPA is mainly produced by the placenta and may thus contribute to the inhibition of fibrinolysis that occurs during pregnancy. It is also synthesized in monocytes and epidermal cells, but is not usually found in the plasma of non-pregnant subjects. It is detectable in plasma from about the eighth week of pregnancy, rising to a peak at around 33 weeks and falling only slowly after delivery, the half-life being around 24 hours. Paradoxically, levels are often low in pre-eclampsia due to placental insufficiency. The inhibitory action of PAI-2 involves its Arg380–Thr381 residues and it is more effective against uPA than tPA, although for both the potency is at least 10-fold less than that of PAI-1.

Similarly, other protease inhibitors such as α_1-antitrypsin, C1-esterase inhibitor, α_2-antiplasmin (see below), α_2-microglobulin and protease nexin 1 (one of a group of cell membrane-bound heparin-potentiated serpins) also neutralize tPA, but at a rate that is probably too slow to be of physiological significance. However, α_2-microglobulin is thought to be the major inhibitor of the SK–PLG complex.

Inhibitors of plasmin

As they do with thrombin and tPA, a number of the broad-spectrum inhibitors contribute to neutralization of plasmin. By far the most potent plasmin inhibitor is the serpin α_2-antiplasmin, a single-chain glycoprotein synthesized by the liver, which has a half-life of about 60 hours and shows considerable sequence identity with AT and α_1-antitrypsin. Its physiological importance is supported by the fact that a congenital deficiency (known as Miyasato disease) is associated with a clinically significant bleeding disorder due to uncontrolled fibrinolytic activity, and that levels are reduced in DIC and during thrombolytic therapy.

α_2-Antiplasmin

This serpin is the predominant plasmin inhibitor. It forms a stable complex with plasmin, in which the protease is completely inactivated. The reaction appears to involve the cleavage by plasmin of a specific Leu–Met bond in the inhibitor, exposing the reactive loop Arg364–Met365 peptide bond to the serine-active centre on the light chain of plasmin. Plasmin-modified α_2-antiplasmin can also bind to native Glu-PLG and to fibrin, the latter reaction being mediated by FXIIIa-mediated cross-linking. Thus, in addition to inactivating preformed plasmin, α_2-antiplasmin retards fibrinolysis by reducing PLG activation and by 'masking' the lysine-binding sites through which plasmin(ogen) interacts with fibrin. However, these subsidiary mechanisms for inhibiting fibrinolysis are to some extent overcome by any Lys-plasmin(ogen) present, which has a higher affinity than Glu-PLG for fibrin, and is thus less susceptible to the action of α_2-antiplasmin.

In plasma (as opposed to on fibrin strands), although the concentration of PLG (~2 μmol/L) exceeds that of α_2-antiplasmin

(~1 μmol/L), basal fibrinogenolytic activity is minimal because the tiny amounts of plasmin normally generated under physiological conditions are rapidly neutralized by the inhibitor. However, in certain pathological conditions (e.g. obstetric emergencies or snake bite) where extreme activation of fibrinolysis occurs, the latter may be swamped. Under these circumstances, other inhibitors, in particular α_2-microglobulin and histidine-rich glycoprotein, may become clinically important. The action of α_2-microglobulin on plasmin is similar to its effect on thrombin. Following the plasmin-induced cleavage of a specific Arg–Leu bond in the inhibitor, the latter forms a complex with the light chain of plasmin. The serine active site of plasmin is not involved and the complex retains weak biological activity, albeit only briefly, until it is removed in the liver. Histidine-rich glycoprotein inhibits fibrinolysis by blocking the lysine-binding sites of PLG, thus preventing its interaction with fibrinogen.

Lipoprotein A

The protein portion of lipoprotein A is termed apo(a). It is synthesized in the liver and circulates in plasma. There is considerable structural homology with PLG, as it possesses both serine protease and kringle domains. It can compete with PLG for binding sites on fibrin(ogen) or tPA, and may also increase PAI-1 expression, both actions potentially inhibiting fibrinolysis. That lipoprotein A has some clinical importance is indicated by the finding that raised levels are associated with an increased incidence of thrombosis.

Thrombin-activatable fibrinolysis inhibitor

In the presence of thrombomodulin, thrombin activates carboxypeptidase B, also called TAFI, and TAFIa in turn inhibits fibrinolysis; this provides another link between coagulation and the fibrinolytic pathway. TAFI removes the C-terminal lysine residues formed by limited plasmin proteolysis of fibrin, removing the binding sites for PLG and tPA. Thus the fibrin cofactor function in PLG activation is reduced, downregulating fibrinolysis. Mice with TAFI knockout have defective wound repair, and data from backcrossing against heterozygous PLG-deficient mice showed that TAFI modulates both fibrinolysis and cell migration *in vivo*.

Selected bibliography

Aird WC (2003) Endothelial cell heterogeneity. *Critical Care Medicine* **31**: S221–S230.

Bajaj MS, Birktoft JJ, Steer SA *et al.* (2001) Structure and biology of tissue factor pathway inhibitor. *Thrombosis and Haemostasis* **86**: 959–72.

Coagulation sequence and structure database. Available at: http://www.coagbase.org/

Dahlback B, Villoutreix BO (2003) Molecular recognition in the protein C anticoagulant pathway. *Journal of Thrombosis and Haemostasis* **1**: 1525–34.

Esmon CT (2000) The endothelial cell protein C receptor. *Thrombosis and Haemostasis* **83**: 639–43.

European Association for Haemophilia and Allied Disorders. Available at: http://www.eahad-db.org/

Huntington JA (2003) Mechanisms of glycosaminoglycan activation of the serpins in hemostasis. *Journal of Thrombosis and Haemostasis* **1**: 1535–49.

Morrissey JH (2001) Tissue factor: an enzyme cofactor and a true receptor. *Thrombosis and Haemostasis* **86**: 66–74.

Ngo JC, Huang M, Roth DA, Furie BC, Furie B (2008) Crystal structure of human factor VIII: implications for the formation of the factor IXa–factor VIIIa complex. *Structure* **16**: 597–606.

Ruggeri ZM (2003) Von Willebrand factor, platelets and endothelial cell interactions. *Journal of Thrombosis and Haemostasis* **1**: 1335–42.

Shen BW, Spiegel PC, Chang CH *et al.* (2008) The tertiary structure and domain organization of coagulation factor VIII. *Blood* **111**: 1240–7.

Weiler H, Isermann BH (2003) Thrombomodulin. *Journal of Thrombosis and Haemostasis* **1**: 1515–24.

CHAPTER 37

The vascular function of platelets

37

Stephen P Watson, Neil V Morgan and Paul Harrison

Centre for Cardiovascular Sciences, Institute for Biomedical Research, College of Medical and Dental Sciences, University of Birmingham, Birmingham, UK

Introduction

Platelets circulate in the bloodstream in a quiescent state, but undergo rapid activation as and when necessary. To achieve this balance, endothelial cells constitutively release nitric oxide (NO) and generate prostacyclin to inhibit platelet activation through elevation of the cyclic nucleotides cGMP and cAMP, respectively. In addition, endothelial cells express the ecto-ADPase CD39 and thrombomodulin. Under normal circumstances, platelets are unable to form stable contacts on endothelial cells, but this can occur on diseased or activated endothelial cells or at sites of disturbed flow, including arterial bifurcations where plaques are commonly formed. Adhesion of platelets to endothelial cells promotes local inflammatory events that can initiate plaque formation, endothelial cell activation and damage. Lesions in the endothelial layer lead to exposure of subendothelial matrix proteins, particularly collagens and membrane surface-expressed tissue factor, which can initiate haemostasis.

Platelets have a much wider role than simply supporting aggregation (Figure 37.1). Platelet dense granules and α-granules are packed with a rich diversity of small molecules and proteins that play fundamental roles in many aspects of haemostasis, including vessel constriction, leucocyte recruitment and vessel repair, as well as in other pathways, including host defence. In addition, platelets release and support the formation of a variety of active lipid mediators including sphingosine 1-phosphate, TxA_2 and lysophosphatidic acid. The ability to generate such a cocktail of biological molecules may reflect the evolutionary relationship of the platelet to the haemocyte in lower organisms, which is involved in both the innate defence system and haemostasis. In higher organisms, two distinct sets of cells perform these functions, namely thrombocytes (or platelets) and leucocytes.

Platelets are a major target in the treatment of individuals at risk of arterial thrombosis. However, careful consideration is required as to whether a patient should receive antiplatelet therapy and for how long, bearing in mind that such treatment always carries a risk of increased bleeding. For the majority of individuals receiving the orally active antiplatelet drugs aspirin or clopidogrel this risk is low, but it is not zero. The decision on whether a patient should receive antiplatelet therapy is based on the net sum of risk factors for arterial thrombosis, such as age, stress, weight, sex, lifestyle, cholesterol level, blood pressure, smoking and previous thrombotic history, and the potential risk from excessive bleeding. It is now accepted that individuals judged to be at medium to high risk of thrombosis should receive some form of antiplatelet therapy (unless otherwise contraindicated), usually low-dose aspirin, and nearly always in combination with other treatments such as statins and blood-pressure-lowering drugs. There is also a small but significant advantage in using clopidogrel and aspirin in combination for the treatment of patients at high risk of thrombosis, but with an increased risk of bleeding. Stronger inhibitors of platelet function, namely blockers of the major platelet integrin αIIbβ3 (GPIIb/IIIa) are used only in acute situations of thrombotic risk in the clinic, for example in cases of unstable angina or angioplasty and stenting, because of their narrow therapeutic window.

Platelet structure and organelles

Platelets are discoid in shape, with dimensions in the human of approximately 3.0 × 0.5 μm and a mean volume of 7–11 fL. This

Postgraduate Haematology, Seventh Edition. Edited by A Victor Hoffbrand, Douglas R Higgs, David M Keeling and Atul B Mehta.

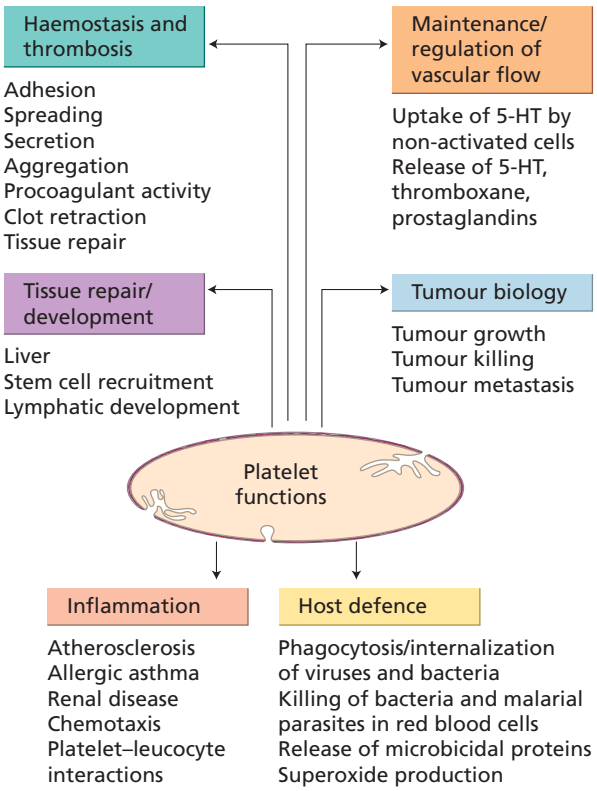


Figure 37.1 Functional roles of the platelet in the vasculature.

shape and small size enables platelets to be 'marginated' by red blood cells to the edge of the vessel, placing them in contact with vascular endothelial cells. The platelet count in humans is usually in the range $150–400 \times 10^9$/L. This relatively high count represents a considerable degree of redundancy for haemostasis, as it is usually when the platelet count reaches as low as 20×10^9/L that major spontaneous bleeds are seen. Low platelet counts are also associated with increased risk of excessive bleeding during major trauma and surgery.

The discoid shape of the platelet is regulated by the platelet cytoskeleton, which consists of a spectrin-based membrane skeleton, circumferential bands of a single microtubule that lie beneath the plasma membrane and a rigid actin filament network that fills the cytoplasm of the cell. The rigid structure and platelet strength is supported by 2 million copies of actin per platelet, of which approximately 40% in a non-stimulated cell are assembled into actin polymers. These polymers connect with each other and with the cytosolic tail of the membrane glycoprotein (GP)Ibα via filamin in a lattice-like structure.

Platelets lack a nucleus and have a short lifespan of 10 days. They have a very limited ability to undergo protein synthesis, as the result of residual mRNA that has been carried over from their precursor cell, the megakaryocyte. However recent evidence suggests that activated platelets can synthesize a small number of key proteins that may be of functional significance. The presence of low levels of platelet mRNA provides an opportunity to perform a limited number of genetic studies, although concern over contamination from nucleated cells, which express several orders of magnitude higher levels of mRNA, necessitates the need for direct confirmation of protein expression. Despite these concerns, the application of both genomic and proteomic approaches have given an unprecedented insight into the protein composition of the platelet, although the challenge remains to identify the function and interplay of the key proteins, as many may be vestigial or have been carried over from the megakaryocyte.

Platelets contain four main types of storage granule, dense granules, α-granules, lysosomes and peroxisomes, and several mitochondria. There are between five and nine dense granules in platelets, which contain high levels of ADP, ATP, polyphosphates, 5-hydroxtryptamine (serotonin) and Ca^{2+}. There are approximately 80 α-granules per platelet and these contain a rich diversity of proteins and membrane receptors that support haemostasis, vascular repair, inflammation and host defence (Table 37.1). Major components of α-granules include clotting factors such as fibrinogen, VWF, factor (F)V, protein S and tissue factor pathway inhibitor (TFPI), the chemokines SDF-1α, PF4,

Table 37.1 Platelet α-granule constituents.

Physiological role	Constituent
Angiogenesis	VEGF-A, VEGF-C, PDGF
Antibodies	IgG
Coagulation cascade	Factor V, fibrinogen, tissue factor pathway inhibitor
Endothelial cell activation	TGF-β
Fibrinolysis	Plasminogen, PAI-1, α_2-antiplasmin
Growth factors	PDGF, FGF, HGF, IGF-1, EGF
Leucocyte recruitment	Chemokines: PF4, RANTES, β-thromboglobulin, ENA-78, SDF-1α
Matrix breakdown	Hydrolytic enzymes MMP-2, MMP-9
Membrane proteins	$\alpha_{IIb}\beta_3$, P-selectin, CD40L
Bacterial killing	Microbicidal proteins
Miscellaneous	Amyloid β-protein precursor, Gas6
Proteases	Protease nexin II
Platelet aggregation	Fibrinogen, fibronectin, VWF
Irreversible aggregation	Thrombospondin (locks fibrinogen bridges between $\alpha_{IIb}\beta_3$)

Examples of platelet α-granule constituents and their physiological roles are shown. Several other molecules have also been reported to be present in α-granules and to be released on activation.

β-thromboglobulin, and RANTES, the growth factors platelet-derived growth factor (PDGF) and vascular endothelial growth factor (VEGF) A and C. Platelet α-granules also express key transmembrane proteins including integrin αIIbβ3, P-selectin (CD62) and CD40L that are only expressed on activated cells.

Platelets are enriched in signalling and cytoskeletal proteins that enable them to undergo dramatic changes in morphology and which enable the aggregate to withstand the high shear forces of the vasculature. Platelets have a network of intracellular membranes known as the dense tubular system that release intracellular Ca^{2+} in response to the second messenger inositol 1,4,5-trisphosphate (IP3). They also have a network of invaginations of the surface membrane, known as the surface-connected canalicular system (SCCS), which increase the surface area of the plasma membrane during platelet spreading. The SSCS also gives rise to membrane tethers that play a vital role in supporting adhesion.

Animal models

Genetically modified mice and other organisms, notably zebrafish, are used to address the *in vitro* and *in vivo* function of specific genes and proteins in platelets (and thrombocytes), but with the caveat of several key differences to human platelets with regard to rheology and protein composition. For example, the mouse genome lacks the gene for the only Fc receptor found on human platelets, FcγRIIA, and mouse platelets do not express one of the major receptors for thrombin on human platelets, PAR-1. There are also important differences in participation of protein isotypes in signalling cascades and various responses. Moreover, the relevance of many of the *in vivo* models (such as tail bleeds and vessel wall injury by ferric chloride or laser ablation) to haemostasis and thrombosis in humans is uncertain and fails to mimic the conditions that give rise to arterial thrombosis, which typically occurs on ruptured atherosclerotic plaques. Nevertheless, the processes that govern platelet activation in mice and other species appear to be shared with humans, and the value of animal models in analysing platelet activation and *in vivo* haemostasis and thrombosis, and thereby establishing a foundation for ongoing and future clinical trials, is immense.

Platelet formation

Platelets are formed from megakaryocytes, which can reach more than 50 μm in diameter. The nucleus undergoes a process known as endomitosis that involves nuclear replication without cellular division, giving rise to DNA ploidy values that range from 4*n* to 128*n*. The reason why endomitosis occurs is not fully understood, but it may simply reflect the need to increase the DNA content to enable the cell to expand its protein synthetic capacity to generate 2000–3000 platelets per megakaryocyte. In addition, it allows cell growth and differentiation to occur without interruption by nuclear and cell division.

The differentiation of bone marrow progenitor cells into megakaryocytes (MK) is regulated by the cytokine thrombopoietin (TPO). The TPO receptor, c-Mpl, which is expressed on stem cells, megakaryocytes and platelets, signals via the JAK family of tyrosine kinases. An acquired mutation in JAK2 is responsible for myeloproliferative disorders (see Chapter 26). Platelet lifespan is determined by an internal molecular clock governed by time dependent degradation of BCL-X_L which normally restrains BAK-1 and BAX from inducing apoptosis. In addition platelets also undergo desialylation as they age resulting in their eventual removal via the hepatic Ashwell-Morell receptor and induction of TPO synthesis. Circulating free levels of TPO are additionally controlled by binding and internalization mediated through platelet and MK c-Mpl. The two feedback mechanisms provide a simple means of tightly controlling the platelet count. Thus, if the platelet count decreases, or platelets are desialylated, the free level of TPO rises and there is an increase in megakaryocytopoiesis and platelet formation. Various TPO mimetics are in clinical use for treatment of various forms of thrombocytopenia.

The last few years have seen considerable advances in our understanding of the events leading to platelet formation. There is recognition that megakaryocyte differentiation occurs in a defined compartment of the bone marrow known as the osteoblastic niche and that megakaryocytes subsequently migrate from this to sinusoidal endothelial cells. At this vascular niche, the megakaryocytes generate long thin processes known as proplatelets, which form platelets at their terminals. The proplatelet arms protrude between bone marrow sinusoidal endothelial cells and release preplatelets (which further divide into platelets) and platelets directly into the bloodstream. Approximately 1 million platelets are released into the blood every second.

Thrombus formation

The events that underlie platelet adhesion and aggregation, and which lead to thrombus formation, at intermediate to high rates of shear (1000–5000/s) of the arteriolar system are discussed below and depicted in Figure 37.2.

Platelet capture and stable adhesion

The initial event that takes place on damage to the vasculature is the tethering or capture of platelets through the interaction of VWF with the GPIb–IX–V complex. VWF and GPIb do not interact unless a conformational change is induced in VWF as a consequence of its binding to collagen and/or elevated shear.

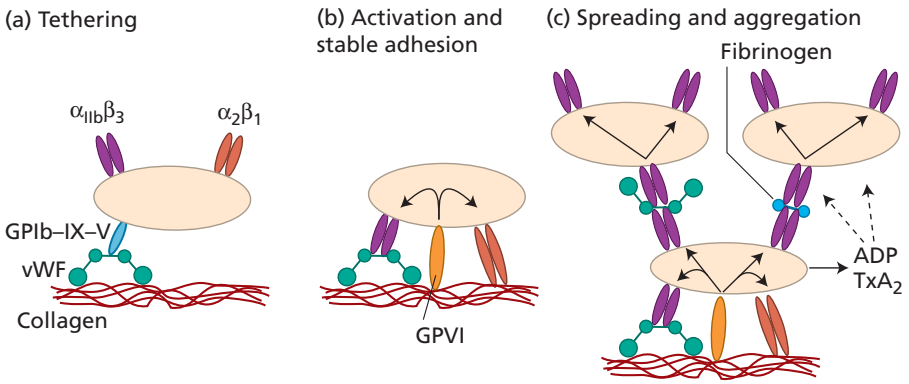


Figure 37.2 Thrombus formation at arteriolar rates of shear. Platelets are tethered by VWF bound to immobilized collagen, which then activates the low-affinity receptor GPVI leading to activation of integrins $\alpha_{IIb}\beta_3$ and $\alpha_2\beta_1$, which bind to VWF/fibrinogen and collagen, respectively. This mediates stable adhesion and further activation of GPVI. The signals from GPVI and the two platelet integrins induce spreading of platelets on the matrix and release of the feedback messengers ADP and TxA_2. VWF and fibrinogen, in combination with ADP, TxA_2 and thrombin, mediate thrombus formation (aggregation) and stabilization (clot retraction). The formation of a procoagulant surface also supports formation of thrombin (not shown).

The fast on-rate of association between VWF and GPIbα enables binding to take place at the intermediate to high shear forces of the arteriolar system. However, a fast off-rate of association means that this interaction is insufficient to give rise to stable adhesion. Thus, platelets translocate in the direction of flow on a surface of VWF unless otherwise activated. High-resolution imaging technology has revealed that platelets translocate in a stop–start manner through sliding rather than rolling. Platelet capture is facilitated by the formation of membrane tethers, which are thin processes of membrane generated from the platelet surface by haemodynamic forces. These membrane tethers can extend as far as 30 μm in length, with much of their length being derived from the SCCS. The formation of membrane tethers helps to sustain adhesion within high-shear environments by minimizing the drag on captured platelets.

The conversion to stable adhesion is dependent on activation of β_1 and β_3 integrins by other receptors, most notably the collagen receptor GPVI and vessel-wall-generated thrombin. This transition can be readily demonstrated *in vitro* by comparing adhesion and aggregate formation when blood flows over immobilized VWF or a VWF/collagen surface, as illustrated in Figure 37.3. The majority of platelets translocate on a surface of VWF

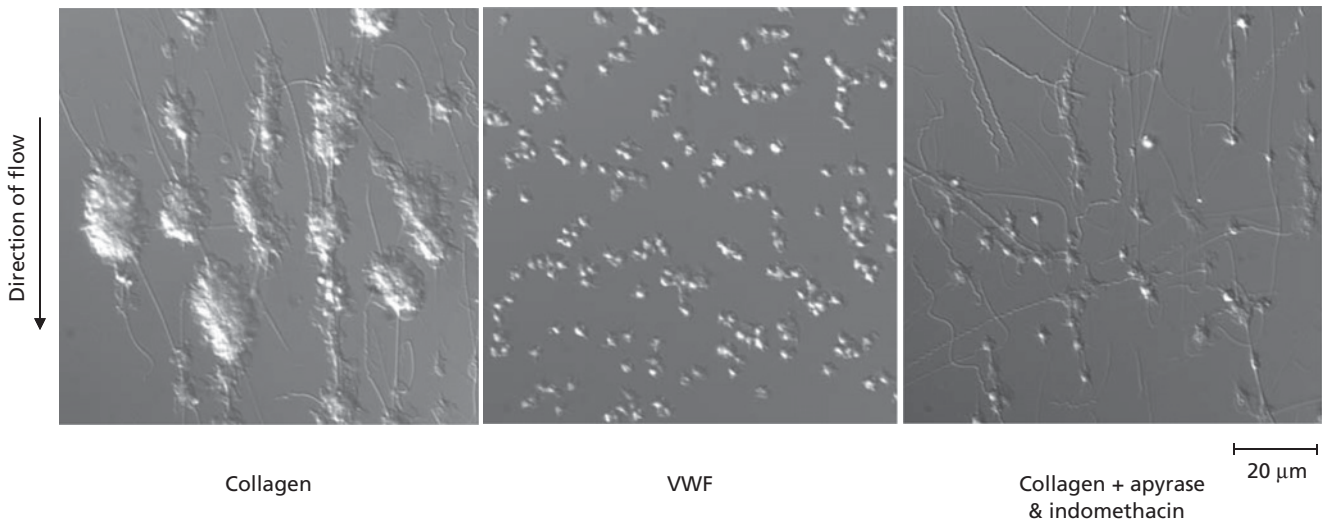


Figure 37.3 Platelet adhesion and aggregation at arteriolar shear. Human blood, anticoagulated with 40 μmol/L P-PACK, was flowed over collagen or VWF at a shear rate of 1000/s for 4 min and rinsed with Tyrode buffer for 5 min. Differential interference contrast images of adherent platelets were recorded. Where indicated, blood was pretreated with indomethacin (10 μmol/L) and apyrase (5 U/mL) to reduce the secondary mediators, TxA_2 and ADP respectively, prior to flow.

and form small stable aggregates only after several minutes. In comparison, platelets undergo rapid stable adhesion and form large aggregates on collagen. This process is critically dependent on the binding of plasma-derived VWF to collagen through its A3 domain, unmasking the binding site for GPIb within the VWF A1 domain, and the release of the positive feedback mediators ADP and TxA_2.

The integrins that mediate stable adhesion on collagen are αIIbβ3 and α2β1 (GPIa/IIa), which bind to immobilized VWF and collagen, respectively. The role of α2β1 in this process is masked by the 50 times higher level of αIIbβ3, which is expressed at approximately 80,000 copies per platelet, with a further 40,000 copies on intracellular α-granules that become exposed on activation. Platelets express three other integrins that have the potential to mediate stable adhesion *in vivo*, α5β1, α6β1 and αvβ3, which bind to fibronectin, laminin and vitronectin, respectively. The capture and recruitment of circulating platelets onto a growing aggregate occurs in a similar way, but in this case β_1 and β_3 integrin activation is mediated by the release of ADP and TxA_2, and the local formation of thrombin. Binding of VWF to integrin αIIbβ3 on the surface of activated platelets mediates tethering of non-activated platelets through GPIb–IX–V. High-resolution imaging studies have revealed that the initial stage involves discoid-shaped platelets and is reversible and independent of ADP and TxA_2. This initial phase therefore serves to recruit platelets into the growing aggregate. This is followed by an irreversible phase that is associated with a marked change in platelet morphology and platelet activation, and which is dependent on released secondary mediators and localized thrombin formation.

There are important differences in the mechanisms that give rise to adhesion and aggregation at the low and very high rates of shear found in the venous system and in stenosed arteries, respectively. At flow rates of <500/s, platelets adhere directly to exposed subendothelial matrix proteins and undergo stable adhesion independent of VWF and GPIb–IX–V. Release of secondary mediators and thrombin generation is sufficient to enable aggregate growth to occur. At shear rates >10,000/s, adhesion and aggregation are mediated entirely by the interaction of soluble and immobilized VWF with GPIb–IX–V, independent of platelet activation.

Spreading

Platelets undergo a characteristic set of morphological changes on contact with a surface, known collectively as spreading, as illustrated in Figure 37.4. Initial shape change or rounding is

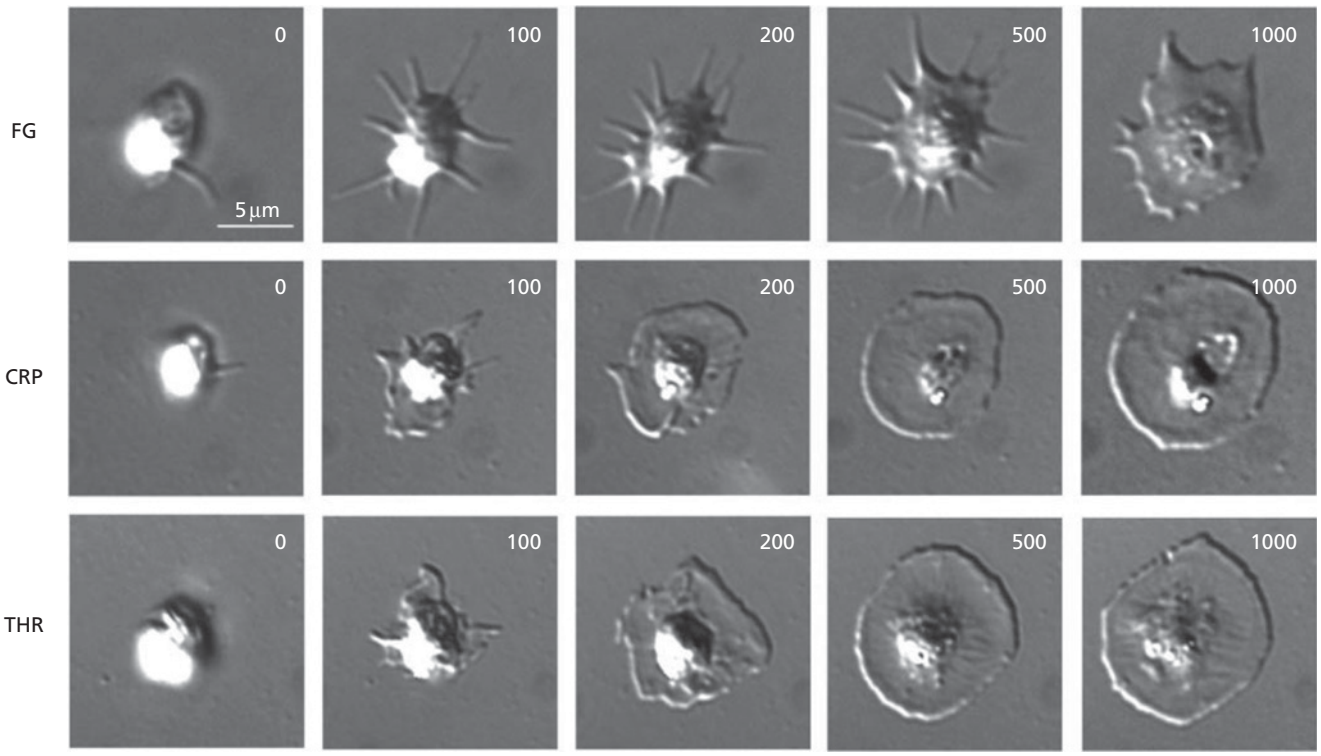


Figure 37.4 Distinct morphological changes in washed human platelets exposed to fibrinogen (FG), collagen-related peptide (CRP) or thrombin (THR). Representative morphology of a single platelet on each surface at specific time points (seconds) are shown. (Source: Thornber *et al*., 2006 [*FEBS J.* 2006; 273: 5032–5043]. Reproduced with permission of Wiley.)

followed by the generation of finger-like projections known as filopodia, which grow from the periphery of the cell, and the subsequent formation of lamellipodia. As these events proceed, granules and organelles are squeezed into the centre of the cell, resulting in a characteristic 'fried egg' appearance. These dramatic changes in morphology are brought about by a powerful severing and reassembly of the actin cytoskeleton. The spreading of platelets and formation of stress fibres helps to secure the platelet and thrombus against the flow and shear forces of the vascular system.

Granule secretion and TxA_2 formation

The secretion of ADP from dense granules and the *de novo* generation of TxA_2 from arachidonic acid, by the action of phospholipase (PL)A_2, play a feedback role in mediating platelet activation. In nearly all cases, activation of platelets by low concentrations of agonists is potentiated by the release of the two feedback agonists (see Figure 37.3). Secretion from platelet α-granules usually occurs concomitantly with that from dense granules. Fusion of α-granules leads to expression of P-selectin on the platelet surface, which is the major ligand for P-selectin glycoprotein ligand (PSGL)-1 on circulating microvesicles and leucocytes. The capture of circulating microvesicles provides a mechanism for further activation of the coagulation cascade on a growing thrombus, while binding to leucocytes contributes to inflammatory events at the vessel wall.

Aggregation

Aggregation is the specific term used for the cross-linking of activated platelets through binding of bivalent or multivalent ligands to integrin αIIbβ3. It is distinct from agglutination which is a passive process mediated, for example, by the antibiotic ristocetin, which induces cross-linking of VWF to GPIb–IX–V. Ristocetin is routinely used to investigate possible defects or gain-of-function mutations in plasma VWF or GPIb, and for investigation of patients with Glanzmann thrombasthenia.

The integrin αIIbβ3 is present in a low-affinity form on non-stimulated platelets, but undergoes a conformational change to a high-affinity form in response to 'inside-out' signals. The subsequent binding of bivalent ligands to the integrin mediates platelet aggregation and adhesion. In turn, these interactions lead to clustering of αIIbβ3 on the platelet surface and the generation of 'outside-in' signals that regulate actin polymerization (and therefore platelet spreading) and other responses in synergy with other platelet receptors.

Fibrinogen, VWF and fibronectin have similar affinities for integrin αIIbβ3, but fibrinogen is considered to be the major ligand because of its much higher concentration in plasma.

Thrombus stabilization

There is increasing recognition of so-called 'late events' that help to stabilize the platelet aggregate or thrombus in the high shear environment. These events involve remodelling of the actin cytoskeleton and the interaction of several platelet membrane proteins with themselves or with each other. The affinity of the latter interactions is low and so these interactions only occur when platelets are brought into close contact with each other.

The ability of blood clots to retract over a course of minutes to hours is termed clot retraction. The platelets are the force-generating components of this response, with integrin αIIbβ3 playing a fundamental role, both by linking cytoplasmic actin filaments to surface-bound fibrin polymers and by generating intracellular signals that, together with those from other receptors, enable myosin to serve as a motor and drive the process. A novel actin-dependent, but fibrin-independent, mechanism of retraction has been recently described that brings newly captured platelets into the aggregate and strengthens the aggregate by reducing the shear forces at the aggregate edge.

The aggregate is further consolidated by the binding of a number of platelet membrane proteins to themselves (homophilic interactions) or to other surface receptors on adjacent platelets (heterophilic interactions). Studies using blocking antibodies and mutant mice have provided evidence that these interactions contribute to the stability of the thrombus, although in many cases it is unclear if this is through the generation of intracellular signals or by strengthening platelet–platelet interaction. There is also evidence that some of these interactions serve to limit the degree of thrombus formation. Examples of receptors that may support these so-called late events are semphorins and ephrins.

Many adhesion proteins are shed during activation, including GPVI, GPIbα, GPV, P-selectin, and semaphorin 4D, and this serves to help prevent unwanted platelet activation and limit thrombus growth.

Procoagulant activity

Activated platelets also provide a negatively charged phospholipid surface for the assembly of two multiprotein complexes that form part of the intrinsic and amplication pathways of coagulation, namely the tenase and prothrombinase complexes. A complex of FIXa and FVIIIa on the negatively charged lipid surface converts FX to FXa (tenase complex), which in turn forms a complex with FVa on the same surface to efficiently convert prothrombin to thrombin (prothrombinase complex). In this way, a large amount of thrombin is generated in the vicinity of the growing aggregate that serves to enhance platelet activation and convert fibrinogen to fibrin.

The platelet procoagulant response, which is mediated by exposure of phosphatidylserine, is elicited only by very powerful platelet agonists or combinations of agonists, and requires sustained entry of Ca^{2+} across the plasma membrane. Rare

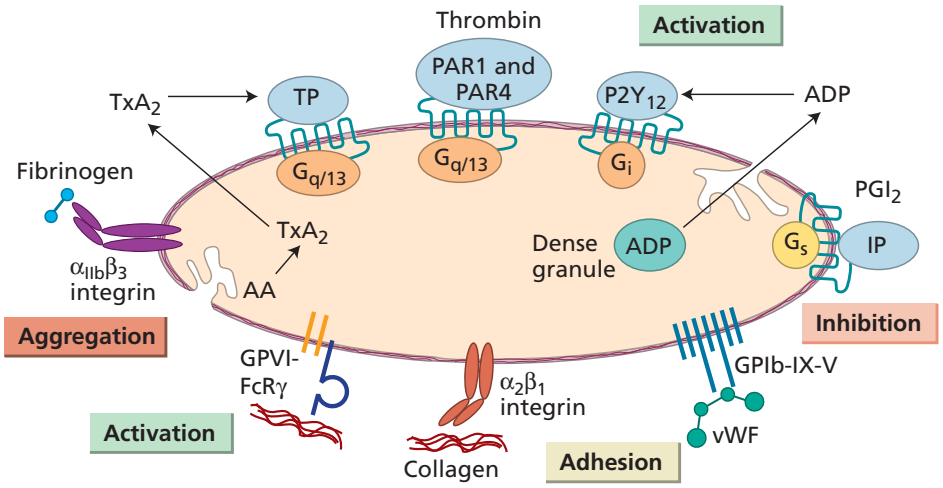


Figure 37.5 Schematic of the major tyrosine-kinase-linked and G-protein-coupled receptors regulating platelets. The major receptors regulating platelet activation and inhibition are shown. The major signalling receptors that mediate platelet activation are the tyrosine-kinase-linked receptor GPVI and the G-protein-coupled receptors for TxA_2, thrombin and ADP. Of this group, the $P2Y_{12}$ ADP receptor is unique in that it synergizes with the receptors for the other agonists to mediate powerful activation. The GPIb–IX–V complex is critical for platelet tethering and integrins $\alpha_2\beta_1$ and $\alpha_{IIb}\beta_3$ for stable adhesion and, in the case of the latter, platelet aggregation. All three of these adhesion receptors also generate weak tyrosine kinase-based intracellular signals of uncertain significance. The prostacyclin receptor (IP) elevates cAMP and, in combination with nitric oxide which elevates cGMP (not shown), mediates powerful platelet inhibition.

patients have been described with a mild bleeding disorder linked to defective procoagulant activity, a condition known as Scott syndrome, and this provides evidence that this pathway plays an important role in thrombus formation. Scott syndrome is caused by mutations in the membrane transporter, transmembrane protein-16F (TMEM-16F). A related disorder, Storkmorken syndrome, describes patients with permanently exposed negatively charged phospholipids on their platelets and is even less frequent.

The activation of platelets by powerful agonists and entry of Ca^{2+} is associated with formation of microvesicles and these too have been proposed as providing a massive increase in the surface area for binding of the tenase and prothrombinase complexes by binding to other surfaces.

Stimulatory receptors and their signalling pathways

The major physiological agonists that mediate platelet activation signal primarily through single transmembrane proteins that regulate Src and Syk family tyrosine kinases or via seven-transmembrane-spanning proteins that are coupled to heterotrimeric G proteins. The former group includes the adhesion molecules collagen, VWF and fibrinogen, and the latter the so-called 'soluble' agonists thrombin, ADP and TxA_2 (Figure 37.5).

The collagen receptor GPVI is the most powerful of the adhesion receptors, signalling through tyrosine phosphorylation of an immunoreceptor tyrosine-based activation motif (ITAM) in its associated FcR γ-chain, leading to formation of an LAT signalosome and activation of PLCγ2. Integrin αIIbβ3 and GPIb–IX–V generate much weaker signals that mediate spreading and activation in synergy with other receptors. Signalling by GPVI and integrin αIIbβ3 occurs in distinct regions of the platelet membrane, with the former signalling in cholesterol-rich membrane domains known as lipid rafts (Figure 37.6). Heterotrimeric G proteins are composed of α and βγ subunits, and take their name from the α subunit, although both subunits regulate effector proteins. Thrombin and TxA_2 receptors activate G_q and G_{13}, which regulate PLCβ and Rho kinase, respectively, whereas the $P2Y_1$ ADP receptor is coupled to G_q. The $P2Y_{12}$ ADP receptor is unique among this group of receptors in that it is coupled to the G_i family of G proteins.

Tyrosine kinase-linked receptors

GPIb–IX–V

The GPIb–IX–V is unique to platelets and consists of four subunits, GPIbα, GPIbβ, GPIX and GPV, with 25,000 copies of the complex on the platelet surface. The principal ligand for GPIb–IX–V is VWF, which consists of a series of multimers that vary in size from 500 to more than 200,000 kDa. The largest multimers of VWF are more effective in mediating platelet tethering

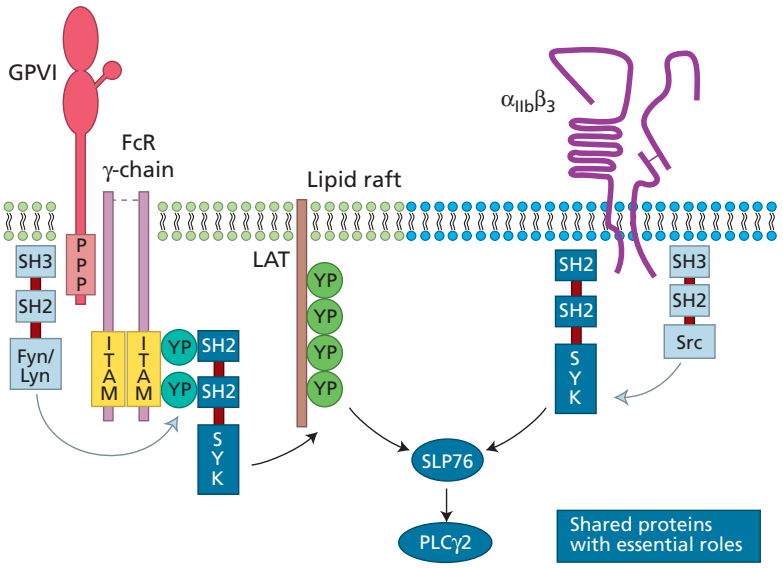


Figure 37.6 Schematic of GPVI and $\alpha_{IIb}\beta_3$ signalling. Activation of the collagen receptor GPVI leads to tyrosine phosphorylation of the FcR γ-chain ITAM by the two Src kinases, Fyn and Lyn. This leads to recruitment of the tandem SH2 domain-containing tyrosine kinase Syk, which mediates tyrosine phosphorylation of the adapter proteins LAT and SLP-76, and PLCγ2. Signalling occurs in cholesterol-rich membrane domains known as lipid rafts. The integrin $\alpha_{IIb}\beta_3$ also signals through Src and Syk tyrosine kinases leading to activation of PLCγ2. However, there are a number of differences to the GPVI signalling cascade, including: (i) Src and Syk binding directly to the β3 cytoplasmic tail independent of an ITAM, (ii) the integrin uses Src, (iii) signalling occurs outside of lipid rafts and (iv) is independent of LAT. These differences may account for the much weaker signal from the integrin relative to that from GPVI.

and platelet activation. Ultra-large VWF is stored and released from endothelial cell Weibel–Palade bodies and is broken down into the smaller multimers in the circulation by ADAMTS-13. Deficiencies in ADAMTS-13 lead to the persistence of ultra-large VWF multimers in the circulation and a life-threatening thrombotic microangiopathy known as thrombotic thrombocytopenic purpura (TTP), which is characterized by VWF-rich platelet aggregates in the skin and vital organs. VWF also interacts with and stabilizes FVIII.

At high rates of shear, multimeric VWF undergoes a change from a globular to an extended form that can reach several microns in length and which has multiple platelet and collagen binding sites. VWF can also self-associate when extended to form very large filaments on the surface of endothelial cells and on collagen. In this extended form, VWF is able to tether to platelets through the GPIb–IX–V complex and thereby support adhesion and aggregation. The critical role of extended VWF in mediating platelet tethering is dependent on the fast on-rate of association of its A1 domain to GPIbα. However, as discussed above, other agonists, such as collagen are required to mediate *rapid* activation of platelet integrins leading to stable adhesion and aggregation.

GPIb–IX–V binds to several other cell-surface and plasma proteins, including integrin αMβ2 (also known as Mac-1 or CR3) and P-selectin of leucocytes. The interaction with αMβ2 and P-selectin contributes to attachment and transmigration of leucocytes through a mural thrombus. Binding to P-selectin has been reported to support rolling of platelets on activated endothelium in conditions of low shear.

Mutations in GPIbα, GPIbβ and GPIX give rise to the rare bleeding disorder Bernard–Soulier syndrome, which is characterized by macrothrombocytopenia.

GPVI and integrin $\alpha_2\beta_1$

Collagen is recognized as the most thrombogenic component of the subendothelial matrix. Nine forms of collagen have been described within the vessel wall, with types I and III being the predominant ones in the deeper layers of the vessel wall and type IV in more superficial layers. The major receptors for collagen on platelets are the immunoglobulin GPVI and the integrin α2β1, which are expressed at 4000–6000 and 1500–2500 copies per platelet, respectively. The majority of vascular collagens activate GPVI and integrin α2β1, with the exception of the sheet-like collagen type IV, which does not interact with either receptor.

Collagen is composed of three helical chains, which interact to form a superhelical structure that is interrupted by non-helical regions. Synthetic collagen peptides based on repeating trimers have been shown to bind selectively to GPVI or integrin α2β1. Peptides that are specific to GPVI have a backbone of GPO residues, whereas peptides that bind to α2β1 have a backbone of GPP residues interspersed with at least one or more GER-containing motifs, such as GFOGER. Many snake toxins target GPVI, including the C-type lectin convulxin, which is a tetramer of α and β subunits and induces powerful activation through clustering.

GPVI forms a complex with the FcR γ-chain and is the major signalling receptor for collagen. The FcR γ-chain contains an ITAM, defined by the sequence $YXXLX_{6-12}YXXL$, which plays a critical role in cell signalling. ITAM receptors signal through sequential activation of Src, Syk and Tec family kinases. Src family kinases, including Fyn and Lyn, mediate phosphorylation of the conserved tyrosines in the ITAM, leading to binding of the tandem SH2 domain Syk family kinase and initiation of a downstream signalling cascade that culminates in activation of PLCγ2. Although GPVI is primarily known as a receptor for

collagen, it is also activated by laminin. Because of the low affinity for GPVI, the interaction with laminin requires prior binding to the integrin α6β1. Laminin is a major component of the basement membrane and may therefore mediate platelet activation following minor injuries that fail to expose subendothelial type I or III collagen or tissue factor.

The primary role of integrin α2β1 is to support platelet adhesion to collagen at sites of damage to the vasculature and to cause a net increase in binding to GPVI in view of the modest affinity of collagen for the immunoglobulin receptor. Integrin α2β1 also mediates weak activation of PLCγ2, although the physiological significance of this is unclear due to the more powerful action of GPVI and the redundancy in signalling with other receptors. The physiological significance of α2β1 in supporting thrombus formation *in vivo* is masked by the much greater level of expression of integrin αIIbβ3.

A small number of patients have been described with inherited or acquired disorders that lead to reduction or abrogation of expression or signalling by GPVI. In general, these patients exhibit mild bleeding, but the presence of additional complications means that it is unclear if this is due solely to the defect in GPVI. Only three patients have been described with an acquired deficiency of α2β1 and they too have other vascular complications.

Integrin αIIbβ3 (GPIIb/IIIa)

Integrin αIIbβ3 is the most abundant protein on the platelet surface and is estimated to comprise approximately 15% of surface proteins. αIIbβ3 binds to several multivalent ligands in the vasculature including fibrinogen, VWF, fibronectin and vitronectin. Although all of these are capable of supporting platelet aggregation, fibrinogen is considered to be the major ligand because of its much higher concentration in plasma. In all cases, the binding of these soluble ligands requires 'inside-out' activation of integrin αIIbβ3 to a high-affinity conformation in the presence of divalent cations.

The clustering of αIIbβ3 generates weak 'outside-in' signals that have been shown to induce platelet spreading, clot retraction and secretion in combination with other platelet agonists such as ADP. The integrin signals through at least two and possibly more pathways. The most thoroughly characterized of these is the activation of Src and Syk tyrosine kinases by direct interaction with the β3 cytosolic tail. These two kinases regulate actin polymerization and PLCγ2.

The discovery that the Arg-Gly-Asp (RGD) sequence in fibronectin mediates its interaction with integrin α5β1 led to the unexpected discovery that small peptides or snake toxins containing this sequence inhibit binding of fibronectin to integrin αIIbβ3. This subsequently resulted in the realization that other matrix proteins use RGD to bind to integrin αIIbβ3, including VWF, vitronectin and CD40 ligand. However, even though it has two pairs of RGD sequences, the major site of interaction of fibrinogen with the integrin is via a KQAGDV sequence at the C-termini of its two γ-chains. Despite this, the binding of fibrinogen can still be blocked by synthetic RGD-containing peptides.

The critical role of integrin αIIbβ3 in aggregation encouraged the development of novel antiplatelet agents to integrin αIIbβ3. A chimeric Fab molecule, abciximab, was introduced into the clinic in 1994 and was later followed by two RGD-based peptide inhibitors, eptifibatide and tirofiban. The three inhibitors have been approved for use in patients undergoing percutaneous coronary interventions involving stenting and in patients with unstable angina. Excessive bleeding and thrombocytopenia are major side-effects for all three of these αIIbβ3-blocking agents and thus limits their use to patients considered to be at high risk of thrombosis within a hospital setting. Unexpectedly, clinical trials with orally active αIIbβ3 antagonists were associated with increased mortality, increased bleeding and occasional severe thrombocytopenia, and were stopped many years ago.

G-protein-coupled receptors

$P2Y_1$ and $P2Y_{12}$ ADP receptors

ADP was reported to activate platelets in the early 1960s, but it has only been since 2000 that its critical role as a positive feedback agonist has finally been established. This delay in recognition arose because ADP is a weak platelet agonist, and it is only through its ability to mediate powerful platelet activation in synergy with Ca^{2+}-mobilizing receptors via the $P2Y_{12}$ receptor that its critical role in thrombus formation is achieved. The latter is illustrated by the clinical efficacy of the $P2Y_{12}$ receptor antagonist, the thienopyridine clopidogrel (Plavix), in the long-term prevention of thrombosis. Clopidogrel is a prodrug that is metabolized by the liver to generate an active metabolite that covalently modifies, and therefore irreversibly inhibits, the $P2Y_{12}$ ADP receptor. The related thienopyridine prasugrel is metabolized more rapidly to the same active metabolite as clopidogrel and therefore has a faster onset of action. Two $P2Y_{12}$ receptor antagonists that do not require prior metabolism and bind non-covalently are cangrelor and ticagrelor. These have the advantage of a faster onset of action. Ticagrelor, but not cangrelor, is available for oral administration and is now used in the long-term treatment of patients at risk of thrombosis. Despite being structurally related to ADP, ticagrelor is a non-competitive antagonist.

ADP stimulates sustained platelet aggregation in Born-aggregometry through the synergy between the $P2Y_{12}$ receptor, which is coupled to the G_i family, and the $P2Y_1$ receptor, which is coupled to G_q. $P2Y_1$ is expressed at low level with only 150 copies per platelet and it has been proposed that its role is to support rapid platelet activation in synergy with $P2Y_{12}$. Consistent with this, arterial, but not venous, thrombus formation is reduced in mice deficient in the $P2Y_1$ receptor or in the presence of a $P2Y_1$ receptor antagonist. However, it is unclear whether the $P2Y_1$ receptor performs a similar function in humans, as

patients with a bleeding disorder associated with a defect in $P2Y_1$ have not been described. A striking feature of the $P2Y_1$ receptor is its ability to undergo rapid desensitization and this, together with its low level of expression, may serve to prevent activation of platelets following exposure to ADP in the absence of other platelet stimuli.

The $P2Y_{12}$ ADP receptor activates the G_i family of G proteins, predominantly $G_{i\alpha 2}$, leading to inhibition of adenylyl cyclase and activation of phosphatidylinositol 3-kinase (PI3K). The latter, in combination with other pathways, underlies its ability to synergize with Ca^{2+}-mobilizing receptors, whereas inhibition of adenylyl cyclase on its own is insufficient to mediate activation. Over 10 gene mutations in $P2Y_{12}$ have been identified in patients with a mild bleeding disorder, consistent with its critical role in supporting platelet activation.

TxA_2 receptor

TxA_2 signals through a single G protein-coupled receptor, the TP receptor, which is coupled to G_q and G_{13} heterotrimeric G proteins. The TP receptor is alternatively spliced in its C-terminus to two isoforms, TPα and TPβ, that signal in the same way. Platelets express only the TPα isoform. TxA_2 has a very short half-life and is rapidly metabolized to the inactive metabolite TxB_2, which prevents its use in functional studies, where the stable mimetic, U46619, is preferred. TxB_2 is routinely used to monitor formation of TxA_2 by radioimmunoassay and U46619 is sometimes used in the clinical evaluation of patients with suspected thromboxane receptor defects. Arachidonic acid, which is a direct substrate for cyclo-oxygenase, is used to check for activity of the TxA_2 pathway.

The clinical importance of the TxA_2 pathway in platelet activation is illustrated by the antithrombotic action of aspirin, which inhibits platelet cyclo-oxygenase. The irreversible nature of the acetylation induced by aspirin, in combination with the inability of platelets to synthesize significant levels of new proteins, means that the effective antithrombotic concentration of aspirin is considerably lower than that required to target cyclo-oxygenase in other cells. Large-scale clinical trials have shown that aspirin reduces the mortality of myocardial infarction by approximately 25% and the number of major vascular events in individuals at risk by about one-third. Several patients with platelet-based bleeding disorders have been shown to have heterozygous, function-disrupting mutations in the thromboxane receptor, with the bleeding believed to be due to the associated presence of one or more further mutations (which have yet to be identified).

PAR-1 and PAR-4 thrombin receptors

Thrombin is among the most powerful of all platelet agonists and induces activation of human platelets through the proteolytic cleavage of PAR-1 and PAR-4, thereby generating tethered ligands which bind to the receptor and promote activation. Synthetic peptides corresponding to the newly exposed receptor-specific sequences induce activation in the absence of receptor cleavage. These thrombin receptor-activating peptides can be used to selectively activate the individual PAR receptors. In humans, it is believed that PAR-1 mediates rapid activation in response to low concentrations of thrombin, whereas PAR-4 is activated by higher concentrations of the protease. PAR-1 and PAR-4 signal through G_q and G_{13} heterotrimeric G proteins. Patients with defects in either of the thrombin receptors have yet to be identified, possibly because such mutations are lethal. An orally active PAR-1 receptor antagonist is now available.

The GPIbα subunit of GPIb–IX–V expresses a high-affinity binding site for thrombin that is believed to position the protease in the vicinity of the PAR-1 and PAR-4 receptors in order to facilitate activation. Intravenous and oral direct inhibitors of thrombin are used in the clinic as anticoagulants.

Other platelet receptors and their ligands

Platelets express a large number and diversity of receptors for a variety of ligands, including adhesive proteins, amines, chemokines, cytokines, lipids, nucleotides, proteases and transmembrane proteins. Many of the receptors are present at such a low level, or are not ordinarily exposed to their ligand, that their physiological role in mediating or potentiating activation may be minimal, although they could potentially play a role in disease. Examples include regulation of human platelets by endothelin-1, Gas6, insulin-like growth factor, leptin, vasopressin, platelet-activating factor and PDGF. In some cases, expression of a receptor on the platelet surface appears to reflect a role in megakaryocyte development/platelet formation rather than in regulating platelet activation. Examples include the SDF-1α receptor, CXCR4 and the TPO receptor c-Mpl. SDF-1α plays a key role in supporting migration of megakaryocytes to the vascular niche in bone marrow and TPO is the major regulator of megakaryocyte growth and development. Both receptors also mediate weak potentiation of platelet activation, although the physiological significance of this is unclear. The function of a number of these platelet receptors is discussed below.

$5\text{-}HT_{2A}$ receptors

Platelets contain very high levels of 5-HT, which is taken up by an active transport mechanism and stored in dense granules. Release of 5-HT induces powerful vasoconstriction, thereby limiting blood loss. 5-HT mediates weak platelet activation through the $5\text{-}HT_{2A}$ receptor, which is coupled to G_q, but there is little evidence to suggest that this plays a major role in supporting haemostasis.

α_{2A}-Adrenoceptor

Adrenaline activates the α_{2A}-adrenoceptor, which is coupled to the G_i family of proteins, predominantly G_z. As is the case for the $P2Y_{12}$ ADP receptor, the α_{2A}-adrenoceptor is able to induce powerful platelet activation in synergy with Ca^{2+}-mobilizing

receptors. Studies in α_{2A}-adrenoceptor knockout mice have provided evidence for a minor physiological role of adrenaline in supporting haemostasis, although it is not clear if this is also the case in humans, even though platelet function tests frequently use adrenaline as a coagonist.

$P2X_1$ ATP receptor

$P2X_1$ is a receptor for ATP, which is released alongside ADP from dense granules. It is the only known ligand-gated ion channel on the platelet surface. Binding of ATP promotes direct entry of Ca^{2+} into the platelet, which occurs over a much more rapid time course than IP3-mediated Ca^{2+} mobilization and subsequent entry of extracellular Ca^{2+} via store-operated Ca^{2+} entry. The $P2X_1$ receptor has been shown to induce platelet shape change, but requires the presence of other receptors to mediate aggregation. The physiological role of $P2X_1$ may be to facilitate rapid activation of platelets in synergy with other agonists. It is intriguing that $P2X_1$ undergoes extremely rapid desensitization and this too may serve to prevent 'unwanted' activation.

FcγRIIA

v Clustering of the low-affinity immune receptor FcγRIIA by immune complexes or via primary and secondary antibodies mediates powerful activation through an ITAM-based pathway that is believed to be similar to that used by the collagen receptor GPVI. FcγRIIA is not thought to have a physiological role in supporting haemostasis, but it mediates activation by immune complexes and is the causative receptor in heparin-induced thrombocytopenia. FcγRIIA is absent from the mouse genome.

CD36

The scavenging receptor CD36 is among the most highly expressed of platelet surface glycoproteins, being present at approximately 20,000 copies per platelet. Despite the fact that it was first identified on platelets more than 30 years ago, it is only recently that it has been implicated as a major player in mediating the increased platelet responsiveness in patients with dyslipidaemia. Oxidized choline glycerophospholipids are present at sufficiently increased levels in the plasma of humans with low levels of high-density lipoprotein to potentiate activation of platelets by threshold concentrations of ADP and other agonists. Moreover, this potentiation is not seen in individuals who do not express CD36 on their platelets, suggesting that they may be protected against the dyslipidaemia-induced prothrombotic state. However, further research is still required to fully evaluate the contribution of platelet CD36 to cardiovascular disease.

CLEC-2

The C-type lectin receptor CLEC-2 mediates powerful activation of platelets in response to the snake venom toxin rhodocytin, or its endogenous ligand podoplanin. The latter is expressed on lymphatic endothelial cells, renal epithelial cells and on type I lung alveolar cells and is upregulated on macrophages at sites of inflammation and on the leading edge of tumours, where it is implicated in cancer metastasis. Mice deficient in CLEC-2 or podoplanin die shortly after birth, and embryos at mid-gestation have blood-filled lymphatics and haemorrhaging in the brain. These defects have been shown to be platelet in origin. CLEC-2 signals through an ITAM-like pathway similar to that used by GPVI and FcγRIIA, although it uses a single rather than a tandem YXXL motif to activate Syk.

CLEC-2 appears to play a minimal role in supporting haemostasis, consistent with the absence of podoplanin in the vasulature. However, platelet CLEC-2 plays a critical role in vascular integrity (i.e. the ability of a healthy endothelial cell layer to stop haemorrhaging), notably at sites of inflammation. CLEC-2 expression is restricted to platelets and to activated dendritic cells.

Second messenger pathways underlying activation

Calcium

The divalent cation Ca^{2+} plays a critical role in mediating platelet activation, usually in combination with other signalling pathways. An increase in Ca^{2+} is brought about by its release from intracellular stores by the action of IP3 and by entry through the plasma membrane, predominantly through store-operated Ca^{2+} entry. The latter has recently been shown to occur via binding of the transmembrane protein STIM-1, which is localized to intracellular stores, to the surface membrane Ca^{2+} channel, Orai-1, with the trigger for entry being the release of Ca^{2+} from the store. The entry of Ca^{2+} through Orai-1 is critical for procoagulant activity and possibly other responses. The released Ca^{2+} is efficiently removed from the cytosol by Ca^{2+}-ATPases in the endoplasmic reticulum (SERCA) and in the plasma membrane (PMCA).

Protein kinase C

PKC refers to a family of structurally similar lipid-regulated protein kinases that is divided into *classical* isotypes regulated by Ca^{2+} and 1,2-diacylgycerol (DG), *novel* isotypes regulated by DG and *atypical* isotypes regulated independently of the two messengers. Human platelets express two classical isotypes, α and β, and three novel isotypes, δ, θ and η. PKCα is the major isoform regulating dense granule and α-granule platelet secretion in mouse platelets, with increasing evidence suggesting that the other PKCs have isotype-specific inhibitory and activatory roles.

DG, together with Ca^{2+}, regulates the Rap1 (also known as Rap1b) GTP exchange factor, CALDAG-GEF1. Rap1 plays a critical role in mediating activation of integrin αIIbβ3, although

this role can be bypassed by high concentrations of powerful ligands. CALDAG-GEF1-deficient mice and patients with defective expression of the exchange factor (leucocyte adhesion deficiency type III) have defective Rap1 activation and a bleeding diathesis.

Phosphatidylinositol 3-kinase

PI3K generates the 3-phosphorylated lipid second messenger phosphatidylinositol 3,4,5-trisphosphate (PIP3) from phosphatidylinositol 4,5-bisphosphate (PIP2). The two 5-phosphorylated lipids mediate translocation of proteins containing one or more pleckstrin homology (PH) domains to the membrane, including PLCβ and PLCγ isoforms, protein kinase B (Akt) and the Tec family kinase Btk, which mediates phosphorylation of PLCγ2. Platelets express several regulatory and catalytic subunits of PI3K, although p110β and p85α/β are believed to the major catalytic and regulatory subunits supporting activation.

The molecular basis of platelet activation

The physiological end points of platelet activation are stable adhesion, aggregation, granule secretion, TxA_2 formation, spreading with stress fibre formation, clot retraction and procoagulant activity. These responses are regulated by the interplay between tyrosine kinase- and G protein-regulated signalling pathways, with evidence of considerable redundancy in their actions. Here we provide brief details on the major mechanisms that underlie each of these events.

Stable adhesion and aggregation

These responses are mediated through activation of integrin αIIbβ3 by 'inside-out' signals from tyrosine kinase and G-protein-coupled receptors. Activation of the integrin is brought about by the binding of talin to the β3 cytoplasmic tail via a FERM domain in its head region. Platelet activation is associated with the recruitment of talin from the cytoplasm to the β3 tail via a process that remains poorly understood, but which is regulated in part by binding of Rap1 to its effector protein RIAM. The interaction of talin with the β3 tail leads to separation of a salt bridge between the αIIb and β3 subunits, and a resulting conformational change in the extracellular region of the integrin from a bent to an extended conformation. The FERM domain-containing protein kindlin-3 has also been shown to bind to the β3 tail and to mediate activation in combination with talin, although the precise details of this interaction remain to be established. Ligand engagement of the integrin leads to clustering and generation of 'outside-in' signals that stimulate actin polymerization and PLCγ2, which support many aspects of platelet activation, including spreading and granule secretion.

Secretion

Platelet dense granule and α-granule secretion is triggered through a synergistic interaction between Ca^{2+}- and PKC-regulated pathways, with both arms being critical for secretion to occur. Granule fusion is orchestrated by a superfamily of proteins termed SNAREs (soluble N-ethylmaleimide-sensitive attachment protein receptors) that form a universal membrane fusion machine. The activity of SNARE proteins is regulated by a series of chaperone proteins that facilitate membrane fusion. PKC phosphorylates several SNARE and chaperone proteins, including Munc-18, syntaxin-4 and SNAP-23, altering their affinity for their binding partners. For example, phosphorylation of Munc-18 interferes with its ability to bind to syntaxin-4 and phosphorylation of syntaxin-4 by PKC inhibits its ability to bind SNAP-23. This allows syntaxin-4 to interact with SNARE proteins on the opposing membrane, promoting fusion and granule secretion. There is evidence, albeit controversial, for differential regulation of the two sets of granules, although the molecular basis of this is not known and in general dense granule and α-granule secretion occur together.

TxA_2 formation

Cytosolic PLA_2 liberates arachidonic acid from membrane phospholipids and is regulated by Ca^{2+} and serine phosphorylation downstream of mitogen-activated protein kinases. Ca^{2+} is critical for activation, whereas phosphorylation causes a relatively small increase in activity and on its own is insufficient to mediate activation. The liberated arachidonic acid is metabolized by cyclo-oxygenase and thromboxane synthase to PGG_2/PGH_2, which immediately convert into TxA_2, and by lipoxygenase enzymes to leucotrienes.

Actin polymerization

Activation of platelets leads to a series of dramatic changes in platelet morphology that helps to secure the thrombus at the site of injury in the high-pressure arteriolar system. In addition, actin polymerization plays an important, but poorly understood role in signalling by platelet surface glycoprotein receptors, including GPVI and integrin αIIbβ3. Actin is assembled into a number of morphologically distinct structures, including filopodia, actin nodules, lamellipodia and stress fibres downstream of small GTP-binding proteins. For example, the small G protein Rac1 drives formation of lamellipodia in platelets through the Arp2/3 complex, while the small G protein Rho regulates formation of actin–myosin stress fibres. Mice deficient in Rac1 or treated with inhibitors of Rho kinase have unstable aggregates, emphasizing the importance of actin polymerization in generating and strengthening the thrombus.

Inhibitory agonists and their receptors

It is essential for platelets to have powerful inhibitory mechanisms that prevent activation in intact healthy vessels and which limit thrombus growth during haemostasis.

The major direct mechanism of inhibition of platelet function is through elevation of the cyclic nucleotides cGMP and cAMP by NO and prostacyclin, respectively. The two cyclic nucleotides mediate their effects in platelets through regulation of cGMP-dependent protein kinase (PKG)-1 and cAMP-dependent protein kinase (PKA). In addition, cGMP regulates the three major platelet phosphodiesterases (PDEs), including inhibition of PDE3, thereby leading to an increase in the level of cAMP. There is therefore cross-talk between the two cyclic nucleotides, even though many of the targets for PKG-1 and PKA are distinct. This is illustrated by the demonstration that PKG-1, but not PKA, mediates inhibition of platelet activation through phosphorylation of the IP3 receptor-associated cGMP kinase substrate (IRAG), which is expressed in a macromolecular complex with PKG-1 and the IP3 receptor type 1. Phosphorylation of IRAG by PKG-1 inhibits IP3-induced Ca^{2+} release and, importantly, targeted deletion of the IP3-binding region of IRAG prevents NO-mediated inhibition of platelet activation, whereas the inhibitory effect of cAMP is retained. In contrast, elevation of cAMP has been shown to lead to inhibition of activation of PLC. The two cyclic nucleotides induce phosphorylation of many substrates in platelets, including VASP, which is used to monitor their activation by western blotting for flow cytometry.

There is evidence that a family of platelet glycoproteins, characterized by the presence of tandem immunoreceptor tyrosine-based inhibitory motifs (ITIMs) in their cytosolic tails, mediate weak inhibition of activation by both tyrosine-kinase-linked and, to a lesser extent, G-protein-coupled receptors. PECAM-1 is the most characterized of this family of proteins and is expressed at a level of approximately 10,000 copies per platelet. A constitutively phosphorylated ITIM-containing transmembrane protein, G6b-B, plays a constitutive role in prevention of signalling by the collagen receptor GPVI and by GPIb. Mouse platelets deficient in G6b-B are thrombocytopenic and have a reduced surface level of GPVI due to constitutive signalling and shedding of the glycoprotein receptor. This effect of G6b-B is mediated by inhibition of Syk signalling by the dual SH2 domain tyrosine phosphatases, SHP-1 and SHP-2.

The endothelial surface also plays a significant role in preventing excessive thrombus growth, again through removal of ADP and thrombin, by release of NO and prostacyclin, and by activation of the fibrinolytic pathway. Additional mechanisms help to limit excessive growth, including shedding of platelet surface glycoprotein receptors, cleavage of intracellular proteins through calpain, and removal of intracellular messengers and reversal of phosphorylation by the action of tyrosine and serine/threonine phosphatases.

Platelet-based bleeding problems

Bleeding can be due to abnormalities in platelet function/number or to extensive antiplatelet medication. Patients with platelet disorders present with a typical mucocutaneous bleeding pattern of variable severity and are often prone to bleeding during surgery and trauma. However, inherited platelet defects are an uncommon cause of bleeding and are difficult to diagnose and manage, with wide variation in practices between laboratories. Before any platelet function test is requested, a full clinical and family history is taken to determine the underlying cause of the bleeding problem. This includes studying the pattern of bleeding, whether it is lifelong or recent, triggered by trauma (e.g. dentistry, surgery or accident), is present within other family members and associated with any medication the patient may be taking. von Willebrand disease (VWD) is the most commonly inherited bleeding disorder, but presents with a platelet-like bleeding pattern and so measurement of VWF and other coagulation factors is performed ahead of platelet function assays. Although bleeding histories are subjective and can vary within a lifetime, they remain an important screening method for diagnosing a potential platelet defect. However, it is also important to exclude acquired defects of platelet function, including those caused by antiplatelet drugs and other clinical disorders. A complete drug history is therefore important, as aspirin and non-steroidal anti-inflammatory drugs are the commonest causes of an acquired platelet defect and testing may have to be deferred and repeated.

Inherited platelet defects can be classified according to the severity of bleeding, with severe disorders such as Glanzmann thrombasthenia and Bernard–Soulier syndrome usually being detected very early in life, while more mild bleeding disorders (e.g. due to $P2Y_{12}$ deficiency) often not being detected for many years until the patient is exposed to surgery or trauma. Indeed, there is overlap between symptoms in milder platelet defects and the normal population, which may be partly related to varying levels of plasma VWF, making the former group difficult to diagnose. Further information on inherited platelet bleeding disorders is given in Chapter 42.

Platelet function testing

Platelet function tests are primarily used to aid in the diagnosis of patients presenting with bleeding problems. A variety of tests can be used to diagnose an underlying cause of the bleeding problem. Normal platelet function is highly dependent on extracellular Ca^{2+} and Mg^{2+} concentrations and so the choice of anticoagulant is important. Most current testing is still performed on citrated blood within a few hours of sampling.

Global tests of platelet function are often initially used as screening tests during the laboratory investigation of individuals

with suspected haemostatic defects. Since global tests of platelet function do not enable specific diagnosis of platelet disorders, they are normally performed as the first part of a two-step strategy that requires further testing with more specialized assays of platelet function to confirm or refute any clinical diagnosis. The most commonly proposed rationale for testing global platelet function as a first-line investigation is exclusion of a platelet function disorder so that further specialized testing can be avoided. For this reason, global platelet function tests are usually initially performed at the same time as global assays of coagulation pathway function (prothrombin time and activated partial thromboplastin time), VWF screening tests (VWF:Ag, VWF:RCo and FVIII:C) and measurement of platelet number with a full blood counter. The most widely performed tests for initial global screening platelet function disorders are currently the template bleeding time and the closure time within the platelet function analyser (PFA)-100.

The bleeding time was the first *in vivo* test of platelet function and is performed by timing the arrest of bleeding from standard-sized cuts made in the skin of the forearm. Most clinicians consider the test to be poorly reproducible, invasive, insensitive (particularly to mild platelet defects) and time-consuming. Despite these drawbacks, the test is surprisingly still widely used. A number of *in vitro* tests that attempt to accurately simulate platelet function have been developed. The most widely used of these is the PFA-100. This simulates high-shear platelet adhesion to foreign surfaces (e.g. collagen) and either monitors the drop in flow rate as an aperture is closed by platelet adhesion/aggregation respectively. The PFA-100 instrument provides optional replacement of the bleeding time as a limited screening test. Although the instrument therefore has good negative predictive value for eliminating severe platelet defects and VWD, it is very important that further diagnostic tests are still performed if clinical suspicion of a platelet defect is strong.

Automated cell counting of whole blood in the modern full blood count investigation is an essential screening test in patients with abnormal bleeding. Modern blood counters can rapidly detect abnormalities in platelet number, platelet size distribution and mean platelet volume (MPV) and provide a reliable method for screening samples from patients, provided good quality control procedures are followed within the laboratory. Normal results will therefore rapidly eliminate thrombocytopenia and anaemia as potential causes of bleeding and ensure that any platelet function tests that are being performed will not be affected by low platelet counts. If any abnormalities in platelet count, MPV or distribution are flagged by the instrument, then a blood smear should be examined to confirm defects in platelet number, size and granule content. Whole-blood morphology may also assist in the diagnosis of platelet disorders by indicating abnormalities such as red cell schistocytes in TTP or neutrophil inclusions and giant platelets in MYH-9-related disorders.

Depending on the results of clinical and laboratory screening of patients, a series of diagnostic platelet function tests are usually performed. Since the 1960s, light transmission aggregometry (LTA) became the gold standard of platelet function testing and has revolutionized our ability to identify and diagnose a wide variety of platelet defects. Conventional Born aggregometers monitor the changes in light transmission that occur in a suspension of platelets in plasma or a physiological buffer that are stimulated with different concentrations of various agonists (e.g. ADP, collagen, adrenaline, ristocetin). Typical parameters recorded from LTA tracings include shape change, lag phase, rate of aggregation, and maximal and final extent of aggregation. These parameters, coupled with the pattern of primary and secondary aggregation responses obtained with different dosages of different agonists, enable the experienced operator to diagnose specific signalling defects. Modern aggregometers are now multichannel, fully computerized, easy to use compact instruments and some can also simultaneously measure ATP secretion levels by luminescence. Measurement of the storage and release of the dense granular nucleotides can also be performed by a variety of alternative methods for confirming either storage and/or release defects, although recent surveys suggest that measurement of secretion is underused. The interpretation of aggregation traces remains challenging due to the feedback actions of ADP and TxA_2. Whole-blood impedance-based aggregometers are now also common and measure the changes in electrical resistance between two electrodes caused by platelet adhesion and aggregation after addition of agonists, but are heavily influenced by platelet count. Flow cytometry provides an exquisite, sensitive and powerful tool for studying and diagnosing various platelet defects. Flow cytometric analysis of platelets is performed in fresh whole blood or in platelet-rich plasma, and the technique can be used with very small fluid volumes, even in thrombocytopenia. This technique is used to determine the copy density of platelet membrane glycoproteins and receptors, and is therefore useful for confirming the absence of various glycoproteins or receptors in disease. Platelet function testing can also be performed and the ability of platelets to degranulate, express activation markers (e.g. P-selectin) and expose procoagulant phospholipids can be studied. A major limitation of the above tests is that they are performed at low shear conditions and therefore do not mimic accurately many of the important physiological processes of platelet adhesion, activation and aggregation that occur at higher shear rates *in vivo*. There is clearly still no gold standard for platelet testing, and indeed there may never be, and LTA remains the most widely used test. However, a variety of screening and diagnostic tests coupled with an accurate bleeding history will always be required to confirm a platelet-based bleeding disorder, simply because of the heterogeneity in the range of disorders encountered coupled with the cross-talk between receptors.

Platelets and thrombosis

There is much recent interest in the possibility of using platelet function tests to reliably detect platelet hyper-reactivity in patients with arterial thrombosis (e.g. myocardial infarction, unstable and stable angina, stroke, transient ischaemic attacks and peripheral vascular disease). This could potentially allow patients to be risk stratified and their treatment adjusted accordingly to potentially improve their clinical outcomes. The increasing interest in this area and clinical discussion of aspirin or clopidogrel 'resistance' has resulted in the development of a variety of instruments that can be useful for testing (e.g. Platelet Mapping System, VerifyNow, Plateletworks, Aspirinworks). The VerifyNow test was originally developed as a whole-blood cartridge-based test to monitor αIIbβ3 blockade because of the narrow therapeutic window and increased bleeding risk associated with these types of drugs. The test has also been adapted to measure aspirin and $P2Y_{12}$ blockade with two other cartridge formulations. Platelet responsiveness to antiplatelet drugs can be easily monitored by the tests mentioned, as well as by LTA, whole-blood aggregometry, PFA-100 and IMPACT. Poor responders or 'resistant' individuals can therefore be identified, although it is still unclear how to manage them and whether there is a clinically relevant relationship between poor response and outcome. Emerging evidence suggests that the true incidence of so-called aspirin resistance is probably very low (when non-compliance is accounted for) and that many studies are measuring platelet hyper-reactivity and not blockade of cyclo-oxygenase type 1. Although recent meta-analyses suggest that this could nevertheless be clinically informative, it is still unclear how to reliably identify these patients, with which test and cut-off (as different tests give different results) and then actually how to manage these patients. Testing is therefore still very much in the clinical research setting. Indeed, International Society for Thrombosis and Haemostasis (ISTH) Scientific and Standardization Committee guidelines suggest that aspirin resistance should not be routinely monitored and the wider question remains whether routine testing of platelet hyper-reactivity and/or drug responsiveness is actually clinically necessary. Certainly there is a spectrum of platelet responsiveness within the normal population and those individuals with a stable or even an acquired hyper-reactive phenotype could theoretically be at higher risk of thrombosis.

Clopidogrel resistance is a different phenomenon as this is a prodrug that requires metabolism by liver cytochrome P450. This results in the generation of an active metabolite that irreversibly inhibits the $P2Y_{12}$ receptor. Differences in the efficiency of metabolism between individuals (related to their P450 genotype) taking standard clopidogrel dosing results in a spectrum of responsiveness and recent meta-analysis suggests that poor responders are indeed prone to increased risk of thrombosis. Again, there is a variety of tests available to monitor the efficacy of clopidogrel, including LTA, whole-blood aggregation (Multiplate), VASP phosphorylation and VerifyNow. One of the problems facing the investigator is not only which test to use, but also the definition of a suitable cut-off defining a non-response and how to effectively manage these individuals. Although there are emerging clinical cardiology guidelines, as yet they do not include routine monitoring of clopidogrel responsiveness.

It is clear that large well-designed prospective trials are required in this area, where patients with hyper-reactive platelets and/or a poor response to aspirin/clopidogrel are randomized to different treatments based on a platelet function test result. Only with this type of data will personalized platelet function testing perhaps become a reality in the future.

Genetics of platelet function disorders

Inherited platelet function disorders (PFDs), associated with normal or reduced platelet counts, are individually rare, but collectively they account for a significant proportion of bleeding diatheses. Inherited platelet disorders can be classified into distinct groups, which include defects in adhesion, receptor signalling, secretion, cytoskeleton, procoagulant activity and production.

Identification of the underlying genetic defects is difficult and often complex in the majority of PFD cases. This is due to the variable clinical expression of the bleeding symptoms and the relative redundancy of known platelet receptor and signalling pathways. Causative mutations for PFDs have therefore only been identified in a small number of patients. For example, only two patients with compound heterozygote defects in the major collagen receptor, GPVI, have been previously described. This is in addition to a homozygous frameshift mutation in multiple Chilean patients, which is likely to be due to a founder effect. The number of patients with mutations in the Gi-coupled P2Y12 ADP receptor is only now into double figures. In most cases, the defect is frequently recessively inherited and the patients are either homozygous or compound heterozygous for P2RY12 mutations. Only three functionally disrupting mutations have been reported in the Gq-coupled thromboxane receptor, of which all patients are heterozygous for the mutation. A significant number of mutations in PFDs are indeed heterozygous, but many are unlikely to cause extensive bleeding in isolation. This is illustrated in some patients with heterozygous P2RY12 defects diagnosed with type 1 VWD, a finding which emphasizes the heterogeneity and polygenic nature of PFDs.

In the last 25 years DNA-based technologies have played a significant role in the characterisation and diagnosis of PFDs. This first began in the early 1990s with the identification of the two genes encoding the platelet integrin αIIbβ3. Mutations in these two genes, ITGA2B and ITGB3, were discovered in patients with Glanzmann thrombasthenia. Since then, cloning

techniques have gone on to scrutinize the domains of genes which behold their function. Currently, the most time- and cost-efficient means of screening genes is whole-exome sequencing (WES). WES enables the detection of genetic variation throughout the protein-coding fraction of the genome. The main challenge in applying WES for gene discovery is in filtering out the large numbers of irrelevant variants (typically in excess of 20,000 per individual). However the 'unbiased' nature of this approach means we are not limited by preconceptions as to likely causative mutations. This powerful technique is illustrated in regard to platelet biology by the recent discovery of a gene, neurobeachin-like 2 (NBEAL2), for Gray platelet syndrome, an autosomal recessive bleeding disorder characterized by large platelets that lack α-granules. More recently, WES has also been utilized for the identification of activating causative mutations in STIM1 and ORAI1 for Stormorken syndrome.

The most recent sequencing technology to emerge is whole-genome sequencing (WGS), which is likely to become the method of choice as costs are reduced and sequence variation databases are improved. Alternatively a more gene-specific targeted platform may be the method of choice to reduce the time to diagnosis, by allowing the capture and sequencing of a panel of genes of relevance to bleeding and thrombotic disorders. Ultimately, DNA-based analysis will play an increasingly important role in the first-line investigation of patients with PFDs. Insights into the underlying mechanisms of platelet function through genetic investigation in patients will, in turn, lead to more effective treatments for bleeding and cardiovascular disease in the future.

Conclusions and future developments

Given that platelets were considered to be just pieces of circulating dust about 150 years ago, they are now one of the most important cells in pathology. Now that the megakaryocytic/platelet genome and proteomes are being defined, many new platelet proteins have been discovered, although many of these appear to have minor or negligible roles in regulating activation. The challenge is to identify which of these are functionally important in the many roles of platelets.

The ultimate goal of platelet research remains the development of improved antithrombotic agents that provide effective treatment without an increased risk of bleeding. Although it may never be possible to fully attain this goal, as we begin to understand more about the events that underlie thrombus formation, there will be increased opportunity for discovery of more rational forms of therapy. New medicines targeted to the very early stages of platelet activation or the amplification phase (as proven with aspirin and clopidogrel) may be more effective in preventing thrombosis without significantly disturbing normal haemostasis.

Acknowledgements

SPW is a British Heart Foundation Chair. Paul Harrison is the Healing Foundation Senior Lecturer. NVM is a HEFCE-funded lecturer in cardiovascular genetics. The authors gratefully acknowledge the support of the British Heart Foundation, Wellcome Trust and BBSRC, and all the past and present members of their laboratories who have significantly contributed to the ideas and thoughts that have formed the basis of this chapter.

Selected bibliography

Bamshad MJ, Ng SB, Bigham AW *et al.* (2011). Exome sequencing as a tool for Mendelian disease gene discovery. *Nature Reviews Genetics* **12**: 745–55.

Coller BS, Shattil SJ (2008) The GPIIb/IIIa (integrin αIIbβ3) odyssey: a technology-driven saga of a receptor with twists, turns, and even a bend. *Blood* **112**: 3011–25.

Daly ME, Leo VC, Lowe GC, Watson SP, Morgan NV (2014). What is the role of genetic testing in the investigation of patients with suspected platelet function disorders? *British Journal of Haematology* **165**: 193–203.

Dawood BB, Lowe GC, Lordkipanidzé M *et al.* (2012) Evaluation of participants with suspected heritable platelet function disorders including recommendation and validation of a streamlined agonist panel. *Blood* **120**: 5041–9.

Gachet C (2008) P2 receptors, platelet function and pharmacological implications. *Thrombosis and Haemostasis* **99**: 466–72.

Ho-Tin-Noe B, Demers M, Wagner DD (2011) How platelets safeguard vascular integrity. *Journal of Thrombosis and Haemostasis* **9**: 56–65.

Jackson, SP (2011) Arterial thrombosis: insidious, unpredictable and deadly. *Nature Medicine* **17**: 1423–36.

Mackman N (2008) Triggers, targets and treatments for thrombosis. *Nature* **451**: 914–18.

Michelson A (2013) *Platelets*, 3rd edn. Academic Press, San Diego.

Watson SP, Lowe K and Finney BA (2014) Platelets in lymph vessel development and integrity. *Advances in Anatomy, Embryology and Cell Biology* **214**: 93–105.

CHAPTER 38

Haemophilia and Von Willebrand Disease

38

Michael A Laffan[1] and K John Pasi[2]

[1]Imperial College School of Medicine, Centre for Haematology, Imperial College, Hammersmith Hospital, London, UK
[2]Barts and The London School of Medicine and Dentistry, Royal London Hospital London, UK

Introduction

The existence of lifelong bleeding disorders and their familial occurrence was noted in the medical literature as early as the sixteenth century. Early writers were impressed by the helplessness of the physician in the face of haemophilic bleeding. Although long recognized, the pathophysiology and genetics of these disorders was not fully understood until the latter half of the twentieth century. Advances in protein chemistry and molecular biology now allow a comprehensive understanding of normal coagulation, the physiological defect in haemophilia and the underlying molecular genetics. Chapter 36 outlines the normal coagulation mechanism. In this chapter, the clinical features and principles of management of the most common inherited bleeding disorders, haemophilia and von Willebrand disease (VWD), are described, together with a summary of the genetic lesions responsible.

Haemophilia

Pathophysiology of haemophilia

Haemophilia A and B are caused by deficiency of factor FVIII and FIX respectively. FVIII and FIX are the two components of the intrinsic tenase (see Chapter 36). Consequently, their absence causes virtually identical patterns of bleeding. Combined with the fact that they are both encoded on the long arm of the X chromosome, they present almost indistinguishable sex-linked clinical syndromes and specific assays are required to determine which is present. In both cases, failure to form the intrinsic tenase complex results in failure to produce the thrombin burst characteristic of normal coagulation. As a result, a loose friable fibrin mesh is produced that is easily dislodged and which has increased susceptibility to fibrinolysis. Fibrinolytic breakdown of the clot is also favoured by the failure of the weak thrombin burst to activate the thrombin-activated fibrinolysis inhibitor (TAFI). The consequent failure to consolidate the primary haemostatic (platelet) plug results in the characteristic bleeding pattern of haemophilia, which is both delayed after trauma and much prolonged. Typical sites of bleeding are joints and muscles. Replacement by intravenous infusion of the deficient factor can normalize the haemostatic mechanism.

Clinical features

Haemophilia A and B affect approximately 1 in 10,000 and 1 in 50,000 live births, respectively, and are equally common in all ethnic groups. The vast majority of patients are male, but haemophilia can occur very rarely in females (see below). The severity and frequency of bleeding is inversely correlated with the residual level of FVIII or FIX. Table 38.1 summarizes this relationship and gives the relative frequency of the categories, based on UK national data. The main load- or strain-bearing joints (ankles, knees and elbows) are most affected, but any joint can be the site of bleeding. If untreated, this intracapsular bleeding causes severe swelling, pain, stiffness and inflammation, which gradually resolves over days or weeks. It is not clear why bleeding in haemophilia shows a predilection for joints, but it has been suggested that low levels of tissue factor (TF)

Postgraduate Haematology, Seventh Edition. Edited by A Victor Hoffbrand, Douglas R Higgs, David M Keeling and Atul B Mehta.

Table 38.1 Haemophilia A: clinical severity.

FVIII (units/dL)	Bleeding tendency	Relative incidence (%)
<1	Severe: frequent spontaneous† bleeding into joints, muscles and internal organs	50
1–5	Moderate: some 'spontaneous' bleeds, bleeding after minor trauma	30
>5–45	Mild: bleeding only after significant trauma, surgery	20

†'Spontaneous' bleeding refers to those episodes in which no obvious precipitating event preceded the bleed. No doubt, minor tissue damage consequent on everyday activities actually initiates bleeding.

expression in synovial tissue may be at least part of the explanation. There is some suggestion that bleeding in severe FIX deficiency is slightly less frequent, but the two disorders were not clearly distinguished until the middle of the twentieth century. Haemophilia B then acquired the alternative name 'Christmas disease' after one of the boys in the first report in 1952. Bleeding into muscles and joints is the hallmark of haemophilia, although the pattern of bleeding described below is now often masked by the use of prophylaxis.

Blood is highly irritating to the synovium and causes an immediate inflammatory reaction in the joint. In the longer term, blood and particularly increased iron deposition promote a chronic proliferative synovitis resulting in overgrowth of friable and highly vascular synovium, which has an increased tendency to further bleeding, thus setting up a vicious cycle. As a result of the vicious cycle of bleeding and synovial hypertrophy, a particular joint frequently becomes the 'target joint' in an individual, whereas other joints may be relatively spared. Blood also has a rapid destructive effect on cartilage that is evident after only a single haemarthrosis. Accumulation of iron in chondrocytes may also contribute to the multifactorial degenerative arthritis, resulting in irregularity of articular contour, thinning of the cartilage, bony overgrowth, subchondral cysts and finally ankylosis (Figure 38.1).

Muscle bleeding can be seen in any anatomical site, but it most often presents in the large load-bearing groups of the thigh, calf, posterior abdominal wall and buttocks. Local pressure effects often cause entrapment neuropathy, particularly of the femoral nerve with iliopsoas bleeding. This causes a common symptom triad of groin pain, hip flexure and cutaneous sensory loss over the femoral nerve distribution. Bleeding into the calf, forearm or peroneal muscles can lead to ischaemic necrosis and contracture.

Haematuria is less common than joint or muscle bleeding in individuals with haemophilia, but severely affected patients may

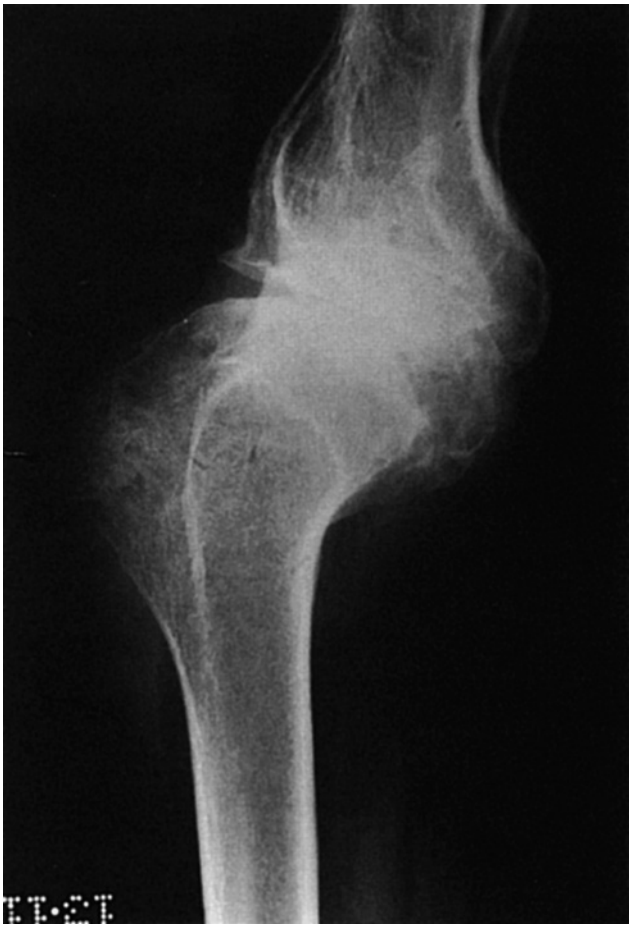


Figure 38.1 Radiograph of knee joint showing advanced haemophilic arthropathy. Note the loss of cartilage, eburnation, deformity subluxation, osteophytes, subchondral cysts and irregularity of joint contours.

have one or two episodes per decade. These may be painless and resolve spontaneously, but if bleeding is heavy, it can produce clot colic. Usually, no anatomical abnormality is found to account for the haematuria on radiological investigation.

Central nervous system bleeding is uncommon, but can occur after minimal head injury and remains a significant cause of death in haemophilia A. Intestinal tract bleeding frequently presents with haematemesis and melaena and should be routinely investigated for peptic ulceration or malignancy. It may also present as obstruction due to intramural haemorrhage.

Oropharyngeal bleeding, although uncommon, is clinically dangerous, as extension through the soft tissues of the floor of the mouth can lead to respiratory obstruction. Bleeding from the tongue after laceration can be very persistent and troublesome due to fibrinolytic substances in saliva and the impossibility of immobilizing the tongue.

Surgery and open trauma invariably lead to dangerous haemorrhage in the untreated individual with haemophilia. There

may be persistence of haemorrhage, often after an initial short-lived period of haemostasis. Clots, if formed, are bulky and friable and break down, with recurrent haemorrhage occurring intermittently over days and weeks. Nowadays this is only seen in patients who are resistant to conventional replacement therapy due to the presence of inhibitors (see below) or when patients with mild or moderate haemophilia present after their first surgical or dental procedure. Bruising is a common feature of haemophilia A, but is usually only of cosmetic significance as it remains superficial and self-limiting. Large extending ecchymoses may occasionally require treatment.

Presentation

If a mother is likely or certain to be a carrier of severe haemophilia, the absence of FVIII or FIX in a cord blood sample will establish the diagnosis in the infant. Mild degrees of deficiency may be more difficult to confirm until later, due to liver immaturity (FIX) or stress response (FVIII). However, haemophilia may be sporadic and in approximately one-third of cases there is no family history. In such cases, haemophilia may come to light in the neonatal period with cephalohaematoma or other bleeding resulting from the trauma of birth. In cultures where early circumcision is the rule, this will cause prolonged haemorrhage. Quite often, the diagnosis is delayed until it is noticed that the infant has many large bruises from hand pressure when being picked up or from minor knocks on the cot. These sometimes cause diagnostic confusion and the erroneous label of 'non-accidental injury' may be applied, with needless psychological trauma to the parents. Soon after the infant starts to walk actively, joint bleeding begins to appear. In other children, excessive bleeding from the eruption of primary dentition or from lacerations prompts performance of diagnostic tests. The median age of diagnosis for severe haemophilia is 8 months. Levels of 0.01 or 0.02 IU/mL (moderate haemophilia) may be sufficient to largely prevent spontaneous bleeding and delay diagnosis until bleeding occurs after trauma or surgery. Mild cases may only present in adulthood when severe trauma or surgery provokes unusual bleeding.

Investigation of coagulation defects and haemophilia

Investigation of a suspected inherited bleeding disorder usually begins with global screening tests of coagulation, which although considerably removed from normal *in vivo* coagulation mechanisms, remain extremely useful in detecting and diagnosing coagulation disorders. The principal routine tests in common use are the prothrombin time (PT) and the activated partial thromboplastin time (APTT); also known as the partial thromboplastin time with kaolin (PTTK) or the kaolin cephalin clotting time (KCCT). These tests require the activity of all the conventional procoagulant factors except FXIII; however, they do not require the presence of platelets (they are performed using platelet-poor plasma and exogenous phospholipid) and do not activate the protein-C–protein-S system.

Prothrombin time

In the PT, coagulation is triggered using a thromboplastin, i.e. tissue factor (TF) plus phospholipid; typically re-lipidated recombinant human TF or occasionally derived from animal tissue (rabbit brain). The plasma is re-calcified and the clotting time recorded. The PT will be prolonged by deficiencies in FVII, FX, FV, FII or fibrinogen. It is therefore normal in haemophilia.

Activated partial thromboplastin time

The APPT is triggered using a negatively charged surface (e.g. kaolin, micronized silica or ellagic acid) to initiate contact activation followed by phospholipid (to mimic platelet membrane) and calcium to allow the coagulation cascade to progress. It is thought that activation of coagulation rarely proceeds by this route *in vivo* and consequently deficiency of the contact factors is not associated with any increased tendency to bleeding. The APTT will be prolonged by deficiencies of FXII, FXI, FIX, FVIII, FX, FV, FII, fibrinogen, prekallikrein and high-molecular-weight kininogen. It is therefore prolonged in haemophilias A and B.

Thrombin time

The thrombin time is performed simply by adding a dilute preparation of thrombin to citrated platelet-poor plasma and recording the clotting time. Prolongation of the thrombin time therefore implies that there is either an inhibitor of thrombin present (most commonly the effect of heparin) or that there is a problem with fibrin cleavage or polymerization. It is extremely sensitive to the presence of unfractionated heparin, but is also prolonged by low or abnormal fibrinogen and hypoalbuminaemia. It is normal in haemophilia.

Inhibitors

The simplest interpretation of a prolonged test of coagulation is that one or more of the factors required for its execution is deficient in the patient's plasma. However, an alternative explanation is that an inhibitor of coagulation, usually an antibody, is present. The presence of an inhibitor is typically demonstrated by mixing normal and test plasma (with a prolonged clotting time) in equal proportions. Correction of the prolonged clotting time by the normal plasma indicates a coagulation factor deficiency and failure to correct indicates the presence of an inhibitor. The mixture may need to be incubated for 2 hours for inhibitors of FVIII to manifest their effect.

When an inhibitor is detected, further tests are performed to determine whether it has specificity for a particular clotting factor or not. Non-specific inhibitors usually fall into the class of antiphospholipid antibodies and are paradoxically associated

with thrombosis rather than bleeding (see Chapter 45). In previously normal patients, specific inhibitors are most commonly directed against FVIII, resulting in an 'acquired' haemophilia. In patients with haemophilia they prevent the correction of the factor deficiency by normal plasma.

Specific factor assays

When initial tests or clinical history suggests factor deficiency, then specific factor assays are performed to determine which is affected. This is most usually done by performing a bioassay in which the ability of the test plasma to correct the defect in plasma lacking a specific factor is measured. Bioassays have the advantage of directly measuring the biological activity of the factor in question, but because they also utilize other plasma components, they are susceptible to numerous interferences and confounding effects. This can be avoided in some cases by using artificial small substrates (chromogenic substrates). This aids automation and avoids interference, but often does not capture all aspects of the molecule's biological activity; for example, mutations in the FIX GLA domain may be missed. Finally, the physical amount of the factor present can be assessed immunologically. Immunological assays are accurate and precise, but do not give information about the functional activity of the molecules that have been measured.

FVIII assays

In most laboratories the FVIII assay is a modified APTT which is often referred to as a 'one-stage assay' because the activation of FVIII and coagulation are performed in a single step. However, it is also possible to construct a 'two-stage assay' in which the FVIII is activated in a first step and its activity measured in a second separate step. Chromogenic FVIII assays also use a two-stage procedure. The two assays normally give similar results, but the distinction is important because some defective FVIII molecules give different results in the two assays. Most commonly in these cases, the one-stage assay gives a higher result than the two-stage, but the patient's bleeding phenotype is more in keeping with the latter (lower) result. Thus the assessment of the severity (and sometimes even the diagnosis) of the haemophilia can be mistaken if the two assays are not carried out and compared. This is referred to as 'one-stage–two-stage discrepancy'. The reverse phenomenon is also encountered, but is less common and appears to be infrequently associated with bleeding. Factor concentrates may also give different results with the two assays, which is important when monitoring therapy.

Laboratory diagnosis of haemophilia

Initial tests show a prolonged APTT, normal PT and thrombin time (if performed) and a normal platelet count. This should prompt specific factor assays of FVIII, FIX and FXI. When a low FVIII result is encountered it should be ensured that this is not the result of von Willebrand disease.

Treatment

Clotting factor concentrates

The discovery in the early 1960s that FVIII was concentrated in cryoprecipitate and the subsequent development of lyophilized concentrates, prepared from many thousands of donors, made home treatment feasible and promised the prospect of a normal life for people with severe haemophilia. Similar concentrates containing the vitamin-K-dependent factors were developed for haemophilia B. However, use of such plasma-derived concentrates was associated with the transmission of hepatitis C virus (HCV), hepatitis B virus (HBV) and HIV. Although these infections were present at only low frequency in the donor population, concentrates were made from many thousands of donations and there was a very high risk that batches of concentrate were infected, especially with HCV. Since 1985, viral eradication methods have been rigorously applied (such as terminal dry-heat treatment, solvent/detergent treatment, nanofiltration and pasteurization) in conjunction with donor selection and plasma pool screening and quarantining, reducing the risk of transmitting these viruses to close to zero. However, there remains the problem of transmission of non-lipid-coated viruses (such as hepatitis A and parvovirus), which can survive solvent/detergent sterilization and the possibility that new viruses will emerge. For these reasons, products are required to undergo more than one inactivation/removal process. The technique of nanofiltration has been used, which can prevent transmission of non-lipid-coated viruses.

It has recently transpired that the prion protein responsible for variant Creutzfeldt–Jakob disease (vCJD) may be transmitted via pooled plasma concentrates, although only a single case has been reported to date. With the ongoing concern about prion transmission through blood and blood products, concentrates are no longer produced from plasma collected in countries with notable vCJD in the donor population. Consequently, the plasma-derived clotting factor concentrates in use at present are high-purity products with a high specific activity and excellent safety.

Recombinant FVIII

The FVIII and FIX genes were cloned in 1984 and 1982, respectively. This has allowed the manufacture of the recombinant proteins in cell culture, which therefore eliminates the risk of disease transmission from donors. Progressive refinements have in many cases eliminated all additional animal protein from the cell culture medium and any protein added as stabilizer in the lyophilized product. Recombinant FVIII can be made either from the wild-type 'full-length' cDNA or from cDNA that has the large B domain deleted. The B domain has been shown to be completely dispensable for the procoagulant function of FVIII and its absence does not change the half-life of the molecule.

Subsequently, recombinant technology has allowed the development of modified bioengineered FVIII and FIX

Table 38.2 Indications and guidelines for factor replacement in haemophilia A and B.

Site of haemorrhage	Optimal factor level (%)	Dose (units/kg body weight)		Duration (days)
		FVIII	FIX	
Joint	30–50	20–30	30–50	1–2
Muscle	30–50	20–30	30–40	1–2
Gastrointestinal tract	40–60	30–40	40–60	7–10
Oral mucosa	30–50	20–30	30–40	Until healing
Epistaxis	30–50	20–30	40–60	Until healing
Haematuria	30–100	25–50	70–100	Until healing
Central nervous system	60–100	50	80–100	7–10
Retroperitoneal	50–100	30–50	60–100	7–10
Trauma or surgery	50–100	30–50	60–100	Until healing

Source: Escobar, 2003 [*Haemophilia* 2003; **9**: 360–7]. Reproduced with permission from Wiley.

molecules with enhanced pharmacokinetics, by coupling them to other molecules such as polyethylene glycol, albumin or the immunoglobulin Fc fragment. Clinical studies have demonstrated both safety and efficacy, with prolongation of plasma half-life by a factor of 1.5 for FVIII and up to 5 for FIX.

Treatment of haemophilia

Since the development of replacement therapy, the goal of treatment for haemophilia patients has been the prevention of haemorrhagic episodes and maintenance of a normal life. The severity and the frequency of haemarthrosis is directly related to the degree of deficiency of the clotting factor, but the precise plasma level needed to prevent development of arthropathy is still unknown. Based on studies from haemophilia carriers and patients with mild haemophilia, we can deduce that a level of 0.05-0.3 IU/mL is adequate to arrest minor bleeds, but inadequate for major surgery or trauma. At least 0.5 IU/mL sustained throughout the healing period is needed for normal wound healing to occur.

After trauma or at the onset of bleeding, therapeutic infusion of replacement factor should be administered as early as possible and to a haemostatic level to stop or prevent haemorrhage. The dosing of factor replacement is still based on theoretical calculations and clinical experience. Guidelines for the treatment of haemorrhagic episodes in haemophilia are given in Table 38.2, with the estimated likely effective level required for haemostasis. If sufficient concentrate is available, it is feasible to restore the FVIII level into the normal range and maintain it there by continuous infusion or frequently repeated infusions.

Formulae based on plasma volume and expected recovery give a rough guide to dosage required to achieve a given level, but where the level is critical, as for surgery or when there is serious bleeding, it should always be checked by assay after administration. On average, FVIII infusion produces a plasma increment of 0.02 IU/mL per unit infused per kilogram body weight and FIX a rise of 0.01 IU/mL. From these observations, a simple formula can be derived:

$$\text{Dose to be infused (IU)} = \text{Weight (kg)} \times \text{Desired rise (IU/mL)} \times 100/K$$

where K is approximately 2 for FVIII concentrate and 0.8–1.0 for FIX concentrate.

Estimating the duration of treatment required is a matter of clinical judgement according to the individual circumstance. Cover for surgery, other than very minor procedures, requires maintenance of normal FVIII levels for at least 1 week, followed by a period at reduced dosage during convalescence. This can be achieved either by repeated bolus injections (every 8–12 hours for FVIII and 12–24 for FIX), paying particular attention to trough levels, which should not fall below 0.5 IU/mL, or by continuous infusion. The half-life of FVIII varies considerably and may be particularly short in children, requiring more frequent dosing. It should also be noted that the doses required during the immediate peri- and postoperative period may be considerably more than expected.

Prophylaxis

There is a long tradition of giving prophylaxis to young boys with haemophilia; prophylaxis was developed as early as 1958 in Malmö, Sweden. The rationale for the prophylactic model was the observation that chronic arthropathy was seen less frequently and less severely in individuals with a factor level of 0.01–0.04 IU/mL.

Today prophylaxis is regarded as the standard of care for all boys with severe haemophilia. Regular prophylactic treatment is often begun at around the age of 1 year, before the onset of joint bleeds; so-called primary prophylaxis. Ideally, FVIII is administered every second day at a dose of 25–40 units/kg. FIX can usually be given every third day. The goal is to adjust the

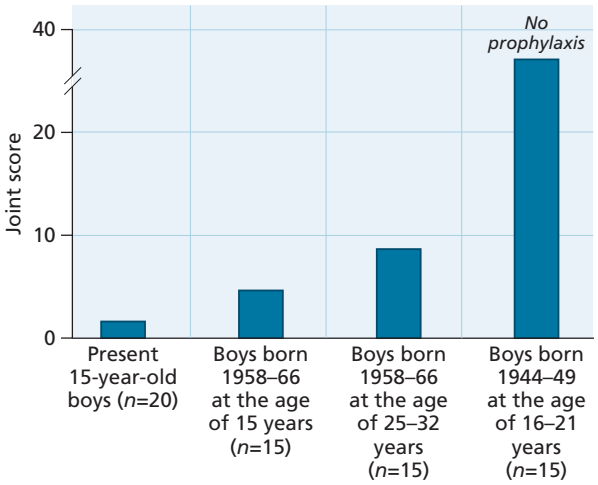


Figure 38.2 Orthopaedic joint scores for patients intensively treated in the present. A group of 15-year-old boys compared with a less intensively treated group at different ages and with patients not receiving prophylaxis (historical control subjects). (Source: United Kingdom Haemophilia Centre Doctors' Organisation, 2003 [*Haemophilia* 2003; **9**: 1–23]. Reproduced with permission of Wiley.)

dose and frequency to maintain a trough level of at least 0.01 IU/mL and to prevent all bleeds. The optimum trough level is not established, but evidence from trials indicates that the risk of bleeding rises as the length of time spent with FVIII levels below 0.01 IU/mL increases. There is good historic evidence that the Malmö model is effective in preventing haemophilic arthropathy. For the youngest cohort of boys born in 1981–90, who began prophylaxis between 1 and 2 years of age with 4000–9000 units/kg annually, both the orthopaedic and radiological scores were zero (Figure 38.2). More recently, a large randomized study comparing prophylaxis with enhanced on-demand therapy showed that after 6 years of therapy, 93% of those on primary prophylaxis had a normal joint score on magnetic resonance imaging (MRI) compared with only 55% on intensive on-demand regimens. The current goal of prophylaxis is to achieve an annualized bleeding rate of as close to zero as possible. Indeed most bleeds seen in individuals on prophylaxis nowadays are secondary to trauma rather than arising spontaneously.

Use of the peripheral vein is the preferred means of delivering prophylaxis. However, if this is not possible then the implantation of an indwelling venous access device (e.g. Port-a-Cath) may be necessary. Another alternative is the fashioning of a small arteriovenous fistula in the arm.

In older patients who have established joint disease or have had previous joint bleeds, prophylaxis is also effective in reducing bleeding episodes and/or slowing progression of joint disease. This 'secondary' prophylaxis may be used continuously or for a limited period to allow the synovial hypertrophy in a target joint to regress, or be targeted to periods when bleeding is more likely (such as a particular activity).

Tranexamic acid

A significant component of the bleeding tendency in haemophilia is the increased susceptibility of the clot to fibrinolysis. Tranexamic acid is a lysine analogue that competes with fibrin for binding to the kringle domains of plasminogen, thus preventing its binding to fibrin and inhibiting fibrinolysis. This anti-fibrinolytic effect makes tranexamic acid a useful adjunctive therapy in haemophilia, especially for bleeding at sites of increased fibrinolytic activity such as the oral mucosa. It can be given orally or intravenously, ideally beginning before surgery, at a dose of 1 g tds (20 mg/kg/day for children, in divided doses). For oral bleeding a mouthwash can be prepared (5% w/v) which is also useful.

Desmopressin

Desmopressin (1-deamino-8-D-arginine vasopressin) can be used to treat mild haemophilia A and provides a non-plasma alternative to concentrates. It has no effect on FIX deficiency. It acts via the V_2 receptors on endothelial cells to stimulate exocytosis of Weibel–Palade bodies, which contain VWF and, in some tissues, also FVIII. In practice, this effect can be used to elevate the plasma FVIII level two- to fourfold above baseline. Desmopressin is a synthetic analogue of vasopressin and retains the antidiuretic action of the natural hormone, but not its vasopressor activity. Patients are advised to restrict their fluid intake after desmopressin. This is especially important in young children in whom unrestricted fluid intake may result in hyponatraemia and convulsions. Desmopressin is contraindicated in elderly patients and those with vascular disease, because arterial thrombosis has been reported following desmopressin in these circumstances.

Desmopressin can correct the haemostatic defect in mild haemophilia A sufficiently to cover minor surgery or treat a minor bleeding episode. A typical regimen would be to give 0.3 μg/kg body weight either by slow intravenous infusion over 20 min or by subcutaneous injection. Desmopressin also releases tissue plasminogen activator and so tranexamic acid is frequently given as well. A high-concentration nasal spray is also available that delivers 150 μg per actuation. The effect is maximal at 30 min after an intravenous dose and 1 hour following a subcutaneous or intranasal dose. The half-life of the endogenously released FVIII is about 8 hours, but may be shorter in some cases and the fall-off should be documented. The dose can be repeated and although tachyphylaxis occurs such that the second dose may produce a rise only 30% of the first, there is often no further drop with succeeding doses. If repeated doses are given over a short interval (e.g. 48–72 hours), the FVIII response to desmopressin injection should be assayed to ensure that the reserves are not exhausted.

Complications of therapy

Inhibitors

Antibodies to FVIII (inhibitors) are most common in patients with severe haemophilia A and are most likely to develop during childhood within the first 25 exposures to FVIII; after which they are infrequent, but do rarely develop after years of exposure, typically after periods of intensive FVIII treatment. In severe haemophilia A, inhibitors are detected in around 25–30% of patients at some stage in their treatment. However, many of these are low-level transient inhibitors that are not a long-term problem. The cumulative incidence of high-responding inhibitors is in the region of 10–15%. The likelihood of an individual developing an inhibitor is strongly dependent on underlying FVIII mutation. Patients with large deletions or stop codons in the light chain appear to be particularly prone to inhibitor development. Approximately 25% of patients with the common intron-22 inversion develop inhibitors.

Inhibitors in moderate and mild haemophilia are far less common. However, some missense mutations causing milder haemophilia, particularly those in the light chain, are strongly associated with inhibitor development. Typically, such patients develop inhibitors when FVIII is administered in large amounts (e.g. severe bleeds) or in association with inflammatory stimuli (e.g. surgery). Because the development of inhibitors is unpredictable, it is important to test for their emergence at clinic visits and before and after operations in all patients that have been exposed to FVIII concentrates.

Antibodies to FVIII frequently have a time-dependent component in their action and an incubation time of 2–4 hours is required to reliably measure the full effect of the antibody. Inhibitors usually inactivate FVIII in a predictable linear manner (type 1 kinetics), but they may have complex equilibria and not inactivate all available FVIII (type 2 kinetics).Type 2 kinetics are seen particularly in mild to moderate and acquired haemophilia. Inhibitors are most commonly measured using the Bethesda assay where 1 Bethesda unit (BU) is the amount of antibody that reduces the FVIII activity in pooled normal plasma by 50% after 2 hours, incubation. Some inhibitors are 'low-responding' and remain at a low or moderate level (<5–10 BU) after further FVIII exposure, whereas in high responders, treatment with FVIII elicits a sharp anamnestic rise after 5–8 days and inhibitor levels may reach hundreds or thousands of BU. Low responders can be treated repeatedly with high doses of FVIII concentrate. However, high responders are refractory to treatment with FVIII concentrate and therefore alternative therapies are required.

Inhibitors in haemophilia B are much less common than in haemophilia A. Only about 2–3% of patients with severe haemophilia B will develop inhibitors, but interestingly many of these inhibitors present with anaphylaxis to infusions of FIX concentrate sometimes after only one or two exposures. Unlike FVIII inhibitors, FIX inhibitors are not time dependent, but can be similarly quantitated by a modified Bethesda assay. The majority of patients who develop inhibitors have large gene deletions or nonsense mutations occurring in the first 20% of the FIX gene. Acute management of bleeding can be achieved with recombinant FVIIa (as with haemophilia A).

Inhibitors: treatment of bleeding

At present the most common approach to achieving haemostasis in patients with inhibitors is to use a 'bypassing agent' that promotes thrombin generation without depending on the intrinsic tenase. The currently available bypassing agents are human recombinant FVII (FVIIa) and a plasma-derived prothrombin complex concentrate that contains activated coagulation factors (factor eight inhibitor bypassing activity, or FEIBA). Activated FVII is an inefficient enzyme when not complexed with TF and does not therefore produce systemic activation of coagulation when infused. Its mechanism of action in haemophilia has been the subject of much debate. It seems most likely that FVIIa binds to the surface of activated platelets at the site of injury and promotes FXa formation there. Recent evidence suggests that platelets may be able to synthesize TF or at least acquire it from circulating microparticles. High concentrations (30–90 nmol/L) of FVIIa are required to generate sufficient FXa and FIIa to achieve haemostasis. The advantages of recombinant FVIIa include viral safety, low systemic activation of coagulation, effectiveness independent of inhibitor titre and an excellent overall safety profile, allowing home use and ease of administration. The main disadvantage is the short 2-hour half-life and therefore the need for frequent administration at standard dose (90 μg/kg), although this may be partially overcome by using a 270 μg/kg dose for mild to moderate bleeds.

Both native and activated prothrombin complex concentrates appear to be able to bypass the need for FVIII and have a role in the treatment of patients with FVIII antibodies. Their efficacy appears to derive from a combination of activated clotting factors and the elevation of prothrombin concentration, which is a major determinant of thrombin generation. The activated prothrombin complex concentrates contain more of the FVIII bypassing activity and are designed specifically for use in patients with FVIII inhibitors. Activated prothrombin complex concentrates (e.g. FEIBA) have proved efficacious in controlling bleeding in many situations, but are of plasma origin and contain some FVIII, which may cause an anamnestic rise in inhibitor titre. On the other hand, the longer dosage interval (given once or twice daily) makes it simpler to use and suitable for prophylactic use. A major problem with both bypassing agents is that there is no reliable laboratory measure of their haemostatic effect.

A number of new agents for achieving haemostasis in the presence of coagulation inhibitors are under development. Porcine FVIII functions well in human plasma, but is frequently neutralized less efficiently by human anti-FVIII antibodies: a recombinant porcine FVIII has been developed and appears effective in clinical trials. Modified FVII molecules with improved activity,

inhibitors of TFPI and a bispecific antibody mimicking FVIII's ability to bind FIXa and FX are also in development.

Treatment of inhibitors: immune tolerance

Bypassing agents are considerably less effective than factor replacement at achieving haemostasis and cannot yet provide reliable prophylaxis. It is therefore highly desirable to eliminate a FVIII or FIX inhibitor, a process referred to as immune tolerance induction (ITI). Basic ITI regimens involve the continued administration of FVIII, and both low-dose (50 units/kg three times per week or alternate days) and high-dose (100 units/kg twice daily) regimens have been used. Neither regimen has been conclusively shown to be superior, but the response is more rapid and bleeding events fewer with the high-dose regimen. Both may need to be continued for 12 months or more. Overall, the success rate is approximately 70%, after which prophylaxis must be maintained to prevent re-emergence of the antibody. The success is greatest in those with a historic peak titre below 200 BU, when treatment is begun after allowing the inhibitor to fall to below 10 BU and where there are no interruptions to the ITI regime. For resistant cases, second-line approaches include the use of a FVIII concentrate containing VWF, combination with anti-CD20, or corticosteroids, alkylating agents, plasmapheresis and immunoglobulin (Malmö regimen). Success is defined as normalization of the FVIII recovery and a FVIII half-life greater than 7 hours.

Immune tolerance for FIX inhibitors is generally less successful and is sometimes precluded by anaphylactic responses or complicated by the development of nephrotic syndrome.

The aging haemophilia population

Although HIV and HCV remain significant complications of therapy, improvements in both haemophilia care and antiviral treatments have resulted in significant increases in median life expectancy for individuals with haemophilia over recent years. Increasing numbers of patients are living beyond 65 years of age. The classic complications of haemophilia, including intracranial haemorrhage, joint disease, and inhibitor development all increase with increasing age, as well as hepatic complications from HCV and HIV/HCV coinfection. In addition, the older patients face the same medical conditions associated with ageing in the general population, such as cardiovascular disease and cancer. Care can become very complex, such as use of antithrombotic therapy for cardiovascular disease in a patient with an inherited bleeding disorder. Multidisciplinary and cross-specialty coordination is key to successful management.

Molecular genetics of haemophilia A

The FVIII gene (Figure 38.3) spans 190 kb of the X chromosome. The protein-coding regions (exons) are separated by 25 introns, some of very large size (e.g. intron 22 is over 35 kb). FVIII is synthesized in endothelial cells, including those within the liver and spleen, which creates the desmopressin-releasable pool. The amino acid sequence contains a triplicated region (A1, A2, A3 in Figure 38.3), whose elements are more than 30% homologous with each other. A second duplicated homology region (C1, C2) is responsible for the phospholipid-binding properties of the molecule. The third type of sequence in the protein is the heavily glycosylated B domain, which is coded entirely within exon 14, connects A2 and A3, and is removed on thrombin activation of FVIII.

The processed mRNA specifies a protein of 2351 amino acids, from which a 19-amino-acid N-terminal leader sequence is cleaved on secretion. The mature plasma protein initially consists of a single chain of 2332 amino acids, but as a result of proteolysis circulates as a series of light-chain and heavy-chain heterodimers with B domains of varying length.

The vast majority (>90%) of FVIII circulates in complex with VWF, in the absence of which its half-life is extremely short. Studies of FVIII have identified the a3 (1648–1689), C1 and C2 domains as important for VWF binding, but it is uncertain how many of these interact directly with VWF because in isolation the recombinant fragments have a much lower affinity. After thrombin activation, the cofactor consists of a heterotrimer: the N-terminal heavy chain corresponding to region A1 and the A2 domain isolated by cleavages at 372/373 and 740/741, which is held by a divalent cation-dependent linkage to the C-terminal light chain corresponding to part of A3 plus C1 and C2 (see Figure 38.3). The B domain is discarded by cleavages at 740/741 and 1689/1690. Thrombin cleavage also results in release of FVIII from VWF.

Protein C inactivates FVIIIa by cleavage at 336/337 and 562/563 but *in vivo* inactivation of FVIIIa is mostly due to spontaneous dissociation of the heterotrimer. The crystal structure of FVIII has recently been reported and several of the interactions mapped including the FIX-binding site on the A2 and A3 domains. The C domains appear to be responsible for phospholipid binding.

Thanks to the development of rapid sequencing methods, the mutations in nearly all patients with haemophilia A can be identified (Table 38.3). Consequently, a large database of mutations has accumulated and is maintained at http://www.factorviii-db.org. These include large deletions and missense mutations, as well as nonsense, frameshift, splicing (affecting mRNA processing) and insertional mutations. Surprisingly, about half of all severe haemophilia A is due to a major inversion (Figure 38.4) that occurs quite frequently during male gametogenesis (approximately 1 in 10,000 spermatozoa).

Molecular genetics of haemophilia B

The FIX gene lies telomeric to the FVIII gene on the X chromosome at Xq27.1 and is considerably smaller, extending for only 34 kb in length. It comprises eight exons producing a 2.8-kb transcript mRNA encoding the protein of 461 amino acids. FIX is

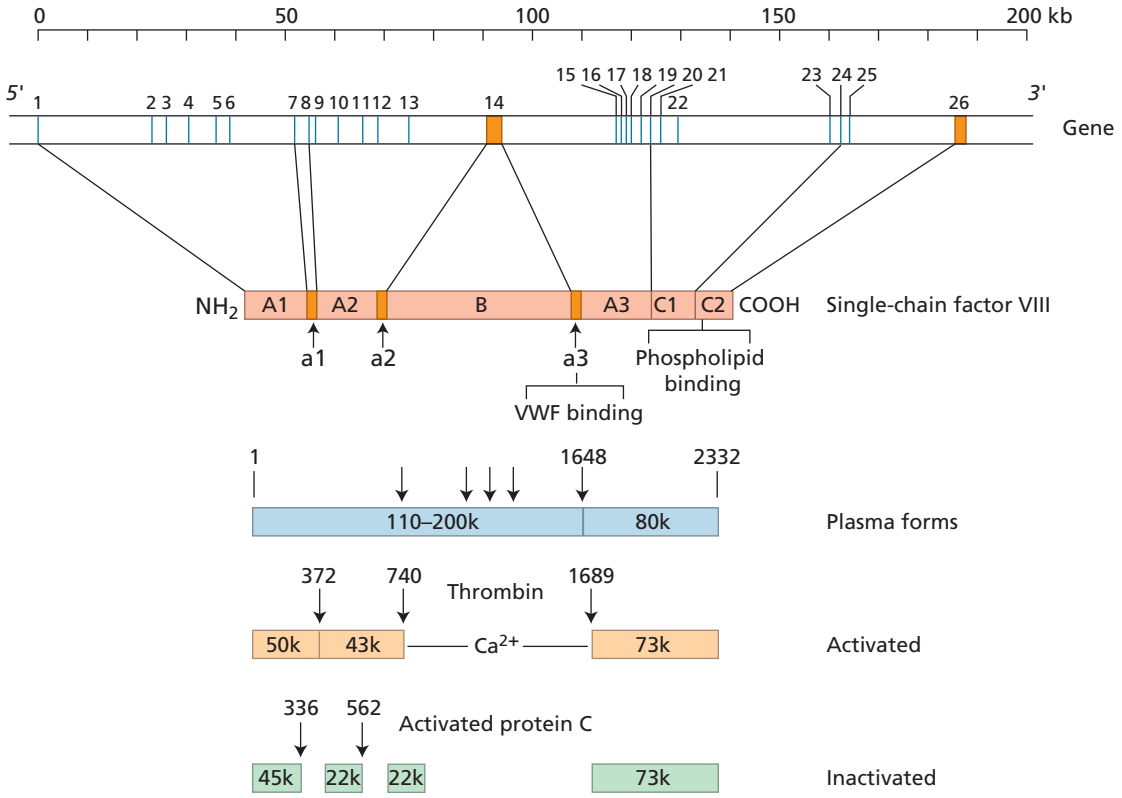


Figure 38.3 The FVIII gene and protein. Top line, scale for gene in kilobase pairs of DNA; second line, location of exons (protein-coding segments) of the FVIII gene shown as solid bars; second to third line, corresponding domains of FVIII gene and protein; fourth line, plasma FVIII is heterogeneous due to cleavage at various points within the B domain during synthesis and secretion, yielding heterodimers in which the heavy chain (N-terminal segment) varies from 740 to 1648 amino acids in length; fifth line, thrombin activation releases the a3 segment from the light chain, the B domain from the heavy chain and cleaves between the A1 and A2 domains; bottom line, inactivation of factor VIIIa by activated protein C is accompanied by cleavages at 336, releasing the a1 domain and 562 in the middle of the A2 domain.

synthesized in the liver and has a domain structure similar to the other serine proteases of the coagulation cascade, comprising a signal peptide, a GLA domain, two EGF domains and a serine protease domain. The single-chain zymogen can be activated by both FXIa and the TF-VIIa complex; in both cases cleavages after Arg 191 and Arg 226 release an activation peptide creating a disulfide-linked two-chain active enzyme. The mutation spectrum responsible for haemophilia B is quite different from that responsible for haemophilia A, primarily because there is no single mutation corresponding to the intron 22 inversion. Instead the majority of mutations responsible for severe haemophilia B are missense mutations rather than nonsense mutations, as for haemophilia A (http://www.factorix.org/). This may account for any difference in phenotype, but again it is large deletions that are associated with the highest risk of inhibitor formation. The normal plasma concentration of FIX (~90 nM) is much higher than that of FVIII (~0.4 nM).

Gene therapy for haemophilia

Haemophilia is an excellent model for gene therapy because the clinical manifestation is the result of a deficiency of the single gene product and only a small amount of protein is required to ameliorate symptoms. Recently, using modified serotypes and engineered self-complementary adenoassociated virus, Phase I studies have achieved prolonged (>2 years) expression of FIX (0.01–0.06 u/mL) with clear therapeutic benefit in some subjects. A few patients required transient immunosuppression due to an immune response to the vector. Further AAV-mediated gene transfer trials in haemophilia B are currently being undertaken, aiming to express the factor IX gene from the liver. Although much more successful than previous approaches, challenges remain that require resolution, including humoral and cellular immunity to the AAV and naturally occurring vector-neutralizing antibodies.

Table 38.3 Mutations responsible for haemophilia A.

Mutation type	No. of mutations observed	Percentage
Intron 22 inversion		
Distal	260	30.7
Proximal	37	4.4
Rare	5	0.6
Total	302	35.7
Intron 1 inversion		
Total	8	1.0
Point mutations		
Missense	323	38.2
Nonsense	79	9.3
Total	402	47.5
Small deletions or insertions		
Small deletions	63	7.5
Small insertions	22	2.6
Combination of insertions and deletions	1	0.1
Total	86	10.2
Large deletions (>50 bp)		
Total	25	3.0
Splice-site mutations		
Intronic	20	2.4
Cryptic	2	0.2
Total	22	2.6

Source: Graw *et al.*, 2005 [*Nature Reviews Genetics* 2005; **6**: 488–501]. Reproduced with permission from Nature Publishing.

FVIII remains a significant challenge due to the large size of the FVIII gene and associated problems accommodating the larger FVIII cDNA, achieving adequate levels of transgene expression and continuing concerns regarding the prevention or risks of FVIII inhibitor development, which is much more common than in FIX deficiency. Improved FVIII expression cassettes have been developed to attempt to address the scientific challenges posed by FVIII.

Although there is currently much focus on AAV-based approaches, following the success of the FIX programme, alternative approaches, such as HSC transduction, and vectors, such as lentivirus, remain in development.

Haemophilia A and B in females

True homozygous haemophilia A is rare and even rarer for haemophilia B, but well described, being due to the marriage of a carrier to an affected male. Severe menstrual haemorrhage

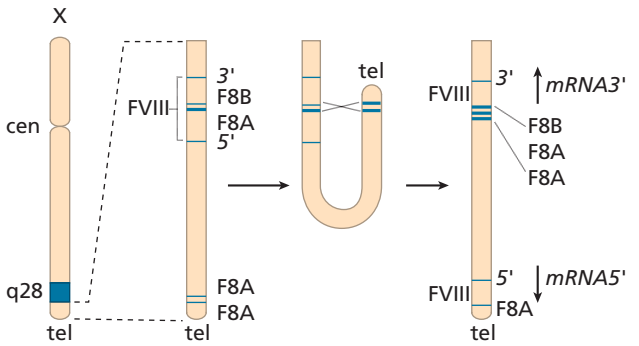


Figure 38.4 How the tip flips: the mechanism of inversion through intron 22. The mutation is responsible for up to half of all cases of severe haemophilia A. During spermatogenesis, at meiosis the single X pairs with the Y chromosome in the homologous regions, but there is nothing to pair with most of the long arm of X. Unfortunately, the possibility of intrachromosomal pairing and crossover exists because there are three copies of a gene designated *F8A*, one lying within intron 22 and two in the opposite orientation situated 400 kb telomeric to the FVIII gene. Cross-over with either the distal or the proximal *F8A* copy divides the FVIII gene in two, such that separately transcribed mRNAs are produced, neither of which encodes functional FVIII.

occurs, but responds to replacement therapy. Haemophilia A and B have been described in females with Turner syndrome.

X-chromosome inactivation allows only one FVIII/FIX gene to function in each female cell. Because the process is random, carriers of haemophilia A or B have, on average, 50% of the mean normal level of the clotting factor, but skewing of the inactivation towards the X chromosome containing the abnormal or the normal FVIII/FIX gene may occur, resulting in a clotting factor level that is either well within the normal range or low enough to produce a mild, or even moderate, haemophilia phenotype. It is important to note that while a low level of FVIII/FIX strongly implies carriership, a normal FVIII/FIX level does not exclude it.

As a result of *de novo* mutation, especially the FVIII intron 22 inversion in spermatogenesis, new cases may present with no family history of haemophilia. In such women, careful investigation is essential to ensure the correct diagnosis is made – haemophilia A carrier, VWD type 1 or VWD type 2N (see below) – due to the reproductive implications. Molecular analysis is often required to clarify.

Current techniques make definitive diagnosis by identification of the causative mutation in patients and carriers straightforward in the vast majority of cases. In identified carriers, antenatal diagnosis can be carried out when pregnant, and although typically only offered in cases of severe haemophilia it can now be performed on free fetal DNA in maternal blood, which may widen its practice. Alternatives include molecular analysis of chorionic villus biopsy at 11–12 weeks' gestation and third trimester amniocentesis. Preimplantation genetic

diagnosis (PIGD) with selection of unaffected embryos for re-implantation has been successfully carried out for haemophilia.

General organization of haemophilia care

As these are relatively uncommon disorders, with effects on patients and families at all stages of life, who require care and support services across the whole field of medicine and social services, it is now accepted that this is best delivered comprehensively by specialized centres. The staff of a major comprehensive care centre will include physicians, nurses, social workers, laboratory scientists and physiotherapists, devoting all or a substantial part of their time to haemophilia care. An orthopaedic surgeon or musculoskeletal rheumatologist prepared to see haemophilic patients regularly in a clinic set aside for their problems is a valuable addition to this team. Access to specialized care for monitoring patients with HIV and HCV infection is still important.

Within the NHS, a service specification has been developed to support and set out the provision of haemophilia services. It is recognized that not every centre can provide every facility and that there should be a wide distribution according to population density. The functions of a centre are to provide 24-hour emergency treatment for haemophilic patients and their families and a full range of diagnostic tests for identifying new patients and monitoring treatment. Full records should be kept of all treatments, whether given in hospital or as home therapy and progress monitored through regular follow-up. Paediatric centres are best located within paediatric units. Genetic counselling, including carrier detection and antenatal diagnosis, must be available for families of patients with haemophilia. All patients should have direct or indirect access to a large centre providing all these facilities and treating a large number of patients. On diagnosis, all patients are issued with a special medical card indicating laboratory test results, inhibitor status, main centre and local centre for treatment. Caring for haemophilic patients and their families is demanding, but rewarding. Despite the setbacks in the 1980s due to virus transmission, the trend continues to be an ever-improving life expectancy and social participation, based on continuing medical progress and multiprofessional input. The younger generation of haemophiliacs is now treated with virus-safe or recombinant concentrates and has escaped both HIV and liver disease, and can lead virtually normal lives.

Acquired haemophilia

This chapter is devoted to the inherited form of haemophilia arising from mutations in the FVIII and FIX genes, but haemophilia can also arise as a result of autoantibodies that neutralize FVIII cofactor activity. This is covered in Chapter 40, p. 758.

von Willebrand disease

Although described in 1926, our understanding of this complex and variable bleeding disorder remains far from complete. The basic defect common to all variants is a deficiency of one or more aspects of VWF functional activity. The abnormality may be quantitative and/or qualitative.

von Willebrand factor (VWF)

VWF is a large plasma glycoprotein encoded by a gene on chromosome 12p13, which was cloned simultaneously by several groups in 1986. The prepro-VWF primary translation product comprises a 22-amino-acid signal peptide, a 741-amino-acid propeptide and the 2050-amino-acid mature subunit. After a recent structural review, the pro-VWF monomer is now described as comprising three types of domains (A,C,D) arranged as follows: NH_2–D1–D2–D′–D3–A1–A2–A3–D4–C1–C2—C3-C4-C5-C6-CK–COOH (Figure 38.5). VWF is produced predominantly in vascular endothelial cells and also in

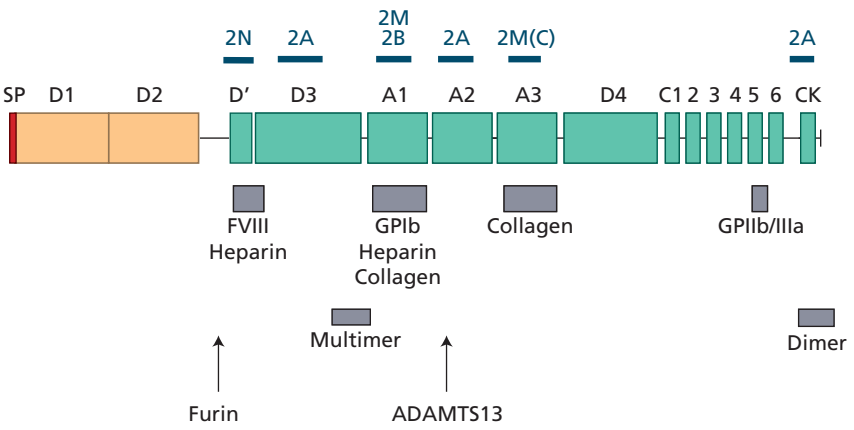


Figure 38.5 The domain structure of the primary VWF translation product. The SP-D1D2D'D3A1A2A3D4C1-C6CK domain structure is shown with the binding properties of the domains indicated below. The corresponding VWD types associated with mutations in the various domains are shown above the domains. Cleavage by Furin removes the propeptide before or soon after secretion. The ADAMTS13 cleavage site in the A2 domain is also shown. SP, signal peptide, CK, cysteine knot. 2M(C) indicates a type 2M pattern caused by defective collagen binding.

megakaryocytes, although plasma VWF derives entirely from endothelial cell production. After translation, the monomer enters the endoplasmic reticulum and the signal peptide is removed. The monomers then dimerize via the C-terminal cysteine knot domain before leaving the endoplasmic reticulum for the Golgi, where the dimers are joined via their N-termini to form a series of multimers with molecular masses ranging from 1000 to 20,000 kDa. During this process the molecule is extensively glycosylated with N- and O-linked glycans that form 20% of its final molecular weight. Importantly, approximately 13% of the N-linked glycans terminate in an ABO blood-group antigen. In endothelial cells, multimerized VWF is then transferred into Weibel–Palade bodies, some of which pass directly to the cell surface to release their contents (constitutive pathway), while others remain in the cell for release in response to inflammatory stimuli (regulated pathway). Platelet VWF does not contribute significantly to plasma VWF and is stored in the α-granules of platelets prior to release upon platelet activation. During synthesis, the VWF propeptide, which is essential for multimerization, is cleaved from the mature molecule, stored with VWF in the Weibel–Palade bodies and released with it into plasma, but is not known to have any further function thereafter. Some VWF is also secreted ablumenally from the base of the endothelial cell where it binds directly to collagen in the subendothelial matrix.

Rapid release of VWF from the endothelial Weibel–Palade storage granules via the regulated pathway can be induced by a number of agonists including thrombin, adrenaline, histamine and vasopressin. This can be used to advantage by using desmopressin to treat mild forms of VWD. Elevation of plasma VWF due to increased synthesis and secretion occurs as part of the acute-phase response to injury, inflammation, infection and neoplasia, and also in pregnancy and hyperthyroidism. The confounding effects of stress and inflammation need to be taken into account when measuring VWF levels for diagnosis.

The two principal functions of VWF are: (i) binding to matrix molecules, particularly collagen, at sites of vascular injury and subsequent capture of platelets to form the primary haemostatic plug and (ii) the stabilization of FVIII in the circulation. The adhesive function of FVIII requires functional binding sites for collagen and the platelet glycoproteins GPIb and GPIIb/IIIa (also known as integrin $\alpha_{IIb}\beta_3$) and is also dependent on the presence of large VWF multimers. The VWF released from endothelial cells into plasma is in the form of ultra-large multimers, which are subsequently cleaved into smaller forms by the action of the plasma metalloprotease ADAMTS-13. This is of critical importance because it is the high-molecular-weight multimers that have the greatest collagen- and platelet-binding activity: absent cleavage results in formation of platelet microthrombi and vascular occlusion (in thrombotic thrombocytopenic purpura; TTP) and excessive cleavage results in increased bleeding (type 2A VWD).

FVIII and VWF (Figure 38.6) circulate together as a complex in which VWF protects FVIII from degradation, so that a deficiency in VWF or a reduction in its ability to bind FVIII may also result in a low plasma level of FVIII. Therefore, deficiency of VWF can give rise to a dual haemostatic defect: reduced plasma levels of FVIII (due to its shorter half-life in the absence of VWF) and a defect in primary haemostasis because of the failure in assisting platelets to adhere to the cut edges of small blood vessels. Clearly, the multistep synthesis, assembly and secretion of VWF and its multiple binding interactions provide many ways in which mutations of the VWF gene can lead to loss of function. In particular, the multimeric structure is susceptible to dominant-negative effects of mutant alleles. The diagnosis of VWD is based on the recognition that for normal VWF function it must: (i) be present in adequate amounts, (ii) have a normal multimeric structure, and (iii) have intact functional domains (binding sites). When any of these properties is only slightly reduced, the presence or absence of a bleeding tendency may also depend on the quality and quantity of the other haemostatic components, particularly platelets and collagen, with which VWF must interact.

Clinical features

The clinical features of VWD vary according to the severity of the deficiency. The most common clinical picture is of a mild to moderately severe bleeding tendency, characterized by bruising, epistaxis, prolonged bleeding from minor cuts, menorrhagia and excessive, but not often life-threatening, bleeding after trauma or surgery. Patients often present for investigation in the second or third decade after prolonged bleeding from dental extraction or surgery has aroused clinical suspicion. Menorrhagia, unexplained by local or hormonal factors, can also be the presenting symptom: an estimated 20% of women with menorrhagia have VWD. The distribution of bleeding in VWD can be explained on the basis that VWF is required for platelet adhesion at high shear rate, which is the condition of flow found in the smallest blood vessels exposed to trauma in skin and mucous membranes. Much less common is autosomal recessive (type 3) VWD, where VWF is completely absent and FVIII:C levels are usually around 0.01 or 0.02 IU/mL. These patients have a bleeding tendency that also has features of severe haemophilia A, with haemarthroses, muscle bleeds and life-threatening haemorrhage after trauma, as well as a propensity to small-vessel bleeding, which is not a feature of haemophilia A.

Mild bleeding symptoms of the type typical of VWD are very common in the general population and thus have a low specificity for the presence of a bleeding disorder. It is now recommended that a bleeding score based upon bleeding history and calculated using a questionnaire is used to assess the likelihood of a significant defect in haemostasis and that the result is borne in mind when interpreting the results of laboratory investigations.

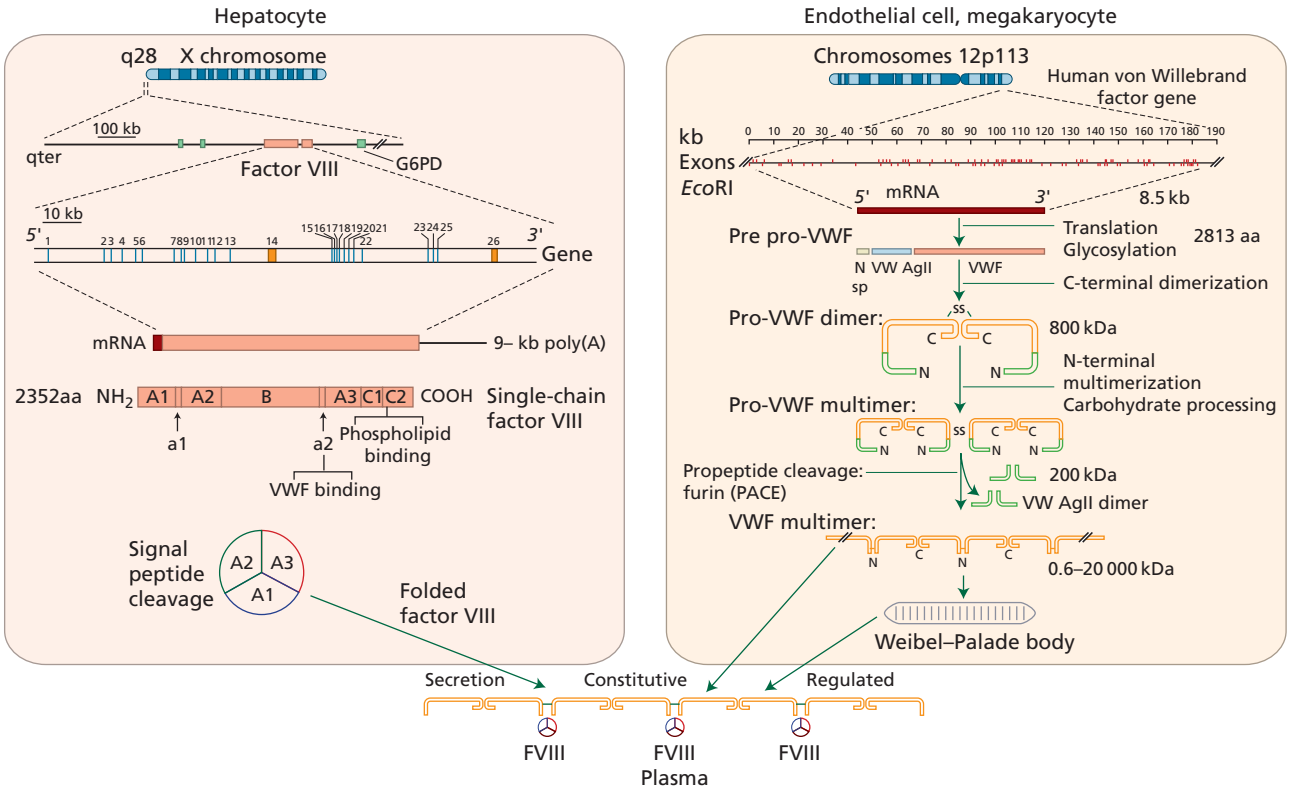


Figure 38.6 Assembly of FVIII–VWF complex. FVIII synthesized by hepatocytes as a single chain is partially proteolysed and (possibly) complexed with VWF prior to release. VWF is synthesized as a single-chain precursor, which dimerizes with disulfide bond formation, then multimerizes with further disulfide exchange, with concomitant loss of a large propeptide segment. The propeptide is essential for multimer formation and probably functions as an acidic disulfide isomerase.

Laboratory diagnosis

Preliminary diagnosis

After obtaining a suggestive personal history and/or a family history, the laboratory diagnosis of VWD rests on assessing both the amount of VWF present (VWF:Ag) and its functional capacity. At present, it is generally possible to assess three important functions:

1 *FVIII binding (VWF:FVIIIB).* Assessed first by a FVIII assay and then if this is reduced, by an assay of VWF FVIII-binding capacity using a modified ELISA-based technique.

2 *Platelet-dependent function (VWF:RCo).* The standard assessment of VWF adhesive functional activity remains the ristocetin cofactor assay, in which dilutions of patient plasma are tested for their ability to promote platelet agglutination in the presence of the antibiotic ristocetin. Automated versions are now available using latex beads.

3 *Collagen-binding function (VWF:CB).* This is performed in an ELISA-based assay in which a well coated with collagen is used to capture VWF.

Measures of VWF:RCo and VWF:CB are both sensitive to the loss of high-molecular-weight multimers, but measure different binding properties of VWF. Thus they should be seen as complementary rather than alternative assays.

Secondary tests for classification of VWD

If a deficiency suggestive of VWD is detected then further tests, particularly multimer analysis and ristocetin-induced platelet aggregation (RIPA) should be performed to allow accurate classification (Figure 38.7, Table 38.4). The classification of VWD has been examined many times and extensive guidelines are available, detailed in the bibliography. What follows is necessarily a limited summary of much work.

Type 1 VWD is defined as a simple quantitative deficiency of VWF, and the VWF present should be functionally normal with a normal multimeric pattern. In practice it is acceptable that the ratio of function to antigen (VWF:RCo/VWF:Ag) should be >0.6 and there should be no significant loss of high-molecular-weight multimers, although on close inspection some subtle abnormalities of multimer pattern may be discerned.

The characteristic of the 2A, 2B and 2M variants is a functional deficiency of VWF activity, which is reduced to below 0.6 of the antigenic measure. In type 2A, this results

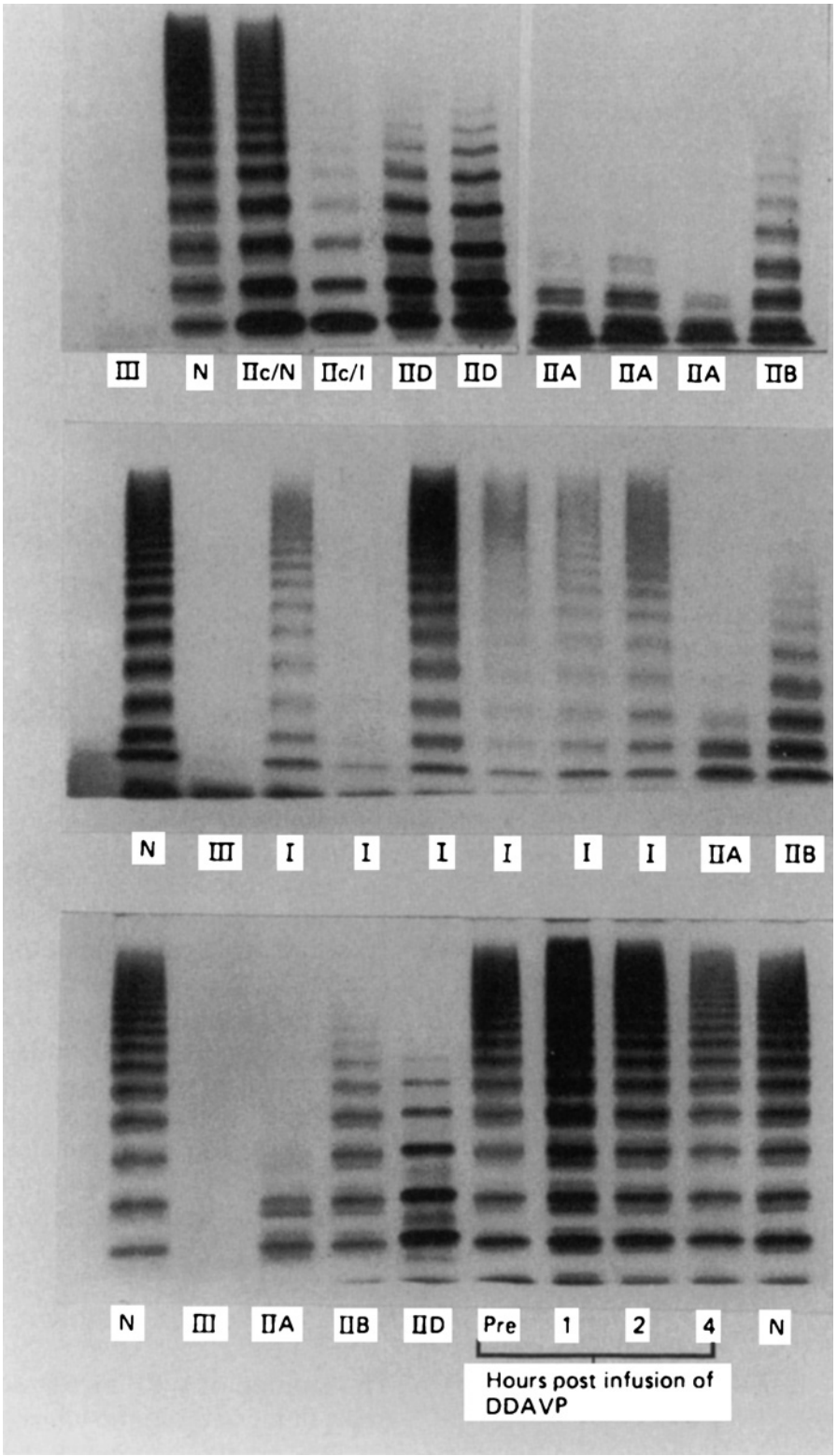


Figure 38.7 Multimer analysis of VWF from patients with VWD and normal control subjects. Note that many cases classified at present as type 2A were previously subdivided according to the details of the abnormality of triplet pattern.

Table 38.4 Classification of von Willebrand disease.

Type	Description	Comments	Inheritance
1	Partial quantitative deficiency of VWF	Includes VWF mutations causing rapid VWF clearance (e.g. VWF Vicenza) and requires function:antigen ratio >0.6	Mostly autosomal dominant inheritance when VWF <0.3 IU/mL. Mutations of *VWF* in kindred with levels >0.3 IU/mL show variable penetrance
2	Qualitative VWF defects		
2A	Decreased VWF-dependent platelet adhesion with selective deficiency of high-molecular-weight multimers	Some controversy exists regarding classification of VWF mutations associated with subtle reductions in HMW multimers	Mostly autosomal dominant
2B	Increased affinity for platelet GPIb	Should be distinguished from platelet type pseudo-VWD (PT-VWD), using either platelet agglutination tests or genetic testing. Cases with normal VWF multimer and platelet count have been described	Autosomal dominant
2M	Decreased VWF-dependent platelet adhesion without selective deficiency of HMW multimers	This also includes defects of VWF collagen binding. May be combined quantitative/qualitative defect	Autosomal dominant
2N	Markedly decreased binding affinity for FVIII	Should be distinguished from mild haemophilia A	Reduced VWF:FVIII binding defects are more commonly identified in a compound heterozygote state with a VWF null allele rather than the classical homozygous form
3	Virtually complete deficiency of VWF	Equivalent to <0.03 iu/mL in most assays	Autosomal recessive, frequent null VWF alleles. Bleeding symptoms in 26-48% of obligate carriers

Source: Laffan *et al*., 2014 [*Br. J. Haematol*. 2014 **167**(4):453–65.] Reproduced with permission of Wiley.

from a lack of high- and intermediate-size VWF multimers, whereas in type 2M a similar loss of platelet-binding activity is seen despite the presence of normal VWF multimeric composition and is caused by mutations in the VWF GPIb-binding site. Loss of high-molecular-weight multimers in type 2A may arise from failures of intracellular processing or from accelerated proteolysis in the circulation due to mutations close to the ADAMTS-13 cleavage site. VWD due to an isolated collagen-binding defect without loss of high-molecular-weight multimers is rare, although a number of cases have been reported. The associated bleeding disorder is mild, perhaps because VWF contains two collagen-binding sites. If multimer analysis is not available, then the loss of high-molecular-weight multimers can be inferred from a drop in both VWF:RCo/VWF:Ag and VWF:CB/VWF:Ag ratios.

In type 2B, VWF binds prematurely with platelets in the circulation. *In vivo* this results in the loss of high-molecular-weight multimers and platelets from the circulation due to formation of platelet aggregates and to increased cleavage of platelet-bound VWF by ADAMTS-13. The abnormality is detected in the laboratory by agglutination of platelets at a low concentration of ristocetin (0.5–0.7 mg/mL) that does not cause agglutination with normal VWF. If normal washed platelets are re-suspended in the patient's plasma, the phenomenon is reproduced, demonstrating that the abnormality is in the plasma. Thus, RIPA should be routinely performed as part of the diagnostic work-up, as it is important to detect this variant (see below), which does not always show up in the other tests. Indeed, a recent analysis showed that many patients with type 2B VWD have normal platelet counts and VWF multimer patterns.

Type 2N VWD is characterized by VWF which has reduced affinity for FVIII (VWF:FVIIIB), but is normal in other respects. Preliminary investigations reveal only a reduced FVIII level (5–35%), which is easily mistaken for mild haemophilia. Inheritance is autosomal recessive, which may provide a clue to diagnosis, but confirmation requires an assay of VWF FVIII-binding capacity or genetic analysis of FVIII and the VWF FVIII-binding site in the D′ domain.

Overall, the majority of patients are found to be type 1, which accounts for 75% of kindreds. Types 2A and 2B are

fairly common, together accounting for about 15% of kindreds. Type 3 is the least frequent and has a prevalence of 1–2 in 10^6. Although classification has proved useful in understanding VWD and can help in planning treatment, some forms remain difficult to classify and may show features of more than one type.

Problems in diagnosis of type 1 VWD

The diagnosis of type 1 VWD implies that the patient has a significantly low level of VWF, which is responsible for an increased tendency to bleed. It is often not easy to be certain of this conclusion for the following reasons:

1 Slightly low levels of VWF are common in the population. Although the lower end of the normal range for VWF in the general population is approximately 0.5 IU/mL, most of the 2.5% of the population who have levels below this have normal haemostasis. In particular, individuals with blood group O have VWF levels on average 30% lower than those with non-O blood groups and the lower end of a 'blood group O normal range' is approximately 0.35 IU/mL. Thus the lower limit of the normal range is not the minimum level required for normal haemostasis. Nonetheless, individuals with blood group O are over-represented in the group with type 1 VWD and there is a continuous relationship between bleeding tendency and VWF level that extends into the normal range without a clear cut-off between normal and abnormal. A firm diagnosis of VWD can be made when VWF function is <0.3 IU/mL but those with levels 0.3–0.5 IU/mL are regarded as having 'low VWF'. This is a risk factor for bleeding, but will lead to a positive bleeding history only when collagen and platelet function are insufficient to compensate.

2 A history of minor bleeding episodes (e.g. easy bruising, epistaxes) is very common in the population and is not a good predictor of bleeding in other circumstances, such as operations. On the other hand, a negative bleeding history may be because the patient has not yet have been exposed to a significant test of haemostasis.

3 As a result of the two points above, a history of, say, easy bruising and slightly low VWF levels will often be found together. However, this does not necessarily mean that the patient has VWD and caution should be exercised in drawing this conclusion, as it has many consequences for the patient.

4 Intercurrent events such as stress, exercise, illness and pregnancy may all elevate the VWF level, making it difficult to be certain that a representative measure of VWF levels has been obtained. It is therefore necessary to perform carefully standardized sets of assays with concordant results on at least two occasions to be sure of the diagnosis and its severity.

5 The family history is often unhelpful, particularly in mild cases where penetrance is weak or variable. This is partly due to the modifying effects alluded to in (1) above and it is now evident that in some families, VWD does not segregate with the VWF gene. A number of other factors, including ABO blood group have been found to modulate VWF levels.

Treatment

Patients with mild or moderate VWD attend infrequently for treatment. The first-line treatment for minor bleeding after local measures have failed in type 1 VWD is desmopressin. This will produce a brisk rise in VWF and FVIII levels (30 min after intravenous infusion) and a shortening of the bleeding time (for details of therapy see discussion of mild haemophilia A). Desmopressin is much less effective in types 2A and 2M, presumably because the patient's released VWF is abnormal and still unable to promote platelet adhesion. Desmopressin is generally regarded as contraindicated in type 2B VWD, as the released abnormal VWF will cause circulatory platelet aggregates to form, with a further fall in the platelet count. Before relying on desmopressin for therapy, patients should be given a test dose in order to determine the response at 30–60 min, because it varies considerably. In addition, the duration of the response is also variable and further samples at 4–6 hours should be taken because some forms of type 1 VWD (e.g. VWF Vicenza) are characterized by a much shortened survival of VWF. A therapeutic trial may also be worthwhile in type 2A, as some families do respond. In type 2N, the FVIII response is of normal magnitude, but is of limited efficacy due to its short duration, emphasizing the importance of making the correct diagnosis. It may still be useful for minor procedures. Tranexamic acid may be effective on its own in some circumstances such as menorrhagia or as a mouthwash for oral cavity bleeding.

In patients in whom desmopressin is ineffective or contraindicated, the next line of treatment is a concentrate containing adequate amounts of functionally active VWF with preservation of the high-molecular-weight multimers. As with all concentrates, the source of plasma and viral inactivation are also important. Recombinant VWF concentrate is not yet licensed, but is in Phase III clinical trials. Cryoprecipitate has been used in these circumstances, but cannot be subjected to viral inactivation procedures and is therefore not recommended as first-line therapy. Factor concentrates will always be required for treatment of type 3 VWD, most type 2 and those type 1 variants with inadequate desmopressin responses. It is important to remember that following infusion of some high-purity VWF with low FVIII content, there is a delay of approximately 12 hours before the level of FVIII rises substantially and, if rapid correction is required, FVIII concentrate should also be given. In general, none of these replacement treatments is reliable in correcting the bleeding time, but this is not necessarily a bar to effective haemostasis. This seemingly paradoxical result may be because they do not correct the deficiency of intra-platelet VWF or because the largest high-molecular-weight multimers are not present. In situations when concentrates fail to stop bleeding, cryoprecipitate and platelet concentrates may prove effective.

Clinical course and complications

Patients with types 1 and 2 VWD lead relatively normal lives, with normal life expectancy. Menstruation is seldom a cause of severe blood loss, although menorrhagia is common. This can usually be managed satisfactorily with antifibrinolytics or by hormonal therapy. The Mirena coil is also effective. If these are not effective, then self-administration at home of desmopressin by intranasal or subcutaneous routes can be useful. During pregnancy, the VWF levels rise spontaneously to the normal or low-normal range in most patients with type 1 VWD. In type 2, the response is unpredictable and should be closely monitored. Type 2B is particularly complicated in pregnancy because the rise in endogenous VWF production exacerbates the thrombocytopenia. Patients with severe type 3 disease have a clinical course resembling severe or moderately severe haemophilia A, including the development of joint damage, and should be considered for prophylaxis. A small number of patients with type 3 disease develop antibodies to VWF, which inhibit its platelet adhesion-promoting property and cause rapid removal from the circulation of infused material. Unlike anti-FVIII antibodies, some anti-VWF antibodies may mediate anaphylactic shock.

Molecular genetics

The cloning of VWF cDNA and its gene has led to progress in identifying the underlying mutations responsible for the various phenotypes. A database of mutations responsible for VWD has been established at http://vwf.group.shef.ac.uk/.

Type 1

In the last few years there have been extensive studies to determine the molecular basis of type 1 VWD and a number of important points have emerged. Firstly, in about 35% of cases reported to have VWD, the disorder does not segregate with the VWF locus and therefore presumably results from a combination of other modifying genes, in particular the ABO locus. This phenomenon is most common in individuals diagnosed with VWD and VWF levels above 0.30 IU/mL and involves a disproportionate number of blood group O individuals. The likelihood of the disorder segregating with the VWF gene and of a causative mutation being identified increases significantly as the VWF level falls and a mutation is found in all cases where the VWF multimer pattern is abnormal. Among those with type 1 VWD and an identified mutation, mutations are widely scattered throughout the molecule.

Type 2

Mutations responsible for type 2 VWD affect a relevant functional site in the molecule and the corresponding portion of the gene. Mutations causing types 2M and 2B both affect the binding site for platelet GPIb on VWF and are found within the A1 domain of the molecule. They are inherited dominantly. As discussed above, type 2A may arise from increased susceptibility to proteolysis by ADAMTS-13 caused by mutations in the A2 domain. Otherwise, type 2A mutations are scattered throughout the gene, but particularly in the propeptide, D′–D3 and the cysteine knot. In general, type 2A is dominant, but occasional recessive forms occur. Mutations causing type 2N are found in the FVIII-binding site in the D′ domain, but because the binding capacity of VWF greatly exceeds the FVIII concentration the inheritance is recessive.

Type 3

Type 3 is characterized by the complete absence of VWF and so arises from inactivating or null mutations in both VWF alleles and is inherited in an autosomal recessive manner. The rare patients who develop antibodies to VWF nearly all have large deletions of their VWF gene.

Pseudo von Willebrand disease (platelet-type)

Several families have been described with a disorder closely resembling type 2B VWD, but in whom mixing experiments show the defect to be in their platelets rather than their plasma. Patients with pseudo VWD (PT-VWD) have moderately reduced levels of VWF:Ag and platelets, with an enhanced response of their platelet-rich plasma to low levels of ristocetin (0.5 mg/mL). The addition of normal cryoprecipitate to their washed platelets causes spontaneous aggregation, whereas

Table 38.5 Disorders associated with acquired von Willebrand syndrome and mechanism responsible.

Disorder	Mechanism	Treatment
Hypothyroidism	Reduced synthesis	Correct thyroid status
Aortic stenosis	Shear stress and increased cleavage	Replace aortic valve
Myeloproliferative	Adherence to platelets and increased cleavage	Restore platelet count to normal
IgG paraprotein	Immune	IVIG, chemotherapy to reduce paraproteins
IgM paraprotein	Immune	Chemotherapy (IVIG ineffective)
Wilms tumour	Adsorption	Tumour removal if possible

IVIG, intravenous immunoglobulin.

the reverse experiment is without effect (compare with type 2B). Missense mutations in platelet membrane GPIb, such that it spontaneously binds higher multimers of VWF, have been shown to be the underlying cause of this autosomal dominant mild bleeding syndrome. Treatment has not been extensively evaluated, but it should probably be with normal platelet concentrates, rather than desmopressin or cryoprecipitate. Thus this syndrome should be excluded by mixing tests as above or by genetic analysis before diagnosing type 2B VWD.

Acquired von Willebrand syndrome

In a number of circumstances, patients may acquire a deficiency of VWF function that is not inherited. The commonest association is with a paraprotein that binds to VWF and accelerates clearance, producing a type 2A or type 3 pattern. Severe aortic stenosis can cause sufficient shear stress on VWF to accelerate ADAMTS13 cleavage and cause an acquired von Willebrand syndrome that resolves immediately after valve replacement. A similar phenomenon occurs in myeloproliferative disease when thrombocytosis is present and excess platelet binding promotes cleavage. VWF synthesis is reduced in hypothyroidism (Table 38.5). Treatment of the underlying disorder will improve the syndrome. If urgent treatment is required, then replacement therapy may be necessary (taking into account accelerated clearance) and IgG paraproteins will usually respond to intravenous immunoglobulin.

Selected bibliography

Bowman M, Mundell G, Grabell J *et al.* (2008) Generation and validation of the Condensed MCMDM-1VWD Bleeding Questionnaire for von Willebrand disease. *Journal of Thrombosis and Haemostasis* **6**: 2062–6.

Collins PW, Chalmers E, Hart DP *et al.* (2013) Diagnosis and treatment of factor VIII and IX inhibitors in congenital haemophilia: (4th edition). UK Haemophilia Centre Doctors Organization. *British Journal of Haematology* **160**: 153–70.

Goodeve A, Eikenboom J, Castaman G *et al.* (2007) Phenotype and genotype of a cohort of families historically diagnosed with type 1 von Willebrand disease in the European study, Molecular and Clinical Markers for the Diagnosis and Management of Type 1 von Willebrand Disease (MCMDM-1VWD). *Blood* **109**: 112–21. [Erratum appears in *Blood* 2008 **111**(6): 3299–300].

James PD, Notley C, Hegadorn C *et al.* (2007) The mutational spectrum of type 1 von Willebrand disease: Results from a Canadian cohort study. *Blood* **109**: 145–54.

Laffan MA, Lester W, O'Donnell JS *et al.* (2014) The diagnosis and management of von Willebrand disease: a United Kingdom Haemophilia Centre Doctors Organization guideline approved by the British Committee for Standards in Haematology. *British Journal of Haematology* **167**: 453–65.

Manco-Johnson MJ, Abshire TC, Shapiro AD *et al.* (2007) Prophylaxis versus episodic treatment to prevent joint disease in boys with severe hemophilia. *New England Journal of Medicine* **357**: 535–44.

Mannucci PM, Schutgens RE, Santagostino E, Mauser-Bunschoten EP (2009) How I treat age-related morbidities in elderly persons with hemophilia. *Blood* **114**: 5256–63.

Nathwani AC, Reiss UM, Tuddenham EG, *et al.* (2014) Long-term safety and efficacy of factor IX gene therapy in hemophilia B. *New England Journal of Medicine*. Available at: http://www.ncbi.nlm.nih.gov/pubmed/25409372. Accessed on 3 September 2015.

Nichols WL, Hultin MB, James AH *et al.* (2008) von Willebrand disease (VWD): evidence-based diagnosis and management guidelines, the National Heart, Lung, and Blood Institute (NHLBI) Expert Panel report (USA). *Haemophilia* **14**: 171–232.

Richards M, Williams M, Chalmers E *et al.*; Paediatric Working Party of the United Kingdom Haemophilia Doctors, O. (2010) A United Kingdom Haemophilia Centre Doctors' Organization guideline approved by the British Committee for Standards in Haematology: guideline on the use of prophylactic factor VIII concentrate in children and adults with severe haemophilia A. *British Journal of Haematology* **149**: 498–507.

Tsui NB, Kadir RA, Chan KC *et al.* (2011) Noninvasive prenatal diagnosis of hemophilia by microfluidics digital PCR analysis of maternal plasma DNA. *Blood* **117**: 3684–91. [Erratum appears in *Blood* 2012 **119**(19): 4577].

CHAPTER 39

Rare inherited coagulation disorders

39

Flora Peyvandi and Marzia Menegatti

Angelo Bianchi Bonomi Hemophilia and Thrombosis Center, Fondazione IRCCS Ca' Granda Ospedale Maggiore Policlinico, Università degli Studi di Milano, Milan, Italy

Introduction

The rare bleeding disorders (RBDs), including the inherited deficiencies of coagulation factors fibrinogen, prothrombin, factor (F)V, combined FV and FVIII, FVII, FX, FXI, FXIII and of multiple deficiency of vitamin-K-dependent factors, are usually transmitted in an autosomal recessive manner. The incidence of RBDs in the general population is low, and homozygous or double heterozygous deficiencies vary from 1 in 500,000 for FVII deficiency to 1 in 2 million for prothrombin (FII) and FXIII deficiencies (Table 39.1). This incidence is also strongly influenced by ethnicity and is significantly increased by a high rate of consanguinity in the population.

Evaluation of the distribution of the RBDs relies on the two largest surveys, aimed at collecting epidemiological data and at providing information to haemophilia organizations and treatment centres involved in reducing or preventing complications of bleeding. One is led by the World Federation of Haemophilia (WFH, http://www.wfh.org/) and the other was developed within the European Network of the Rare Bleeding Disorders (EN-RBD; www.rbdd.eu). The WFH started the collection of data on RBDs in 2004, while the EN-RBD project was developed within the public health program promoted by the Public Health Executive Agency and funded by the European Commission and the Directorate General for Health and Consumers (DG SANCO) in 2007. Comparison of the data obtained by the two surveys shows similar results, confirming that FVII and FXI deficiencies are the most prevalent RBDs (representing about 38% and 27% of the total number of affected patients, respectively), followed by deficiencies of fibrinogen, FV and FX (8–9%), FXIII (~6%) and combined FV and FVIII (~3%), while the rarest disorder was FII deficiency with a prevalence of 1% (Figure 39.1). In addition, the WFH survey shows that half of the available data comes from Europe; this could mean that more effort should be put on setting up more accurate diagnosis and better data collection system in the rest of the world.

Clinical symptoms

Clinical symptoms among patients with RBDs differ significantly from one disorder to another, and from one patient to another, even when suffering from the same disorder. The most common symptoms are mucosal tract bleeding and excessive bleeding at the time of invasive procedures, delivery in women and circumcision in boys (where it is customary). Bleeds that endanger life, such as in the central nervous system (CNS), appear to be less frequent than in haemophilia. Particular care should be paid to women affected by RBDs as they may experience excessive monthly bleeding associated with menstruation. Indeed, menorrhagia is not the only gynaecological problem that women with RBDs are more likely to experience, they are also at risk of increased bleeding in conditions such as haemorrhagic ovarian cysts, endometriosis, endometrial hyperplasia, polyps and fibroids. Pregnancy and childbirth, two important stages in the life of a woman, pose particular clinical challenges in women with RBDs, since information about these issues are very scarce and limited to just a few case reports.

Classification

The rarity of the RBDs limits our knowledge of their natural history. It is now well known that every RBD can have several

Postgraduate Haematology, Seventh Edition. Edited by A Victor Hoffbrand, Douglas R Higgs, David M Keeling and Atul B Mehta.

Table 39.1 General features of autosomal recessive deficiency of coagulation factors.

Deficiency	Estimated prevalence*	Gene (chromosome)
Fibrinogen	1 in 1 million	FGA, *FGB*, *FGG* (all on 4q28)
Prothrombin	1 in 2 million	F2 (11p11–q12)
Factor V	1 in 1 million	F5 (1q24.2)
Combined factor V and VIII	1 in 1 million	LMAN1 (18q21.3–q22) MCFD2 (2p21–p16.3)
Factor VII	1 in 500,000	F7 (13q34)
Factor X	1 in 1 million	F10 (13q34)
Factor XI	1 in 1 million	F11 (4q35.2)
Factor XIII	1 in 2 million	F13A1 (6p24–p25) F13B (1q31–q32.1)
Vitamin-K-dependent coagulation factors	Reported in <50 families	GGCX (2p12) VKORC1 (16p11.2)

*Including dysfunctional proteins.

bleeding symptoms, ranging from minor post-trauma bleeding to the most severe episodes, which may appear at birth or later in life. In some deficiencies the residual coagulant activity level is directly proportional to the risk of developing a haemorrhage, but this is not true for all types of deficiencies. This information has been finally made available, in recent years, thanks to the published scientific literature and registries/databases aimed at collecting data on different aspects of single deficiency. Among these, the European Network on Rare Bleeding Disorders (EN-RBD), based on a cross-sectional study using data from 489 patients and involving 13 European treatment centres, for the first time evaluated the correlation between the coagulant residual plasma activity level and clinical bleeding severity in each RBD. Clinical bleeding episodes were classified into four categories of severity, based on the location and the potential clinical impact, as well as the trigger of bleeding (spontaneous, after trauma or drug induced). By means of linear regression analysis, this study found a strong association between coagulant activity level and clinical bleeding severity for fibrinogen, combined FV + VIII, FX and FXIII deficiencies. A weak association with clinical bleeding severity was present for FV and FVII deficiencies, while coagulation-factor activity level of FXI did not predict clinical bleeding severity. From the same study it is also clear that the minimum level to ensure complete absence of clinical symptoms is different for each disorder, leading to the conclusion that RBDs should not be considered as a single class of disorders, but instead studies should focus on the evaluation of specific aspects of each single RBD (see Table 39.2).

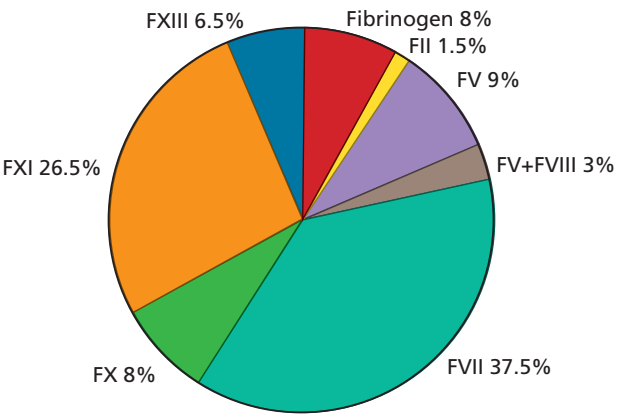


Figure 39.1 Worldwide distribution of rare bleeding disorders derived from World Federation of Haemophilia (WFH) and European Network of the Rare Bleeding Disorders (EN-RBD).

Laboratory diagnosis

The laboratory diagnosis of RBDs is currently carried out by means of coagulation screening tests, such as the activated partial thromboplastin time (APTT) and the prothrombin time (PT), applied to subjects reporting a clinical and family history of bleeding. A prolonged aPTT with a normal PT is suggestive of FXI deficiency (after exclusion of FVIII, FIX and FXII deficiencies). The reverse pattern (normal aPTT and prolonged PT) is typical of FVII deficiency, whereas the prolongation of both tests directs further analysis towards possible deficiencies of FX, FV or prothrombin. All coagulation tests that depend on the formation of fibrin as the end point are necessary to evaluate fibrinogen deficiency; hence, beside PT and aPTT, thrombin time (TT) has to be performed. Abnormal results of screening tests should be followed by mixing studies (50:50) to exclude the presence of a natural inhibitor. If there is a correction of the PT or PTT, this indicates a factor deficiency and specific coagulation assays should be performed in order to diagnose the specific factor involved; no correction indicates the presence of an inhibitor. Factor antigen assays are not strictly necessary for diagnosis and treatment, but are necessary to distinguish type I from type II deficiencies; these measurements are very important in fibrinogen or FII deficiency, where normal antigen levels and reduced coagulant activity (dysfibrinogenaemia and dysprothrombinaemia) can be associated with a risk of thrombosis. The standard laboratory clotting tests (PT, aPTT, fibrinogen level, platelet counts) are normal in FXIII deficiency; therefore its diagnosis is established by the demonstration of increased clot solubility in 5 M urea, dilute monochloroacetic acid or acetic acid. However, this method, qualitative and poorly standardized, detects only severe FXIII deficiency (with activity <5%), thus leading to a possible under-diagnosis. FXIII activity

Table 39.2 General recommendations for treatment and long-term prophylaxis in patients with recessive coagulation disorders.

	On-demand			Reported dose schedule for successful long-term prophylaxis			
Deficient factor (plasma half-life)	**Recommended trough levels**	**Product and dosages**	**Recommended trough levels (to maintain patient asymptomatic) after publication of the EN-RBD results**	**Products**	**Dose**	**Frequency**	**Notes**
Fibrinogen 2–4 d	0.5–1 g/L	Cryoprecipitate (5–10 bags) SD-treated plasma (15–30 mL/kg) Fibrinogen concentrate (50–100 mg/kg)	1 g/L	Cryoprecipitate Fibrinogen concentrate	1 unit 3 units 30–100 mg/kg	3/week every 7–10 d every week	Afibrinogenaemic patients with recurrent life-threatening bleedings or undergoing surgeries
Prothrombin 3–4 d	20–30%	SD-treated plasma (15–20 mL/kg) FIX concentrate and PCC (20–30 units/kg)	>10%	PCC	25–40 units/kg	1/week	–
Factor V 36 h	10–20%	SD-treated plasma (15–20 mL/kg)	10%	SD-treated plasma	30 mL/kg	2/week	–
Factor V + VIII FV 36 h FVIII 10–14 h	10–15%	As for FV	40%	No data			–
Factor VII 4–6 h	10–15%	FVII concentrate (30–40 mL/kg) PCC (20–30 units/kg) rFVIIa (15–30 µg/kg every 4–6 hours)	>20%	FFP pdFVII rFVIIa	10–15 mL/kg 10–50 units/kg 15–30 µg/kg	2/week 3/week 2–3/week	Prevention of bleeding during surgery or in children with recurrent haemarthrosis or CNS bleeding

(*Continued*)

Table 39.2 (*Continued*)

	On-demand			Reported dose schedule for successful long-term prophylaxis			
Deficient factor (plasma half-life)	**Recommended trough levels**	**Product and dosages**	**Recommended trough levels (to maintain patient asymptomatic) after publication of the EN-RBD results**	**Products**	**Dose**	**Frequency**	**Notes**
Factor X 40–60 h	10–20%	SD-treated plasma (10–20 mL/kg) PCC (20–30 units/kg) FX/FIX concentrate (10–20 units/kg)	>40%	PCC FX/FIX concentrate	30–40 units/kg 20–40 units/kg	2–3/week 1–2/week	Patients with recurrent life-threatening bleedings or undergoing surgeries
Factor XI 50 h	15–20%	SD-treated plasma (15–20 mL/kg) FXI concentrate (15–20 units/kg)	15–20%		No data		–
Factor XIII 9–12 d	2–5%	Cryoprecipitate (2–3 bags) SD-treated plasma (3 mL/kg) FXIII concentrate (50 units/kg for high haemorrhagic events)	30%	Cryoprecipitate SD-treated plasma FXIII concentrate	2 units 15–20 mL/kg 10 units/kg	every 3 weeks every 4–6 weeks every 4–6 weeks	Highly recommended in severe patients
Vit.K-dep.		Vitamin K (10 mg i.v. or s.c. for minor bleeding)	No data available	Vitamin K	10 mg	1/week orally	–

d, days; h, hours; PCC, prothrombin complex concentrate; pdFVII, plasma-derived FVII; rFVIIa, recombinant activated FVII; SD, solvent/detergent.

should also be determined quantitatively by the measurement of ammonia released during the transglutaminase reaction or the incorporation of radioactive amines into proteins. In this case, the plasma blanking procedure is mandatory to avoid FXIIIa-independent ammonia release that could lead to incorrect results in the low-activity range below 5–10%.

Molecular diagnosis

Molecular diagnosis is based on the search for causative mutations in the genes encoding the corresponding coagulation factors. Exceptions are combined deficiency of FV and FVIII, caused by mutations in genes encoding proteins involved in FV and FVIII intracellular transport (*MCFD2* and *LMAN1*), and combined deficiency of vitamin-K-dependent proteins (FII, FVII, FIX and FX), caused by mutations in genes that encode enzymes involved in post-translational modification and in vitamin K metabolism (*GGCX* and *VKOR*). Information on already identified mutations causing RBDs are easily traceable from the mutation database on the ISTH website (http://www.isth.org/?MutationsRareBleedin). Missense mutations are the most frequent gene abnormalities, representing 50–80% of all identified mutations, except for LMAN1 variants, where the most frequent mutations are insertions/deletions (50%). Insertion/deletion mutations represent 20–30% of the gene variations of the fibrinogen, FV, MCFD2 and FXIII genes and less than 15% of the remaining coagulation-factor gene mutations. Splicing and nonsense mutations comprise 5–15% of all identified mutations in all coagulation factors, with a maximum rate of 20% in the LMAN1 gene. Variants located in the 3′ and 5′ untranslated regions of the genes are the least frequent types of mutation (less than 5%) found only at the fibrinogen, FVII, FXI and FXIII loci. Despite significant advances in our knowledge of the genetic basis of the RBDs, in 5–10% of patients affected with severe clotting-factor deficiencies, no genetic defect can be found. In these patients, the use of next generation sequencing (NGS) might help to identify novel pathways in coagulation disorders. Over the past 20 years the literature on molecular aspects of RBDs based on naturally occurring mutations is much increased. However, the early work in this area, and especially studies that have highlighted potential genotype/phenotype correlations have been conducted in mouse models. The results of these studies showed that the complete absence of almost all coagulation factors led to a severe clinical phenotype often incompatible with life or with reaching adulthood.

Global haemostasis tests

Tests investigating global haemostatic capacity, such as the thrombin generation test and thromboelastography have recently raised the interest of the scientific community, since standard coagulation tests, such as PT or APTT, note only the integrity of the coagulation process, and not its speed and extent. Moreover, factor assays performed using these tests are limited by their sensitivity at very low levels. Therefore global haemostasis tests are believed to provide more accurate evaluation of an individual's *in vivo* haemostatic state and response to treatment and to be better suited to predicting clinical phenotype because they can more effectively assess the rate and total thrombin generated, whole blood clot formation and/or fibrin polymerization. They have been recently used to evaluate haemostasis in patients with RBDs and haemophilia and, in particular, they are now becoming a new laboratory strategy to determine the effectiveness of therapies and to monitor treatment in various bleeding situations in RBDs, in particular in those deficiencies, such as FXI, where standard assays fail to determine the bleeding risk, even though the analytical conditions of the reaction have not yet been optimized or standardized. The ability to reproduce reliable thrombin generation measurements should be facilitated by the use of standardized preanalytical and analytical procedures.

Treatment

Treatment of RBDs is a difficult task, since information on their clinical management is often scarce, and replacement therapy of coagulation factors may require the prescription of unlicensed products that are not readily available. The main treatments are represented by non-transfusional adjuvant therapies (antifibrinolytic amino acids, oestrogen/progestogen) and replacement therapy of the deficient coagulation factor. Unfortunately, information on the safety and efficacy of the few available products is sparse, and experience in their optimal use significantly more limited than that available for haemophilia. The avoidance of transmission of blood-borne infectious agents is the primary requisite in the choice of replacement material. Solvent/detergent-treated plasma is an important source of replacement that is recommended in the majority of these disorders; virus-inactivated concentrates, when commercially available, are also safe, but expensive, especially for developing countries. Non-virus-inactivated plasma and cryoprecipitate should be avoided if possible. Virally inactivated factor concentrates are available for several deficiencies and are preferred when virally inactivated plasma is not available or repeated infusions are needed, causing fluid overload, as may occur during surgery or in cases of bleeding in the CNS. An updated registry of the available clotting factor concentrates is published by the WFH (9th edition, http://www.wfh.org). A patient's personal and family history of bleeding are important guides for management. The conventional treatment for RBDs is given on an 'on demand' basis and is administered as soon as possible after onset of bleeding; however, in clinically severe cases, a

prophylactic treatment should be considered. Dosages and frequency of treatment depend on the minimal haemostatic level of the deficient factor, its plasma half-life and the type of bleeding episode. The choice of prophylaxis is related to the frequency of bleeding, the risk of severe spontaneous bleeding and the risk of long-term disabilities associated with the occurrence of bleeding, despite on-demand treatment. General recommendations for treatment on demand and prophylaxis are summarized in Table 39.2.

Fibrinogen deficiency

Fibrinogen deficiency is heterogeneous and two main phenotypes can be distinguished: plasma and platelet levels of the protein are either not measurable or very low, leading to afibrinogenaemia and hypofibrinogenaemia. In other cases, low clottable fibrinogen contrasts with normal or moderately reduced fibrinogen antigen leading to dysfibrinogenaemia and hypodysfibrinogenaemia. Fibrinogen is produced in the hepatocyte from three homologous polypeptide chains, Aα, Bβ and γ, which assemble to form a 340-kDa hexamer. The three genes encoding fibrinogen Bβ (*FGB*), Aα (*FGA*) and γ (*FGG*), ordered from centromere to telomere, are clustered in a region of approximately 50 kb on chromosome 4 in humans.

Dysfibrinogenaemic and hypofibrinogenaemic patients are usually asymptomatic, while those with afibrinogenaemia have a bleeding tendency that usually manifests in the neonatal period, with 85% of cases presenting umbilical cord bleeding. Bleeding may also occur in the skin, gastrointestinal tract (GI), genitourinary tract or the CNS, while persistent damage to the musculoskeletal system and resulting handicap is less frequent. A milder symptom such as epistaxis is also frequent. Women may suffer from menometrorrhagia, but some have normal menses. First-trimester abortion is common in women with afibrinogenaemia, but less common in those with dysfibrinogenaemia. The conventional treatment is episodic and on-demand, but effective long-term secondary prophylaxis with administration of fibrinogen every 7–14 days has been described after CNS bleeds. However, high-level replacement, e.g. during pregnancy, should be moderated by the fact that thromboembolic events can occur. Depending on the country of residence, patients receive fresh-frozen plasma (FFP), cryoprecipitate or fibrinogen concentrate, the last being the treatment of choice because it is virally inactivated and thus safer than cryoprecipitate or FFP.

Prothrombin deficiency

Prothrombin deficiency is perhaps the rarest inherited coagulation disorder, with a prevalence of about 1 in 2 million. As with fibrinogen deficiency, two main phenotypes can be distinguished: hypoprothrombinaemia (both activity and antigen levels are low) and dysprothrombinaemia (normal synthesis of a dysfunctional protein); hypoprothrombinaemia associated with dysprothrombinaemia was also described in compound heterozygosis. FII, a vitamin-K-dependent glycoprotein synthesized by the liver, is the zymogen of the serine protease α-thrombin and is encoded by a gene of approximately 21 kb located on chromosome 11.

Severe deficiency of FII, a condition in which plasma levels are <5% (homozygous or double heterozygous patients), is always characterized by severe bleeding, including prolonged post-injury bleeding, mucosal bleeding and subcutaneous and muscle haematoma. Conversely, GI bleeding episodes were reported in only a few cases. Dysprothrombinaemia manifests as a variable bleeding tendency that is usually less severe than true deficiency, while heterozygous subjects are usually asymptomatic although, occasionally, excessive bleeding after surgical procedures can be observed. In homozygous women, menorrhagia is frequent.

Replacement therapy is needed only in severe patients, in case of bleeding or to ensure adequate prophylaxis before surgical procedures. Since no FII concentrate exists, FFP, prothrombin complex concentrate (PCC) or both are used for treating patients. In severe clinical settings, higher levels of prothrombin may be achieved with PCCs without the risk of potential volume overload induced by FFP. However, most PCCs contain other vitamin-K-dependent coagulation factors, which could potentially induce thrombotic complications.

Factor V deficiency

FV has a double role in the coagulation process: it is a protein cofactor required by the prothrombinase complex for thrombin generation, but also contributes to the anticoagulant pathway by downregulating FVIII activity. The majority of individuals with FV deficiency are characterized by concomitant deficiency of FV activity and antigen levels (type I deficiency), but approximately 25% have normal antigen levels, indicating the presence of a dysfunctional protein (type II deficiency). FV is mainly secreted by hepatocytes, but there is some evidence that it can also be synthesized *in vivo* in the megakaryocyte/platelet lineage. FV protein is encoded by a large (80 kb) and complex (25 exons) gene located on chromosome 1. Severe deficiency typically presents early in life, nonetheless, it is clinically heterogeneous and, sometimes, despite severe deficiency, patients may not bleed as much as expected. Observations made in the past few years point to an important role for platelet FV that provide new insight into this discrepancy. Megakaryocytes can synthesize FV; however, the majority of platelet FV is endocytosed from the plasma pool by megakaryocytes. Following endocytosis, FV is modified intracellularly; these changes to platelet FV appear to provide the cofactor with unique physical and functional characteristics, which render it more procoagulant compared with its plasma counterpart. Platelet degranulation and

release of platelet FV at the site of vascular injury is thought to be a critical contributor to the local FV concentration. Furthermore, there is some evidence that, because platelet FV is locally released in high concentrations, it is less susceptible to inhibition and may function normally in haemostasis.

Symptomatic patients usually present with umbilical stump bleeding, skin and mucosal tract haemorrhages; epistaxis and menorrhagia are relatively frequent, even in patients with measurable FV levels. Postoperative and oral cavity haemorrhages are common, but not fully predictable, as these symptoms also occur with plasma levels as high as 5–10%. Haemarthroses and haematomas occur in only 25% of patients, and life-threatening bleeding episodes (GI and CNS bleedings) are rare. In patients with mild to moderate deficiency, therapy with antifibrinolytic drugs is sufficient to control epistaxis, menorrhagia or other non-life-threatening mucosal bleedings. Menorrhagia can also be managed directly using oral contraceptives, progestin-containing intrauterine devices, endometrial ablation or hysterectomy. Replacement therapy with FV can still be administered only through FFP, preferably virus inactivated, because no FV concentrate is currently available on the market. However, a FV concentrate has been recently developed for clinical use in deficient patients and preclinical studies are currently being performed for the orphan drug designation application to the European Medicine Agency (EMA) and the Food and Drug Administration (FDA), in order to make it available on the market as soon as possible.

Combined deficiency of factor V and factor VIII

Combined FV and FVIII deficiency (F5F8D) is characterized by concomitantly low levels (usually 5–20%) of the two coagulation factors, both as coagulant activity and as antigen. It appears that the concomitant presence of two coagulation defects does not enhance the haemorrhagic tendency observed in each defect separately. In 1988, F5F8D was causally associated with mutations in the *LMAN1* gene, encoding lectin mannose-binding protein (previously named *ERGIC-53*), a 53-kDa type 1 transmembrane protein that acts as a chaperone in the intracellular transport of both factors. In 2003 another locus associated with the deficiency was identified in about 15% of affected families, with no mutations in *LMAN1*. This was named the *MCFD2* gene, encoding multiple coagulation factor deficiency (MCFD)2 protein, which acts as a cofactor for LMAN1, specifically recruiting correctly folded FV and FVIII in the endoplasmatic reticulum. Recent studies have failed to identify additional components of the LMAN1–MCFD2 receptor complex, supporting the idea that F5F8D might be limited to *LMAN1* and *MCFD2*. The phenotypes associated with mutations in *MCFD2* and *LMAN1* are indistinguishable and manifested only by deficiencies of FV and FVIII, although a selective delay in secretion of the cargo protein procathepsin C has been observed in HeLa cells overexpressing a dominant-negative form of LMAN1.

Symptoms are usually mild, with a predominance of easy bruising, epistaxis and bleeding after dental extractions. Menorrhagia and postpartum bleeding have also been reported in affected women. More severe symptoms, such as haemarthroses and umbilical cord bleeding, are observed very seldom and GI and CNS bleeding have been reported in only a few patients. Soft-tissue haematomas are unusual. Because of the mild clinical pattern, bleeding episodes are usually treated on demand and do not require regular prophylaxis. These cases can be treated with desmopressin, as the low FVIII is usually of more significance than the low FV. More severe cases may need to be treated with FVIII concentrate and/or FFP as sources of both FV and FVIII may be needed when their different plasma half-lives (FV 36 hours, FVIII 10–14 hours) have to be taken into account.

Factor VII deficiency

FVII deficiency is the most common autosomal recessive coagulation disorder (1 in 500,000). A typical feature of this disease is its clinical heterogeneity, which ranges in severity from lethal to mild, or even asymptomatic forms. FVII is synthesized in the liver and is encoded by the FVII gene (*F7*) located on chromosome 13, 2.8 kb upstream of the FX gene. Plasma levels of FVII coagulant activity (FVII:C) and FVII antigen (FVII:Ag) are influenced by a number of genetic and environmental factors (sex, age, cholesterol and triglyceride levels) and it is also well known that FVII levels are modulated by *F7* polymorphisms. The majority of patients have concomitantly low levels of FVII:C and FVII:Ag, but several cases are characterized by normal or low borderline levels of FVII:Ag, contrasting with lower levels of FVII:C.

The severity of symptoms of FVII deficiency is variable: some patients do not bleed even with a very low level of FVII:C, while others with similar levels of coagulant activity have frequent bleeding. Life- or limb-endangering bleeding manifestations are relatively rare, the most frequent symptoms being epistaxis and menorrhagia. However, CNS bleeding was also reported to have high incidence (16%) in a series of 75 patients, the study concluded that the greatest risk factor for development of bleeding was trauma related to the birth process. Thrombotic episodes have also been reported in 3–4% of patients with FVII deficiency, particularly in the presence of surgery and replacement treatment, but spontaneous thrombosis may also occur; therefore it can be inferred that patients affected with FVII deficiency are not protected from thrombosis.

A number of therapeutic options can be offered to patients with FVII deficiency, including FFP and PCCs (however, concentration's of FVII in FFP and PCC are low and therefore these are not the best treatments), plasma-derived FVII concentrates and recombinant FVIIa, which has to be considered the

optimal replacement therapy because it can be used at very low dose (10–20 μg/kg). Prophylaxis has been a debated issue in FVII deficiency, particularly in those patients with a severe bleeding history. However, the short half-life of infused FVII makes it difficult to establish a standardized protocol for prophylaxis.

The occurrence of frequent menorrhagic symptoms is almost invariably associated with chronic anaemia associated with iron deficiency in women with the severe form of FVII deficiency. Pregnancy itself does not require special precautions and uncomplicated delivery is possible without prophylaxis. However, all cases of postpartum bleeding reported so far occurred in women with low FVII coagulant activity (<15 units/dL) not receiving prophylaxis. Therefore, delivery should occur under the coverage of a short-term replacement.

Factor X deficiency

FX is a glycoprotein that plays a pivotal role in the coagulation cascade, being the first enzyme in the common pathway of thrombin formation. FX is mainly synthesized by the liver and is encoded by the *FX* gene (*F10*), comprising 22 kb and located on chromosome 13, a few kilobases downstream of *F7*. The clinical phenotypes of FX deficiency are characterized by concomitantly low levels of coagulant activity and antigen, or low coagulant activity contrasting with normal or low borderline antigen values.

In FX deficiency the bleeding tendency may appear at any age, although the more severely affected patients (FX coagulant activity <1%) present early in life with umbilical-stump, CNS or GI bleeding. Patients with severe deficiencies also commonly experience haemarthroses and haematomas. The most common bleeding symptoms reported at all levels of severity are epistaxis and menorrhagia.

Data from the UK Haemophilia Centre Doctors' Organisation (UKHCDO) registry shows that the proportion of patients with this deficiency who require treatment is much higher than that of other rare coagulation deficiencies. Heterozygous patients have also been reported to have bleeding after delivery that required treatment. A recently developed freeze-dried human coagulation FIX and FX concentrate with specified amounts of FX (and FIX) has facilitated prophylaxis in patients with FX deficiency. A clinical trial investigating the pharmacokinetics of a new high-purity FX concentrate has been also recently completed (ClinicalTrials.gov identifier: 00930176) and is ongoing in children in order to be presented to the EMA. Therapeutic options for the control of menorrhagia are both medical (e.g. antifibrinolytics, oral contraceptives, levonorgestrel intrauterine device and clotting factor replacement) and surgical (e.g. endometrial ablation and hysterectomy). Even if FX plasma concentrations increases in pregnancy, women with the severe form of the deficiency and a history of adverse pregnancy outcomes (such as abortion, placental abruption or premature birth), may benefit from continuous replacement therapy.

Factor XI deficiency

FXI deficiency is characterized by a decrease in the functional activity of this plasma protein, usually accompanied by correspondingly low FXI antigen levels. More rarely, normal or borderline levels of plasma FXI antigen are associated with a dysfunctional form of the protein. The estimated prevalence of the severe FXI deficiency in most populations is about 1 in 1 million, but is reported to be much higher in Ashkenazi Jews, heterozygosity being as high as 8%.

FXI is mainly synthesized in the liver, although tiny amounts of transcript can also be detected in megakaryocytes and platelets. The protein is encoded by the *FXI* gene, comprising 23 kb and located on chromosome 4. At variance with other RBDs, whose pattern of inheritance is autosomal recessive, in some cases of FXI, missense mutations were shown to exert a dominant negative effect through heterodimer formation between the mutant and wild-type polypeptides, resulting in a pattern of dominant transmission. Moreover, the existence of a platelet FXI transcript, originating from the skipping of exon 5, was hypothesized for a long time, but never confirmed by subsequent work. Nonetheless, recent data support the view that FXI transcripts undergo alternative splicing, leading to the synthesis of FXI isoforms whose physiological role and importance still need to be demonstrated.

The relationship between FXI levels in plasma and the bleeding tendency is not as clear-cut as for other coagulation factor deficiencies as also shown by the EN-RBD results. The phenotype of bleeding is not correlated with the genotype, but with the site of injury. When a site of injury with local high fibrinolytic activity is involved (e.g. urogenital tract, oral cavity after dental extraction or tonsillectomy), the risk of bleeding is increased (49–67%) in comparison to sites with less local fibrinolytic activity (1.5–40%). Usually, patients with severe FXI deficiency (1% or less) are mildly affected and most bleeding manifestations are injury-related. Surprisingly, patients with low, but detectable, levels of FXI (1–5%) are also mild bleeders, so that clinical phenotypes are not strikingly different in these two groups. The most frequent symptom is oral and postoperative bleeding, which occurs in more than 50% of patients. Women with FXI deficiency are prone to excessive bleeding during menstruation, but case series of women affected by severe FXI deficiency showed that 70% of pregnancies were uneventful, despite no prophylactic treatment.

Treatment is based on the use of antifibrinolytic agents, FFP and FXI concentrate. Recombinant FVIIa was successfully used in surgeries, but care should be taken to reduce the risk of complications such as thrombotic events (especially with FXI concentrate), volume overload, allergic reactions and

development of inhibitors. Future studies are needed to confirm the role of rFVIIa in surgery and FXI deficiency. Patients with a partial deficiency in FXI who have no bleeding history do not require prophylactic treatment.

Factor XIII deficiency

FXIII is a transglutaminase and functions to cross-link the α and γ fibrin chains, resulting in a stronger clot with an increased resistance to fibrinolysis. The plasma factor consists of two catalytic A subunits (FXIII-A) and two carrier B subunits (FXIII-B). FXIII-A is synthesized in cells of bone marrow origin, while FXIII-B is produced in the liver. The corresponding genes are located on chromosomes 6 and 1. FXIII deficiency is, together with prothrombin deficiency, the rarest of the recessively transmitted coagulation factor deficiencies (1 in 2 million). In inherited FXIII deficiency, plasma levels of FXIII-A measured as functional activity or immunoreactive protein is usually extremely reduced, whereas the FXIII-B subunit is reduced, but always at measurable levels. Patients with FXIII-A deficiency have a bleeding tendency that is usually severe, particularly because of the early onset of life-threatening symptoms such as umbilical-cord and CNS bleeding occurring in up to 80 and 30% of patients, respectively. Ecchymoses, haematomas and prolonged bleeding following trauma are also typical. Haemarthroses and intramuscular haematomas may appear unexpectedly, although less frequently. In women of reproductive age, miscarriage and intraperitoneal bleeding are often reported. All these symptoms usually lead to an early diagnosis, so that patients who survive are often treated prophylactically, starting early in life.

This approach to treatment is rendered simple and feasible by the fact that plasma levels of FXIII of 2–5% are sufficient to prevent severe bleeding, and that the long *in vivo* half-life of the factor (11–14 days) makes it possible to infuse cryoprecipitate or concentrates at intervals of 1 month or longer. The recent EN-RBD study showed that patients with FXIII deficiency need to reach a FXIII:C >30% to remain completely asymptomatic. Prophylactic infusions of FXIII are recommended in FXIII-A-deficient women who are pregnant, in order to prevent fetal loss.

Because of the risk of blood-borne diseases, FFP and cryoprecipitate are less satisfactory for treatment or prophylaxis, and FXIII concentrate is recommended whenever available. Plasma-derived FXIII has been used for several years and shown to be safe and effective; however, a new recombinant FXIII-A_2 concentrate (rFXIII-A_2) has become available and a Phase III clinical trial (ClinicalTrials.gov identifier:00713648) has recently been completed, proving that rFXIII is safe and effective in preventing bleeding episodes in patients with congenital FXIII-A subunit deficiency. The rFXIII was recently approved for the treatment of FXIII-A deficiency in Australia, Canada, the European Union, Switzerland and the USA.

Only a few cases of inherited FXIII-B deficiency have been reported to date, and only 16 causative mutations have been identified; FXIII-B deficiency bleeding symptoms appear milder than in FXIII-A-deficient patients.

Vitamin-K-dependent coagulation factors deficiency

Vitamin-K-dependent coagulation factors such as FII, FVII, FIX and FX require γ-carboxylation of glutamic acid residues at their Gla domains to enable binding of calcium and attachment to phospholipid membranes. The γ-carboxylation is catalysed by hepatic γ-glutamyl carboxylase (GGCX) and its cofactor, reduced vitamin K (KH_2). During the reaction, KH_2 is converted to vitamin K epoxide (KO), which is recycled to KH_2 by the vitamin K epoxide reductase (VKOR) enzyme complex. Heritable dysfunction of GGCX or the VKOR complex results in the secretion of poorly carboxylated vitamin-K-dependent coagulation factors, leading to combined deficiency of the vitamin-K-dependent clotting factors. The γ-carboxylation of glutamic acid residues is also required for activity of the anticoagulant factors protein C, protein S and protein Z, and although protein S and protein C levels are low in VKCFD, there are no reports of venous or arterial thrombosis. Thus, the effect of VKCFD is clearly in the bleeding area. The GGCX and VKOR proteins are encoded by two corresponding genes: *GGCX* (13 kb, 15 exons) located on chromosome 2 and the unusually small *VKORC1* (5126 bp, three exons) located on chromosome 16; the latter gene was so named because of evidence suggesting that VKOR is a multisubunit complex.

VKCFD is a rare autosomal recessive bleeding disorder that often presents during infancy, although the routine administration of vitamin K may delay the diagnosis of VKCFD in neonates. Severe symptoms, such as intracranial haemorrhage or umbilical stump bleeding, are usually associated with factor levels below 5 units/dL. Severely affected children may also present with skeletal abnormalities such as nasal hypoplasia, distal digital hypoplasia, epiphyseal stippling and mild conductive hearing loss. Older patients can present with easy bruising and mucocutaneous or postsurgical bleeding.

Treatment with oral or parenteral vitamin K should be started as soon as possible in all patients at diagnosis. However, some patients show insufficient response, and there are limited data on the effectiveness of prophylaxis with weekly administration of vitamin K_1 10 mg. In fact, massive parenteral doses of vitamin K do not always correct FII, FVII, FIX and FX activities, and there is clear biochemical evidence that the molecules are not fully carboxylated by such treatment. In this group of patients, factor replacement by virally inactivated FFP could also be used in cases of acute bleeding episodes or before surgery. PCC is the alternative choice when a rapid increase in clotting factor levels is necessary.

Concluding remarks

Rarity of RBDs limits the in-depth analysis of their clinical and laboratory features and limit their knowledge, in particular, among general physicians increasing the risk of misdiagnosis and fatal consequences. Therefore, the establishment of global networks of treatment centres is mandatory to improve the amount of collected data, to strengthen the power of the analysis deriving from such data and to improve and standardize laboratory assay to define the minimum residual level of coagulation factor to ensure a normal haemostatic process. The resulting definition of the clinical end points will facilitate the design of appropriate clinical trials for the evaluation of the safety and efficacy of novel drugs. Multinational multicentre data collection for long-term postregistration surveillance remains an important milestone in the study of RBDs. To this aim, two important programs have already been set up by Universal Data Collection (UDC) and the European Haemophilia Safety Surveillance system (EUHASS) to adequately register and monitor treatment and complications in treated patients. In addition, a web-based database, PRO-RBDD (http://eu.rbdd.org/), was recently set up to prospectively collect clinical and laboratory data of patients with fibrinogen and FXIII deficiencies in order to evaluate bleeding frequency and management, as well as consumption of treatment products and related complications.

Acknowledgements

We would like to thank all our colleagues who have helped us to develop issue 35 of *Seminars of Thrombosis and Hemostasis*, 2009 dedicated to 'Rare bleeding disorders' and all who updated it in the 2013 issue; their revision of specific types of deficiency helped us to more accurately revise our present chapter.

Selected bibliography

Bolton-Maggs PH, Perry DJ, Chalmers EA *et al*. (2004) The rare coagulation disorders: review with guidelines for management from the United Kingdom Haemophilia Centre Doctors' Organization. *Haemophilia* **10**: 593–628.

Chitlur M (2012) Challenges in the laboratory analyses of bleeding disorders. *Thrombosis Research* **130**: 1–6.

Kadir RA, Davies J (2013) Hemostatic disorders in women. *Journal of Thrombosis and Haemostasis* **11**(Suppl. 1): 170–9.

Lee CA, Chi C, Pavord SR *et al*. (2006) The obstetric and gynaecological management of women with inherited bleeding disorders: review with guidelines produced by a taskforce of UK Haemophilia Centre Doctors Organization. *Haemophilia* **12**: 301–36.

Muszbek L, Bagoly Z, Cairo A, Peyvandi F (2011) Novel aspects of factor XIII deficiency. *Current Opinion in Hematology* **18**: 366–72.

Palla R, Peyvandi F, Shapiro AD (2015) Rare bleeding disorders: diagnosis and treatment. *Blood* **125**: 2052–61.

Peyvandi F, Kunicki T, Lillicrap D (2013) Genetic sequence analysis of inherited bleeding diseases. *Blood* **14**: 3423–31.

Peyvandi F, Palla R, Menegatti M *et al*. (2012) Coagulation factor activity and clinical bleeding severity in rare bleeding disorders: results from the European Network of Rare Bleeding Disorders. *Journal of Thrombosis and Haemostasis* **10**: 615–21.

Ruiz-Saez A (2013) Occurrence of thrombosis in rare bleeding disorders. *Seminars on Thrombosis and Hemostasis* **39**: 684–92.

CHAPTER 40

Acquired coagulation disorders

40

Peter W Collins[1], Jecko Thachil[2] and Cheng-Hock Toh[3]

[1]School of Medicine, Cardiff University, University Hospital of Wales, Cardiff, UK
[2]Department of Haematology, Manchester Royal Infirmary, Manchester, UK
[3]Roald Dahl Centre, Royal Liverpool University Hospital, Liverpool, UK

Introduction

Acquired disorders of haemostasis are a heterogeneous group of conditions with varied and often complex aetiologies. Patients may have multiple and overlapping causes for bleeding and distinct conditions often have similar pathophysiologies. In order to manage acquired haemostatic failure it is important to understand the mechanisms by which haemostatic disturbance occurs and how these apply in different ways to a variety of conditions.

Systemic disease may present to haematologists with symptoms of bleeding or bruising or abnormalities of coagulation tests. Alternatively, patients with known disorders may need haematological input to manage symptoms or at the time of invasive procedures. Assessment of patients with symptoms of bleeding or bruising requires clinical review and laboratory investigation. Serial laboratory testing is often required because the haemostatic disturbance may evolve rapidly. The possibility of a congenital disorder should be considered.

Haematologists are also often involved in the management of acutely bleeding patients who may have complex medical conditions and in whom bleeding is usually multifactorial. Critical to optimizing the management of these patients is a well-organized multidisciplinary team approach. Hospitals should have systems in place to respond rapidly to massive haemorrhage and hospital transfusion committees should establish protocols for these medical emergencies. Laboratories should be geared to producing accelerated full blood count (FBC) and coagulation results. The rapid supply of blood products and efficient transfer of these to the patient is vital to success and 'Fire drills' to test responses to emergencies are good practice.

Tests of coagulation and point-of-care testing

Routine tests of haemostasis

To investigate a patient suspected of an acquired haemostatic defect, a platelet count is required and examination of the blood film is often useful. Routine coagulation tests include the prothrombin time (PT), activated partial thromboplastin time (APTT), thrombin time (TT) and Clauss fibrinogen level. Derived fibrinogen assays should not be used because results are likely to be misleading.

If screening tests are abnormal, correction studies help to distinguish factor deficiencies from inhibitors. Comparison of TT with the reptilase time will establish whether abnormal results are due to heparin. Measurement of fibrin breakdown products, such as D-dimers, is required for investigation of possible disseminated intravascular coagulation (DIC). Further measurement of individual coagulation factor and von Willebrand factor (VWF) levels may be required. Assessment of platelet function may be useful in some circumstances, through platelet aggregation studies, platelet nucleotides or a bleeding time, although these results may be difficult to interpret in the context of an acquired haemostatic defect. The role of global platelet function analysers remains to be defined, but these are available in many hospitals.

Thromboelastometry

Tests based on thromboelastometry are point-of-care tests that assess the viscoelastic properties of whole blood under low-shear

Postgraduate Haematology, Seventh Edition. Edited by A Victor Hoffbrand, Douglas R Higgs, David M Keeling and Atul B Mehta.

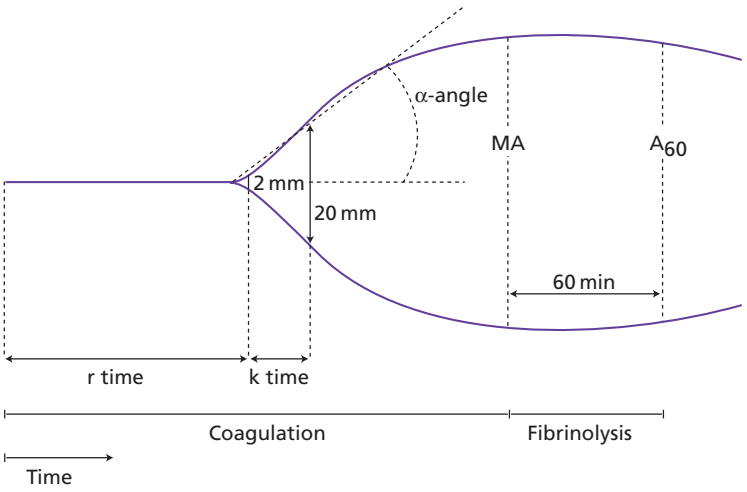


Figure 40.1 Representative thromboelastographic trace.

conditions (Figure 40.1). Their main advantage is the assessment of the different components of clot formation, including platelets and coagulation factors, and rapid availability of results. They are used to guide blood product administration in patients undergoing liver transplantation and cardiopulmonary bypass. Other reported uses include the assessment of coagulation in liver disease, neonates, obstetrics and trauma, although more studies are required to validate their role. The two main instruments are TEG and ROTEM, which measure the viscoelastic properties of an evolving clot in similar, but distinct ways. Results are displayed graphically and show the developing and dissolving clot. Different activators are available and heparinase cups can be used to investigate heparinized samples. Tests that correlate with fibrinogen levels are available. If thromboelastometry is used, hospitals should have agreed algorithms for interpretation and infusion of blood products.

Thrombin generation assays

These tests measure the amount of thrombin generated over time and may reflect an individual's coagulation potential. They are not available in routine practice. Assays are initiated by the addition of a trigger, usually tissue factor (TF), to recalcified plasma in the presence of phospholipids. Thrombin is detected through cleavage of a chromogenic or fluorogenic substrate. When low concentrations of TF are used, contact activation must be inhibited by corn trypsin inhibitor. Several variations of these assays have been developed by varying the concentrations of TF, corn trypsin inhibitor, phospholipid and protein-C-sensitizing reagents such as soluble thrombomodulin. Thrombin generation tests have been studied in the context of haemophilia, including monitoring of inhibitor bypass agent therapy. Other studies have examined thrombophilia and liver disease and monitoring of anticoagulant therapy and warfarin reversal. The clinical utility of many of these potential applications remains to be demonstrated.

APTT biphasic waveform

The multichannel discrete analyser generates an optical profile charting changes in light transmittance during clot formation on the routine APTT. In contrast to the sigmoidal appearance of a normal APTT waveform, a 'biphasic' appearance correlates with acquired haemostatic dysfunction (Figure 40.2). The reliability of this method has been validated in several reports and correlates with the diagnosis of overt DIC according to the International Society for Thrombosis and Haemostasis (ISTH) criteria. This biphasic response is due to calcium-dependent complex formation between C-reactive protein and very low-density lipoprotein, which has been shown to increase thrombin generation.

Activated clotting time

This test is used to monitor anticoagulation with heparin during cardiopulmonary bypass (CPB) surgery. It uses whole blood and has a linear response at the high concentrations of heparin used during CPB. The APTT cannot be used because plasma is unclottable using this method at these heparin concentrations. Measured using a specialized analyser, the activated clotting time (ACT) reference range varies according to the method and usually falls within 70–180 s. During heparinization for CPB, the goal is to exceed 400–500 s. Off-pump cardiac surgery has been described using less anticoagulation and lower ACT reference ranges of 200–300 s.

Disseminated intravascular coagulation

DIC is characterized by the loss of localization or compensated control of intravascular activation of coagulation. Arising from diverse causes (Table 40.1), its pathology can manifest systemically and contribute to a worse prognosis for the patient. Uncoupling of the highly regulated balance between procoagulant, anticoagulant, profibrinolytic and antifibrinolytic processes can

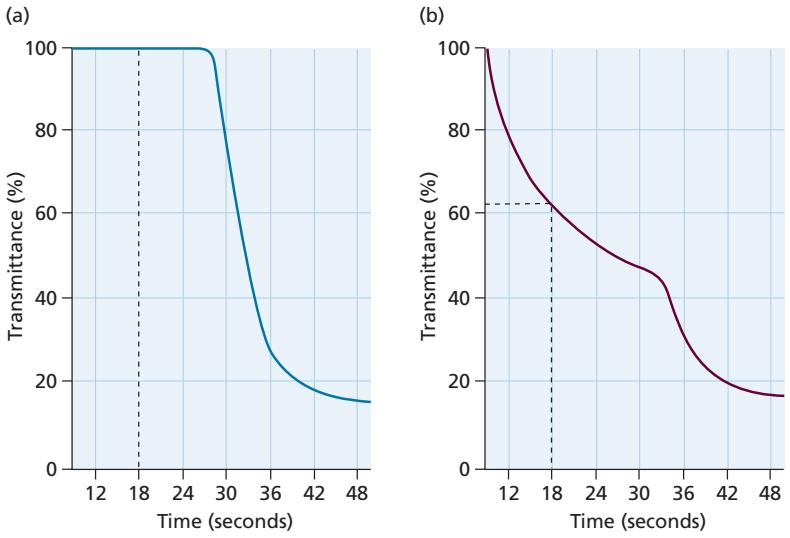


Figure 40.2 Representative normal and biphasic waveforms. Activated partial thromboplastin time waveform patterns: (a) normal; (b) biphasic. Dashed lines denote transmittance level at 18 s (TL18) as the quantitative index: (a) TL18 = 100; (b) TL18 = 63.

Table 40.1 Clinical conditions associated with disseminated intravascular coagulation.

Sepsis/severe infection
Potentially any microorganism, including Gram-positive and Gram-negative bacteria, viruses, fungi and rickettsial infections
Malaria and other protozoal infections
Trauma
Serious tissue injury
Head injury
Fat embolism
Burns
Malignancy
Solid tumours
Haematological malignancies (e.g. acute promyelocytic leukaemia)
Obstetric complications
Placental abruption
Amniotic fluid embolism
Pre-eclampsia
Intrauterine fetal demise
Vascular abnormalities
Giant haemangiomas (Kasabach–Merritt syndrome)
Large vessel aneurysms (e.g. aortic)
Severe toxic or immunological reactions
Snake bites
Recreational drugs
Severe transfusion reactions
Transplant rejection
Miscellaneous
Severe pancreatitis
Heat stroke
ABO mismatch transfusion

result in simultaneous bleeding and microvascular thrombosis at different vascular sites (Figure 40.3). Removal of the inciting cause is the best means of restoring haemostatic control, but may not be possible. Sepsis and trauma cause more than half the cases of DIC.

Pathophysiology

There are several important themes in the pathophysiology of DIC (Figure 40.3): first is the central role played by the generation of thrombin; second, mechanisms that perpetuate thrombin generation become pathogenic in its dissemination; third, parallel and concomitant activation of inflammation and fourth, the importance of the endothelial microvasculature in this process.

Excess bleeding in DIC is partly attributable to the depletion of coagulation factors and platelets. However, other factors may contribute, including abnormal platelet function and hyperfibrinolysis. Conversely, microvascular thrombosis can be precipitated by reduced levels of circulating anticoagulant proteins as well as loss of receptors, such as thrombomodulin. Endothelial dysfunction can also lead to the depletion of nitric oxide and result in uninhibited platelet activation.

Thrombin generation *in vivo*

The TF pathway plays the major role in initiating thrombin generation. In sepsis, infecting micro-organisms induce TF expression on monocytes and other inflammatory cells, while in trauma thromboplastin-like substances can be released from injured tissues such as the brain. During obstetric complications, the placenta and amniotic fluid act as rich sources of TF, while malignant cells release products with TF-like activity. Although the TF pathway is considered more important in thrombin generation, the contact pathway through factor (F)XII activation contributes to the pathological state by activation of

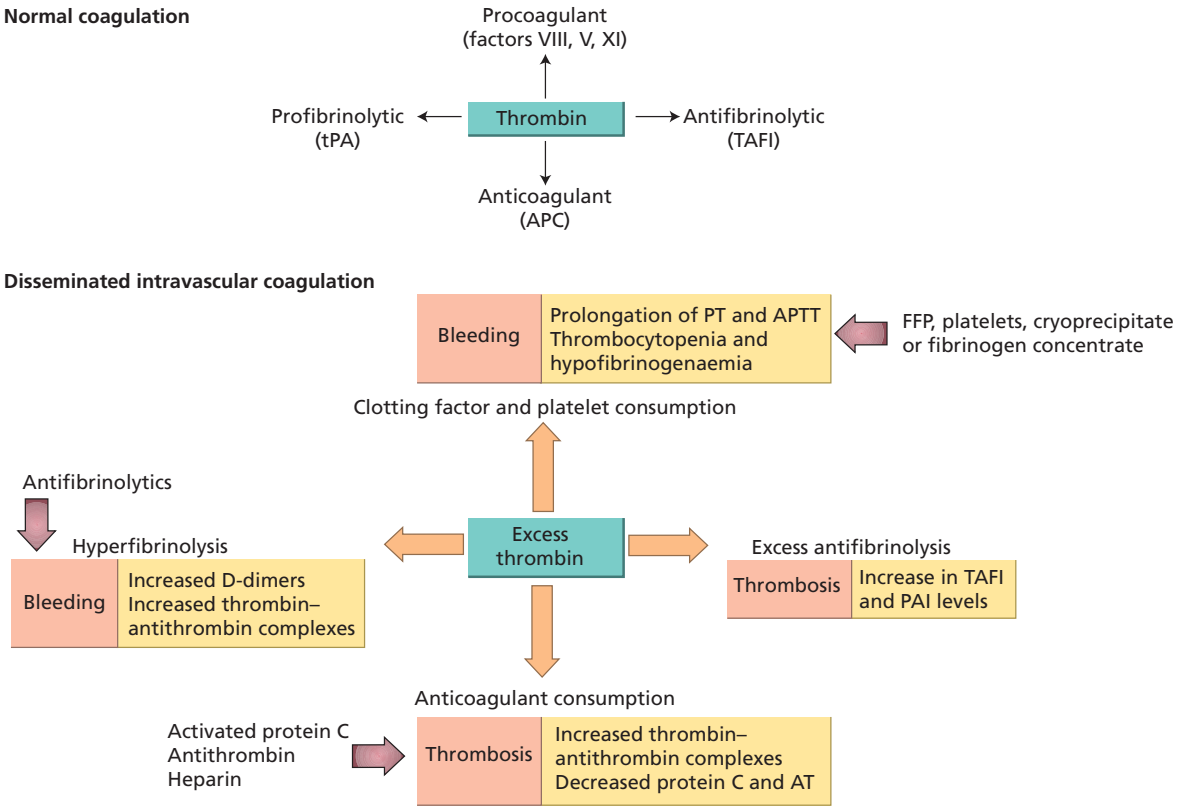


Figure 40.3 The changes in disseminated intravascular coagulation are compared with the normal coagulation. The excess thrombin generation in DIC leads to either bleeding or thrombosis (left) based on the predominant coagulation change (right) which has occurred. The therapeutic intervention in each setting is given by the arrow towards each double box. TAFI, thrombin activatable fibrinolysis inhibitor; PAI, plasminogen activator; AT, antithrombin.

the kallikrein–kinin system, causing vasodilatation, hypotension and activation of fibrinolysis.

Mechanisms for disseminating and sustaining thrombin generation

When the inciting insult is persistent or severe, the amount of thrombin generated becomes continuous and excessive. This can lead to consumption and depletion of coagulation and anticoagulant factors. Increased exposure of negatively charged phospholipid surfaces facilitates the assembly and enhances the rate of coagulation reactions. Such surfaces, mainly rich in phosphatidylserine, are provided by externalization of the inner leaflet of cell membranes upon activation and apoptosis. Cell damage also leads to the generation of microparticles from platelets, monocytes and endothelial cells that increase the circulating surface area for coagulation reactions. Phospholipid surfaces are also provided by very low-density lipoprotein, which can increase several-fold in severe sepsis to further enhance and sustain thrombin generation. Together, these mechanisms promote a spatially and temporally expanded response that is the hallmark of DIC.

Links between inflammation and coagulation

Once activated, the inflammatory and coagulation pathways interact to amplify the response. Cytokines and proinflammatory mediators such as tumour necrosis factor (TNF), interleukin (IL)-1 and high mobility group box protein (HMGB)-1 induce activators of coagulation. Thrombin and other serine proteases interact with protease-activated receptors (PARs) on cell surfaces to promote inflammation. When this process escapes the well-developed local checks, it results in a dysregulated response that fuels a vicious cycle. An important role for the complement system has also been increasingly identified in DIC. Thrombin can convert C5 to C5a, while the mannan-binding lectin pathway triggers coagulation by converting prothrombin to thrombin.

Endothelial cell activation and dysfunction

Dysfunction and failure of the endothelium beyond the host adaptive response can lead to the development of DIC. The degree to which this occurs and the dominance of thrombotic or bleeding sequelae depend on genetic and other host-related factors. Damage to and activation of the endothelium downregulates and depletes its anticoagulant receptors, such as

thrombomodulin and endothelial protein C receptor, exposes subendothelial collagen that binds activated platelets, and releases plasminogen activator inhibitor (PAI)-1, which inhibits fibrinolysis and ultra-large VWF multimers, which increase platelet aggregation.

Neutrophil extracellular traps and histones

In contrast to the well-known modes of cell death like apoptosis or necrosis, a unique type of neutrophil death has been recently described known as NETosis. This process, where neutrophils release extracellular traps (NETs), can aid in the removal of pathogens. At the same time, NETs can activate coagulation by assisting in the extracellular delivery of tissue factor, degrade tissue factor pathway inhibitor, stimulate platelets and cause aggregation, and also cause endothelial damage. In other words, a beneficial antimicrobicidal function in septic states can turn to become deleterious, causing disseminated coagulation. Another pathogenic player in DIC is the most abundant protein in the nucleus, histones. High levels of histones have been detected in the plasma of experimental animals, and human patients with sepsis and trauma, common causes of DIC. Extracellular histones released after cell damage can cause massive thromboembolism associated with consumptive coagulopathy and are major mediators of death in sepsis. It can also cause acute lung injury in trauma models and may be a useful therapeutic target.

Clinical features

The perturbed coagulation associated with DIC can manifest clinically at any point in a spectrum varying from bleeding to thrombosis. Although bleeding, ranging from oozing at venepuncture sites to major gastrointestinal or intracranial haemorrhage, is the archetypal and most obvious manifestation of DIC, organ failure due to microvascular thrombosis is more common and often unrecognized. For example, in meningococcal septicaemia and rarely pneumococcal infection, thrombosis of the adrenal vessels can lead to adrenal insufficiency and Waterhouse–Friderichsen syndrome.

Diagnosis

DIC is a clinicopathological syndrome and, as such, there is no single laboratory test that can confirm or refute the diagnosis. In clinical practice, the diagnosis is usually made by a combination of routinely available coagulation tests in a clinical situation where DIC is suspected. DIC must be differentiated from other acquired disorders of haemostasis (Table 40.2). The typical findings are a prolonged PT and APTT, elevated products of fibrin breakdown (e.g. D-dimer), thrombocytopenia and reduced fibrinogen. However, results within the normal range for these tests do not exclude a significant consumptive coagulopathy, because the acute-phase response shortens the APTT and increases fibrinogen. A fall in platelet number within the normal range may be significant.

It is important to recognize that DIC is a dynamic process and thus interpreting a series of laboratory tests over time is more relevant than looking at a single set of results. The ISTH Sub-Committee of the Scientific and Standardization Committee on DIC has recommended the use of a scoring system for overt DIC (Table 40.3). This has been prospectively validated, indicating

Table 40.2 Differential diagnosis of disseminated intravascular coagulation (DIC).

Condition	Similarities	Differences
Liver disease	Bleeding common PT, APTT abnormal Platelet count low Fibrinogen low	D-dimer usually normal* FVIII levels not affected
Microangiopathic haemolytic anaemia (e.g. HELLP, TTP)	Microthrombi common Platelet count low	Bleeding uncommon Coagulation tests normal†
Hyperfibrinolysis	PT, APTT abnormal Fibrinogen low	Platelet count normal
Catastrophic antiphospholipid antibody syndrome	PT, APTT abnormal Platelet count low	Fibrinogen not low D-dimer normal‡
Massive transfusion	PT, APTT abnormal Fibrinogen low Platelet count low	D-dimer normal

*Unless additional disorders which increase the D-dimer coexist.
†Unless coexisting DIC.
‡Some antibodies can affect the D-dimer results.
HELLP, haemolysis, elevated liver enzymes, low platelet count (syndrome); TTP, thrombotic thrombocytopenic purpura.

Table 40.3 ISTH Sub-Committee of the Scientific and Standardization Committee on DIC recommended scoring system for overt disseminated intravascular coagulation (DIC).

1 Risk assessment: does the patient have an underlying disorder known to be associated with overt DIC?
If yes, proceed
If no, do not use this algorithm
2 Order global coagulation tests (platelet count, PT, fibrinogen, soluble fibrin monomers or FDPs)
3 Score global coagulation test results
Platelet count ($\times 10^9$/L): >100, score 0; <100, score 1; <50, score 2
Elevated fibrin-related marker (e.g. soluble fibrin monomers or FDPs): no increase, score 0; moderate increase, score 2; strong increase, score 3
Prolonged PT (s): <3, score 0; >3 but <6, score 1; >6, score 2
Fibrinogen (g/L): <1 g/L score 1, fibrinogen ≥1 g/L score 0
4 Calculate score
5 Score ≥5 compatible with overt DIC. Repeat scoring daily

FDPs, fibrin degradation products; PT, prothrombin time.

a very high sensitivity and specificity. A strong correlation between an increasing DIC score and mortality has been demonstrated. The optical light transmittance profile of the APTT, referred to as the biphasic waveform, has also been shown to correlate well with the overt DIC score. The biphasic waveform occurs independently of prolongation in the clotting times in patients with DIC.

Treatment

The mainstay of treatment of DIC is to remove the underlying cause. However, DIC often continues after appropriate treatment for the underlying condition. Supportive therapy with blood products may be necessary. Although the efficacy of blood product replacement has not been proven in randomized controlled trials, it is a biologically rational option to replace both thrombin-promoting and thrombin-opposing proteins, particularly when there is significant depletion of these factors in a patient who is either bleeding or at risk of bleeding. In those patients not bleeding, transfusion of platelets or plasma in patients with DIC should not be undertaken based on laboratory results. There are no clinical or experimental data to suggest that platelet or plasma transfusions worsen the thrombotic process.

In patients with DIC and bleeding or at high risk of bleeding (e.g. after invasive procedures) and a platelet count less than 50×10^9/L, transfusion of platelets should be considered, especially because platelet function may be impaired. In bleeding patients with DIC and prolonged clotting times, fresh-frozen plasma (FFP) should be administered. If fluid overload is an issue, prothrombin complex concentrates (PCCs) could be considered, but these will only correct vitamin-K-dependent components. Activated PCCs may precipitate or worsen DIC. Severe hypofibrinogenaemia (<1.5 g/L) requires the use of fibrinogen concentrate or cryoprecipitate to correct.

A recent large trial in patients with severe sepsis showed a non-significant benefit of low-dose heparin on mortality and suggested that this is continued in patients with DIC and abnormal coagulation parameters, in the absence of overt bleeding. Notably, these patients are at highest risk of venous thromboembolism due to immobility, recent surgery and a proinflammatory state. However, the role of heparin remains controversial and unproven. A prophylactic dose of low-molecular-weight heparin is given to the high-risk patients with no evidence of active bleeding.

Since reduction of natural anticoagulants is often observed, trials using these agents have been conducted. Although activated protein C was noted to be beneficial in the initial stages, the latest evidence has shown no clear benefit and the use of APC for DIC is not recommended anymore. Similar results were found with antithrombin, although this agent continues to be used in Japan. Initial promising results have been shown with thrombomodulin, but further studies are required. Studies from Japan have demonstrated that, in comparison with unfractionated heparin, recombinant thrombomodulin helped with DIC resolution rates in patients with infection and haematological malignancy.

Haemostatic dysfunction in acute promyelocytic leukaemia

Acute promyelocytic leukaemia (APL), in particular the microgranular variant (AML-M3v), is associated with major coagulation disturbance, including DIC in at least 80% of cases. Although the introduction of all-*trans* retinoic acid (ATRA) and arsenic trioxide as differentiation agents has markedly reduced the rate of early haemorrhagic death and almost 90% of patients are cured, the 10% mortality from bleeding complications has not improved. Bleeding does not correlate with the clotting parameters, but with a high white cell count (see also Chapter 20).

The cause remains poorly understood and probably relates to enhanced proteolysis, including fibrinolysis and disruption of endothelial barrier integrity. APL blast cells express TF and can also stimulate the production of inflammatory cytokines, which can further amplify TF levels and promote thrombosis. Cancer procoagulant, a cysteine proteinase detected on APL blast cells, can activate FX independent of FVII. Both these procoagulants are noted to be progressively reduced once patients have been commenced on ATRA. ATRA has also been demonstrated to inhibit vascular endothelial growth factor (VEGF) production,

which indirectly limits TF production. Annexin II expression is enhanced on the surface of blast cells and can act as a cell surface receptor for plasminogen and its activator, tissue plasminogen activator (tPA). Elastase, cathepsin-G and proteinase-3 are present in the granules of the APL blasts and can directly degrade fibrinogen, α_1-antitrypsin, C1 esterase inhibitor and VWF, inducing the loss of high-molecular-weight multimers.

Management of the haemostatic abnormalities revolves around supportive care, including platelet transfusions to maintain a count of at least 30×10^9/L (or 50×10^9/L if the patient is bleeding) and the adequate replacement of fibrinogen to at least 1.5 g/L. The use of antifibrinolytic agents is supported by some evidence, but remains controversial. Although markers of coagulation activation and fibrinolysis fall rapidly and completely following the start of ATRA therapy, there appears to be a slower resolution of procoagulant markers (up to 30 days). This may partly explain the clinical observation of thromboembolic events occurring in patients on ATRA, and thus prophylactic use of low-molecular-weight heparin or other anticoagulants should be considered once bleeding manifestations have settled.

Vitamin K and related disorders

Vitamin K metabolism

Vitamin K ('K' denoting *Koagulation*) is a group of lipophilic and hydrophobic vitamins that are needed for the post-translational modification of proteins. Vitamin K is a cofactor for vitamin-K-dependent carboxylase, an enzyme that catalyses the carboxylation of glutamic acid (Glu) residues in several proteins (Figure 40.4). This reaction results in vitamin K 2,3-epoxide, which is recycled to reduced vitamin K by vitamin K epoxide reductase (VKOR), the enzyme inhibited by coumarin anticoagulants.

The Gla domains of the vitamin-K-dependent coagulant and anticoagulant proteins (FII, FVII, FIX, FX and proteins C and S) allow calcium-dependent binding to negatively charged phospholipids of activated cell membranes. This enhances the rate of the enzymatic reactions. In the bones, carboxylation of the Gla protein osteocalcin is essential for incorporation of calcium into hydroxyapatite crystals. Retrospective studies suggest that long-term therapy with coumarin-based anticoagulants can affect vertebral bone density and fracture risk.

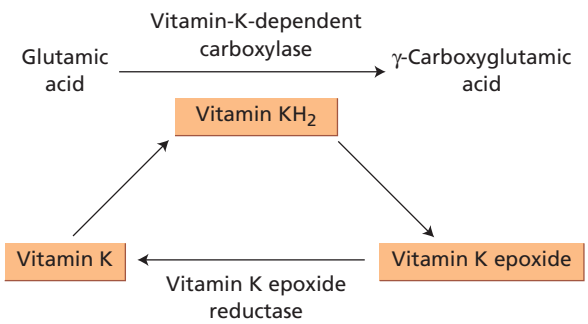


Figure 40.4 Vitamin K metabolism.

Vitamin K deficiency

Vitamin K deficiency can occur at any age, but is more common in infants because vitamin K does not cross the placenta, breast milk has low levels and there is low colonic bacterial synthesis. The clinical presentation of vitamin K deficiency in infants is with bleeding. It was previously called 'haemorrhagic disease of the newborn' but is now termed vitamin K deficiency bleeding (VKDB). Supplementation of vitamin K is necessary to reduce the risk of bleeding, especially in exclusively breast-fed babies. A single intramuscular injection of vitamin K 1 mg prevents VKDB; however, if oral replacement is used, prolonged administration is required, although the exact dose is controversial.

In adults, vitamin K deficiency is uncommon due to recycling of the vitamin and an adequate gut flora. However, a poor dietary intake in combination with antibiotic therapy can cause deficiency. Other causes of vitamin K deficiency include malabsorption and cholestatic liver disease (poor enterohepatic circulation).

Vitamin-K-dependent coagulation factors in healthy full-term infants are about half of normal adult values. Adult values are reached by about 6 months, except for protein C, which does not reach adult levels until adolescence. Coagulation tests should be compared with age-matched reference ranges to distinguish physiological and pathological deficiencies.

VKDB in young children has been classified into early, classical and late types. The clinical features are shown in Table 40.4. The diagnostic criteria for VKDB include a PT more than four times control in the presence of at least one of: (i) a normal platelet count, normal fibrinogen level and absent fibrinogen degradation products, (ii) normalization of coagulation tests after parenteral vitamin K administration or (iii) the presence of proteins induced by vitamin K absence or antagonism (PIVKA) in plasma. The presence of PIVKA without a coagulation deficit is a marker of subclinical vitamin K deficiency.

The treatment of a non-life-threatening bleed is with vitamin K_1 (phytomenadione) given slowly intravenously. A dose of 1–2 mg is enough to fully correct the deficiency in infants aged up to 6 months, higher doses offering no advantage in efficacy or speed of reversal. There have been reports of anaphylactic reactions although these are rare. The PT should improve within 48 hours and if this has not occurred after three doses, continuation of vitamin K in most cases is unlikely to help. Oral replacement may be used if there is no active bleeding and absorption is normal. Patients with severe bleeding should be treated with FFP or PCC.

Table 40.4 Features of vitamin K-deficiency bleeding (VKDB).

	Early VKDB	Classical VKDB	Late VKDB
Presentation	First 24 hours of life	Days 2–7 of life	Day 8 to 6 months of life
Cause	Infants of mothers on vitamin-K-inhibiting drugs	Inadequate feeding	Cholestasis or malabsorption syndromes
Clinical presentation	Severe, with intracranial and intrathoracic bleeds	Milder, with gastrointestinal or umbilical bleeds being common	Severe, with high incidence of intracranial bleeds with 'warning bleeds' in up to one-third
Incidence of intracranial haemorrhage	High	Rare	Very high

Poisoning following the accidental ingestion of large doses of warfarin or compounds with similar properties is suggested by poor correction of PT or International Normalized Ratio (INR) with normal doses of vitamin K. Serum assays of warfarin concentration can be undertaken for confirmation.

Haemostatic disturbance in liver disease (Figure 40.5)

Almost all procoagulant factors, natural anticoagulants and inhibitors of coagulation are synthesized in the liver. The liver is also involved in the clearance of activated clotting factors. The effect of liver disease on haemostasis is therefore complex and a balance between procoagulant and antithrombotic changes (Figure 40.5; Table 40.5). Routine coagulation tests such as PT and APTT are commonly prolonged in liver disease, but are only sensitive to decreases in procoagulant factors and hence are liable to over-estimation of the haemostatic defect. Patients with liver disease have similar thrombin generation to healthy control individuals, when tests are performed in the presence of thrombomodulin to make them sensitive to protein C deficiency.

Acute hepatitis

Patients with acute hepatitis are often thrombocytopenic (platelets 100–150 × 10^9/L). The mechanism for this may be immune, due to concurrent hypersplenism or DIC. Platelet function may be impaired in acute hepatitis, although this is unlikely to be of clinical significance. In acute liver disease, biosynthesis of clotting factors is impaired and this may be reflected in a prolongation of coagulation tests. The PT and FV, FVII, antithrombin and protein C levels are the most sensitive to hepatic biosynthetic dysfunction. The PT and FVII level predict survival in acute liver disease. The plasma fibrinogen may be raised in acute hepatic disturbance as part of an acute-phase

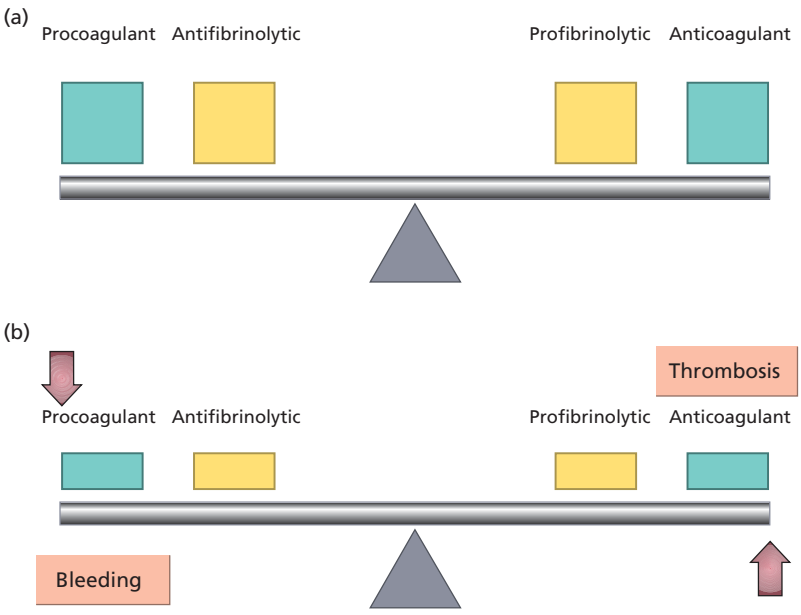


Figure 40.5 Haemostatic disturbance in liver disease. The haemostatic system in liver disease (bottom) compared with normal haemostasis (top). In normal healthy adults, the procoagulant and anticoagulant systems and the profibrinolytic and antifibrinolytic systems form a balance that is not easily disturbed. In patients with liver impairment, all four components are variably reduced and additional stimuli like infections or renal failure can upset the balance very easily, leading to either bleeding or thrombosis.

Table 40.5 Causes of increased bleeding tendency or coagulation abnormalities in liver disease.

Decreased coagulation factors
Decreased synthesis
Vitamin K deficiency
Increased clearance
Abnormal coagulation factors due to hypocarboxylation
Decreased platelets
Sequestration in spleen or liver
Thrombopoietin deficiency
Destruction due to toxins or antiplatelet antibodies or toxins
Toxic effect of alcohol on marrow
Dysfunctional platelets
Altered platelet arachidonic acid metabolism
Defective signal transduction
Storage pool deficiency
Abnormalities of the platelet glycoprotein
Increased platelet inhibitors nitric oxide and prostacyclin
Hypofibrinogenaemia
Impaired synthesis
Loss into extravascular spaces (ascites)
Increased catabolism
Loss due to massive haemorrhage
Dysfibrinogenaemia: abnormal sialic acid content
Abnormal fibrinolysis
Increased release and decreased clearance of tissue plasminogen activator
Decreased synthesis of α_2-antiplasmin, plasminogen and thrombin-activated fibrinolysis inhibitor
Infections
Endogenous heparinoids: anti-FXa activity
Renal failure
Sinusoidal endothelial cell dysfunction
Hypothermia

response and a low fibrinogen is associated with a poor prognosis. A pattern of reduced fibrinogen and increased fibrin degradation products seen in acute liver disease may be due to hyperfibrinolysis.

Chronic liver disease

Low platelet count and abnormal platelet function are also common in chronic liver disease. Mild to moderate thrombocytopenia is noted in up to 30% of patients with chronic liver impairment and in up to 90% of patients with terminal liver disease. In patients with cirrhosis, this is frequently related to hypersplenism, secondary to portal hypertension. In addition, dietary problems, such as folate deficiency and the direct toxic effect of ethanol on megakaryocyte function, may contribute to thrombocytopenia. A more profound thrombocytopenia in patients with hepatitis C virus (HCV) infection may be autoimmune, although the response to steroids is poor; however, approximately half of adult patients treated with interferon-α responded with a rise in platelet count. More recently, thrombopoietin receptor agonists have been shown to increase platelet counts to more than 100×10^9/L in the majority of patients with cirrhosis associated with HCV infection, in a dose-dependent manner.

A variety of platelet function abnormalities has also been demonstrated in liver disease, whilst VWF may be increased tenfold and has been shown to support platelet adhesion despite reduced functional capacity. However, desmopressin did not show any efficacy in reducing blood loss in patients undergoing partial liver resection or liver transplantation.

Coagulation factor and natural anticoagulant synthesis is affected by chronic liver disease. FIX is usually reduced less than FII, FVII, FX and protein C. The levels of these factors, especially FVII, have been demonstrated to fall proportionately with increasing severity of disease, with levels of FVII shown to be an independent predictor of survival (level below 34% associated with 93% mortality). Chronic liver impairment also leads to vitamin K deficiency, mainly by decreased absorption from the gut as a result of reduced bile salt secretion in cholestasis (parenteral vitamin K responsive) and decreased synthesis in parenchymal disease (parenteral vitamin K non-responsive). The PT has been incorporated into prognostic indices of chronic liver disease, such as the Child–Pugh and Mayo End-Stage Liver Disease scores. Despite the prolongation of clotting tests such as PT and APTT, bleeding after liver biopsy or other potentially haemorrhagic procedures in patients with cirrhosis is rare, suggesting that these tests are not on their own indicative of the haemorrhagic tendency. The levels of anticoagulant proteins, antithrombin and proteins C and S are reduced in a similar manner to other clotting factors, thus potentially providing a net balance in haemostasis. It is common practice to use the INR instead of the PT to access patients with liver disease. However, the INR is standardized for patients taking vitamin K antagonists and its use in patients who have liver disease would require determination of an international sensitivity index (ISI) relevant to liver disease and recently, attempts have been made to derive ISI_{liver} and so INR_{liver}.

Most patients with stable chronic liver disease have normal or increased fibrinogen, although in advanced disease fibrinogen falls. This may be due to impaired synthesis, loss into extravascular spaces (ascites), increased catabolism or massive haemorrhage. Some patients develop dysfibrinogenaemia due to increased activity of sialyltransferase expressed by immature hepatocytes generated during hepatic injury, which leads to low-molecular-weight fibrinogen with abnormal α-chains and higher sialic acid content. Dysfibrinogenaemia is reflected by a prolonged TT and/or reptilase time, with the fibrinogen antigen level reduced less than the Clauss fibrinogen. Cirrhosis can also be associated with accelerated fibrinolysis due to decreased clearance of tPA, and decreased synthesis of α_2-antiplasmin

and thrombin-activatable fibrinolysis inhibitor (TAFI). Also, the levels of plasminogen (a fibrinolytic) are reduced and those of PAI-1 (an antifibrinolytic) are high in liver disease, ultimately providing a balance for the fibrinolytic system. However, hyperfibrinolysis has been found to be a predictor of the first episode of upper gastrointestinal bleeding in cirrhotic patients with portal hypertension. Life-threatening haemorrhage in patients with liver disease is often related more to portal hypertension (the localized coagulation problem is termed 'accelerated intravascular coagulation and fibrinolysis') than to net dysfunction of the clotting cascade, wherein the varices play a major part. Renal failure is also common in advanced liver disease (hepatorenal syndrome), which can impart a bleeding risk.

The management of a bleeding patient with hepatic dysfunction depends on the site of bleeding and the haemostatic dysfunction. Clotting function should be assessed by means of the PT, APTT, TT, and fibrinogen and D-dimer levels, although the results should be interpreted in the clinical context before transfusions are undertaken. Vitamin K should be administered intravenously (10 mg daily for 3 days) to aid biosynthesis of vitamin-K-dependent factors. The widespread practice of managing the coagulopathy of liver disease with transfusions of FFP or platelets should only be considered in the presence of active bleeding. Large volumes of FFP may be required and this may present a management problem in patients with hepatic disease who are at risk of fluid overload. Prothrombin complex concentrates are increasingly being used in this setting to minimize the risk of fluid overload, although these products are associated with a theoretical risk of thrombosis. Platelet transfusion may be necessary, although platelet recovery may be reduced because of hypersplenism or immune-mediated destruction. There are no evidence-based guidelines to establish safe coagulation levels for liver biopsy and similar procedures and as such it is important to allow these only when the benefits outweigh the risks (transjugular route is safer than percutaneous for liver biopsy). Correction of the coagulopathy is particularly important before the placement of an intracranial pressure transducer in patients with hepatic encephalopathy. Spontaneous intracranial bleeding has been reported in such patients, though rarely, and this remains one of the principal concerns regarding abnormal coagulation in liver disease. In patients with evidence of increased fibrinolysis, antifibrinolytic drugs such as tranexamic acid should be considered.

Liver transplantation

The changes in the balance of haemostasis seen during liver transplantation are complex and multifactorial and can be due to surgical and non-surgical causes. Traditionally, the process of liver transplantation is split into three stages: stage I, the preanhepatic stage, which ends with the occlusion of the recipient's hepatic blood flow; stage II, the anhepatic phase, which ends with the re-perfusion of the donor liver; and stage III, the reperfusion and neohepatic period. The risk of bleeding in the preanhepatic stage is related directly to the preoperative haemorrhagic risk related to the underlying liver disease. A reduction in procoagulant factors may be seen, especially if large blood losses necessitate transfusion, leading to the dilution. During the anhepatic stage, studies have reported enhanced fibrinolytic activity. Reperfusion of the liver during the postanhepatic phase is associated with a dramatic increase of fibrinolysis in almost three-quarters of patients. Usually, hyperfibrinolysis subsides within an hour, but in a damaged donor liver, sustained increased fibrinolytic activity may be observed. The endothelium of the donor liver is an important source of tPA; the ischaemic damage to the graft during preservation may explain the dramatic increase in plasminogen activators. After re-perfusion, the release of heparin-like compounds has also been shown. In the postoperative period, a reduction in platelet count is related to platelet activation and blood loss, in addition to low thrombopoietin levels. The levels of this hormone increase significantly on the first day after liver synthetic function is restored, with the platelet count normalizing in two weeks.

There is little evidence that one can predict the transfusion requirement of a patient by examining the preoperative coagulation results. The profound coagulation abnormalities seen in fulminant hepatic failure should be corrected preoperatively; however, correction of a mild coagulopathy in a patient with chronic liver disease is less likely to be of benefit. During surgery, the aim is to transfuse blood components to prevent the development of intractable coagulopathy. Near-patient tests of haemostasis performed on whole blood are often used in theatre to complement the tests performed in the coagulation laboratory. Many departments use algorithms based on coagulation and near-patient testing to guide FFP, red cell, cryoprecipitate and platelet transfusion during the procedure. Aprotinin can be administered to counteract the increase in fibrinolysis observed.

Hypercoagulability in liver disease

Tipping of the balance in the complex interplay between endogenous procoagulants and the anticoagulant system in liver disease can lead to a hypercoagulable state, in contrast to the common assumption of an 'autoanticoagulated state'. The hypercoagulable events may be clinically evident, as with portal vein thrombosis and venous thromboembolism, but in the microvasculature can contribute to portopulmonary hypertension, liver fibrosis, thrombosis of extracorporeal circuits and progression of non-alcoholic steatohepatitis to cirrhosis. Hypercoagulation in liver disease may be related to poor flow, endothelial dysfunction and vasculopathy associated with a chronic inflammatory state, increased levels of FVIII and VWF, or decreased synthesis of the naturally occurring anticoagulant proteins. Hepatic fibrogenesis may be caused by tissue ischaemia and direct thrombin-mediated stellate cell activation by PAR-1 cleavage.

Haemostatic disturbance in renal disease

Bleeding was a frequent cause of morbidity and mortality in patients with renal failure before the advent of dialysis. Common manifestations are gastrointestinal bleeding from angiodysplasia and peptic ulcers, as well as prolonged bleeding from skin puncture sites. Subdural haematomas and haemorrhagic pericarditis are also seen, although bleeding at the time of renal biopsy is rare. Bleeding is seen despite normal or elevated circulating levels of coagulation factors, suggesting that platelet abnormalities are the likely cause. This is supported by the finding of a prolonged bleeding time and reduced platelet aggregation to various agonists. Severe thrombocytopenia (platelets $<50 \times 10^9$/L) secondary to renal failure is rare and its presence should suggest concomitant conditions such as HCV infection or vasculitis.

Platelet dysfunction with renal failure is multifactorial and can be divided into: (i) intrinsic platelet defects, (ii) abnormal interaction of platelets with the endothelium, (iii) effects of uraemic toxins, (iv) effects of anaemia on platelets and (v) dialysis-related. Intrinsic platelet abnormalities reported in association with uraemia are shown in Table 40.6. The interaction between uraemic platelets and the endothelium can be markedly reduced, due partly to impaired VWF binding to platelets and increased endothelial production of nitric oxide, an inhibitor of platelet aggregation. Uraemic patients with prolonged bleeding times also have raised prostaglandin (PG)I_2 levels.

Anaemia-related haemostatic dysfunction is primarily a result of reduced displacement of platelets to the vessel wall by red cells. Dialysis can improve bleeding symptoms. Antiplatelet drugs, antibiotics (especially those that can accumulate in renal failure) and heparin may contribute to the bleeding risk.

The bleeding time better correlates with clinical bleeding than tests of renal function, although measurement is rarely clinically useful. Adequate dialysis will improve symptoms and the bleeding time in most patients. Intravenous, subcutaneous or intranasal administration of desmopressin in conjunction with antifibrinolytics has also been used successfully to control uraemic bleeding. Although the response is brief (4–8 hours), the rapid onset of action is beneficial. Care is required to avoid hyponatraemia. Desmopressin is also thought to improve platelet function by the enhanced release of endothelial VWF multimers and also by increasing the levels of platelet glycoprotein (GP)Ib/IX.

Correction of anaemia with recombinant human erythropoietin so that the haematocrit exceeds 30% improves platelet interaction with the vessel wall and has also been shown to increase the number of reticulated platelets. Other potentially beneficial effects of recombinant human erythropoietin include improved platelet intracellular calcium mobilization, increased expression of GPIb and improved platelet signal transduction.

Table 40.6 Platelet abnormalities reported in uraemia.

Decreased GPIb complexes
Reduced serotonin and ADP in the granules
Increased levels of cyclic AMP
Defective ristocetin-induced platelet aggregation
Abnormal mobilization of free cytoplasmic calcium in response to agonists
Reduced release of arachidonic acid from membrane phospholipids
Decreased conversion of arachidonic acid to thromboxane A_2
Abnormal dense-granule and α-granule secretion
Abnormal cytoskeletal assembly
Deficient tyrosine phosphorylation

Pregnancy-related haemostatic dysfunction

Normal pregnancy is associated with physiological changes in haemostasis, with increased levels of procoagulant proteins such as fibrinogen, VWF and FVIII and a fall in anticoagulants such as protein S. These changes partly contribute to the increased risk of venous thromboembolism during pregnancy.

Obstetric haemorrhage is the most common cause of severe pregnancy-related morbidity. The cause of bleeding is often multifactorial, with a combination of physical and acquired haemostatic defects. Consumption leads to depletion of platelets and coagulation factors, particularly fibrinogen. A particularly aggressive consumptive coagulopathy is seen with amniotic fluid embolus which is often fatal. Consumptive coagulopathy may also be triggered by placental abruption, intrauterine fetal death, pre-eclampsia and acute fatty liver of pregnancy. Once bleeding and resuscitation has started, a dilutional coagulopathy may exacerbate the haemostatic failure and contribute to the ongoing bleeding.

Treatment requires rapid recognition of bleeding and a coordinated response from obstetricians, midwives, anaesthetists and haematologists. Physical methods to control bleeding include the use of oxytocins to contract the uterus, B Lynch sutures, intrauterine balloon tamponade and uterine artery embolization. Assessment of haemostatic failure requires an urgent FBC and a coagulation screen that includes a Clauss fibrinogen because the PT and APTT may be normal despite a low fibrinogen. Thromboelastometry is becoming more commonly used because laboratory coagulation tests are often too slow to manage patients adequately. Blood product replacement is often required before the results of the blood test are available and empirical replacement therapy may be necessary. In patients with severe haemostatic failure, a standard dose of FFP (15 mL/kg) is unlikely to be adequate to correct the haemostatic defect and the use of larger volumes should be anticipated. Platelets should be maintained above 75×10^9/L and fibrinogen above at least 1.5–2 g/L (although the appropriate level is

debated) if bleeding is ongoing. Multiple studies have shown that a fibrinogen less than 2 g/L is associated with progression to severe postpartum haemorrhage. Some centres use fibrinogen concentrate in obstetric haemorrhage to rapidly correct this clotting factor. Fibrinogen concentrate is not licensed in the UK and clinical studies are ongoing to assess its role during obstetric bleeding. Some clinicians advocate the use of rFVIIa if haemostasis cannot be secured by standard methods, but it will not be useful in the presence of low fibrinogen or thrombocytopenia and its use is empirical.

As soon as bleeding has been controlled the patient is likely to be at high risk of venous thromboembolism; this should be assessed and appropriate thromboprophylaxis started.

Haemostatic dysfunction associated with cardiopulmonary bypass surgery

Coronary artery bypass graft (CABG) surgery is associated with excessive bleeding in about 10% of cases. Re-operation for bleeding is required in 3–10% of cases, with a mortality of about 30%. Bleeding is more common with revision procedures and prolonged time on bypass. A surgically correctable source of bleeding may be found. Bleeding due to haemostatic disturbances can be patient related or CPB related. Antiplatelet, anticoagulant or fibrinolytic drugs, particularly at the time of emergency CABG after stenting procedures, commonly contribute to perioperative bleeding. Despite its antiplatelet activity, aspirin has not been demonstrated to cause increased blood loss after CABG, and has been shown to reduce the risk of death by one-third.

Haemostatic abnormalities associated with CPB are related to the interaction of blood with the extensive non-endothelial bypass surfaces and re-transfusion of pericardial blood. All coagulation factors except factor VIII are reduced during CPB. CPB decreases both the number and function of platelets. Platelet counts fall by 25–60% within 15 min of first passage of blood through the primed CPB circuit, in association with a prolonged bleeding time. The platelet count rarely falls below 100×10^9/L and persistent or profound thrombocytopenia should prompt consideration of alternative causes, including heparin-induced thrombocytopenia.

The observed changes in platelet function in patients undergoing CPB are due to modification of membrane components with loss of receptors, decrease in granule contents, aggregation to fibrinogen adsorbed onto the bypass circuit, mechanical trauma, exposure to hypothermia and heparin. Following CPB, platelet function returns to normal within 1 hour, although the platelet count may take several days to normalize. Coagulation factors fall because of haemodilution, which can be exacerbated by use of cell-salvage systems. Fibrinolysis is also enhanced by returned blood from the pericardiotomy suction.

Blood component administration in patients with excessive bleeding related to CABG is often empirical, with replacement of FPP and platelets. Near-patient whole-blood testing has been used to assess haemostasis and fibrinolysis and has been shown to help rationalize blood product usage. The off-licence use of PCCs and fibrinogen concentrates has also been reported. Randomized studies do not support the use of rFVIIa and it has been associated with increased thrombotic adverse events.

Aprotinin has been used to reduce bleeding during CPB by inhibition of fibrinolysis. Double-blind studies have consistently shown its effectiveness in reducing blood loss in comparison with tranexamic acid, α-aminocaproic acid and desmopressin. However, recent reports have shown a doubling of the risk of renal failure requiring dialysis in patients undergoing complex coronary artery surgery and a 55% increase in cardiovascular and cerebrovascular events. The use of aprotinin is currently contentious and under intense debate.

Haemostatic dysfunction associated with trauma

Coagulopathy associated with trauma is multifactorial and the important initiators include tissue injury, loss of procoagulant factors and platelets, physiological and therapeutic haemodilution, inflammation, hypothermia and acidosis. The localization of coagulation to the sites of injury (in contrast to DIC of sepsis) has made some experts use the term 'acute coagulopathy of trauma-shock'. Tissue damage in trauma initiates coagulation via TF, with head injury causing release of brain-specific thromboplastins into the circulation. Activation of the thrombomodulin/protein-C pathway and reduced TAFI have been described.

Massive blood loss leads to deficiencies of haemostatic factors and up to half of the fibrinogen and one-third of platelets may be lost before treatment of the injury has begun. Physiological filling of the vascular space with fluid from cellular and interstitial spaces, along with the administration of intravenous fluids and plasma-poor red cells, causes haemodilution to worsen the coagulopathy. Plasma expanders are also associated with anticoagulant effects such as reduced VWF levels and their use should be limited.

The presence of an abnormal coagulation test on arrival in the emergency department correlates with the severity of injury and the mortality rate. In one study an abnormal PT and APTT increased the risk of mortality by 6.3-fold and 10.7-fold, respectively.

Guidelines on triggers for blood product support are evolving and different specialties have produced variable advice; an example is given in Table 40.7. Many trauma centres use thromboelastometry to guide blood product replacement. Recent observations, predominantly based on experience of battlefield trauma, have prompted recommendations for aggressive and early plasma and platelet replacement with a ratio of 1:1:1 for red cells, plasma and platelets. This strategy has been reported

Table 40.7 Guide to blood product replacement in massive blood loss.

1 Control bleeding using surgical and/or radiological interventions
2 Restore an appropriate circulating blood volume
3 Control exacerbating factors: hypothermia and acidosis
4 Blood product support as detailed below

Red cells
Use O-negative red cells first
If no record of red cell antibodies, ABO- and Rh-compatible cross-matched blood should be available within 30 min (maximum 45 min)
Replace red cells as required to maintain circulating blood volume
Use blood warmer to avoid hypothermia

Fresh-frozen plasma
Transfuse FFP to prevent coagulopathy to maintain PT and APTT <1.5 times normal
If PT and APTT >1.5 times normal FFP at doses >15 mL/kg will be required

Platelet transfusion
One to two adult doses after 1.5–2 blood volume replacement (equivalent to 8–10 bags of red cells)
Transfuse at 75×10^9/L to maintain platelet count $>50 \times 10^9$/L

Fibrinogen
Maintain at least >1.5 g/L
Cryoprecipitate (dose: two donation pools)
Fibrinogen concentrate (50 mg/kg will increase fibrinogen by about 1g/L)

to improve outcome in non-randomized studies, but definitive evidence is lacking. Issues such as hypothermia and acidosis need to be addressed by using prewarmed fluids and extracorporeal warming devices and appropriate resuscitation. The role of rFVIIa remains controversial.

Coagulopathy in massive blood loss

Massive blood loss is usually defined as one blood volume (about 5 L in an adult) in 24 hours or 50% blood volume in 3 hours or more than 150 mL/hour. The associated coagulopathy is complex and multifactorial. It is partly dilutional, as blood is replaced with volume expanders and red cells, but DIC, hypoxia, hypothermia and acidosis contribute in some clinical situations. Haemostatic management should be to maintain a platelet count of at least 75×10^9/L, a threshold that is likely to be reached after a two-blood-volume transfusion. A higher platelet count of about 100×10^9/L will be needed in major trauma or central nervous system bleeding. Fibrinogen should be maintained above at least 1.5 g/L, a level that is likely to be reached after about a 1.5-blood-volume transfusion, assuming no significant additional consumptive coagulopathy.

Haemostatic therapy is to infuse platelets and FFP in volumes likely to maintain a safety margin above these critical levels (see also Chapter 13). Hypofibrinogenaemia unresponsive to FFP may require cryoprecipitate or fibrinogen concentrate. It is important to predict ongoing dilution and consumption of coagulation factors and to replace these expectantly if bleeding continues. Regular measurement of FBC and coagulation screens (including Clauss fibrinogen) are important for monitoring replacement treatment, although often these are not available rapidly enough to guide initial management and empirical treatment may be required in the early phases (see Table 40.7).

Bruising

Purpura simplex (normal easy bruising)

Distinguishing normal from pathological bruises may be difficult and individuals have variable thresholds for presenting for medical review. If bruising is not associated with other symptoms suggestive of a bleeding disorder and routine tests of coagulation (including VWF levels and platelet number) are normal, the patient can be reassured. Drugs such as non-steroidal anti-inflammatory agents or selective serotonin reuptake inhibitors are sometimes implicated.

Non-accidental bruising

Bruising is a common feature of non-accidental injury in both children and adults. Bruises that affect unusual sites, are in different stages of maturation or shaped like a hand or instrument should raise concern. Any bruising in a non-mobile baby should be investigated. In the case of children, appropriate liaison with child safeguarding agencies should be undertaken, although the possibility of an underlying congenital or acquired bleeding disorder should also be thoroughly investigated. Self-harm may also present with bruising and should be suspected if the pattern of bruising is atypical. Deliberate ingestion of anticoagulants and long-acting vitamin K antagonists is also possible.

Senile purpura (atrophic or actinic purpura) and steroid-related purpura

Bruising is more common in elderly people due to atrophy of subcutaneous tissues and loss of collagen and elastin fibres in subcutaneous tissues. Blood vessels in the skin can be broken by minor trauma or shearing forces and typically purpura are seen on hands and forearms. If routine tests of platelet number and coagulation are normal, no further investigation is required. Long-term steroid use is also associated with bruising secondary to atrophy of collagen fibres supporting blood vessels.

Painful bruising syndrome (psychogenic purpura)

Diamond–Gardner syndrome is a rare condition most commonly seen in women with psychological problems. It is also called painful bruising syndrome and psychogenic purpura. It presents with painful ecchymotic lesions, mostly on the extremities and the face and rarely on less accessible parts of the body. The skin lesions can be preceded by paraesthesia or pain and are usually reported after surgery or minor trauma. The usual tests for haemostasis reveal no abnormalities. The cause of the bruising is poorly understood, although Diamond and Gardner suggested the possibility of a local reaction to the patient's own red cell stroma and gave it the name 'autoerythrocyte sensitization syndrome'.

Scurvy

Scurvy is caused by a lack of vitamin C (ascorbic acid), which is required for the hydroxylation of prolyl and lysyl residues in the formation of mature collagen. Vitamin C deficiency renders the collagen unable to self-assemble into rigid triple helices, resulting in blood vessel fragility and poor wound healing. The characteristic clinical finding is perifollicular haemorrhage but large ecchymoses can occur on the legs and trunk. Follicular hyperkeratosis and corkscrew hairs are other cutaneous features. Bleeding gums are common. Scurvy is more common in people with poor nutrition, especially alcoholics. The diagnosis is clinical, although plasma or leucocyte vitamin C can be measured. A prompt response to vitamin C treatment is also diagnostic.

Inherited disorders of collagen and elastic fibres

Ehlers–Danlos syndrome (EDS) is a group of dominantly inherited disorders of connective tissue characterized by lax hyperelastic skin, joint laxity and poor wound healing. There are six subtypes and subtype IV is associated with severe varicosities and spontaneous rupture of major blood vessels. Bruising is seen in 90% of patients and has also been reported in people with lax joints who do not fulfil the criteria of EDS. Although mild platelet and coagulation factor dysfunction has been reported in association with EDS, they are likely to be coincidental. Some patients have a prolonged bleeding time and this can respond to desmopressin. Other disorders such as osteogenesis imperfecta and Marfan syndrome may also be associated with bruising.

Haemostatic dysfunction associated with vasculitis

The vasculitides are a heterogeneous group of disorders characterized by inflammation within blood vessel walls of different organs. It is classified by the size of the blood vessel involved and can be associated with both thrombosis, for example Behçet's disease (see Chapter 45), and haemorrhage. The classic presentation of Henoch–Schönlein purpura is palpable purpura without thrombocytopenia or coagulopathy, arthritis/arthralgia, abdominal pain and renal disease. Henoch–Schönlein purpura is more commonly associated with bleeding than thrombosis.

Arteriovenous malformations

Hereditary haemorrhagic telangiectasia

Hereditary haemorrhagic telangiectasia (HHT) is a dominantly inherited, highly penetrant, familial disease in which telangiectasia develop on the skin and mucous membranes. Associated large arteriovenous malformations may occur in solid organs. International consensus diagnostic criteria have been developed based on epistaxis, mucocutaneous telangiectasia, visceral lesions and an affected first-degree relative.

Mutations in the *ENG* gene (encoding endoglin) and *ACVLR1* gene (encoding activin receptor-like kinase 1) have been described. These genes encode proteins that are involved in signalling by the transforming growth factor (TGF)-β superfamily. Increased levels of VEGF have been demonstrated in HHT. Genetic testing may be useful to inform asymptomatic family members or in people in whom the diagnosis is unclear.

Cerebral, pulmonary and neurological involvement is described more commonly with *ENG* mutations, whereas liver involvement has been associated with *ALK1* mutations, although members of the same family can have variable phenotypes. Typical telangiectatic lesions are red 0.5–3 mm spots that blanche with pressure and are often noted on the skin of face and lips and mucosal surfaces of the tongue and mouth. Symptoms range from cosmetic and recurrent epistaxis and mouth bleeding to iron deficiency and intracranial haemorrhage. Emboli, including septic emboli, may occur and antibiotic prophylaxis at the time of invasive procedures such as dental extraction is recommended. Occasionally bleeding can occur from internal organs without visible lesions.

Virtually all internal organs can be affected. The lungs develop pulmonary arteriovenous fistulae in about 30% of patients, leading to haemoptysis, sometimes catastrophic, hypoxia, dyspnoea, pulmonary hypertension, high-output cardiac failure and clubbing. Transient ischaemic attacks, cerebral infarction and systemic emboli due to right-to-left shunting occur in about half of the patients with pulmonary arteriovenous fistulae. Pulmonary arteriovenous fistulae may worsen during pregnancy. Liver involvement is present in the majority of patients and, though often asymptomatic, can lead to cirrhosis, portal hypertension and high-output cardiac failure. Conjunctival telangiectasia may cause benign bleeds, while more serious retinal bleeds can occur. Intracranial haemorrhage is associated with arteriovenous fistulae, but embolization, including septic emboli, may also cause cerebral symptoms.

Options for symptomatic lesions include local treatment, such as laser therapy of skin and mucosal telangiectasia and embolization of arteriovenous malformations in lung or nervous system. Systemic treatment with oestrogens and tranexamic acid may reduce bleeding symptoms. Experimental treatments include thalidamide and bevacizumab to inhibit VEGF.

Kasabach–Merritt syndrome

Kasabach–Merritt syndrome is a rare condition in which a vascular tumour (Kaposiform haemangioendothelioma) is associated with thrombocytopenia, hypofibrinogenaemia and bleeding, which can be life-threatening. The thrombocytopenia in Kasabach–Merritt syndrome is presumed to be due to platelet adhesion to the abnormally proliferating endothelium. This leads to activation of platelets with secondary consumption of clotting factors, including VWF, resulting in systemic haemostatic failure. Excessive flow and shear rates secondary to arteriovenous shunting also contribute and microangiopathic haemolysis may be seen. The platelet count is often less than 20×10^9/L and the platelet half-life dramatically reduced. Intralesional bleeding can cause rapid enlargement of the haemangioma and can worsen the consumptive coagulopathy. Intralesional thrombosis may rarely occur, causing spontaneous resolution of some lesions. Management involves supportive care and, if possible, removal of the lesion. Steroids, interferon, chemotherapy with vincristine, radiotherapy and antiangiogenic agents have been tried.

Microthromboembolic disease

Cholesterol embolism

Cholesterol embolism is a complication of widespread atherosclerotic disease. Rupture of an atherosclerotic plaque can occur spontaneously or, more commonly, at the time of vascular surgery or invasive procedures. Anticoagulant or fibrinolytic therapy can weaken thrombi that usually prevent the release of cholesterol crystals. The characteristic presentation is with small limb-vessel occlusion with well-preserved peripheral pulses, which may occasionally lead to gangrene. Another common skin finding is livedo reticularis, where the cutaneous venous plexus becomes visible because of increased amounts of desaturated venous blood. Organ involvement from cholesterol emboli is dependent on the vascular supply, with renal ischaemia being the commonest. The diagnosis is often missed unless the occurrence of the clinical features is related to the triggering procedure. Fundoscopy can reveal the presence of retinal cholesterol crystals (Hollenhorst plaques) in about 25% of cases. Biopsy of the cutaneous lesion or the affected organ is required for definitive diagnosis. The prostacyclin analogue iloprost has been used successfully in painful cutaneous necrotic lesions and renal insufficiency in a single report.

Fat embolism syndrome

Fat embolization is characterized by release of fat droplets into the systemic circulation after a traumatic or iatrogenic event. Its incidence depends on the bone involved (fractures of the femoral shaft >tibia or fibula >neck of femur), isolated or multiple fractures, age (10–40 years) and gender (occurs more in males). Although more common after lower limb fractures, fat embolism syndrome can also occur after liposuction, bone marrow harvesting, total parenteral nutrition, sickle cell crisis and pancreatitis. Classically, it presents with the triad of respiratory distress, mental status changes and petechial rash 24–48 hours after pelvic or long-bone fracture. The rash is pathognomonic and is seen usually on the conjunctiva, oral mucous membranes and upper body, possibly due to embolization of fat droplets accumulating in the aortic arch. The diagnosis is made clinically because laboratory and radiographic diagnosis is non-specific and inconsistent. Thrombocytopenia is common, due to platelet activation and consumption into the thrombi. Management is supportive to ensure haemodynamic stability. Aspirin and corticosteroids have also been shown to be helpful, although the use of heparin is controversial.

Warfarin-induced skin necrosis

Warfarin-induced skin necrosis is a rare condition that is due to microvascular thrombi provoked by a transient imbalance between procoagulant and anticoagulant factors. It most commonly affects the breasts, buttocks and thighs. It may occur on initiation of warfarin, especially in patients with heterozygous protein C or protein S deficiency because of the relatively short half-life proteins compared with prothrombin. It is also seen in association with heparin-induced thrombocytopenia. It is prevented by bridging the initiation of warfarin with heparin and avoiding high loading doses. Treatment requires discontinuation of warfarin and starting therapeutic heparin. Intravenous infusions of protein C concentrate may also be used in the short term.

Haemostatic dysfunction associated with paraproteinaemia and amyloidosis

Paraproteinaemia

Bleeding and thrombotic complications both occur in association with paraproteinaemias, although abnormalities in laboratory tests are found much more frequently than clinical effects. Bleeding is more common in Waldenström macroglobulinaemia and amyloidosis than multiple myeloma. Among myeloma patients, bleeding is most common in those with IgA paraproteinaemia. The circulating paraprotein in these conditions can: (i) cause hyperviscosity and lead to arterial and retinal bleeds due to abnormal wall shear stress; (ii) inhibit or increase

clearance of FVIII and VWF leading to acquired von Willebrand disease; (iii) impair platelet aggregation; (iv) inhibit fibrin polymerization and (v) have heparin-like anticoagulant function that can be reversed by protamine. Management of these conditions is directed towards the underlying cause. Plasma exchange is of help as a temporary measure.

Amyloidosis

The pathophysiology of bleeding associated with systemic amyloidosis is multifactorial. The type of amyloidosis and the pattern of organ involvement are important determinants of the haemorrhagic tendency (Chapter 30). AL amyloidosis, where liver and spleen involvement is frequent, is the commonest type associated with bleeding. FX deficiency is seen most commonly; however, decreased levels of all factors, including VWF, have been reported. The coagulation factor deficiencies are thought to be due to adsorption onto amyloid fibrils. Abnormal fibrin polymerization and hyperfibrinolysis can also contribute to bleeding. Prolongation of the TT is the most common abnormality, seen in up to 90% of patients. A prolonged PT and APTT suggest FX deficiency. Abnormal coagulation screening tests correct with normal plasma.

Platelet dysfunction may be seen, due to the binding of the amyloid light chains to the platelet membrane. Deposition of amyloid fibrils in the blood vessel wall and perivascular tissue may lead to impaired vasoconstriction and vessel fragility. This is exemplified by cerebral amyloid angiopathy, which can lead to intracerebral haemorrhage, especially in elderly non-hypertensive individuals. The microvascular involvement also explains the 'raccoon eyes' (bilateral periorbital purpura from coughing or prolonged inverted positioning for lower gastrointestinal procedures) and the 'pinch purpura' (skin pinching leading to purpura). Several treatments have been described, including factor replacement with FFP or PCCs, platelet transfusion, desmopressin and rFVIIa. Severe FX deficiency can be difficult to manage and first-line therapy is with PCCs or FFP.

Acquired inhibitors of coagulation factors

Inhibitory autoantibodies to all coagulation factors have been described, although those against FVIII and VWF are most common.

Acquired haemophilia A

Acquired haemophilia A has an annual incidence of about 1.5 per million and usually affects older patients with a median age of about 75 years. It affects males and females equally, except in younger patients where there is a female preponderance associated with pregnancy. Acquired haemophilia A presents with a typical bleeding pattern that is distinct from that seen in congenital haemophilia. Widespread subcutaneous bleeding is common, as are muscle bleeds and gastrointestinal and genitourinary bleeding. Neurovascular compression may be limb-threatening. Haemarthroses are relatively uncommon. Fatal bleeding, such as intracranial, pulmonary, gastrointestinal and retroperitoneal, occurs in about 3–8% of cases (Table 40.8). Patients remain at risk of severe bleeding until the inhibitor has been eradicated, even if they present with mild bleeding. Acquired haemophilia A is associated with an underlying autoimmune (systemic lupus erythematosus or rheumatoid arthritis), malignant, lymphoproliferative or dermatological (pemphigoid) disease or pregnancy in about half of cases (Table 40.9). If acquired haemophilia A presents in association with pregnancy, it may recur in subsequent pregnancies and, if antenatal, the fetus may be affected.

Early diagnosis and urgent treatment of bleeding are key to successful management. The diagnosis is suggested by the clinical presentation and an isolated prolonged APTT. FVIII inhibitors are time and temperature dependent and in mixing studies normal plasma must be incubated for 1–2 hours otherwise the diagnosis may be missed. The diagnosis is confirmed

Table 40.8 Bleeding symptoms in patients with acquired haemophilia A*.

Site of bleeding	All bleeds at presentation (%)	Bleeds requiring treatment (%)
	N = 172	*N* = 65
Subcutaneous/skin	81	23
Muscle	45	32
Subcutaneous only	24	Not applicable
Gastrointestinal/ intra-abdominal	23	14
Genitourinary	9	18
Retroperitoneal/ thoracic	9	5
Other	9	23
Postoperative	0	11
Joint	7	2
None	4	Not applicable
Intracranial haemorrhage	3	0
Fatal	9	No data

*Patients had often had multiple bleeds.
Source: Data from Collins *et al*. (2007) and [Morrison *et al., Blood* **81**,1513–20, 1993].

Table 40.9 Diseases associated with acquired haemophilia A.

	Green and Lechner (1981)	Morrison et al. (1993)	Collins et al. (2007)
Number of patients	215	65	150
Idiopathic (%)	46	55	63
Collagen, vascular and other autoimmune diseases (%)	18	17	17
Malignancy (%)	7	12	15
Skin diseases (%)	5	2	3
Possible drug reaction (%)	6	3	0
Pregnancy (%)	7	11	2

Source: Data from [Green and Lechner, *Thrombosis and Haemostasis* **45**: 200–203, 1981], [Morrison *et al.*, *Blood* **81**,1513–20, 1993] and Collins *et al.* (2007).

by a low FVIII and positive Bethesda titre (Figure 40.6). In some cases diagnosis is complicated because all the intrinsic factors may be low due to inhibition of FVIII in the factor-deficient plasma used to assay other intrinsic factors. Dilution will result in increased levels of the non-specifically reduced factors while FVIII remains low. It is important to exclude a lupus anticoagulant because this can also be associated with an apparently low FVIII and an APTT that does not correct with normal plasma. Acquired FVIII inhibitors have complex kinetics, so that not all FVIII is inhibited in a Bethesda assay. It is therefore often not possible to measure the inhibitor titre accurately. The FVIII level and inhibitor titre are not predictive of the severity of the bleeding.

Treatment involves minimizing invasive procedures, protecting the patient from trauma, treatment of bleeds and eradication of the inhibitor. At present, the available therapies to treat bleeds are with bypassing agents, namely rFVIIa or the activated PCC FEIBA (factor eight inhibitor bypassing activity) and are equally efficacious. Multiple doses may be required to control bleeds and prevent recurrence. There is a risk of thrombosis associated with bypassing agents, particularly in

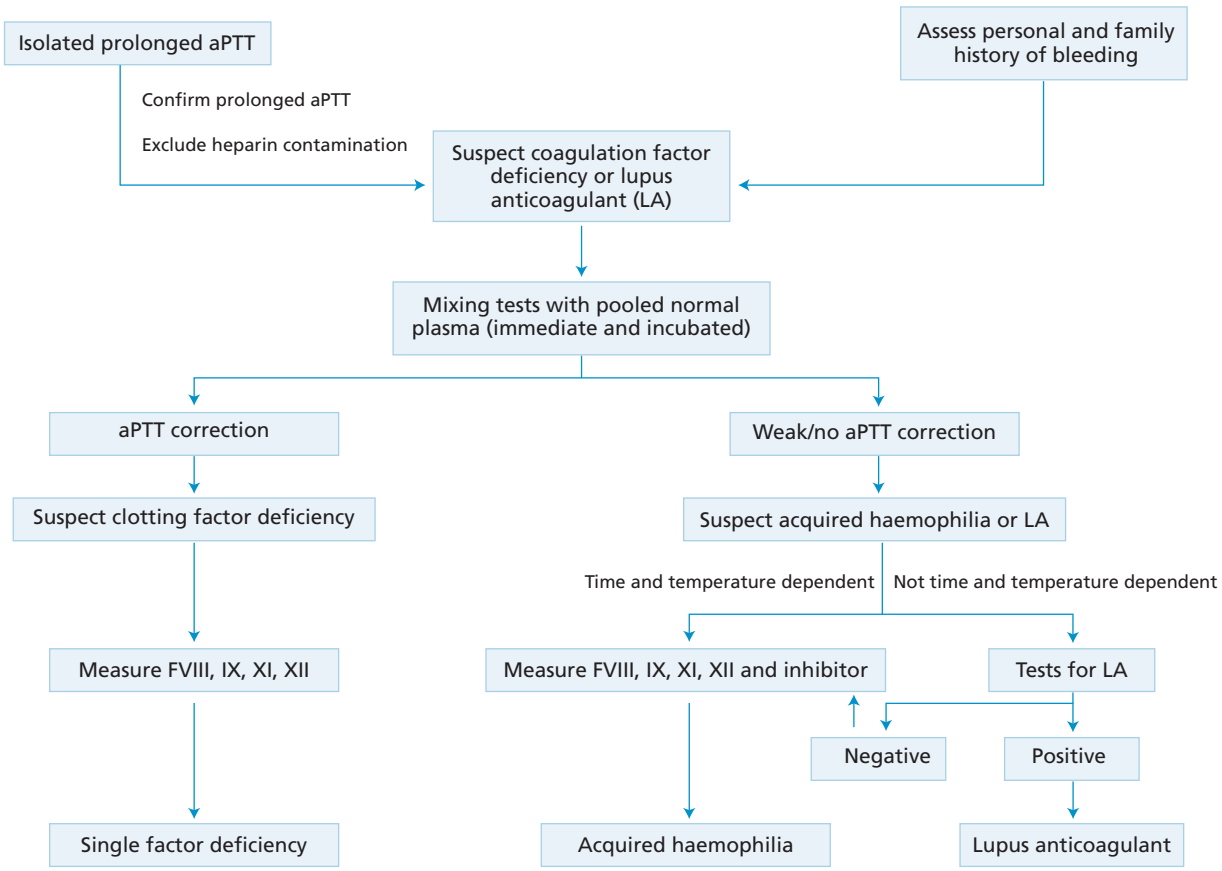


Figure 40.6 Diagnostic algorithm for an isolated prolonged APTT and possible acquired haemophilia A.

elderly patients. FVIII is unlikely to be efficacious in acquired haemophilia A, although mild bleeds in patients with low-titre inhibitors and measurable FVIII may respond to desmopressin. Initial trials with recombinant B-domain-deleted porcine factor have been encouraging and this agent is likely to be available soon. If standard haemostatic treatment with bypassing agents fails, FVIII and immunoadsorption may be used, although this treatment modality is only available in a few specialized centres.

Immunosuppression should be started as soon as the diagnosis is made and is usually with steroids alone or steroids plus a cytotoxic agent such as cyclophosphamide or azathioprine. A response is usually seen after 2–3 weeks, but full remission takes a median of about 5 weeks. If first-line therapy fails, rituximab, ciclosporin or combinations of cytotoxic agents may be used. Intravenous immunoglobulin does not increase the response rate to other immunosuppressive agents.

Acquired von Willebrand syndrome

Acquired von Willebrand syndrome may be associated with an autoantibody, typically in the context of monoclonal gammopathy of undetermined significance, Waldenström macroglobulinaemia and other lymphoproliferative diseases, myeloproliferative disease and systemic lupus erythematosus. The antibody leads to either rapid clearance or functional inhibition of VWF. Low VWF:Ag, VWF:RCo and VWF:CB are found, and measurement of the VWF propeptide may be useful in demonstrating rapid clearance. Mixing studies may not demonstrate an inhibitor.

Low VWF levels are also seen in hypothyroidism and increased VWF is associated with hyperthyroidism. Increased proteolysis of VWF in high-shear environments such as leaking cardiac values leads to a syndrome of acquired type 2A VWS. VWF has also been reported in association with malignancy and it is suggested that it may be adsorbed onto malignant cells. Hydroxyethyl starches used as plasma expanders have been associated with bleeding secondary to VWF deficiency. Treatment options include desmopressin, VWF concentrates and, in patients with an inhibitory antibody, high-dose intravenous immunoglobulin.

Acquired factor V deficiency

Acquired inhibitors to FV may arise spontaneously but are usually associated with exposure to topical thrombin preparations that have trace amount of bovine FV. Laboratory findings are of prolonged PT and APTT that do not correct with normal plasma. Patients may respond to FFP and, in resistant cases, platelets may be useful because surface-associated FV appears to bypass the inhibitor.

Acquired protein S deficiency

Autoantibodies to protein S have been reported in association with infection, particularly chickenpox. Patients present with skin necrosis and DIC.

Prothrombin deficiency associated with lupus anticoagulant

Although lupus anticoagulants are typically associated with thrombosis, occasionally cross-reactivity with prothrombin can lead to a bleeding disorder. Acquired deficiencies of other procoagulant and antifibrinolytic proteins have occasionally been described. Bleeding manifestations are variable and some patients with markedly abnormal laboratory tests do not bleed.

Selected bibliography

Collins PW, Hirsch S, Baglin TP *et al*. (2007) Acquired hemophilia A in the United Kingdom: a 2-year national surveillance study by the United Kingdom Haemophilia Centre Doctors' Organisation. *Blood* **109**: 1870–7.

De Paepe A, Malfait F (2004) Bleeding and bruising in patients with Ehlers–Danlos syndrome and other collagen vascular disorders. *British Journal of Haematology* **127**: 491–500.

Eby C (2009) Pathogenesis and management of bleeding and thrombosis in plasma cell dyscrasias. *British Journal of Haematology* **145**: 151–63.

Falanga A, Rickles FR (2003) Pathogenesis and management of the bleeding diathesis in acute promyelocytic leukaemia. *Best Practice and Research in Clinical Haematology* **16**: 463–82.

Huth-Kuhne A, Baudo F, Collins P *et al*. (2009) International recommendations on the diagnosis and treatment of patients with acquired haemophilia A. *Haematologica* **94**: 566–75.

Levi M, Toh CH, Thachil J, Watson HG (2009) Diagnosis and management of disseminated intravascular coagulation. *British Journal of Haematology* **145**: 24–33.

Levy JH, Dutton RP, Hemphill JC 3rd, Shander A *et al*. (2010) Multidisciplinary approach to the challenge of hemostasis. *Anesthesia and Analgesia* **110**: 354–64.

Lisman T, Leebeek FW, de Groot PG (2002) Haemostatic abnormalities in patients with liver disease. *Journal of Hepatology* **37**: 280–7.

Roberts I, Shakur H, Coats T *et al*. (2013) The CRASH-2 trial: a randomised controlled trial and economic evaluation of the effects of tranexamic acid on death, vascular occlusive events and transfusion requirement in bleeding trauma patients. *Health Technology Assess* **17**(10): 1–79.

Wada H, Thachil J, Di Nisio M *et al*. (2013) Guidance for diagnosis and treatment of DIC from harmonization of the recommendations from three guidelines. *Journal of Thrombosis and Haemostsis* **11**(10): 761–7.

Zangari M, Elice F, Fink L, Tricot G (2007) Hemostatic dysfunction in paraproteinemias and amyloidosis. *Seminars in Thrombosis and Hemostasis* **33**: 339–49.

CHAPTER 41

Congenital platelet disorders

41

Maurizio Margaglione[1] and Paul RJ Ames[2]

[1]Medical Genetics, Department of Clinical and Experimental Medicine, University of Foggia, Italy
[2]Department of Haematology, Haemostasis and Thrombosis, St George's Hospital, London, UK

Introduction

A normal haemostatic system is characterized by a sequence of local events that culminates in spontaneous arrest of bleeding from a traumatized blood vessel followed by sealing of any vascular discontinuity. Three closely linked biological systems are involved: platelets, blood vessels, and coagulation proteins. Platelets are anucleate cells that derive from cytoplasmic fragmentation of bone marrow megakaryocytes with a life span in the bloodstream of about 7–10 days. Platelets have a complex ultrastructure comprising a multitude of molecules and the malfunction of any of these may promote a specific disease (Figure 41.1). In the resting state platelets normally do not interact with endothelial cells or other blood cells, but do so when the vessel wall is disrupted. Adhesion to exposed subendothelial components allows platelet activation and release of active substances from intracellular organelles such as adenosine diphosphate (ADP) that amplify and propagate platelet activation and aggregation into a plug that fills in any voids in a ruptured blood vessel. At this primary haemostasis stage platelets transiently stop bleeding from damaged blood vessels and provide the surface onto which plasma coagulates into a fibrin mesh that strengthens the platelet plug, the so-called secondary haemostasis stage.

Clinically, abnormalities of platelet adhesion and aggregation present with a history of spontaneous and/or easy bruising and a prolonged skin bleeding time. Abnormal bleeding due to deficiencies of platelets shows characteristic features distinct from those seen in disorders of plasma coagulation factors (Table 41.1). The occurrence of skin and mucosal bleeding in patients with platelet disorders as opposed to deeper bleeding in patients with clotting factor disorders is a useful clinical point, but it must be remembered that the activities of platelets and coagulation factors occur almost simultaneously. Bleeding associated with platelet abnormalities manifests as haemorrhages from small vessels. Petechiae usually develop on the skin and the visible mucous membranes, but they may be distributed throughout the body, including internal organs. Characteristically, bleeding resulting from platelet diseases is immediate and transient, tends to stop with local pressure and does not recur when the pressure is removed. When this pattern of bleeding occurs in the neonatal period, in infancy and in childhood, a congenital platelet disorder should be suspected. However, the disease may be clinically silent and the patient may enter adult life before bleeding occurs.

The family history may be of great importance, providing a characteristic pattern of inheritance, but it should be remembered that a negative family history does not exclude an inherited platelet abnormality, i.e. the family history is usually negative in autosomal recessive traits. A comprehensive medical history and a careful clinical examination of the patient presenting with a haemorrhagic disorder are crucial for the correct diagnosis, and will dictate the subsequent choice of laboratory tests that may lead to a better identification, of the platelet abnormality. Congenital defects of platelets may give rise to bleeding syndromes of varying severities and are difficult to classify because of the rarity of many forms, the extreme heterogeneity and incomplete knowledge about a variety of diseases. Table 41.2 classifies congenital platelet disorders into two main groups: thrombocytopenias and thrombocytopathies. Each group is further divided according to specific criteria based on functional

Postgraduate Haematology, Seventh Edition. Edited by A Victor Hoffbrand, Douglas R Higgs, David M Keeling and Atul B Mehta.

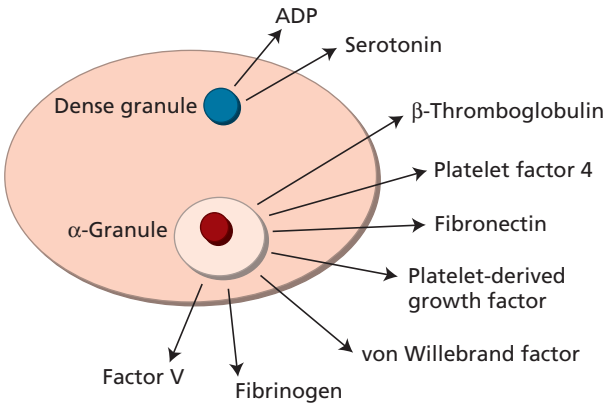


Figure 41.1 In platelets, granules are the storage sites for substances that are important for the haemostatic process. The α-granules contain proteins involved in adhesion (fibronectin, VWF), cell–cell interaction (P-selectin) and in promoting coagulation (FV, PF4), whereas the content of dense granules is important for recruiting additional platelets (ADP, serotonin).

and biochemical defects. Such a classification, as others, should be viewed as temptative, because some disorders may be characterized by multiple pathogenetic factors, whereas some others are grouped together for convenience of classification rather than on their known physiopathology.

Thrombocytopenias

Thrombocytopenia defines a subnormal number of platelets in the circulating blood, usually below 100×10^9/L. Acute thrombocytopenia is the most frequent cause of severe bleeding and the risk of haemorrhage is inversely proportional to the platelet count, with spontaneous bleeding occurring frequently at a platelet count below 20×10^9/L. When automated methods are used, pseudothrombocytopenia can be observed. Different mechanisms can cause a false low platelet reading. This may happen in patients with a wide variety of clinical disorders. Non-technical factors inducing pseudothrombocytopenia include paraproteinaemias, cold agglutinins, giant platelets, previous contact of platelets with foreign surfaces (i.e. dialysis membrane), lipaemia and EDTA-induced platelet clumping. The possibility of a pseudothrombocytopenia must be ruled out through manual counting and/or examination of an adequately stained blood film before concluding that a patient has true thrombocytopenia.

Table 41.1 Main specific clinical differences between diseases of coagulation factors and platelet disorders.

	Disorders of	
Findings	**Coagulation**	**Platelets/vessels**
Onset of bleeding	Delayed after trauma	Spontaneous or immediately after trauma
Mucosal bleeding	Rare	Common
Petechiae	Rare	Characteristic
Deep haematomas	Characteristic	Rare
Ecchymoses	Large and solitary	Small and multiple
Haemarthrosis	Characteristic	Rare
Bleeding from superficial cuts and scratches	Minimal	Persistent; often profuse
Sex of patient	80–90% male	Equal

Table 41.2 Classification of congenital platelet disorders.

a	Thrombocytopenias	
	1 Non-inherited	• Drugs and chemical agents • Isoimmune thrombocyptopenia • Infiltration of bone marrow • Infections • Other causes
	2 Inherited	• Reduced platelet size • Normal platelet size • Increased platelet size
b	Thrombocytopathies	
	1 Disorders of platelet adhesion function	
	2 Disorders of platelet signalling transduction functions	
	3 Disorders of platelet aggregation functions	

Non-inherited congenital thrombocytopenia

Congenital thrombocytopenia not ascribed to inherited causes may be the result of different pathogenetic mechanisms: most likely it is due to deficient marrow platelet production or enhanced platelet destruction.

Drugs and chemical agents

The maternal use of drugs and chemical agents during pregnancy may cause thrombocytopenia by suppressing platelet production, by damaging platelets directly, or by inducing the formation of platelet antibodies. Maternal ingestion of ethanol, thiazides, chlorpropamide, tolbutamide, oestrogens, and other drugs may selectively suppress thrombopoiesis. Alkylating agents, antimetabolites or chemotherapeutic drugs capable of crossing the placenta may lead to severe thrombocytopenia in newborns as the result of a predictable suppression of the bone marrow. The maternal ingestion of quinine, quinidine,

hydralazine, or selected antibiotics known to induce immune-mediated thrombocytopenia may induce congenital immune-mediated thrombocytopenia. In these newborns, thrombocytopenia usually improves rapidly and very rarely gives rise to severe or fatal haemorrhages.

Alloimmune thrombocytopenia

Immune thrombocytopenia may also result from the placental transfer of platelet antibodies formed as the result of active immunization of the mother by fetal platelet isoantigens if the fetus has inherited a paternal platelet-specific antigen that induces antibody formation (see also Chapter 50). Transplacental passage of the maternal antibody, usually of the IgG isotype, may induce severe thrombocytopenia in the fetus. This usually occurs when the mother has PlA1-negative platelets and the fetus carries PIA1-positive platelets. Newborns may show petechiae and purpura, or more severe bleeding at the time of birth, or may appear normal at the delivery and then manifest severe bleeding within the first postnatal week. In this case, thrombocytopenia usually improves within one month, but severe and fatal intracranial haemorrhages can occur.

Bone marrow infiltration

Congenital thrombocytopenia caused by bone marrow infiltration is extremely rare, being limited to cases of disseminated reticuloendotheliosis and congenital leukaemia. Thrombocytopenia with or without associated myeloid and erythroid depression occurs in children with numerous infiltrative disorders, including solid tumours, myelofibrosis, Gaucher's disease, Niemann–Pick disease and the mucopolysaccharidoses.

Infection

Thrombocytopenia, usually mild, but sometimes very severe, is commonly seen in infected newborns: several mechanisms are probably responsible, impaired platelet production as a result of invasion of megakaryocytes, the destruction of circulating platelets, and the formation of antigen–antibody complexes may explain many instances of thrombocytopenia associated with viral infections. Maternal infection with toxoplasma, cytomegalovirus, rubella, herpes viruses or hepatitis varicella, as well as recent maternal vaccinations (rubella), may induce congenital thrombocytopenia.

Other causes

Children born to women with chronic idiopathic thrombocytopenic purpura may have congenital thrombocytopenia due to autoantibodies in the mother crossing the placenta and binding to fetal platelets. Other causes of congenital thrombocytopenia due to increased platelet consumption or destruction include common maternal disorders (pre-eclampsia, systemic lupus erythematosus or other autoimmune diseases, especially if the woman has antiphospholipid antibodies). Finally, a moderate to severe thrombocytopenia has been observed, associated in giant cavernous haemangioma, first described by Kasabath and Merritt, in which the consumption of platelets occurs primarily within the tumour. Thrombocytopenia is often found associated with coagulation abnormalities typical of disseminated intravascular coagulation and the severity tends to parallel the size of the vascular tumour.

Inherited thrombocytopenias

Several inherited diseases may manifest with thrombocytopenia, as an accompanying feature of a generalized bone marrow failure (Fanconi's anaemia) or metabolic diseases causing marrow infiltration (i.e. Gaucher's disease). These and other complex clinical syndromes (i.e. Noonan's syndrome, CATCH 22 syndrome etc.) will not be discussed in this chapter. Inherited thrombocytopenias are very rare. In some of these a thrombocytopathy may coexist with thrombocytopenia (i.e. Bernard–Soulier syndrome, Wiskott–Aldrich, etc.). These diseases will be discussed in the appropriate section according to the most prominent defect. Helpful criteria to classify inherited thrombocytopenias are platelet size and the presence (syndromic) or the absence (non-syndromic) of other clinical features (Table 41.3)

Inherited thrombocytopenias with reduced platelet size

The Wiskott–Aldrich syndrome (WAS) is a rare (1/250,000) X-linked recessive disorder characterized by eczema, susceptibility to infections associated with defects in cellular and humoral immunity, and thrombocytopenia with reduced platelet size (see also Chapter 14). The gene responsible for WAS maps to Xp11.1 and codes for a protein, WASp, expressed only in haemopoietic-derived cells and involved in the transduction of the signals from the receptor to the actin cytoskeleton. Intermittent bleedings, recurrent bacterial and viral infections, and progressive eczema occur during the first months of life. Death at an early age commonly results from intracranial haemorrhage, infection or lymphoreticular malignancy. A number of kindreds have been reported in whom X-linked thrombocytopenia (XLT) occurred alone or in association with partial manifestations of WAS. Also XLT is caused by a mutation within the WAS gene. Patients with XLT mainly suffer from an isolated bleeding tendency.

Inherited thrombocytopenias with normal platelet size

Congenital deficiency of megakaryocytes is a rare form of thrombocytopenic purpura in the newborn and may occur with skeletal, renal or cardiac malformations. Isolated

Table 41.3 Main clinical characteristics associated with syndromic thrombocytopenia.

System/organ	Clinical features	Disease
Skin	Eczema	WAS, Jacobsen/Paris–Trousseau
	Albinism	Hermansky–Pudlak
		Chediak–Higashi
Skeletal	Upper extremities	
	Bilateral absence of the radii with the presence of both thumbs	TAR
	Limited pronation and supination of the forearm	Amegakaryocytic thrombocytopenia with ulnar synostosis
	Facial abnormalities	Jacobsen/Paris–Trousseau
	Hand abnormalities	Jacobsen/Paris–Trousseau
	Lower extremities	
	Various anomalies	TAR
	Hip dysplasia	Amegakaryocytic thrombocytopenia with ulnar synostosis
Pulmonary	Fibrosis	Hermansky–Pudlak
Cardiac	Heart defects	Jacobsen/Paris–Trousseau
	Septal defects	TAR
Renal	Haematuria, proteinuria	MYH9-related disease
	Congenital malformations	TAR
Gastrointestinal	Milk-protein allergy	TAR
	Granulomatosis colitis	Hermansky–Pudlak
Auditory	High-frequency sensorineural hearing loss	MYH9-related disease
Eyes	Cataracts	MYH9-related disease
Neurologic	Mental retardation	Paris–Trousseau
	Ataxia, intellectual disability	Chediak–Higashi
Immunologic	Autoimmunity	WAS, Chediak–Higashi
	Immunodeficiency	WAS, Chediak–Higashi

congenital amegakaryocytic thrombocytopenia (CAMT) is an autosomal recessive syndrome leading to bone marrow aplasia later in childhood. This disorder is associated with abnormalities in the expression or function of the thrombopoietin receptor c-mpl, and a series of mutations in the c-mpl gene have been identified. Most commonly, associated skeletal anomalies are present. Bilateral agenesis of the radius is the most commonly associated abnormality (thrombocytopenia with absent radius (TAR) syndrome). The TAR syndrome is an autosomal recessive disease characterized by severe, even fatal, haemorrhagic manifestations. The molecular basis of the syndrome is associated with alterations of the RBM8A gene and often involves a microdeletion on chromosome 1q21.1. In some infants, the ulna and humerus may also be absent and other skeletal abnormalities may occur. In addition, a proximal radioulnar synostosis, syndactyly and other skeletal abnormalities are reported in association with an autosomal dominant congenital amegakaryocytic thrombocytopenia in patients carrying heterozygous mutations of the HOXA11 gene. Less commonly, patients manifest cardiac and other minor defects. An autosomal dominant, moderate form of thrombocytopenia with a usually mild or absent bleeding tendency has been associated with mutations in the ANKRD26 gene (ANKRD26-related thrombocytopenia) and with haematological malignancies, in particular acute myeloid leukaemias. The Schulman–Upshaw syndrome, which is caused by mutations in the ADAMTS13 gene, is characterized by thrombotic thrombocytopenic purpura with neonatal onset (congenital microangiopathic haemolytic anaemia), thrombocytopenia and frequent relapses, and response to fresh plasma infusion.

Inherited thrombocytopenias with increased platelet size

Thrombocytopenias with increased platelet size are characterized by the presence of an increased platelet volume with a reduced number of platelets, megathrombocytopenia, and are the commonest inherited forms of thrombocytopenias. Among them, a series of diseases have been characterized at the molecular level, whereas other clinical entities still await a clear

identification of the molecular defect. Giant platelets and moderate thrombocytopenia are most frequently found in certain populations of Mediterranean extraction and also may be associated with other inherited or congenital syndromes, such as MYH9-related disease, Bernard–Soulier syndrome, stomatocytosis, etc.

MYH9-related disease is a rare autosomal dominant disease characterized by the presence in almost all patients of giant platelets with thrombocytopenia and basophilic inclusions (Döhle's bodies) within granulocytes. Large Döhle's bodies are seen in granulocytes from peripheral blood and bone marrow, and most patients show mild neutropenia without increased susceptibility to infection. About 50% of patients have significant thrombocytopenia deriving from an ineffective thrombopoiesis, but life-threatening haemorrhage is rare. The platelets of patients suffering from MYH9-related disease show not only a greatly increased volume, twice as normal, but most of them display bizarre morphology and hypergranularity. The condition includes a group of autosomal dominant diseases, in the past known as May–Hegglin anomaly and Sebastian, Fetchner and Epstein syndrome, according to the presence of hearing loss, glomerulonephritis and cataract. Then, because of the high variability of the clinical phenotype, also in members belonging to the same family, these eponyms have been abandoned and the term MYH9-related disease has been proposed independently of the presence of Döhle-like bodies and non-haematologic features. The molecular basis of these syndromes has been elucidated. All these clinical entities are caused by mutations that occur within a gene (MYH9) located on the long arm of the chromosome 22 (22q12-13). This gene encodes for the heavy chain of non-muscle myosin IIA (NMMHC-IIA), a protein involved in the contractile activity of the cytoskeleton. On the whole, all diseases may be grouped as MYH9-related disease.

Gray platelet syndrome (GPS) is a rare disorder inherited as an autosomal recessive trait because of mutations within the NBEAL2 gene, although a dominant transmission has been shown, and characterized by large platelets with a selective deficiency in the number and content of α-granules. Owing to this, platelets are either markedly hypogranular or agranular and display a deficiency of α-granule proteins, such as fibrinogen, von Willebrand factor, thrombospondin, β-thromboglobulin and platelet factor 4. A missense mutation at position 759 in the GATA-1 gene, inducing an amino-acid change (Arg216Gln), which segregates in a X-linked GPS, has been recently identified, suggesting that the GATA-1 gene is required for platelet α-granule regulation. Thrombocytopenia is usually pronounced and severe bleeding may occur.

Mediterranean macrothrombocytopenia is an asymptomatic disorder with moderate isolated thrombocytopenia and large platelets inherited as an autosomal dominant trait. The condition is characterized by mild or no clinical manifestations and normal bone marrow megakaryocytosis, platelet survival and *in vitro* platelet functions. Platelets from some patients suffering from Mediterranean macrothrombocytopenia show a reduced expression of the GPIb-IX platelet receptor, and heterozygous mutations within the GPIbα or Ibβ genes have been described. Therefore, in these patients Mediterranean macrothrombocytopenia may be viewed as a heterozygous form of Bernard–Soulier syndrome (see below). The remaining patients having Mediterranean macrothrombocytopenia do not show a reduction of the content of the GPIb-IX platelet receptor and the pathogenesis of this form needs further clarification.

Type 2B von Willebrand disease is a thrombocytopenia resulting from enhanced platelet destruction. Patients suffering from this syndrome show a qualitative abnormality of plasma von Willebrand factor (VWF) such that VWF binds inappropriately to circulating platelets. This is due to mutations within exon 28 of the VWF gene that give rise to a VWF with increased affinity for the platelet receptor GPIb-IX. Clearance of the resulting VWF-platelet complexes leads to thrombocytopenia and the selective loss of the largest VWF multimers from plasma. In general, the degree of thrombocytopenia is moderate and bleeding is of variable severity. A mutation in the VWF gene has been identified in a kindred previously labelled as Montreal platelet syndrome, a very rare platelet disorder characterized by giant platelets, reduced platelet count and spontaneous *in vitro* platelet aggregation: this kindred has now been re-classified as type 2B von Willebrand disease. Type 2B von Willebrand disease should be distinguished from the rare pseudo-von-Willebrand disease (or platelet-type von Willebrand disease). The defect in this autosomal dominant condition results from gain of function mutations in the gene encoding for the GPIbα subunit of the platelet receptor, increasing the affinity for VWF resulting in spontaneous binding of platelets to VWF with shortened platelet survival and thrombocytopenia. The reasons why most patients suffering from platelet-type von Willebrand disease show variable enlarged platelets remain unclear.

Jacobsen and Paris–Troussau syndromes are other very uncommon causes of macrothrombocytopenia. The Jacobsen syndrome is a contiguous gene syndrome inherited as an autosomal dominant trait characterized by a chromosomal deletion encompassing the long arm of the chromosome 11 (11q23.3-11q24.2). Paris–Trousseau syndrome identifies patients who present with macrothrombocytopenia, abnormal platelet function and dysmegakaryocytes in the bone marrow. A heterozygous deletion of the FLI1 gene has been described in unrelated children with deletions of 11q23 and Paris–Trousseau thrombocytopenia. Interestingly, in patients suffering from Jacobsen syndrome without thrombocytopenia, the chromosomal deletion did not include the FLI1 gene, suggesting that deletion of this gene is specific for the thrombocytopenia that occurs in most patients with Jacobsen and Paris–Trousseau syndromes.

Thrombocytopathies

Platelets participate in haemostasis by adhering to exposed elements of the subendothelial matrix. They then spread onto the subendothelial surface, becoming activated, release the content of their storage organelles and aggregate to each other. Abnormalities in any of these stages, adhesion, activation, secretion and aggregation may give rise to congenital disorders of platelets. Patients suffering from any of these diseases usually show a bleeding diathesis with a prolonged bleeding time and a normal platelet count. The most accurate test for monitoring platelet activation is light transmission aggregometry, which uses platelet-rich plasma or washed platelets, and monitors the change in light transmission upon agonist addition and simultaneously dense granule secretion. However, diagnosing a defect in platelet function is often difficult, because of need of specific expertise, and an experienced investigator is required. A further major challenge is the wide overlap with the response of healthy subjects. In up to 40% of patients presenting with a platelet function defect light transmission aggregometry fails to correctly identify the thrombocytopathy.

Disorders of platelet adhesion

The interaction of platelet receptors with elements of the subendothelium, collagen, fibronectin and blood components allows platelet adhesion to the subendothelium, but it also occurs through the bridging effect of VWF (Figure 40.2). A series of receptors have been identified on the platelet surface that interact with one or more of the previous elements. The most important receptors are the GPIb-IX and the α2-β1 (previously known as GPIa-IIa). The α2-β1 is one of receptors on the platelet surface that binds collagen and is a member of the integrin β-1 subfamily. Different receptors for collagen are GPIV and GPVI. Binding of molecules to these receptors leads to subsequent binding to other receptors, which serves to reinforce adhesion and to generate intracellular signals, such as calcium mobilization and protein phosphorylation.

Bernard–Soulier Syndrome (BSS) is a bleeding disorder characterized by giant platelets seen on the blood smear, mild or moderate thrombocytopenia and prolongation of the skin bleeding time disproportionate to the thrombocytopenia. BSS is a recessively inherited autosomal disorder due to abnormalities of the GP Ib-IX-V receptor complex and consanguinity is common in reported kindreds. The frequency of BSS has been estimated to be approximately 1 case in 1 million people. Bleeding may be severe and fatal haemorrhages may occur. Cutaneous haemorrhages, muscular and visceral bleedings are common. Epistaxis and menorrhagia may be difficult to control. Haemarthrosis has also been reported. Platelet counts range from as low as 20×10^9/L to near normal and, on the peripheral blood film, over 80% of the platelets are usually larger than 2.5 μm, often up to 8.0 μm, in diameter. The number of bone marrow megakaryocytes is usually normal. Patients presenting with BSS show an absent platelet agglutination in response to ristocetin (in the presence of human VWF), normal aggregation, ATP secretion and thromboxane (TX) B_2 formation in response to a variety of aggregating agents, and delayed response to thrombin. Because of the defective binding with VWF, platelets in BSS have substantial reduction in their ability to adhere to sites of vascular injury where subendothelial VWF becomes exposed. As a consequence, plug formation, the primary haemostatic response, is impaired, and increased and prolonged bleeding occurs. Variable levels of the glycoprotein (GP) Ib-IX-V complex have been detected on the platelet surface of patients with BSS, some patients with BSS exhibiting near normal GP Ib-IX-V amounts (variant type of BSS). In spite of the difference in glycoprotein content, clinical bleeding problems, and platelet functional and morphologic abnormalities, these patients were indistinguishable from the classical BBS phenotype. The entire cDNA sequences encoding the protein chains comprising the GP Ib-IX-V receptor complex have been obtained, allowing studies on the molecular basis of the syndrome. This complex consists of four proteins: the disulfide-linked α- (135 kD) and β-chains (25 kD) of GPIb, and the non-covalently associated subunits GPIX (22 kD) and GPV (82 kD). They all share structural and functional features suggesting a common evolutionary origin. Different transcripts encode the four polypeptidic chains and, with the exception of that of GPIbβ, genes show continuous (intron-depleted) open reading frames. In addition, each element contains one or more homologous 24-amino-acid leucine-rich glycoprotein repeats. The genetic heterogeneity in the glycoprotein content of BSS patients shows that multiple molecular abnormalities may lead to a similar clinical disorder, and implies that BSS may be the result of defects within the subunits that hamper the coordinate expression of the complex on the platelet membrane. In this respect, BSS would resemble abnormalities of other multisubunit complexes, in which a defect in a single subunit prevents the assembly and the surface expression of the complex. In addition to quantitative and qualitative abnormalities of the GPIbα gene, the recognized BSS phenotype has also been documented in patients with detrimental mutations within platelet GPIX and GPIbβ genes.

A defect of one of the platelet receptors for collagen has been described in very few patients. These patients showed mild bleeding disorders and a selective impairment, at a variable extent, in collagen response, adhesion to subendothelial surfaces and collagen-induced platelet aggregation. Platelets have two major receptors for collagen, α2β1 integrin, with a major role in adhesion of platelets to subendothelial surfaces, and GPVI, mainly involved in platelet activation. Thus, it is conceivable that collagen binding to platelets occurs through a multistep mechanism, involving first the attachment of platelets to exposed collagen of the subendothelium in flowing blood by means of the

α2β1 receptor and then platelet activation through a second receptor, GPVI.

Disorders of platelet signalling transduction

Platelets that adhere to the subendothelial surface become activated and begin the production or release of several intracellular messengers that modulate a series of platelet responses, such as calcium mobilization, protein phosphorylation and production of arachidonic acid. Activated platelets also release substances stored in their granules, some of which act in the recruitment of additional platelets and lead to the formation of the primary haemostatic plug (see Figure 41.1). Several signalling mechanisms are involved in events that govern platelet responses, starting from platelet adhesion to injured blood vessels and leading to secretion and aggregation. Available evidence suggests specific abnormalities in platelet signalling mechanisms may underlie a platelet dysfunction. The term platelet signalling disorders defines a group of heterogeneous abnormalities in platelet secretion and signal transduction. Congenital defects of platelet signalling mechanisms are put together for convenience of classification rather than on the basis of knowledge in the pathophysiology of specific diseases. Patients suffering from these defects represent the vast majority of subjects presenting with inherited thrombocytopathies. Many platelet receptors are coupled to G-proteins. P2Y1 and receptors for thromboxane A2, and thrombin act via Gαq. Receptors for prostaglandins activate, whereas P2Y12 inhibits adenylate cyclase through Gs and Gi2, respectively. Platelet activation by means of binding to receptors gives rise to hydrolysis of phosphoinositide by phospholipase C, leading to the formation of inositol triphosphate, which, in turn, functions as a messenger to release calcium from intracellular stores. In addition, hydrolysis of phosphoinositide leads to the formation of diacylglycerol, which activates protein kinase C. The activation of the protein kinase C is thought to play a major role in platelet secretion and in the activation of the αIIb-β3 complex. Although defects in phospholipase C activation, calcium mobilization and protein phosphorylation have been suggested in several patients, few patients have been reported with deficiency in platelet Gαq, Gs hyperfunction and reduced expression of phospholipase C-β2. Following stimulation, platelet phospholipase A2 mobilizes the arachidonic acid from the phospholipid pool. Then, the arachidonic acid is metabolized by cyclo-oxygenase and thromboxane synthase to form thromboxane A2, a strong platelet-aggregating agent, which is necessary for a secretion response.

A defect in arachidonic acid mobilisation and thromboxane A2 production has been identified in some patients. Individuals with this defect showed an abnormal aggregation and secretion in response to a series of stimulating factors, but a normal production of thromboxane A2 in response to arachidonic acid. Several patients have been reported with a deficiency of cyclo-oxygenase and showing a slightly prolonged bleeding time and impaired platelet aggregation. In addition, a few patients show a defect in thromboxane A2 formation, presenting with a variable bleeding diathesis.

Disorders of platelet aggregation

Platelet aggregation may be defined as the interaction of activated platelets with one another and occurs after adhesion of platelets to injured blood vessel walls. A series of factors are capable of inducing platelet aggregation and may be classified in primary and secondary platelet aggregating agents. Primary aggregating substances are those factors, such as ADP, epineprhine and thrombin, able to directly induce platelet aggregation independently of their ability to release intraplatelet ADP or to induce the production of prostaglandins. Secondary aggregating factors are substances that induce aggregation of platelets through their ability to provoke the release reaction of ADP or the synthesis of prostaglandins. According to this, disorders due to an impairment of platelet aggregation may be classified into defects of primary aggregation, Glanzmann's thrombasthenia, selective impairment of platelet receptors (ADP, epinephrine), and defects of secondary aggregation, storage pool disease, selective impairment of platelet receptors (thromboxane, collagen). *In vitro*, a series of substances are employed to challenge platelets and the manner they respond to these stimuli may be helpful to identify specific thrombocytopathies (Table 41.4).

Glanzmann's thromboasthenia (GT) is a bleeding diathesis marked by prolonged bleeding time, normal platelet count and absence of platelet aggregation in response to platelet agonists ADP, collagen, arachidonic acid and thrombin. Platelet agglutination induced by ristocetin and VWF is normal. This congenital bleeding disorder is associated with an impaired or absent clot retraction. GT is one of the less common of the congenital bleeding disorders (prevalence 1/1,000,000) and is transmitted as an autosomal recessive trait with consanguinity reported in affected kindreds. The clinical features are those expected with platelet dysfunction: easy and spontaneous bruising, mucosal membrane bleeding, subcutaneous haematomas and petechiae. Rarely, patients suffer from intraarticular bleeding with resultant haemarthroses. Fatal haemorrhages have been reported. Quantitative or qualitative (variant GT forms) abnormalities of the platelet αIIbβ3 integrin (also known as the glycoprotein complex IIb-IIIa) have been shown to be responsible for this disorder. αIIb and β3 subunits are prominent integral components of the platelet membrane that form heterodimers containing specific sites for platelet-to-platelet cohesion. The αIIbβ3 integrin serves as platelet receptor for fibrinogen, fibronectin, vitronectin and von Willebrand factor. In addition, the αIIbβ3 integrin modulates, to some extent, calcium influx, cytoplasmic alkalinization, tyrosine kinase phosphorylation and clot retraction. Clinical heterogeneity of GT has been stressed on the basis of platelet

Table 41.4 Platelet response to aggregating agents in different thrombocytopathies.

Disorder	ADP	Epinephrine	Collagen	Arachidonic acid	Ristocetin
Bernard–Soulier	Normal	Normal	Normal	Normal	Absent
Pseudo-vWD	Normal	Normal	Normal	Normal	Increased at low doses
ADP receptor defect	Impaired	Impaired	Impaired	Impaired	Present
Epinephrine receptor defect	Normal	Impaired	Normal	Normal	Present
Collagen receptor defect	Normal	Normal	Impaired	Normal	Present
Defect of signal transduction	Variable impairment	Variable impairment	Variable impairment	Variable impairment	Present
Glanzmann's thrombasthenia	Absent	Absent	Absent	Absent	Present
δ-SPD	Impaired	Impaired	Impaired	Variable	Present
Thromboxane receptor defect	Impaired	Impaired	Impaired	Impaired	Present

function testing or using crossed immunoelectrophoresis, Western blot, flow cytometry and fibrinogen binding. Type I Glanzmann's thrombasthenia is characterized by the lack of surface-detectable αIIbβ3 complex and a profound defect in platelet aggregation and clot retraction. At variance with the type I, platelets of patients suffering from type II Glanzmann's thrombasthenia have detectable, but markedly reduced, amounts of the αIIbβ3 receptor on the surface, usually from 10 up to 20% of normal values. Platelets show sufficient amounts of receptors to allow for micro-aggregate formation, although there is still a profound defect in the ability to form large aggregates. The clot retraction is only moderately impaired. In addition, a series of patients with a variant form have been described who presented near normal levels of the αIIbβ3 complex, which is dysfunctional in that platelets, when activated, can neither aggregate nor bind fibrinogen. An extreme variability in the clinical symptoms is present, even among patients with similar degrees of platelet abnormality and prolongation of the bleeding time. In general, no aggregation abnormalities are detected in heterozygotes, but a decreased amount of the $\alpha_{IIb}\beta_3$ integrin has been reported, platelet content being approximately one-half of the normal amount.

Defects of platelet ADP receptors have been characterized in a few patients and all suffered from a bleeding diathesis. On the platelet surface, different types of ADP receptors have been identified, two G-protein-coupled receptors, P2Y1 and P2Y12. The P2Y1 receptor is responsible for the shape change of platelets and transient platelet aggregation, giving rise to centralization of platelet granules and formation of filopodia, while the P2Y12 receptor is involved in the amplification of the response and in the stabilization of platelet aggregates through the full activation of the αIIbβ3 integrin. The P2Y12 receptor defect is inherited as an autosomal recessive trait and most patients so far identified were born from consanguineous parents. In these patients a blunted platelet aggregation in response to ADP, with a retained shape change, has been reported. Only one patient has been briefly reported with a defect of the P2Y1 receptor. This patient showed impaired platelet aggregation in response to ADP and other agonists.

The selective impairment of epinephrine receptors, α2 adrenergic receptors (α2AR), has been associated with bleeding. A number of individuals with an impaired aggregation response to epinephrine and a congenital defect of platelet α2AR receptors have been described, and some of them presented a history of easy bruising.

A selective impairment of the thromboxane A2 receptor has been described in several patients with a mild lifelong disorder, consisting of mucosal bleeding and easy bruising, and a defective platelet aggregation in response to several agents, but not thrombin. An autosomal dominant inheritance has been suggested in some cases.

Secondary aggregation disorders are more frequent than primary aggregation disorders and the most common in this category are the storage pool deficiency (SPD) syndromes. SPD syndromes may be classified in a system that takes into account the content of both dense and α-granules (see Figure 41.1). The term δ-SPD identifies patients who show low platelet content of dense granules only. Patients with deficiency of both types of granules are designed as αδ-SPD. Finally, patients presenting with reduced or absent platelet content of α-granules, but normal levels of dense granules (α-SPD) are designed as patients with Gray platelet syndrome. The disorder is heterogeneous and the term SPD includes a group of disorders having as their common feature a diminution in secretable substances stored in platelet granules. A storage pool disease is found as an associated defect in most of the patients carrying other rare syndromes, such as Wiskott–Aldrich syndrome and thrombocytopenia with absent radius syndrome. However, the majority of patients presenting with an SPD have no associated diseases and are otherwise normal. The clinical features of this type of secondary aggregation disorder are those expected with a platelet function defect and consist of easy and spontaneous bruising, mucocutaneous

haemorrhages, haematuria and epistaxis. Patients with δ-SPD or αδ-SPD usually have absent ADP and epinephrine-induced secondary aggregation waves, although the primary waves are present. Patients with δ-SPD may present a severe or partial granule deficiency. Collagen-induced aggregation is absent or markedly reduced, whereas ristocetin-induced agglutination is normal. In the absence of other congenital abnormalities or an associated α-granule deficiency, SPD is inherited as an autosomal dominant trait. The forms of dense-SPD coincident with other congenital abnormalities, such as TAR, are usually inherited as autosomal recessive traits, or X-linked, as in the case of WAS. An SPD is usually present in Hermansky–Pudlak and Chédiak–Higashi syndromes.

Hermansky–Pudlak syndrome (HPS) is a rare autosomal recessive inherited disorder characterized by the presence of oculocutaneous albinism, the absence of platelet dense granules and infiltration of ceroid-pigmented reticuloendothelial cells in the lung and the colon, leading to pulmonary fibrosis and inflammatory bowel disease. HPS results from the abnormal formation of intracellular vesicles. The impaired function of specific organelles indicates that the causative genes encode proteins operative in the formation of lysosomes and vesicles. HPS shows a genetic heterogeneity and so far, mutations in eight orthologous genes have been found in HPS patients: *HPS1, AP3B1* (*HPS2*), *HPS3, HPS4, HPS5, HPS6, DTNBP1 (HPS7),* and *BLOC1S3 (HPS8)*.

The Chédiak-Higashi syndrome (CHS) presents with a variable degree of oculocutaneous albinism and a poor resistance to respiratory and cutaneous infections. The infections are generally fatal during infancy or in early childhood, but patients may also die of a chronic lymphohistiocytic infiltration known as the accelerated phase during the second or third decades of life. CHS is extremely rare and is inherited as an autosomal recessive trait. Unlike HPS, CHS does not seem to display locus heterogeneity and the only gene proved to cause CHS, LYST, is located on chromosome 1q42.1-42.2. Findings typical of an SPD are present in both HPS and CHS: bleeding of mucosal, epistaxis and spontaneous soft tissue bruising.

Quebec platelet disorder is a rare autosomal dominant bleeding disease associated with abnormal proteolysis of α-granule proteins; platelets show increased stores of urokinase plasminogen activator and genetic studies show a tandem duplication of the urokinase plasminogen activator gene, PLAU. Intraplatelet generation of plasmin is thought to trigger degradation of stored α-granule proteins, increasing risks for a number of bleeding symptoms, including delayed-onset bleeding after haemostatic challenges that responds only to fibrinolytic inhibitor therapy.

Treatment

As a rule, there is no specific treatment for the vast majority of congenital platelet disorders. In congenital thrombocytopenias due to maternal use of drugs or chemical agents, thrombocytopenia recovers after few days to few weeks, as in the case of thiazide drugs. In surviving thrombocytopenic newborns infected with rubella or cytomegalovirus, platelet levels may return to normal after several months. In newborns with alloimmune congenital thrombocytopenia, platelets return to normal values in 14 to 21 days. Only severe cases need to be treated with washed platelets, corticosteroids or exchange transfusions. In general, principal treatments in patients with congenital platelet diseases are general measures aiming at avoiding bleeding and the use of supportive therapeutic approaches for controlling haemorrhages. However, since types and severity of bleeding vary in different patients, therapeutic approaches have to be individualized.

General measures

Education of patients is of great importance. Patients and their parents have to be instructed to avoid trauma. Regular dental care may be helpful to prevent gingival bleeding. Drugs that impair platelet functions, such as acetylsalicic-acid-containing medications, should be avoided. In women, the use of oral contraceptives can be applied to prevent menorrhagia. Local measures, such as the application of firm pressure in the case of epistaxis, will usually suffice in the event of mild bleeding.

Drugs

Antifibrinolytic agents are useful as an adjunctive measure to prevent and control mild or moderate bleeding and may stop menorrhagia and other mild bleeding manifestations from mucous membranes, such as epistaxis, e.g. tranexamic acid, either orally at a dose of 15–25 mg/kg t.d.s. or intravenously at a dose of 10 mg/kg t.d.s. in the case of more serious bleeding.

Desmopressin (1-desamino-8-D-arginine vasopressin or DDAVP) is a mainstay of the therapy of patients with congenital platelet defects. Depending on platelet defect, administration of desmopressin may shorten the bleeding time. Intravenous, subcutaneous or intranasal administration of desmopressin increases factor VIII and VWF transiently by releasing them from storage sites and its use has been suggested to be of value in some patients presenting with congenital platelet defects, such as BSS, MYH9-related disease, GPS, SPD and Glanzmann's thrombasthenia. The drug is usually administered at a dosage of 0.3 μg/kg in 50 mL of saline by slow intravenous infusion (over 30 min to avoid possible hypotensive effects) or subcutaneously. In general, since the response to desmopressin varies among patients, but is constant in each patient, normalization or improvement of the closure times of the PF100 or of bleeding time pre and post a test dose of desmopressin may identify those patients who may benefit from the drug.

Recombinant activated factor VII (rFVIIa) has proven effective in the treatment of bleeding in haemophilic patients with

Table 41.5 Genes involved in congenital platelet disorders.

Disorder	Gene	Locus	Inheritance
Thrombocytopenias and small platelets			
Wiskott–Aldrich syndrome	WAS	Xp11.23-p11.22	XL
X-linked thrombocytopenia	WAS	Xp11.23-p11.22	XL
Thrombocytopenias and normal-sized platelets			
Congenital amegakaryocytic thrombocytopenia	c-MLP	1p34	AR
Congenital amegakaryocytic thrombocytopenia and radioulnar synostosis	HOXA-11	7p15-p14.2	AD
Thrombocytopenia with absent radii	RBM8A	1q21.1	AR
Schulman–Upshaw syndrome	ADAMTS13	9q34	AR
ANKRD26-related thrombocytopenia	ANKRD26	10p2	AD
Thrombocytopenias and large platelets			
May–Hegglin anomaly and Sebastian, Epstein Fetchner syndromes	MYH9	22q11.2	AD
Mediterranean macrothrombocytopenia	GPIbα	17p13	AD
Gray platelet syndrome	NBEAL2	3p21.1	AR
Jacobsen and Paris–Troussau syndromes	FLI1 (Large deletion)	11q23-ter	AD
Diseases mainly affecting platelet functions			
Bernard–Soulier syndrome	GPIbα	17p13	AR
	GPIbβ	22q11.2	AR
	GPIX	3q21	AR
Pseudo-vWD	GPIbα	17p13	AD
ADP receptor defect	P2Y12	3q24-q25	AR
	P2Y1	3q25	AR
	P2X1	17p13.3	AR
Epinephrine receptor defect	α2AR	10q24-q26	AR
Collagen receptor defect	α2	5q23-q31	AR
Glanzmann's thrombasthenia	αIIb	17q21.32	AR
	β3	17q21.32	AR
Thromboxane receptor defect	TXBA2	19p13.3	AR
Hermansky–Pudlak	HPS1	10q23.1-q23.2	AR
	AP3B1 (HPS-2)	5q14.1	
	HPS3	3q24	
	HPS4	22q11.2-q12.2	
	HPS5	11p15-p13	
	HPS6	10q24.32	
	DTNBP1 (HPS7)	6p22.3	
	BLOC1S3 (HPS8)	19q13	
Chediak–Higashi	CHS1/LYST	1q42.1-q42.2	AR
Quebec platelet disorder	PLAU duplication	10q24	AD
Thromboxane receptor defect	TXBA2	19p13.3	AR

inhibitors. The administration of rFVIIa in patients with congenital platelet disorders, such as BSS, SPD and Glanzmann's thrombasthenia may stop spontaneous bleeding and/or may prevent bleeding during and after surgical procedures, though arterial and venous thromboembolism have been described after its use. Recombinant FVIIa is currently approved in Europe for prophylaxis and treatment of bleeding in patients with Glanzmann's thrombasthenia with antibodies to GP IIb-IIIa and/or HLA, and past or present refractoriness to platelet transfusion. The recommended dose is 90 μg/kg as a minimum of three intravenous boluses at 2-hour intervals.

A second generation of synthetic thrombopoietin agonists, including romiplostim and eltrombopag, have shown great potential for their use in the treatment of forms of thrombocytopenia, including immune thrombocytopenia purpura and chemotherapy-induced thrombocytopenia. It is conceivable that

thrombopoietin agonists would be helpful in congenital thrombocytopenia. Safety and efficacy data on eltrombopag was recently reported also for the management of a major surgical procedure in patients with MYH9-related disease.

Platelet transfusions

Platelet transfusions are employed to control severe haemorrhage in patients with thrombocytopenia or thrombocytopathies. Spontaneous as well as post-traumatic bleeding is common at a platelet count below 20×10^9/L and is extremely likely at a platelet count below 5×10^9/L. Extensive clinical experience has shown that control of bleeding is possible upon achievement of an adequate increase in the circulating platelet count. However, the coexistence of platelet dysfunctions has to be taken into account to correctly calculate the dosage of platelets required. Platelet transfusions are effective in controlling bleeding, but may be responsible for transmission of infectious diseases, febrile reactions or development of alloimmunization. In patients with thrombocytopathies, platelet transfusions should be employed only for the treatment of severe bleeding because of the risk of alloimmunization. The occurrence of platelet alloimmunization is more frequent in patients lacking a membrane glycoprotein, such as in BSS and Glanzmann's thrombasthenia. Alloantibodies develop because they recognize the missing proteins in transfused platelets as foreign and, in turn, may induce refractoriness to platelet transfusions.

Other measures

Splenectomy has generally no effect in congenital thrombocytopathies and in thrombocytopenias but has proven to be effective in patients with inherited thrombocytopathies. Allogenic bone marrow transplantation may provide, in theory, an effective cure for inherited disorders involving platelet count or functions restoring a normal megakaryocytopoiesis. Indeed, it represents the only definitive therapy for patients with CAMT and remains the therapeutic modality of choice for classic WAS. Bone marrow transplants have been successfully performed with complete correction in patients with amegakaryocytic thrombocytopenia with radioulnar synostosis, BSS and severe Glanzmann's thrombasthenia. However, risks of such a drastic procedure may overcome those related to bleeding tendency and, therefore, are rarely required in patients suffering from congenital platelet disorders.

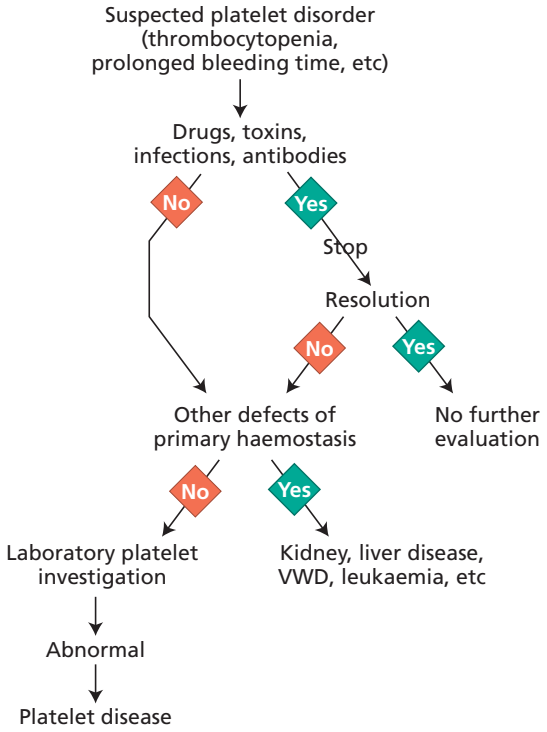


Figure 41.2 Schematic algorithm for the initial screening of patients with congenital platelet disorders.

Conclusions

Over the last few years, a series of studies have improved our understanding of the pathogenesis of congenital platelet disorders: several responsible genes have been identified and several patients have been characterized at the molecular level, allowing a more accurate comprehension of inherited thrombocytopenias and thrombocytopathies (Table 41.5). Careful evaluation of the personal and family history, physical examination and appropriate laboratory work-up are of major value for the evaluation of a patient presenting with bleeding due to congenital platelet disorders (Figure 41.2). Using this approach, in many instances it is possible to correctly identify the platelet defect. However, despite recent advances in knowledge, in most of the patients with a congenital bleeding disorder and impairment of platelet functions the underlying molecular mechanisms are still unknown. The future challenge will be to further our understanding of congenital platelet disorders in order to design specific tools for the diagnosis, prevention and treatment of these bleeding disorders. Gene therapy promises to be a successful cure for congenital platelet disorders. Hopefully, improvements in such approaches will allow treatment of the majority of patients suffering from severe congenital platelet disorders.

Selected bibliography

Alamelu J, Liesner R (2010) Modern management of severe platelet function disorders. *British Journal of Haematology* **149**: 813–23.

Balduini CL, Pecci A, Noris P. (2013) Diagnosis and management of inherited thrombocytopenias. *Seminars in Thrombosis and Hemostasis* **39**: 161–171.

Balduini CL, Savoia A (2012) Genetics of familial forms of thrombocytopenia. *Human Genetics* **131**: 1821–32.

Bolton-Maggs PHB, Chalmers EA, Collins PW *et al*. (2006) A review of inherited platelet disorders with guidelines for their management on behalf of the UKHCDO. *British Journal of Haematology* **135**: 603–33.

Dawood BB, Lowe GC, Lordkipanidze M *et al*. (2012) Evaluation of participants with suspected heritable platelet function disorders including recommendation and validation of a streamlined agonist panel. *Blood* **120**: 5041–9.

Masliah-Planchon J, Darnige L, Bellucci S (2013) Molecular determinants of platelet delta storage pool deficiencies: an update. *British Journal of Haematology* **160**: 5–11.

Sillers L, Van Slambrouck C, Lapping-Carr G. (2015) Neonatal thrombocytopenia: etiology and diagnosis. *Paediatric Annals* **44**: e175–180.

CHAPTER 42

Primary immune thrombocytopenia

42

Drew Provan and Adrian C Newland

Queen Mary University of London, London, UK

Introduction

Primary immune thrombocytopenia (ITP) is an autoimmune bleeding disorder that affects both children and adults. Until fairly recently, it was considered an autoantibody disorder in which platelets, opsonized with antiplatelet antibody, were removed prematurely by the reticuloendothelial system. More recently, studies have shown that in many patients there is a relative platelet under-production that contributes to the thrombocytopenia.

Many patients have no clinical problems, but bleeding may occur. ITP is unpredictable in its clinical course. To date, treatment has aimed at reducing the platelet destruction, mainly through immunosuppression. However, new thrombopoietin receptor agonists (TRAs) have been developed that enhance bone marrow platelet production. These recent targeted treatments are likely to be associated with less toxicity than older therapies.

Clinical features

The two principal forms of ITP, paediatric and adult, differ in their underlying cause and presentation. In children, ITP may follow a viral illness or immunization. The profound thrombocytopenia may be associated with extensive petechiae, purpura and bruises. There may also be bleeding from mucous membranes such as the nose or mouth. Despite the severity of the clinical features, most children need little treatment, and undergo spontaneous remission in the majority of cases. Around 15% will develop chronic ITP.

In adults there is generally no prodromal illness and the patient may be aware of petechiae or excessive bruising, and seek medical attention. ITP is often diagnosed by chance, for example during hospital admission for surgery or when a full blood count (FBC) is checked for other reasons.

ITP, particularly in adults, is heterogeneous. Many patients suffer few clinical problems related to their thrombocytopenia, while others have major bleeding from the outset. The platelet count alone appears to be an unreliable predictor of outcome, and the clinical symptoms and signs of ITP are influenced by patient age, general health, comorbidities, medication and many other factors. In addition, there may be an acquired platelet dysfunction caused by antibody binding to an important region of the glycoprotein molecules on the platelet surface. Autoantibodies reacting with glycoprotein (GP)IIb/IIIa may affect platelet aggregation, and anti-GPIb/IX autoantibodies can impair platelet adhesion to the subendothelial matrix, causing unexpectedly severe bleeding for the level of platelet count. Other autoantibodies such as antiphospholipid antibodies occur in up to 30% of patients with ITP, and these may affect platelet and vascular function. However, in general, unlike the thrombocytopenia that accompanies bone marrow failure, serious bleeding is not common in patients with ITP.

Reaching a consensus on terminology

Until recently, comparing data from clinical studies in ITP has been made extremely difficult due to the inconsistent way in which terminology has been used. For example, the term 'refractory' has been interpreted in different ways. For some it implies the postsplenectomy patient who has failed to respond or who

Postgraduate Haematology, Seventh Edition. Edited by A Victor Hoffbrand, Douglas R Higgs, David M Keeling and Atul B Mehta.

Table 42.1 New Consensus Terminology definitions.

Immune thrombocytopenia	Primary (old term 'idiopathic') Secondary (SLE, etc.)
Diagnosis	Requires platelet count $<100 \times 10^9$/L
'Acute' and 'chronic' replaced by	'Newly diagnosed', 'persistent' and 'chronic' (see Figure 42.1)
Severe ITP	Where there is clinically relevant bleeding irrespective of platelet count
Refractory	Failed splenectomy or relapse *and* severe ITP or risk of bleeding

has relapsed. Others interpret the term to mean those who do not respond well to treatment irrespective of splenectomy status. Similarly, deciding what constitutes a 'complete response' to therapy differs from study to study, making comparisons of drug treatments fraught with problems.

In order to address this, a group of experts held a consensus meeting and agreed on the terminology that should be used in ITP. These have been published and will be used in all drug trials, and in the clinic, from now on. The key terminologies are listed in Table 42.1.

Chronic ITP implies disease that has been present for 12 months or more from diagnosis. Previously, the term 'chronic' applied to individuals who had ITP for 6 months or longer. However, because patients may remit between 6 and 12 months, the term 'chronic' has now been redefined to refer to those individuals whose ITP has been present for 12 months or greater (Figure 42.1).

Pathophysiology

ITP in adults is a typical organ-specific autoimmune disease and, in common with other autoimmune diseases, there is a skewing of the Th1/Th2 response towards Th1 (proinflammatory). These are disorders in which there are antibodies or cells (B cells, T cells, antigen-presenting cells, or others) that react against self-antigens. These may cause disease if the target tissue is damaged. The binding of antibodies to target cells results in clearance of the antigen from the body. The normal adaptive response, such as that seen in microbial infection, results in complete removal of the non-self-antigen. In autoimmune disease, however, there is continual production and incomplete clearance of the antigen, leading to perpetuation of the immune attack. In ITP the antigen is platelet glycoprotein, found on megakaryocytes and platelets.

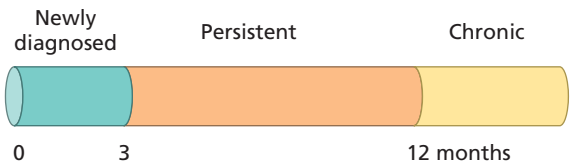


Figure 42.1 Stages of ITP using the new Consensus Terminology. After diagnosis until 3 months the patient is described as having newly diagnosed ITP. If the ITP persists beyond 3 months, it is termed persistent ITP and once the 12-month milestone has been reached the disease is described as chronic ITP.

ITP is multifactorial

From studies of other autoimmune diseases, it is quite clear that ITP has a multifactorial basis, and that loss of tolerance to a self-antigen alone is not sufficient to generate the autoimmune disorder. Patients probably require: (i) a specific set of genetic determinants, such as polymorphisms within major histocompatibility complex (MHC), CTLA4 or other genes, (ii) dysregulation of the immune response (involving dendritic cells, T or B cells, or all three) and (iii) an environmental 'trigger', which may be infectious, for example a viral infection. Autoimmune disease arises only when all these determinants are present in an individual at the same time. This is further reinforced by the observation that self-reactive lymphocytes are commonly found in normal individuals.

Antiplatelet antibodies and their targets

The autoantibodies involved in ITP are generally IgG, but IgA and IgM autoantibodies have been reported. Opsonized platelets are removed prematurely by the reticuloendothelial system via an Fc-dependent mechanism. In addition, the autoantibodies may impair megakaryocyte growth and development, platelet release and may also induce apoptosis of megakaryocytes. The overall result of this is failure of platelet production.

There are various assays for measuring antiplatelet antibodies. Looking for the presence of platelet-associated IgG is of no value since this is found in non-immune as well as immune thrombocytopenia. More sophisticated assays, such as the monoclonal antibody-specific immobilization of platelet antigens (MAIPA) assay, has greater specificity (90%) albeit with low sensitivity (50–65%). Using MAIPA, platelet-associated IgG and antigen capture assays, several platelet antigens have been characterized. These include GPIIb/IIIa ($\alpha_{IIb}\beta_3$, the fibrinogen receptor) and GPIb/IX (the von Willebrand receptor), which appear to be the most frequently involved. Less commonly, GPIa/IIa, GPIV and GPV are involved.

The role of *Helicobacter pylori* in the development of ITP

This Gram-negative microaerophilic bacterium is the main cause of gastritis and peptic ulcer disease. It has also been implicated in the development of gastric adenocarcinoma and mucosa-associated lymphoid tumours, and in some autoimmune disorders. A number of studies have shown improvement in platelet counts in ITP patients positive for *H. pylori* following eradication of the bacterium. However, the data from different studies are conflicting, with some centres showing a very high rate of response to eradication, while others have a low rate. Further support for the involvement of *H. pylori* in ITP comes from studies which have shown that there is a reduction in the level of antiplatelet antibodies in plasma following eradication of the bacterium.

How *H. pylori* may initiate or perpetuate ITP is not known. Possibilities include molecular mimicry, where there is cross-reactivity between the antibody, the bacterium and platelet antigens.

T cells may also be involved

The role of T cells in the development of ITP is becoming clearer. That T cells are involved has been known for some time, since the autoantibodies in ITP are predominantly of the IgG subclass, and isotype switching from IgM to IgG requires T-cell help. However, some 40% of patients with chronic ITP have no detectible autoantibodies yet are thrombocytopenic and appear to have true ITP. $CD8^+$ T cells have been linked to the pathogenesis of many autoimmune diseases, such as type 1 diabetes, and recently it was shown that ITP patients have a direct $CD8^+$ T cell-mediated cytotoxicity that induces platelet destruction. At present, however, it is unknown whether cell-mediated platelet destruction contributes to the severity of disease or to the difficulty of treatment in some patients with ITP. However, recent studies by Wadenvik's group in Sweden suggests that in some patients with ITP, T cells may interact directly with platelets, inducing platelet lysis (Figure 42.2).

Thrombopoietin levels in ITP

Since thrombopoietin (TPO) is the principal growth factor involved in platelet production, it might be expected that TPO levels would be raised in patients with ITP. Plasma TPO levels are undoubtedly elevated in patients with aplastic anaemia, chemotherapy-induced thrombocytopenia and other marrow failure syndromes. However, in ITP, levels of TPO are normal or only modestly elevated (Figure 42.3). The reason for this has been poorly understood until fairly recently.

The liver produces TPO at a constant rate and this binds to TPO receptors on platelets and megakaryocytes. The free

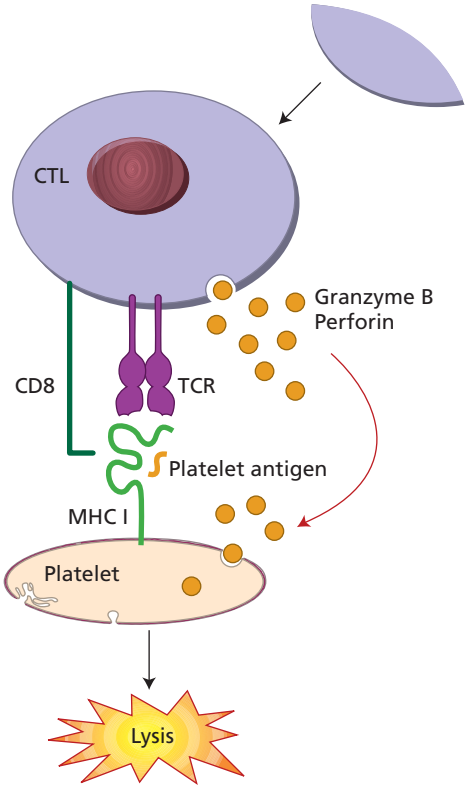


Figure 42.2 Model of cell-mediated cytotoxicity in chronic ITP. In the case of chronic ITP in the active phase, cytotoxic lymphocytes release toxic contents, such as granzyme B and perforin, and platelet lysis occurs.

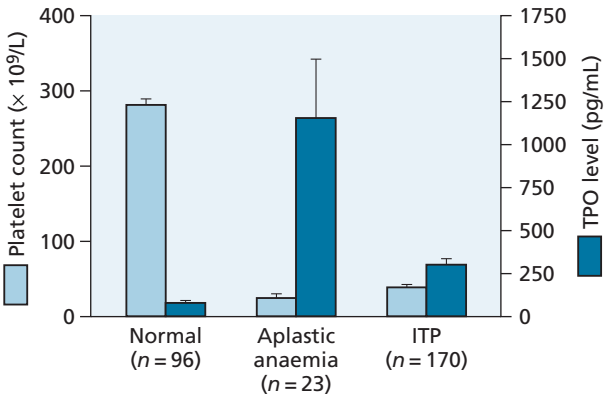


Figure 42.3 In normal individuals serum thrombopoietin (TPO) levels are low. In aplastic states, due to the absence of megakaryocytes and platelets (which bind TPO), TPO levels are high. In patients with ITP, the platelet count is low and TPO levels are normal or modestly elevated. This is believed to be due to absorption of free TPO by the increased megakaryocyte and platelet mass which is found in ITP. (Source: Nichol, 1998 [*Stem Cells* 1998; **16** (Suppl. 2): 165–75]. Reproduced with permission of Wiley.)

Table 42.2 Potential causes of thrombocytopenia in ITP.

- Autoantibody opsonization of platelets leading to destruction by the reticuloendothelial system
- Autoantibody opsonization of megakaryocytes, with inhibition of megakaryocyte growth, differentiation and platelet release
- Autoantibody-induced megakaryocyte apoptosis
- Relative thrombopoietin deficiency
- Molecular mimicry and immune complex formation
- T-cell direct lysis of platelets

(unbound) TPO is able to stimulate the bone marrow to generate sufficient platelets to counterbalance the natural daily platelet losses through platelet senescence. The key here appears to be the free TPO component. The greater this level, the more the bone marrow is driven to produce increased numbers of platelets. If the megakaryocyte mass and platelet turnover are increased, such as in ITP, TPO is bound to the receptors on these cells, leaving little free TPO to stimulate undifferentiated bone marrow cells in patients with ITP. There is therefore a relative lack of TPO in patients with ITP, which compounds the platelet losses that occur through increased destruction of platelets by the reticuloendothelial system.

In summary, ITP is much more complex than originally believed, with several mechanisms leading to thrombocytopenia (Table 42.2).

Natural history of ITP

This has not been systematically studied. It is not clear which patients require treatment, whether treatment has any major beneficial effect on the patient, whether patients live longer or whether treatment alters the natural history of the disease. For the majority of patients, ITP is a fairly minor disorder. Serious bleeding is not common and clinical sequelae of the disease are generally absent in patients with platelet counts above 30×10^9/L. There have been a few studies published looking at clinically relevant bleeding and platelet count. In normal individuals, there is a correlation between the platelet count and bleeding. The lower the platelet count, the more likely is bleeding. In ITP this is not the case, and most patients with ITP show little bleeding until the platelet count drops below 20×10^9/L. One study showed that no grade 4 bleeding occurred in patients with platelets greater than 10×10^9/L.

Diagnosis

Despite significant advances in our understanding of ITP, the diagnosis remains one of exclusion in both paediatric and adult ITP. A thorough history should be obtained, looking for diseases that might cause thrombocytopenia. The patient should be asked about bleeding during previous surgery or dentistry. The sites of bleeding should be determined.

A full physical examination should be normal, apart from the expected clinical signs associated with thrombocytopenia. In addition to the 'dry purpura' of ITP, there should be an examination of the mucous membranes, including the mouth and optic fundi looking for the presence of retinal haemorrhage. Enlargement of the liver, spleen or lymph nodes suggests an alternative diagnosis.

The laboratory investigation of ITP is straightforward. The FBC should confirm the isolated thrombocytopenia. There may be a degree of iron deficiency anaemia, but this should be in proportion to the clinical history. There should be no abnormal white blood cells or red cell fragments in ITP. If the latter are found, the underlying diagnosis may be thrombotic thrombocytopenic purpura, a disorder that is far more serious than ITP and which requires urgent treatment. Table 42.3 outlines the most

Table 42.3 Investigation of suspected ITP.

Basic evaluation of ITP
- Patient history
- Family history
- Physical examination
- Full blood count
- Peripheral blood film
- Blood group (Rh) and reticulocyte count
- Direct antiglobulin test
- Quantitative immunoglobulin measurement
- Bone marrow examination*
- *Helicobacter pylori*
- HIV
- HCV

Tests that may be useful in selected cases
- Antiphospholipid antibody (including anticardiolipin and lupus anticoagulant)
- Antithyroid antibody and thyroid function
- Pregnancy test
- Antinuclear antibodies
- Viral PCR for parvovirus and cytomegalovirus

Tests of unproven or uncertain benefit
- Glycoprotein-specific antibody
- Thrombopoietin assay
- Reticulated platelet count
- Indirect platelet-associated IgG
- Bleeding time
- Platelet survival study
- Serum complement levels

*The need for bone marrow remains contentious and the indications for performing this test are discussed in Table 42.4.

Table 42.4 Indications for bone marrow examination in ITP.

- Failure to respond to, or relapse following, first-line therapy
- Presence of atypical clinical or laboratory features
- Age >60 years
- Before splenectomy (but see text)

useful tests in the diagnosis of ITP, along with other tests that are either of limited or no value.

Bone marrow examination

Previously, one of the mainstays of diagnosis was assessment of the bone marrow. The rationale behind this was the possibility of a patient having a marrow disorder such as leukaemia, lymphoma or infiltration. However, studies to date have shown quite clearly that in patients with isolated thrombocytopenia, and with no atypical symptoms or signs, no cases of bone marrow pathology were detected. For this reason, recent guidelines do not recommend performing a bone marrow examination in ITP. If patients fail to respond to, or relapse following, first-line treatment, then a bone marrow examination should be carried out. Similarly, if there is any hepatomegaly, splenomegaly, lymphadenopathy or any clinical or laboratory feature suggesting the presence of a disease other than ITP, a bone marrow examination should be performed. Bone marrow examination should also be performed if the patient is above the age of 60 years, since myelodysplasia becomes more likely, and myelodysplastic syndrome may resemble ITP. Finally, a bone marrow examination should possibly be carried out if splenectomy is being contemplated (Table 42.4).

Cytologically, the bone marrow in ITP generally shows normal development and maturation of all cell lines. Megakaryocyte numbers are typically normal or increased. During the clinical trials of the TRAs some patients developed increased reticulin fibrosis in the bone marrow. Subsequent reviews of patients with ITP who have not received any treatment have shown that bone marrow reticulin is increased in around two-thirds of patients. The increased reticulin seen in the TRA studies is most likely due to stimulation of pre-existing reticulin.

Management

Major haemorrhage is the most feared event in patients with ITP. Over the years, attempting to prevent this by treating all patients with a platelet count below 30×10^9/L has led to over-treatment. However, catastrophic bleeding is not commonly seen and the main aim of treatment today has shifted away from trying to normalize the platelet count towards finding a strategy that will allow the patient to achieve a 'safe' platelet count. What constitutes a safe platelet count will vary from patient to patient, and also each patient with ITP will have differing requirements throughout their disease history. For example, a young patient with platelet count of 20×10^9/L should be safe to work and carry out normal day-to-day activities, but would not be deemed safe if he or she wished to ski or undergo surgery. The safe platelet count would therefore be higher during these haemostatic challenges. The one-size-fits-all approach has failed to work over the years and has caused increased morbidity and mortality in patients due to the adverse effects associated with many of the treatments in current use. When considering ITP treatments, we need to separate these into two major approaches: short-term and long-term treatments. These are discussed later.

There are many current treatments for ITP, although most have not been approved (licensed) for this indication. Those that do have a licence for ITP include corticosteroids, intravenous immunoglobulin (IVIg), intravenous anti-D and more recently the TRA class of drugs. All the others, of which there are many, have been used for many years, but they have never undergone formal study and are not specifically approved for ITP.

Many treatments are of relatively low efficacy when used alone (around 30%), and most treatments, apart from the TRAs, address solely platelet destruction. They do this through a variety of different mechanisms, but the net result is immunosuppression. The original rationale behind this was that in order to reduce the degree of autoantibody production, platelet opsonization and destruction by the reticuloendothelial system, the entire immune response should be suppressed. To some extent this works, although the treatments are purely palliative and once the treatment stops the platelet count drops to baseline levels once again. The disadvantage of immunosuppression is that, as well as reducing the level of antiplatelet antibody or diminishing the effectiveness of the reticuloendothelial system, normal immunity is suppressed. This can lead to infections and other complications. One study from 2001 showed quite clearly that, of the patients with ITP who died, at least half succumbed to infection caused by the immunosuppressive agents.

Traditionally, treatments have been grouped into first line, second line, third line and so on, depending on the disease stage in which they are used. These terms are not particularly useful clinically, but do provide a guide as to which treatments to use initially and which should be used later.

First-line therapies comprise corticosteroids, IVIg and anti-D. These three treatments work fairly quickly (within 24–48 hours) and their efficacy rates are high at around 70–80%. The disadvantage of these and most of the other treatments is that the platelet count generally drops once the treatment is stopped. These treatments are seldom curative, and must be considered palliative therapies. In addition, IVIg and anti-D are pooled blood products.

If patients fail to respond to first-line treatment they may then be given a second-line drug, such as azathioprine, mycophenolate mofetil or a TPO receptor agonist (Table 42.5). As with first-line drugs, second- and third-line treatments are palliative.

Table 42.5 Treatment options after first-line therapy.

Non-approved second-line treatments
- Azathioprine
- Dexamethasone
- Methylprednisolone
- Ciclosporin
- Mycophenolate mofetil
- Cyclophosphamide
- Danazol
- Dapsone
- Vincristine
- Rituximab

Surgical second-line treatment
- Splenectomy

Approved second-line treatments
- Romiplostim
- Eltrombopag

Splenectomy

Splenectomy is a surgical second-line option for patients who have relapsed after treatment with a first-line drug, although many clinicians prefer to defer splenectomy until at least 12 months from diagnosis. The procedure increases the lifespan of antibody-coated platelets and may reduce antibody production. Responses are seen in about two-thirds of patients, who achieve a normal platelet count, and the response is often sustained with no additional therapy for at least 5 years. Even patients who do not have a complete response may still expect a partial or transient increase in platelet count. Around 14% of patients do not respond and approximately 20% of responders will relapse months or years later.

Postoperative complications of splenectomy

The complication rates of splenectomy are 12.9% with laparotomy and 9.6% with laparoscopic splenectomy. The mortality rate is 1.0% with the open laparotomy procedure and 0.2% with laparoscopy.

Venous thromboembolism (VTE) may occur following splenectomy, particularly if the platelet count rises to 1000×10^9/L or more. ITP itself carries a small increased risk of VTE, although the exact level of risk is not known, but there may be an additive VTE effect in patients with ITP who undergo splenectomy. ITP patients should be evaluated for this risk and appropriate thromboprophylaxis should be provided.

Predicting the response to splenectomy

Clinicians have long sought possible predictors of response to splenectomy. Potential predictors have included response to oral corticosteroids, IVIg and others. Indium-labelled autologous platelet scanning appears to be the most sensitive predictor of response to splenectomy. However, indium scanning is not available in every hospital, but is limited to only a few centres worldwide.

Prevention of infection after splenectomy (see also Chapter 17)

Splenectomized patients are at long-term risk for opportunistic postsplenectomy infection (OPSI) with encapsulated bacteria. OPSI is often rapidly progressive and has a poor outcome. The main organisms responsible for sepsis in splenectomized patients are pneumococci (50%), meningococci and *Haemophilus influenzae*.

In order to prevent OPSI, patients should be given prophylactic polyvalent pneumococcal vaccine at least 4 weeks prior to, or 2 weeks after, splenectomy, with re-vaccination according to local guidelines. *H. influenzae* b and meningococcal C conjugate vaccinations should also be administered prior to splenectomy. Re-vaccination in susceptible individuals for *H. influenzae* should be administered according to local guidelines.

In the UK, the Chief Medical Officer advises the long-term use of oral antibiotics in the form of phenoxymethylpenicillin 250 mg twice daily, or equivalent, or erythromycin 250 mg twice daily if the patient is allergic to penicillin. However, the need for lifelong antibiotic prophylaxis remains unproven.

Short-term treatment

This refers to treatments lasting 2 weeks or more, which can be used when patients have a sudden drop in their platelet count leading to an unacceptably low platelet count, or when patients face haemostatic challenges such as planned surgery or dentistry, or other high-risk activity, where the main risk is excessive bleeding. Patients are generally started on treatment 1 or 2 weeks prior to the event, with regular monitoring of platelet counts. Table 42.6 shows the target platelet counts for the most common procedures.

Long-term treatment

This is required when patients have an unsupported platelet count that is too low for normal day-to-day activities. Generally, these patients will have failed other therapies or, if they

Table 42.6 Target platelet counts for procedures.

Procedure	Target platelet count
Dentistry (non-invasive)	Any
Dentistry (fillings, local anaesthesia)	$\geq 30 \times 10^9$/L
Minor surgery	$\geq 50 \times 10^9$/L
Major surgery	$\geq 80 \times 10^9$/L

responded to corticosteroids, they require unacceptably high doses to maintain a safe count. In this setting the aim is not to achieve a specific count, but rather to move the platelet count into a 'safe' zone. This will vary from patient to patient. In general, once platelets exceed 15–20 × 10^9/L the risk of bleeding is low. Older individuals, or those with comorbidities or previous bleeding, will generally require a higher platelet count.

As we have already seen, most of the existing treatments used for long-term platelet control have known toxicities. The aim with long-term treatment is to use a treatment that has acceptable levels of toxicity, but which achieves the desired platelet count.

Patients with refractory ITP

The term 'refractory' has caused confusion over the years due to a lack of agreement about whether the term should be used only for patients who have suffered a failed splenectomy. The new Terminology Consensus document published in 2009 defines refractory ITP as that where the patient has undergone splenectomy which has failed or relapsed, and who has 'severe' ITP (clinically relevant bleeding) or a risk of bleeding. Some 20% or more of adult patients fall into this category. Clinically, patients with refractory ITP pose a major challenge since, by definition, they are resistant to many of the treatments in current use. They often have low platelet counts, in addition to bleeding that is difficult to control since the disease is unresponsive to conventional therapies.

Is drug treatment needed?

Many adults with chronic refractory ITP are able to tolerate severe thrombocytopenia (platelet count as low as 10 × 10^9/L) relatively well, with near-normal quality of life. For patients who fail to respond to standard therapies and who require treatment, a limited number of options are available. The risks of continuing therapy should be discussed and evaluated with the patient and compared with the benefits that treatment may provide. Patients should be assessed for other possible causes of their thrombocytopenia, including drug-induced, infection, inherited thrombocytopenia and myelodysplastic syndromes. Instead of receiving further ineffective treatments, many patients choose to live with lower platelet counts and, providing they remain free of bleeding, a watching brief can be kept. If any bleeding does occur, the patient can be treated with rescue therapy such as IVIg in combination with intravenous corticosteroids or cyclophosphamide, aimed at raising the platelet count rapidly.

Rituximab

This chimeric monoclonal antibody targets the CD20 antigen on B cells and is licensed for B-cell non-Hodgkin lymphoma and rheumatoid disease. There have been several studies of its use in ITP and other autoimmune haematological disorders. However, there are no randomized controlled studies and the drug has not been approved for use in ITP. Nonetheless, rituximab does appear to be effective in ITP, even when patients have failed multiple therapies. The largest review of rituximab in ITP was conducted by Arnold and colleagues, whose data suggest that around 60% of patients will respond to rituximab and approximately 40% of all patients will achieve a complete response. The responses are bimodal, with an early phase within 1–2 weeks and a second peak at about 6–8 weeks. Response duration varies from 2 months in partial responders to 5 years or more in about 15–20% of initially treated patients.

There are safety concerns with rituximab and it should be used with caution in any patient who has positive hepatitis B core antibody status, since hepatitis B may be re-activated. Adverse events associated with rituximab are usually mild or moderate, with a low incidence of infections. However, B- and T-cell repertoires do not return to normal, even when B-cell numbers have normalized. Severe infectious complications, including *Mycoplasma pneumoniae* and *Aspergillus niger*, echovirus, papovavirus and cytomegalovirus, have been reported. There are also reports of over 50 cases of progressive multifocal leucoencephalopathy (PML) associated with rituximab treatment in patients with lymphoma and connective tissue disorders.

In terms of dosing, instead of using standard lymphoma doses (375 mg/m^2 weekly for 4 weeks), lower doses of rituximab (100 mg IV weekly for 4 weeks) have been shown to be equally effective.

Combination chemotherapy

This form of treatment may be effective for some patients with chronic refractory ITP. Recently, Tao and colleagues investigated a combination comprising intravenous cyclophosphamide on days 1–5 or 7 and prednisone on days 1–7, combined with vincristine on day 1, and one of the following: azathioprine on days 1–5 or 7 or etoposide on days 1–7. The overall response rate in the 31 patients treated was 67.9%, including a complete response in 41.9%; the therapy was well tolerated. Other combinations, such as CHOP and COP, have been used. However, the toxicities and side-effects of these therapies make them unpopular with both clinicians and patients.

TPO receptor agonists

Because of the relative lack of TPO seen in ITP, as discussed earlier, there is the potential to use exogenous TPO in order to stimulate the bone marrow to generate more platelets. Early experience with recombinant human TPO (rhTPO) and recombinant human megakaryocyte growth and development factor

(rhMGDF) yielded mixed results. The drugs appeared to be effective in raising the platelet count, but cross-reacting antibodies to the drugs neutralized endogenous TPO, making patients more thrombocytopenic than before the drug was given. For this reason, development of rhTPO and rhMGDF ceased. More recently, rather than develop TPO receptor agonist drugs based on endogenous TPO, where there would be significant sequence homology and the potential for neutralizing antibodies to develop, pharmaceutical companies decided to create drugs with no sequence homology to the native TPO protein. Two TPO receptor agonists are now approved: one is a peptibody called romiplostim, and the other, eltrombopag, is a small non-peptide molecule. Both have undergone rigorous trials in adult patients with ITP both before and after splenectomy. The efficacy rates are encouraging in all patients, and since they do not involve immunosuppression and are not corticosteroid-based, they appear to offer many advantages over the drugs in current use.

From pivotal randomized clinical trials, responses to romiplostim were seen in 88% of non-splenectomized patients and in 79% of splenectomized patients. Concomitant ITP medication was reduced or stopped in 79% of subjects and there was a reduction in rescue medication. From clinical trials of eltrombopag, efficacy was reported at 79%, again with a reduction or cessation of concomitant medication in 59%, with reduction in bleeding reported at 76%. Both of these agents are now licensed for use as second-line treatment for ITP, in keeping with the recent Consensus document, although within the EU their label indication is for patients who have failed splenectomy or in whom splenectomy is contraindicated.

ITP in children

Children with ITP are treated based on clinical bleeding, quality of life and not just platelet count. Severe bleeding is uncommon in children with ITP, even when the platelet count is very low ($<10 \times 10^9$/L). The incidence of intracranial haemorrhage (ICH) in children with ITP is around 0.1–0.5%. As with adult ITP, there is no method to predict which children will develop ICH, and this complication has been reported even in children receiving treatment for ITP. In general, however, ICH is commoner in patients with refractory or resistant disease than at initial presentation. Investigation of childhood ITP is similar to that in adults, although the incidence of *H. pylori* infection is much lower and there is no need to screen for this.

'Watch and wait' policy

Around 70% of children with acute ITP do not have significant bleeding symptoms and can be managed without drug therapy. Admission to hospital should be reserved for children with clinically significant bleeding, such as severe epistaxis or gastrointestinal or other bleeding. Most children with minor, mild or moderate symptoms can be safely managed in the outpatient setting. The emphasis is towards treating symptoms rather than platelet counts, since the platelet count is a surrogate marker and does not accurately predict clinical outcome.

General measures for persistent and chronic ITP in children

The management of children with persistent/refractory ITP is much the same as for newly diagnosed ITP. Many children stabilize with an adequate platelet count ($>20–30 \times 10^9$/L) and have no symptoms unless injured. In children who are less than 10 years of age at diagnosis, spontaneous remission is likely to occur, and expectant management can continue depending on the risk of bleeding and the degree of activity restriction of the child. The onset of menstruation may be problematic and can be managed with antifibrinolytic agents and the combined contraceptive pill.

The optimal management of chronic ITP in childhood is not known. Treatment should be tailored to each individual child and situation, and therapy should be effective and it should not carry more risk than the untreated condition.

Treatment options in childhood ITP

Children with severe bleeding symptoms should be treated. Treatment should also be considered in children with moderate bleeding or those at increased risk of bleeding.

First-line treatment in children

The majority of children require no medical treatment, and can be managed using a 'watch-and-wait' policy. If treatment is deemed to be required, options include IVIg, anti-D and corticosteroids. The other drugs used in adult ITP are not suitable for childhood disease due to their toxicities and long-term sequelae.

IVIg raises the platelet count in more than 80% of children and does so more rapidly than steroids or no therapy. Transient side-effects are seen in 75% of children. Intravenous anti-D immunoglobulin can be given to Rh(D)-positive children as a short infusion and is useful in the outpatient setting. Mild extravascular haemolysis is common, with a mean drop in haemoglobin of 10 g/L. Like IVIg, anti-D is a pooled blood product, but it is derived from a smaller donor pool. The safety profile of anti-D is good with no reported transmission of infection to date. However, there have been reports of severe intravascular haemolysis leading to disseminated intravascular coagulation, most of which have been fatal. Prednisolone at a dose of 1–2 mg/kg daily for a maximum of 14 days is effective in around 75% of children (platelet count $>50 \times 10^9$/L) within 72 hours. More prolonged high-dose corticosteroid regimens are associated with increased toxicity. There are data to suggest that higher doses of prednisolone (4 mg/kg) for shorter courses may be more effective and associated with fewer side-effects. TRAs are currently not licensed for use in children.

ITP in pregnancy (see also Chapter 50)

Thrombocytopenia occurs in 5% of pregnancies, but most of these cases are gestational rather than immune-mediated. The diagnosis of ITP involves the exclusion of other causes of thrombocytopenia during pregnancy. Patient history, physical examination, blood count and blood film examination are used as in non-pregnant patients. The work-up of a pregnant patient with ITP is essentially the same as that of a non-pregnant patient. There are some conditions specific to pregnancy and these should be considered when investigating a pregnant patient with thrombocytopenia (Table 42.7).

Table 42.7 Causes of maternal thrombocytopenia in pregnancy.

- Gestational thrombocytopenia
- Pre-eclampsia, HELLP syndrome, disseminated intravascular coagulation
- Folate deficiency
- Massive obstetric haemorrhage
- Acute fatty liver

HELLP, haemolysis, elevated liver enzymes, low platelets.

Laboratory investigation of ITP in pregnancy

ITP in pregnancy, as with other types of ITP, is a diagnosis of exclusion. All investigations carried out in pregnancy are aimed at excluding conditions that may result in thrombocytopenia. Some disorders such as thrombotic thrombocytopenic purpura and HELLP syndromes require urgent diagnosis and treatment since the mortality rates are high (Table 42.8).

Table 42.8 Investigation of suspected ITP in pregnancy.

- Coagulation screening (prothrombin time, PT)
- Activated partial thromboplastin time (APTT)
- Fibrinogen assay
- Liver function tests including bilirubin, albumin, total protein, transferases, γ-glutamyl transferase and alkaline phosphatase
- Antiphospholipid antibodies, including anticardiolipin antibodies and lupus anticoagulant
- SLE serology
- Review of the peripheral blood film
- Reticulocyte count
- Bone marrow examination: not required to make the diagnosis of ITP in pregnancy

Management of ITP in pregnancy

There should be close collaboration between the obstetrician, haematologist, obstetric anaesthetist and neonatologist. Treatment is largely based on the risk of maternal haemorrhage. Throughout the first two trimesters, treatment is initiated when the patient is symptomatic and/or when platelet counts fall below 20×10^9/L, or when it is necessary to produce an increase in platelet count to a level considered safe for procedures such as obstetric delivery or epidural anaesthesia. Patients with platelet counts of $20–30 \times 10^9$/L or more do not require routine treatment.

Delivery

A platelet count above 50×10^9/L is generally acceptable for standard vaginal or caesarean delivery. For epidural anaesthesia a platelet count of 75×10^9/L is required, although this is not evidence-based.

Treatment options in pregnancy

The primary treatment options for maternal ITP are similar to those of other adult ITP patients, namely corticosteroids and IVIg. There is limited evidence for the use of intravenous anti-D, splenectomy and azathioprine. Other agents, such as vinca alkaloids, rituximab, danazol, most immunosuppressive drugs (apart from azathioprine) and TRAs should not be used in pregnancy because of possible teratogenicity.

Corticosteroids

These are the most cost-effective option. Prednisolone is initially given at a dose of 10–20 mg/day and may then be adjusted to the minimum dose that produces a haemostatically effective platelet count. At least 90% of the administered dose of prednisolone is metabolized in the placenta by 11β-hydroxlase, but high doses may have an effect on the fetus.

Intravenous immunoglobulin

If prolonged high-dose steroid therapy is required or significant side-effects occur, or a rapid platelet increase is required, IVIg should be considered. The conventional doses of IVIg and likely response rates are similar to those seen in non-pregnant patients.

Options for maternal ITP refractory to first-line treatment

As with non-pregnant adults, combining first-line treatments in the refractory patient may be appropriate in the weeks prior to delivery. High-dose methylprednisolone (1000 mg), possibly in combination with IVIg or azathioprine, has been suggested as a treatment for pregnant patients refractory to oral corticosteroids or IVIg. Azathioprine has been used for many years in the post-transplant setting and this drug appears to be safe in pregnancy,

though the response is generally slow. Splenectomy in pregnancy is rarely required, but if essential is best carried out in the second trimester.

Management of the neonate (of mothers with ITP)

ITP in the neonate (from mothers with ITP) accounts for 3% of all cases of thrombocytopenia at delivery. The fetal or neonatal platelet count cannot be reliably predicted from the maternal platelet count. After delivery, a cord blood platelet count should be determined in all cases. Intramuscular injections (such as vitamin K) in the fetus should be avoided until the platelet count is known. Those infants with subnormal counts should be observed clinically and haematologically, as the platelet count tends to fall further to a nadir between days 2 and 5 after birth.

Selected bibliography

Megakaryopoiesis

Kaushansky K (2005) The molecular mechanisms that control thrombopoiesis. *Journal of Clinical Investigation* **115**: 3339–47.

Autoimmune disease

Arnold DM, Nazi I, Toltl LJ *et al*. (2015) Antibody binding to megakaryocytes in vivo in patients with immune thrombocytopenia. *European Journal of Haematology*. March, 1–6.

Bao W, Bussel JB, Heck S *et al*. (2010) Improved regulatory T cell activity in patients with chronic immune thrombocytopenia treated with thrombopoietic agents. *Blood* **116**: 4639–45.

Maloy KJ, Powrie F (2001) Regulatory T cells in the control of immune pathology. *Nature Immunology* **2**: 816–22.

Neschadim A, Branch DR (2015) Mouse models of autoimmune diseases: immune thrombocytopenia. *Curr Pharm Des*. **21**: 2487–97.

ITP general

Cuker A, Prak ET, Cines DB (2015) Can immune thrombocytopenia be cured with medical therapy? *Seminars in Thrombosis and Hemostasis*. **41**: 395–404.

Semple JW, Provan D (2012) The immunopathogenesis of immune thrombocytopenia: T cells still take center-stage. *Current Opinion in Hematology* **19**(5): 357–62.

Provan D, Stasi R, Newland AC *et al*. (2010) International consensus report on the investigation and management of primary immune thrombocytopenia. *Blood* **115**(2): 168–86.

Toxicities of current treatment

Portielje JE, Westendorp RG, Kluin-Nelemans HC, Brand A (2001) Morbidity and mortality in adults with idiopathic thrombocytopenic purpura. *Blood* **97**: 2549–54.

Development of novel TPO receptor agonists

Kuter DJ, Bussel JB, Newland A *et al*. (2013) Long-term treatment with romiplostim in patients with chronic immune thrombocytopenia: safety and efficacy. *British Journal of Haematology* **161**(3): 411–23.

Kuter DJ, Macahilig C, Grotzinger KM, *et al*. (2015) Treatment patterns and clinical outcomes in patients with chronic immune thrombocytopenia (ITP) switched to eltrombopag or romiplostim. *International Journal of Hematology* **101**: 255–63.

Hallam S, Provan D, Newland AC (2013) Immune thrombocytopenia: what are the new treatment options? *Expert Opinion on Biological Therapy* **13**: 1173–85.

Rizvi H, Butler T, Calaminici M, *et al*. (2015) United Kingdom immune thrombocytopenia registry: retrospective evaluation of bone marrow fibrosis in adult patients with primary immune thrombocytopenia and correlation with clinical findings. *British Journal of Haematology* **169**: 590–94.

Rodeghiero F, Ruggeri M (2015) Treatment of immune thrombocytopenia in adults: the role of thrombopoietin-receptor agonists. *Seminars in Hematology* **52**: 16–24.

CHAPTER 43

Thrombotic thrombocytopenic purpura and haemolytic–uraemic syndrome (congenital and acquired)

43

Pier Mannuccio Mannucci, Flora Peyvandi and Roberta Palla

Bianchi Bonomi Hemophilia and Thrombosis Center, IRCCS Cà Granada, Ospedale Maggiore, Milan, Italy

Historical introduction

Thrombotic thrombocytopenic purpura (TTP) was first described in 1924 by Moschcowitz in a 16-year-old girl who died after an acute illness presenting with a pentad of signs and symptoms (anaemia, thrombocytopenia, fever, hemiparesis and haematuria). Post-mortem examination found disseminated thrombi in the terminal circulation of several organs, mainly composed of platelets. Over the next three decades, other cases were described, mainly but not exclusively in women, occurring in isolation (idiopathic) or in association with various diseases or conditions (Table 43.1). It was understood that thrombocytopenia was caused by platelet consumption due to their widespread deposition in microvascular thrombi, and that anaemia was due to massive intravascular haemolysis after mechanical fragmentation (schistocytosis) of red cells forced by blood flow to pass through partially occluded vessels.

Thirty-one years after Moschowitz, Gasser described a syndrome that we called haemolytic–uraemic syndrome (HUS). In common with TTP, HUS was characterized by microangiopathic haemolytic anaemia, consumption thrombocytopenia and microvascular thrombosis, but differed because it occurred with minimal neurological symptoms, but severe signs of renal damage. Subsequently, it became apparent that a clear distinction between the two syndromes was often difficult. Although thrombocytopenia, anaemia and ischaemic symptoms due to widespread thrombus formation in the terminal circulation of several organs were rather consistent features, the prevalence of neurological over renal symptoms could not clearly differentiate TTP from HUS and vice versa. The term thrombotic microangiopathy (TMA) was then proposed, intended to emphasize the common pathology of the two syndromes, with no implication on the prevalence of neurological versus renal symptoms.

In the early 1980s, a major breakthrough strengthened the unitarian TMA terminology. Even though it had been postulated for a long time that massive thrombus formation was due to the presence in the circulation of substance(s) that aggregated platelets intravascularly, the putative aggregating agent had remained elusive. In 1982, Moake and others demonstrated, first in TTP and subsequently also in HUS, that patient plasma contained highly thrombogenic forms of the multimeric glycoprotein von Willebrand factor (VWF), a major adhesive moiety contained in endothelial cells, platelets and plasma. Abnormal VWF multimers of particularly high molecular weight (ultra-large) bind more avidly to the platelet glycoprotein (GP)Ib and aggregate platelets in conditions of high shear stress. Accordingly, it was postulated that ultra-large VWF is the aggregating agent involved in the formation of occlusive thrombi in the terminal circulation, both in TTP and HUS.

It remained to be explained why ultra-large multimers, normally absent, were present in patient plasma. Even though the deficiency or dysfunction of one or more enzymes disposing of them physiologically was postulated, it was not until the late 1990s that Furlan, Tsai and their associates independently showed that the link between ultra-large VWF multimers and TTP was a metal ion-dependent plasma

Postgraduate Haematology, Seventh Edition. Edited by A Victor Hoffbrand, Douglas R Higgs, David M Keeling and Atul B Mehta.

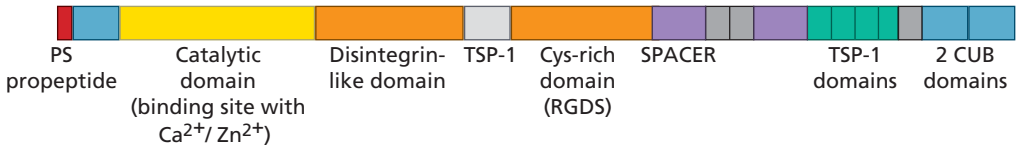


Figure 43.1 Domain structure of ADAMTS13. TSP1: thrombospondin 1; RGDS: Arg-Gly-Asp-Ser; CUB: Complement components C1r/C1s, Urinary epidermal growth factor, Bone morphogenetic protein-1.

Table 43.1 Conditions and diseases associated with TTP.

• Pregnancy and post partum
• Infections (particularly HIV)
• Drugs (quinine and quinidine, ticlopidine, clopidogrel, ciclosporine, interferon-α, statins)
• Chemotherapy (mitomycin, cisplatin, gemcitabine)
• Allogeneic bone marrow transplantation
• Connective tissue disorders (lupus erythematosus and scleroderma)
• Cardiac surgery

metalloproteinase of 190 kDa, identified in 2001 by Zheng and colleagues as a new (the 13th) member of the ADAMTS (A Disintegrin And Metalloprotease with ThromboSpondin-1 repeats) family of metalloproteases (Figure 43.1). The only known physiological function of this protease, present in plasma and platelets, is to regulate the size of VWF by disposing of ultra-large multimers as soon as they are secreted from endothelial cells into plasma (Figure 43.2), thereby preventing heightened platelet aggregation and thrombus formation. Most importantly, both the groups of investigators made the seminal observation that ADAMTS13 was deficient in the plasma of TTP patients, but measurable in normal amounts in those with HUS. This observation challenged the unitarian theory of TTP and HUS as different clinical manifestations of the same pathological process (TMA), and generated the paradigm that TTP is due to low plasma levels of the VWF-cleaving protease, not involved in HUS pathogenesis. This attractive paradigm was not fully sustained by the next progress of knowledge, because not all cases of TMA diagnosed as TTP owing to the prevalence of neurological symptoms have low or undetectable levels of ADAMTS13. On the other hand, even if most cases diagnosed as HUS due to the prevalence and severity of renal symptoms have normal plasma levels of ADAMTS13 (particularly cases typically occurring in association with diarrhoea), there are atypical HUS cases characterized by low or undetectable protease levels. Tools to accurately differentiate one disease from the other are not yet readily available. On the other hand, a differential diagnosis is important in order to allow for the prompt initiation of the most appropriate therapy of these life-threatening diseases (plasma therapy in TTP, the complement inhibitor eculizumab in HUS, particularly in the atypical form).

Thrombotic thrombocytopenic purpura

TTP is a rare disease, with an estimated yearly incidence of less than 5 cases per million in all racial groups. Recently, greater awareness and perhaps improved diagnostic facilities give the impression that incidence is increasing. Even though both sexes

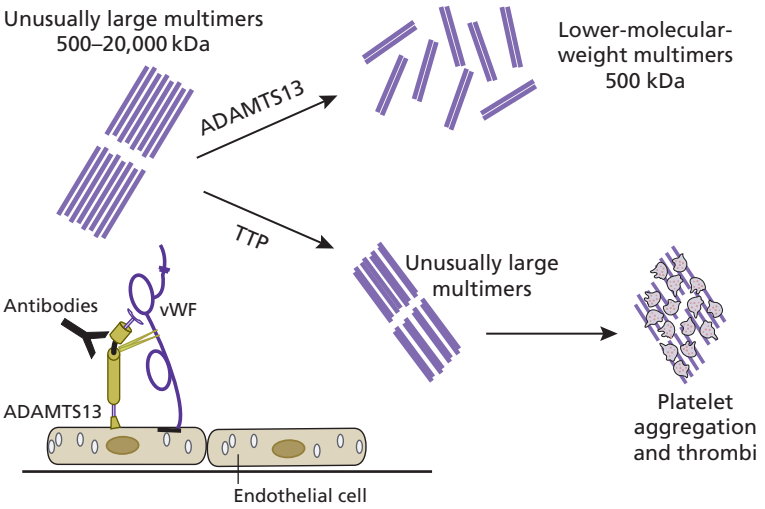


Figure 43.2 Pathophysiology of TTP. VWF is secreted from endothelial cells in the form of ultra-large multimers that anchor to endothelial cell surfaces and are also released into the circulation. ADAMTS13 cleaves a Tyr–Met bond in the A2 domain of VWF, severing the ultra-large multimers. Failure of cleavage leads to the persistence in plasma and onto endothelial cells of ultra-large multimers, which aggregate platelets, especially in conditions of high shear forces.

may be affected, the syndrome is definitely more frequent in women (two-thirds of cases).

Congenital TTP, caused by mutations in the ADAMTS13 gene, is inherited as an autosomal recessive trait and is often, but not exclusively, clinically manifest at birth or during early childhood. *Acquired TTP* is mainly due to autoantibodies that inactivate or bind ADAMTS13. The most common physiological or pathological conditions associated with autoimmune TTP are pregnancy, infections, immunomediated diseases and the use of drugs such as ticlopidine and clopidogrel. There are also forms secondary to metastatic tumors, organ transplantation (particularly allogeneic bone marrow and solid organ transplants) and the use of such drugs as ciclosporin, mitomycin and interferon-α. Mortality was very high (80–90%) until plasma exchange therapy was introduced, and is still high (10–20%) despite the dramatic improvement due to the adoption of this transfusional procedure. Because of the variability of presenting symptoms and associated comorbid conditions (see Table 43.1), cases of TTP may be initially seen by a variety of physicians other than haematologists, such as neurologists, nephrologists, oncologists and gynaecologists, and this sometimes hinders the prompt recognition of the syndrome, an essential requisite for optimal management.

Pathology and pathogenesis

The pathological basis of TTP is the widespread formation in the microcirculation of platelet thrombi, associated with relatively little endothelial cell injury and fibrin formation, but with abundant intrathrombus VWF. Microthrombi are found in several organs (mainly brain, kidney, myocardium, adrenal gland, digestive tract and pancreas) and lead to multiorgan failure, whereas grossly detectable thrombi in large arterial and venous blood vessels are usually lacking. The current mechanistic model implies that endothelial cells, activated by varied and often unidentified triggering agents, secrete larger amounts of ultra-large, uncleaved VWF which, in the presence of ADAMTS13 deficiency, aggregates platelets directly in conditions of high fluid shear stress, leading to massive intravascular platelet aggregation, ischaemic organ damage, consumptive thrombocytopenia and mechanical haemolytic anaemia. As mentioned above, two main mechanisms cause ADAMTS13 deficiency in TTP: mutations in the gene that encodes the protease and the acquired development of anti-ADAMTS13 autoantibodies.

Congenital TTP

Congenital TTP is very rare and represents no more than 5% of all TTP cases. It usually occurs early after birth or during early childhood, although in relatively rare instances the genetic disease may become manifest in adulthood. Congenital TTP is caused by homozygous or double heterozygous mutations in the *ADAMTS13* gene (located on chromosome 9q34) that affect protein secretion or function. To date, more than 150 mutations have been identified. Analysis of mutation location on the *ADAMTS13* gene reveals no evident hotspots, even though 75% of the missense mutations are in the portion of the gene that encodes the N-terminal domains of the molecule. Most of the identified mutations are missense; the remaining are nonsense, frameshift and splicing mutations generating truncated forms of the protease. Of the 20–30% mutations analysed by *in vitro* expression studies, the majority affect ADAMTS13 biosynthesis, intracellular trafficking and secretion and/or proteolytic activity.

Notwithstanding the detection of severe ADAMTS13 deficiency, the clinical presentation of congenital TTP is variable, with between-patient differences in terms of age of disease onset, clinical severity, number of relapses and need for long-term prophylactic replacement therapy. The type of mutation may account for some of this heterogeneity and influences the severity of the clinical phenotype: mutations in the part of the gene encoding the N-terminal domains of the protease may be associated with more severe disease, in terms of earlier age of first TTP episode and higher annual rate of recurrent episodes. Pertaining to the potential triggers of acute episodes, conditions such as pregnancy and infections are believed to act by inducing an increase in VWF levels exceeding the levels that the residual active protease can degrade, leading to the precipitation of acute TTP episodes.

Acquired TTP

Autoimmune TTP accounts for the great majority of clinical cases and mainly occurs in adults, often in the absence of any obvious trigger. However, it can also arise secondary to other conditions: immunomediated diseases, infections, drug intake, pregnancy, sepsis, tumours and bone marrow transplantation. The higher incidence of autoimmune TTP in specific ethnic groups such as Afro-Caribbeans, as well as the onset of idiopathic TTP in two monozygotic twins both developing anti-ADAMTS13 antibodies, argue for a genetic predisposition for this acquired disease. Autoimmune TTP is due to antibodies that inhibit the proteolytic activity of ADAMTS13 and/or bind the protease to accelerate its plasma clearance through opsonization and/or other yet unclear mechanisms. Anti-ADAMTS13 are usually of the IgG type, although in a few cases autoantibodies were of the IgA and/or IgM isotype. Epitope mapping of anti-ADAMTS13 showed that in the majority of cases (88%) they are polyclonal (IgG4), with a primary epitope in the Cys-rich spacer domain of the protease, a domain essential for efficient VWF-cleavage activity and crucial for the interaction between VWF and the protease. However, cases of autoimmune TTP are also observed in patients with IgG molecules that react with the TSP1 2–8 and CUB domains of ADAMTS13. It is unclear why patients with sustained autoimmune ADAMTS13 deficiency in plasma develop clinical symptoms and signs only sporadically.

Not all patients with acquired TTP, as diagnosed by clinical criteria (consumptive thrombocytopenia and mechanic

haemolytic anaemia) have severe ADAMTS13 deficiency (<10% of protease activity) and some may have measurable or even normal plasma levels of activity. Furthermore ultra-large VWF multimers are not constantly detected in patient plasma, perhaps because there is an imbalance between their plasma release from endothelial cells and heightened binding to platelets, so that larger multimers may be even lacking in plasma.

Clinical and laboratory findings

The presence of both consumptive thrombocytopenia and haemolytic anaemia is essential for a diagnosis of TTP. Platelet count is often very low in the acute phase, with values of 30×10^9/L or less. Signs of mechanical haemolytic anaemia (haematocrit usually <20%) include the presence of schistocytes on peripheral blood smears, reticulocytosis, high indirect serum bilirubin, low or unmeasurable haptoglobin and negative direct Coombs test. High serum lactate dehydrogenase (LDH), usually in excess of 1000 U/L, a sensitive albeit non-specific sign of red cell destruction and tissue necrosis due to microthrombotic organ ischaemia, is an excellent parameter to monitor disease evolution. Neurological symptoms (coma, stroke, seizures or focal signs such as motor deficits, diplopia and aphasia) typically fluctuate in presentation and severity as a result of the ongoing formation and dissolution of thrombi in the cerebral microcirculation. Other symptoms or signs (such as headache, blurred vision, ataxia or mental status changes) are less typical. Gastrointestinal and cardiac systems are also frequent. Even though only a minority of patients with TTP have serum creatinine levels higher than 2.0 mg/dL, signs of renal involvement such as microscopic haematuria and proteinuria are frequent. High fever is not constant, despite the presence of this symptom in the original described disease. There is no or little alteration of coagulation and fibrinolysis, even though D-dimer levels may be moderately raised. The main presenting clinical features and signs in acute TTP are in Table 43.2.

Table 43.2 Presenting clinical symptoms and signs in acute TTP.

Thrombocytopenia	Epistaxis, bruising, petechia, haematuria, gastrointestinal bleeding
Neurological symptoms – often fluctuating and variable	Confusion, headache, paresis, aphasia, dysarthria, visual problems, encephalopathy, coma
Fever (>37.5 °C)	Not constantly present
Jaundice	Resulting from microangiopathic haemolytic anaemia
Renal impairment	Proteinuria, microhaematuria
Cardiac	Chest pain, heart failure, hypotension
Gastro-intestinal tract	Abdominal pain

ADAMTS13 testing

The possibility of measuring ADAMTS13 in plasma is based on the availability of several laboratory assays of enzyme activity, based on the degradation by patient plasma ADAMTS13 of purified, plasma-derived or recombinant VWF multimers or of synthetic VWF peptides followed by the direct or indirect detection of VWF cleavage products. A number of variables may interfere with assay results:

- All assays measure ADAMTS13 activity in static conditions and fail to reflect the *in vivo* physiological blood flow conditions necessary for optimal protease activity.
- Added denaturating agents (i.e. guanidine HCl or 1.5 mol/L urea) are required to promote the susceptibility of VWF multimers to cleavage by ADAMTS13 in the absence of shear.
- The use of shorter peptides as FRETS instead of full-length VWF in enzyme immunoassays may fail to detect a few inherited ADAMTS13 defects, since they measure metalloprotease activity independently of the exosite interactions with other VWF domains. Similar problems may occur when ADAMTS13 autoantibodies develop against these exosites.
- Activity assays are only available in specialized laboratories and emergency testing is limited, even though the results of assays using synthetic VWF peptides may be made available in less than one day.

Immunoenzymatic assays employing different monoclonal and polyclonal antibodies to measure plasma antigen levels of ADAMTS13 are also available, but are of limited clinical value, because they measure only the immunoreactive protein, not the protease activity of ADAMTS13.

With regard to the laboratory detection of autoantibodies, because most of them inhibit the enzymatic activity of ADAMTS13, they can be titrated *in vitro* by classical mixing studies, using serial dilutions of plasma mixtures from patients and normal individuals. Direct antibody Western blotting or ELISA assays are also available.

Clinical applications of ADAMTS13 testing

Whereas the original reports by Tsai, Furlan and their colleagues indicated that all patients with acquired TTP had severe ADAMTS13 deficiency, subsequent studies have shown that only a proportion of patients, ranging from 13 to 100%, who develop TTP have severe ADAMTS13 deficiency. The current prevailing opinion is that while undetectable or very low plasma levels of enzymatic activity (<10%) unequivocally establish a diagnosis of inherited or acquired TTP, not all patients appropriately diagnosed with TTP on the basis of clinical and laboratory signs and symptoms have severe protease deficiency in plasma. The cases more frequently associated with normal or moderately reduced ADAMTS13 levels (10–40% of normal) are those

secondary to other diseases or conditions, such as allogeneic bone marrow transplantation, HIV infection, chemotherapy and metastatic cancer. ADAMTS13 antigen assays have limited clinical utility, even though they are sometimes helpful in recovery studies following replacement therapy with ADAMTS13 (from plasma or clinical concentrates), because in this context they tend to be more sensitive than activity assays. The measurement of anti-ADAMTS13 helps to distinguish congenital from acquired TTP and offers a guide to disease prognosis and treatment with immunosuppressive agents (see below). Because some degree of discrepancy may exist between functional assays using full length VWF or short peptide substrates, the use of an alternative assay might be considered when ADAMTS13 results and clinical observations do not agree. The same type of ADAMTS13 assay should be used throughout therapeutic monitoring. The type of method should always be indicated when reporting clinical values of ADAMTS13.

Predictive value of ADAMTS13 assays

After remission from an acute episode of TTP, approximately one-third or more patients develop chronic recurrent disease. Carried out to predict the risk of relapse by using clinical and laboratory markers obtained at the time of the acute episode and/or during remission, prospective and retrospective studies have shown that patients who present with undetectable ADAMTS13 activity and detectable anti-ADAMTS13 during the acute episode and/or during first remission are more likely to experience recurrences. Pertaining to the subclass of anti-ADAMTS13 autoantibodies, the presence of IgG2 at diagnosis was associated with a higher risk of mortality, and cardiac and renal involvement, although the risk of relapse is the same regardless of the subclass. Other studies have suggested that high IgA, IgG1 and IgG3 anti-ADAMTS13 titres may be associated with more severe acute episodes and a higher risk of mortality.

All in all, in order to diagnose TTP in the acute phase of the disease, it is not essential to assay ADAMTS13 and to find very low or undetectable plasma levels. After having excluded other TMAs (see below), patients presenting with normal or moderately reduced ADAMTS13 can still be appropriately diagnosed as TTP. The decision to implement plasma therapy (infusion in patients with inherited disease, exchange in acquired disease) does not warrant the availability of ADAMTS13 values in real time. However, ADAMTS13 testing (plasmatic ADAMTS13 activity and anti-ADAMTS13) should be considered important adjuncts to the clinical diagnosis of TTP and are recommended to guide disease prognosis and treatment outcome.

Differential diagnosis with HUS

Typical cases of STEC-HUS are usually easily distinguishable from TTP owing to the occurrence of diarrheal prodromes after the ingestion of contaminated food or water and to the presence of severe symptoms of renal impairment (see below). It is sometimes more difficult to distinguish TTP from atypical HUS, except for the prevalence of neurological symptoms in the former and more severe renal failure in the latter (see more below). Other issues of differential diagnosis are concerned with conditions that present themselves as TMAs, all being characterized by consumptive thrombocytopenia and microangiopathic haemolytic anaemia (Table 43.3). Therefore, accurate evaluation of the underlying conditions, symptoms and biomarkers is fundamental for the differential diagnosis and hence appropriate therapy.

Table 43.3 Clinical conditions other than TTP and HUS presenting as thrombotic microangiopathies.

- Disseminated intravascular coagulation
- Pre-eclampsia/HELLP syndrome
- Scleroderma/systemic lupus erythematosus
- Catastrophic antiphospholipid antibody syndrome
- Evans syndrome (autoimmune thrombocytopenia and haemolytic anaemia)
- Heparin-induced thrombocytopenia
- Disseminated malignancy

Because ADAMTS13 deficiency is not a constant finding in patients with clinically diagnosed TTP, what is then the usefulness of ADAMTS13 testing in the differential diagnosis between TTP and other TMAs or HUS? Complete deficiency of the protease clearly directs the diagnosis towards TTP. However, because the results of ADAMTS13 testing may take a long time to be available, efforts were made to identify other pretreatment laboratory data that help to predict which subjects are more likely to have severely deficient ADAMTS13 activity. Pretreatment platelet count lower than 30×10^9/L, and serum creatinine levels higher than 2.0 mg/mL help to predict severely deficient ADAMTS13 activity (<10%), and thus to validate the initial clinical management with plasma exchange (PEX). In the presence of a higher platelet count and serum creatinine at presentation, atypical HUS should be considered in patients presenting with TMA characterized by measurable ADAMTS13 activity and a poor response to PEX therapy (see further on).

Because a dysregulation of the complement system has been associated with the development of atypical HUS, the use of markers of complement activation might help in the differential diagnosis. However, low C3 and normal C4 serum levels are not detected in all patients with atypical HUS, limiting the usefulness of these simple and largely available tests in the diagnosis. The search and identification of defects in the genes encoding complement proteins (factor H, factor I, mutant membrane cofactor protein MCP and other rarer proteins) is not strictly necessary to diagnose atypical HUS, not only because gene analysis is time-consuming, but also because no mutation is identified in at least half of the cases with atypical HUS. The

complement system may be abnormally activated also in some TTP patients, presenting with mutations/polymorphisms in complement genes and activated complement proteins. Furthermore, VWF has been shown to interact with complement factor H, potentially linking ADAMTS13/TTP to complement /HUS. All in all, it is presently difficult to use biomarkers of complement activation to accurately differentiate these two disorders.

Differential diagnosis with other thrombotic microangiopathies

TMAs can also be observed in other contexts such as disseminated intravascular coagulation (DIC), the HELLP syndrome in pregnant women, the so-called catastrophic antiphospholipid syndrome, pre-eclampsia and eclampsia, cancer patients and transplantation.

DIC can be distinguished from TTP by markedly increased D-dimer and, in decompensated cases, hypofibrinogenaemia and prolonged prothrombin and activated partial thromboplastin times. Differential diagnosis from TTP also involves other pregnancy-related TMAs such as pre-eclampsia and eclampsia, pregnancy being one of the most frequent clinical associations of TTP. In the latter, malignant hypertension is less frequent, renal damage is not as severe, whereas the degree of anaemia and thrombocytopenia is more severe. Tests of coagulation and fibrinolysis are usually abnormal in pre-eclampsia or eclampsia, albeit less markedly than in DIC. Abnormally high serum transaminases differentiate TTP from the so-called HELLP (Haemolysis, Elevated Liver enzymes and Low Platelets) syndrome of pregnancy. Sometimes connective tissue disorders such as systemic lupus erythematosus and scleroderma present with widespread microvascular thrombosis, mechanical haemolytic anaemia, consumptive thrombocytopenia and fluctuating neurological symptoms. This situation may also occur in the so-called catastrophic antiphospholipid antibody syndrome. These conditions may be distinguished from TTP because laboratory tests such as antinuclear and anticardiolipin antibodies and lupus-like anticoagulant give positive results. Another element of differential diagnosis is the presence of venous and arterial thromboses in large blood vessels, at variance with TTP and HUS.

Evans syndrome, due to the concomitant presence of antiplatelet and antierythrocyte antibodies, is distinguished by a positive Coombs test, lack of schistocytes and the usual absence of end-organ ischaemic symptoms. On the other hand, ischaemic manifestations due to thrombus formation are frequent features of heparin-induced thrombocytopenia, but TTP can be distinguished not only by lack of exposure to this drug, but also by the presence of haemolysis and schistocytosis. Disseminated malignancy is sometimes associated with consumptive thrombocytopenia and microangiopathic haemolytic anaemia. Until further investigations exclude or confirm the presence of metastatic cancer, it is not easy to distinguish TTP from this TMA.

The role of ADAMTS13 has been investigated in these TMAs and several other unrelated conditions and in physiological states, such as neonatal state, pregnancy, localized and metastatic cancer, liver cirrhosis, inflammatory states, postoperative period, uraemia, sepsis and autoimmune diseases: in the majority of studies ADAMTS13 activity was found to be reduced in patients compared with controls, but, at variance with TTP, at measurable plasma levels. Possible explanations for low but measurable ADAMTS13 are the following: (i) because VWF levels are increased in these patients, ADAMTS13 activity may be reduced due to consumption, (ii) because lower-molecular-weight forms of ADAMTS13 were identified in the plasma of patients with sepsis-induced DIC, ADAMTS13 may be inactivated by proteases such as granulocyte elastase, thrombin and plasmin, (iii) protein synthesis in the liver may be impaired in these patients, particularly in those with sepsis-induced DIC or HELLP, (iv) inflammatory cytokines (such as IL-6) may affect VWF proteolysis by ADAMTS13.

Natural history

In approximately two-thirds of the cases, TTP occurs only once (acute sporadic TTP). However, in more than one-third of them, the disease tends to recur at varying time intervals after remission from the acute episode, more often during the first year. Recurrence, defined as the re-emergence of clinical and laboratory signs and symptoms of TTP after remission of the initial acute episode, has a spectrum of presentations ranging from a single relapse to several episodes. Recurrences usually present with less severe clinical manifestations, perhaps because patients become alert to the first signs and symptoms of their disease. The chronic recurrent forms may have a genetic basis or be associated with the persistence of autoantibodies, whereas the forms associated with malignancy or transplantation usually present only as acute episodes with a low propensity to recur (in part because of their high mortality rate).

Treatment

The modern management of TTP started with the serendipitous observation that plasma infusion improved the clinical course of TTP. Plasma exchange (PEX) subsequently appeared to be more effective than plasma infusion. In 1991, the results of a prospective randomized clinical trial definitively established the greater efficacy of PEX over infusion, the clinical response rate being 78% in the former compared with 49% in the latter (mortality rates, 22% and 37%, respectively). Only recently, after the discovery of the role of VWF and its protease in the pathogenesis of TTP, these empirical treatments have found a rationale. Plasma exchange acts mainly by removing anti-ADAMTS13 autoantibodies, the most frequent mechanism of TTP. Plasma exchange

Table 43.4 Steps in the management of TTP.

- Clinical diagnosis (based on consumptive thrombocytopenia, mechanical haemolytic anaemia, presence of schistocytes in the blood film, high serum LDH)
- Plasma infusion (30 mL/kg) until plasma exchange can be started
- Daily plasma exchanges (3–6 L/day) until platelets are higher than 150×10^9/L for at least 2 days and LDH is normal
- Adjuvant treatments: prednisone 1–2 mg/kg per day
- Red cell transfusion as needed
- Platelets transfusion contraindicated

Table 43.5 Suggested immunomodulation in refractory antibody-mediated TTP.

- Prednisone: 1–2 mg/kg/die for at least 15 days, then tapering over 30 days
- Vincristine: 2 mg on day 1, then 1 mg on days 3, 6 and 9
- Azathioprine (1 mg/kg/daily, generally 100–150 mg/day)
- Cyclophosphamide (600 mg/mq/pulses)
- Ciclosporin (2–3 mg/kg/daily, divided doses)
- Immunoadsorption on staphylococcal protein A
- Splenectomy
- Anti-CD20 (rituximab): 375 mg/mq weekly for 4 weeks

also helps to replace the deficient protease, infusion alone being probably sufficient in congenitally ADAMTS13-deficient cases with no associated autoantibody.

A scheme summarizing the steps to be followed in the treatment of acute TTP is in Table 43.4. Treatment with plasma should be initiated as soon as the clinical diagnosis of TTP is suspected. Unfortunately, critically ill TTP patients are often first admitted to hospitals without facilities for PEX; in these circumstances, daily infusion of large amounts of fresh-frozen plasma (FFP) (30 mL/kg) should be started promptly, taking at the same time measures meant to avoid volume overload (see Table 43.4). Patients are best assisted in intensive care units, where renal failure, coma, seizures and such severe side-effects of PEX as congestive heart failure and catheter-related bleeding, thrombosis and infections can be optimally managed. At least 1.5 plasma volume should be exchanged daily until platelets rise to 150×10^9/L or more, LDH is normal and schistocytes are no longer present on blood films. Daily treatment with PEX must be continued for at least 2 days after full remission has been obtained (see Table 43.4). Tapering PEX (over 1–2 weeks) seems to reduce the risk of early relapse. Even though a clinical response is usually achieved after 5–10 days of PEX, daily PEX for up to 1 month is sometimes necessary. Cryoprecipitate-depleted plasma can be used instead of FFP if this fraction is available and convenient, but there is no evidence that this VWF-poor plasma is more effective than intact plasma. Plasma treated with virus-inactivation methods (e.g. solvent/detergent) should be preferred if possible because it decreases the risk of infections associated with patient exposure to plasma from a large number of donors.

In congenital TTP, replacement therapy may consist not only of plasma infusion, but also of virally inactivated, intermediate-purity plasma factor VIII-VWF concentrates containing relatively large amounts of ADAMTS13, such as 8Y (BPL; BioProducts Laboratory, Elstree, Herts) and KoateVR-DVI (Kedrion Biopharma, USA), which have a small infusion volume and can be given in the outpatient or home setting. 15–30 U/kg of 8Y has been used with success, although there is no standardized content of ADAMTS13 in these concentrates. Despite the fact that ADAMTS13 has a half-life of only 2–3 days, the clinical effect of plasma (15–20 mL/kg) or BPL 8Y are such that infusions every 2 weeks are sufficient to supply enough ADAMTS13 to maintain low but measurable plasma levels of the protease and thus avoid platelet consumption. Since relapse is not a constant finding, it is not firmly established which patients should undergo regular prophylaxis and which peak and trough ADAMTS13 levels are necessary to attain remission of the acute episode and prevent recurrence. There is as yet no definite evidence that patients with congenital ADAMTS13 deficiency treated with plasma or plasma products develop alloantibodies toward this protease, nor that an anamnestic response occurs in patients with autoantibodies.

In association to PEX, the most frequent immune pathogenesis of acquired TTP may be tackled by adjuvant immunomodulating agents (Table 43.5). Steroids are usually the first choice at the time of acute presentation (prednisone 1 mg/kg daily, methylprednisolone 1000 mg daily). The same treatment may be considered in the chronic recurrent form of TTP due to the persistence of autoantibodies. In some cases, usually in patients with a suboptimal response or clinical deterioration in spite of 5–6 days of PEX or with a chronic relapsing disease, second-line immunosuppression may be considered using such drugs as rituximab, ciclosporin, cyclophosphamide or azathioprine. Nowadays, the first choice is rituximab, an anti-CD20 antibody, which is efficacious in patients with primary refractory or relapsing TTP by blocking the production of anti-ADAMTS13 antibodies through the depletion of B lymphocytes. Generally administered once weekly at 375 mg/m^2 for four consecutive weeks, immediately after the apheresis procedure, approximately 95% of patients have a complete clinical and laboratory response, including a normal ADAMTS13 level and disappearance of anti-ADAMTS13 antibodies. Mild reactions can be controlled by premedication with steroids, antihistamines and analgesics, more serious complications being relatively uncommon. Relapses occur in approximately 10% of patients after intervals ranging from as little as 9 months to 4 years. Some clinicians suggest that rituximab should be even considered up front in conjunction with standard PEX therapy at the time of the

acute presentation of TTP, but this approach is not universally accepted. Despite the fact that ciclosporin can induce endothelial damage and cause a TTP-like picture, recent studies indicate a beneficial effect of this drug in relapsing or resistant TTP, by inducing durable remissions with normal recovery of ADAMTS13 activity and disappearance of ADAMTS13 inhibitors. Many other immunosuppressive agents have been used in the past (vincristine, cyclophosphamide), but they are usually considered third-line options, due to the risk of severe side effects and less established efficacy. Splenectomy is an alternative option in patients with chronic recurrent forms during remission, but carries a high risk of death in severely ill patients with acute TTP. The clinical efficacy of antiplatelet agents in TTP is unproven, and platelet transfusion is generally contraindicated.

A number of new therapeutic agents are being developed for the management of acute TTP: (i) an anti-VWF nanobody, caplacizumab, which inhibits the interaction between ultra-large VWF and platelet GpIb-IX-V, has achieved clinical proof-of-concept in a Phase II study, (ii) because of the emerging role for the complement cascade in TTP pathogenesis, eculizumab (see below) and other complement system modulators are under evaluation, (iii) a recombinant preparation of ADAMTS13 has been shown to improve the defective proteolysis of VWF *in vitro*, but this product is not yet available for clinical use.

Out of the acute phase, some patients with acquired TTP may require immunomodulation to prevent relapse. In the remission phase this approach is usually indicated in those with recurrence, but also in the presence of risk factors for recurrence (pregnancy, elective surgery, associated active autoimmune disorders, etc.) associated with persistent and severe ADAMTS13 deficiency and anti-ADAMTS13 antibodies. In these cases, the first choice is usually rituximab.

Haemolytic uraemic syndrome

The more typical and frequent form of Shiga-toxin mediated HUS (STEC-HUS) is acquired and occurs acutely and sporadically after gastrointestinal infections with toxin-producing bacteria, particularly in infants and young children, but also in adults. Atypical HUS can be familial, due to the inherited deficiency or dysfunction of complement components and may have a sporadic or chronic recurrent course. It occurs mainly in infants or young children, but at least one-third of cases are in adults, often in association with triggers such as delivery or drug intake.

STEC-HUS

This acute syndrome almost always occurs as a single sporadic episode, heralded by bloody diarrhoea from 2 days to 2 weeks before the onset of symptoms. Besides consumptive thrombocytopenia and mechanical haemolytic anaemia that are in common with TTP, the most peculiar symptoms are severe renal failure with oligoanuria. Serum creatinine and urea are definitely much more abnormal than in TTP, while thrombocytopenia is usually less severe. Signs of compensated DIC, with relatively high plasma levels of D-dimer, are another distinctive feature. Neurological involvement is uncommon, but coma and seizures may occur in association with uraemia and severe hypertension.

Aetiology

The most common bacterial agent that causes the prodromal gastrointestinal infection is *Escherichia coli* type 0157:H7 or O:104 or, less frequently, *Shigella dysenteriae* serotype I. These and other more rarely involved infectious agents produce exotoxins called verocytotoxin, shiga toxin or shiga-like toxins (Figure 43.3). These toxins, after absorption from the gastrointestinal tract into the blood, bind to the glycosphingolipid membrane receptor globotriaosylceramide, which is particularly dense in the glomerular capillary endothelial cells. The toxin–receptor complex is endocytosed and thereby causes cytolysis and extensive endothelial swelling and desquamation, which in turn engenders massive formation of microthrombi in the renal microvasculature, mainly composed of fibrin, but also rich in platelets.

Pathology and pathogenesis

The pathology of STEC-HUS is characterized by more extensive endothelial injury, but less VWF and more fibrin in thrombi than in TTP. The behaviour of plasma VWF is also different, because ultra-large multimers are detectable much more seldom during the acute phase of the disease, perhaps because multimers leaking in excess into plasma from damaged endothelial cells bind avidly to GPIb on the platelet membrane and are thereby removed from plasma, particularly those of larger size. There is agreement that ADAMTS13 is normal or only mildly reduced in the plasma of these patients (Table 43.6).

Natural history

In the great majority of cases, STEC-HUS is self-limiting, with much less tendency than TTP to early or late relapses. The acute episode is often accompanied by severe renal failure, so that temporary haemodialysis is frequently required. Residual renal dysfunction is common after the acute episode and end-stage renal disease requiring maintenance dialysis or kidney transplantation is needed in approximately one-quarter of patients.

Treatment

During the diarrhoeal prodrome, antimotility agents are contraindicated, because they favour the permanence of bacterial toxins in the gastrointestinal lumen and their passage into blood. The role of antibiotics is controversial, but in practice they are

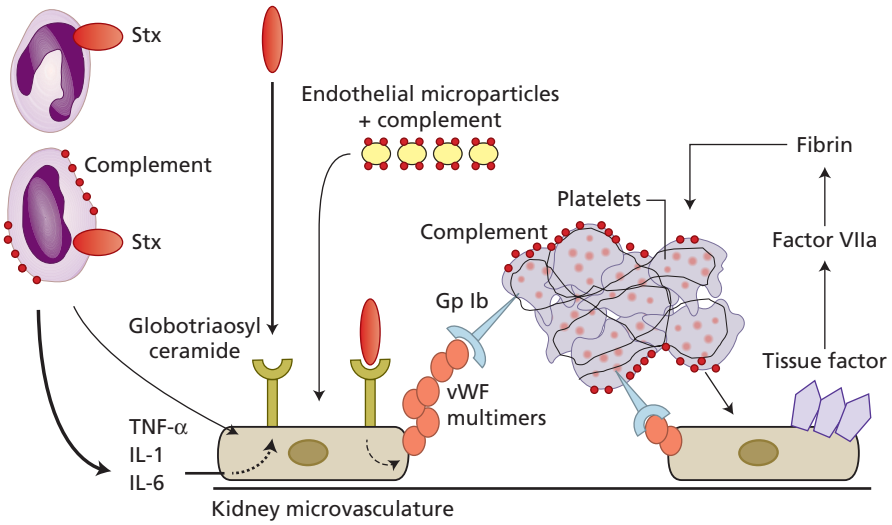


Figure 43.3 Pathophysiology of STEC-HUS. Verocytotoxin, shiga toxin or shiga-like toxins, after absorption from the gastrointestinal tract into the blood, bind to the glycosphingolipid membrane receptor globotriaosylceramide on the glomerular capillary endothelial cells. The toxin–receptor complex is endocytosed and causes cytolysis and endothelial swelling and desquamation, which leads to massive formation of microthrombi in the renal microvasculature, mainly composed of fibrin, but also rich in platelets. Gp Ib: glycoprotein Ib; IL: interleukin; Stx: Shiga toxin; TNF: tumour necrosis factor.

often used, with a strong indication at least for patients with evidence or suspicion of sepsis or *Shigella dysenteriae* infection. In cases of severe renal failure, early institution of dialysis is essential in order to reduce sequelae and death. PEX is usually not recommended in patients with STEC-HUS. A possible forthcoming approach to prevention is vaccination against vero/shiga toxins in children living in endemic areas.

Atypical HUS

The term atypical HUS (aHUS) is used for TMAs characterized by severe renal failure associated with the dysregulation of the alternative complement pathway. Given the key role of complement in aHUS, some now use the term aHUS to be synonymous with complement HUS. In this disease there is an excessive activation of C3 convertase, its partial consumption and production of C3a and C3b, which are deposited on the endothelial cell surface causing cell damage and destruction. C3a and C3b also activate other cells, such as platelets and leucocytes, causing inflammation and thrombosis (Figure 43.4). aHUS is a primarily prothrombotic condition that particularly affects the glomerular vessels. However, thrombosis can occur almost everywhere, mainly in organs such as the brain, lung and gastrointestinal tract. The attacks can be triggered by infections, pregnancy, drugs or trauma. In the last 20 years, dysregulation of the complement, mainly but not exclusively of genetic origin, has been linked to the development of aHUS. Approximately half of the patients have mutations in one of the complement regulatory proteins (factor H, factor I, MCP or factor B) resulting in excessive C3 activation with the consequences described above (Table 43.7). The type of gene mutation impacts the prognosis, patients with mutations in the factor H gene having the poorest prognosis in terms of short- and long-term renal function and relapse risk, even after transplantation (see below). On the other hand, patients with MCP mutations have a better prognosis. A few cases of aHUS have also been described to develop in association with acquired autoantibodies against factor H. Given that nearly half of patients with aHUS have no identified complement abnormalities, much remains to be learned pertaining to disease mechanisms and causes other than complement defects may be the culprit.

Table 43.6 Clinical applications of ADAMTS13 testing in TTP and aHUS.

Stage	ADAMTS13 deficiency (<10%)	Diagnostic implications
Acute phase	Yes	TTP
	No	Typical and aHUS/TTP cannot be excluded
Remission	Yes	Risk of relapse

ADAMTS13 levels are usually normal or only mildly reduced in the thrombotic microangiopathies listed in Table 43.2.

Laboratory diagnosis

Low serum C3 levels are a simple and widely available method for diagnosing affected patients and detecting family members at risk, but several patients diagnosed with aHUS fail to demonstrate this alteration. While mutations in complement proteins are important in the development of aHUS they are not sufficient, because for the actual precipitation of the disease triggering factors leading to increased complement activation are needed, such as infections, surgery, pregnancy or immune-mediated diseases.

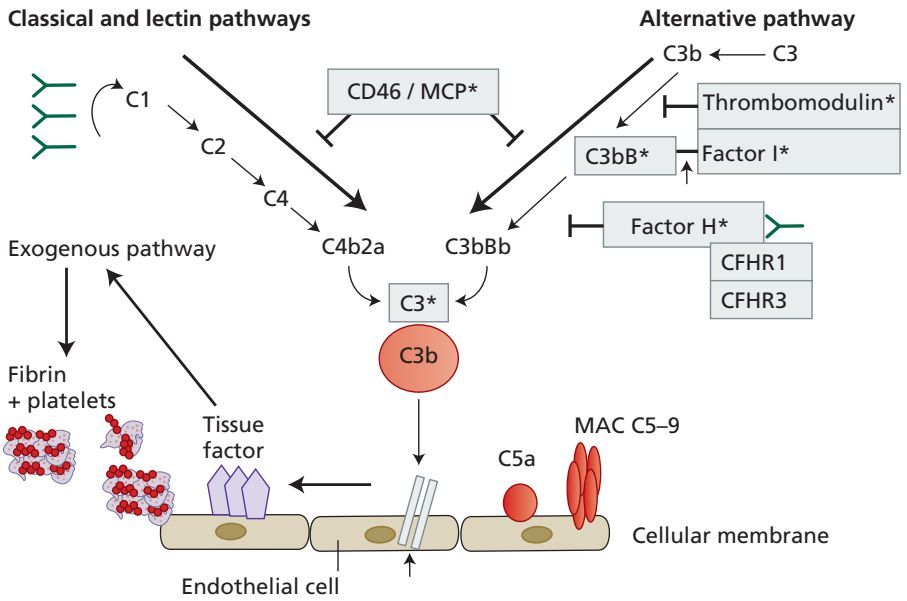


Figure 43.4 Pathophysiology of aHUS. An excessive activation of C3 convertase, its partial consumption and production of C3a and C3b, that are deposited on the endothelial cell surface cause cell damage and destruction. C3a and C3b also activate other cells, such as platelets and leucocytes, causing inflammation and thrombosis. CFHR, complement factor H-related; MAC, membrane attack complex; MCP, membrane cofactor protein, *mutations are reported in genes encoding these complement factors.

Table 43.7 Main complement factors and regulators and their function.

Complement factor or antibody	Function	Frequency of gene mutations/antibodies
CFH	• Plasma molecule that recognizes C3 and mediates cofactor activity for CFI • Directly accelerates the decay of C3-convertase	20–30
CFI	• Plasma serine protease • Cleaves C3b into inactive C3b	4–10
C3	• Its activation is required for both classical and alternative complement activation pathways • Its proteolytic cleavage produces C3a and C3b; C3a is an anaphylotoxin and the precursor of some cytokines and C3b serves as an opsonizing agent	5–10
Thrombomodulin	• Transmembrane protein • Involved in the generation of TaFI, which cleaves C3 and C5a	5
CFB	• Plasma serine protease • Its proteolytic cleavage produces the noncatalytic chain Ba and the catalytic subunit Bb • The active subunit Bb associates with C3b to form the alternative pathway C3 convertase • Bb is involved in the proliferation of preactivated B lymphocytes, while Ba inhibits their proliferation	1–2
MCP	• Integral transmembrane protein that binds C3b and serves as a cofactor for CFI	10–15
CFH antibodies		3–6

CFH, complement factor H; CFI, complement factor I; CFB, complement factor B; MCP, membrane cofactor protein.

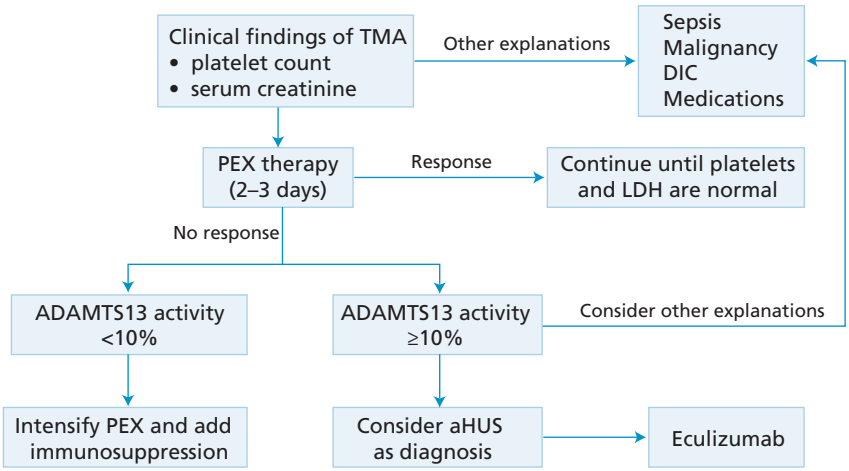


Figure 43.5 Management of patients with a suspected TMA. (Source: Cataland S and Wu, 2014. Adapted with permission of the American Medical Society.)

Natural history

aHUS is a disease with a poor prognosis, because until recently at least one-third of patients died or progressed to end-stage renal insufficiency despite the use of PEX. aHUS is often a permanently ongoing systemic disease, because after recovery from the acute illness patients may suffer from other acute episodes if exposed to clinical triggers (infection, surgery, pregnancy) owing to their persistent inability to regulate complement activity. Very recent advances in treatment let us hope that this gloomy situation will substantially improve in the next few years (see below)

Treatment

Until recently treatment of aHUS was based on plasma therapy (PEX or infusion). The rationale of this therapeutic approach was the replacement of deficient or dysfunctional complement proteins, but also the removal of activated complement factors and anti-factor H autoantibodies when present. Often plasma therapy has to be combined with haemodialysis to control oligoanuria and uraemia.

PEX therapy should be started as soon as possible in patients presenting with a suspected diagnosis of aHUS (Figure 43.5). The response to PEX is typically re-assessed at days 2–3, with the goal to establish whether or not there is haematologic improvement, and improvement or stabilization of end-organ damage. In patients responding to PEX, this therapy is continued daily until normalization of those surrogate markers for ongoing microthrombotic disease such as the platelet count, haemoglobin and LDH, and then stopped without prolonged tapering. Markers of kidney damage (serum creatinine) and neurologic findings and symptoms, when present, are of great importance. A poor response to PEX can be defined as lack of haematologic response, but also by worsening kidney insufficiency, despite daily PEX therapy. In these cases PEX should be discontinued and therapy with eculizumab should be considered.

Eculizumab, a humanized monoclonal antibody recently approved for the treatment of aHUS, binds to complement protein C5, blocks the formation of the membrane attack complex and is an effective life- and kidney-sparing therapy for many patients with aHUS. This product has been prospectively evaluated for efficacy and safety in many aHUS patients with a broad spectrum of disease features (patients recently diagnosed, patients with long disease duration and patients who had already undergone transplant or on continuous dialysis). Therapy did often result not only in haematologic improvement, but also, and most importantly, in the marked improvement of renal function, even in cases with advanced disease. Despite some data suggesting that the more rapid initiation of therapy with eculizumab is associated with greater improvements in renal function, its initiation prior to a window of 2–3 days of PEX is not advised due to the difficulty of differentiating with certainty aHUS from TTP at the time of presentation. In cases where the diagnosis of aHUS is known at the time of acute presentation (such as, for instance, for recurrences), first-line therapy with eculizumab should be considered instead of PEX.

Even though eculizumab is usually well tolerated, all patients should receive a meningococcal vaccine because of the increased risk for this infection in treated patients. Intravenous eculizumab is initiated at a dose of 900 mg weekly for the first 4 consecutive weeks, and then beginning on week 5 maintenance therapy is administered every other week at a dose of 1200 mg. These are the presently recommended dosing schedules, but perhaps in the future dosages of this extremely expensive therapy may be reduced. Eculizumab is often associated with dramatic clinical responses in patients refractory to PEX and, most importantly, such a high degree of recovery of renal function that at least 80% of patients could discontinue dialysis. These findings are in contrast with the previous findings that at least one-third of patients did die or progress to end-stage renal disease, despite PEX. It is of extreme importance to avoid permanent dialysis and renal transplantation, because the results of the

latter are sometime poor and appear to correlate with the cause of complement dysregulation. This disease has not recurred in the transplanted kidneys of patients with MCP gene mutations, presumably because the new kidney produces protective levels of endogenous MCP. In contrast, kidney transplantation alone did not correct the deficiency of plasma factors made in the liver, so that the outcome was poor in patients with factor H or factor I mutations owing to disease recurrence. However, a formidable approach such as combined liver and kidney transplantation may lead to a favourable long-term outcome for recurrent HUS associated with factor H gene mutations.

Concluding remarks

In the last few years, TTP and HUS, known for several decades to clinicians and pathologists for their dire consequences, have witnessed spectacular improvements in our knowledge and development of new diagnostic criteria. Most importantly, improved therapeutic options have led to a marked decrease in mortality and morbidity, particularly for atypical HUS following the availability of eculizumab. In TTP, the pivotal role of VWF and of its major proteolytic enzyme ADAMTS13 in inducing microvascular platelet thrombi is evident, even though not all the variants of TTP completely fit this disease mechanism. The development of laboratory methods to measure ADAMTS13 in plasma has demonstrated that the majority of TTP cases have very low or unmeasurable levels of this protease. On the other hand, severe ADAMTS13 deficiency is not found in all patients who meet the clinical criteria for a diagnosis of TTP (Table 43.6), questioning the appropriateness of the currently available functional assays performed in non-physiological static conditions. Trigger(s) are usually needed for the occurrence of full-blown TTP, because patients with congenital and acquired low levels of ADAMTS13 may not manifest clinical disease for long periods of time. The clinician confronted with the treatment of these seriously ill patients should be cognizant that plasma therapy, mainly plasma exchange, remains the treatment of choice for patients with clinically diagnosed TTP. An important, not completely resolved therapeutic problem is the secondary prevention of the chronic recurrent forms. Various immunomodulatory treatments have been tried with promising results, particularly rituximab.

STEC-HUS is now well characterized from an aetiological standpoint, so that it can be specifically recognized and properly treated. The dire consequences of atypical HUS are dramatically improved by the current availability of a complement inhibition therapy with the humanized monoclonal antibody eculizumab, that is able not only to induce a dramatic therapeutic response in many patients unresponsive to plasma therapy, but also the recovery of renal function, in contrast with the poor prognosis for many patients treated only with plasma-based therapy. Despite much progress, there are numerous problems left to solve and, given the rarity of TTP and HUS, clinical research must be necessarily multicenter and multidisciplinary.

Selected bibliography

Cataland S and Wu HM (2014) How I treat: the clinical differentiation and initial treatment of adult patients with atypical hemolytic uremic syndrome. *Blood* **123**: 2478–84.

Hie M, Gay J, Galicier L *et al.*; French Thrombotic Microangiopathies Reference Centre (2014) Preemptive rituximab infusions after remission efficiently prevent relapses in acquired thrombotic thrombocytopenic purpura. *Blood* **124**: 204-10.

Hovinga JA, Vesely SK, Terrell DR, Lämmle B, George JN (2010) Survival and relapse in patients with thrombotic thrombocytopenic purpura. *Blood* **115**: 1500–11.

Lotta LA, Wu HM, Mackie IJ *et al.* (2012) Residual plasmatic activity of ADAMTS13 is correlated with phenotype severity in congenital thrombotic thrombocytopenic purpura. *Blood* **120**: 440–8.

Mannucci PM (2015) Understanding organ dysfunction in thrombotic thrombocytorpelic purpura. *Intensive Care Medical* **41**: 715–18.

Meri S (2013) Complement activation in diseases presenting with thrombotic microangiopathy. *European Journal of Internal Medicine* **24**: 496–502.

Moake JL, Rudy CK, Troll JH *et al.* (1982) Unusually large plasma factor VIII: von Willebrand factor multimers in chronic relapsing thrombotic thrombocytopenic purpura. *New England Journal of Medicine* **307**: 1432–5.

Peyvandi F, Palla R, Lotta LA *et al.* (2010) ADAMTS13 assays in thrombotic thrombocytopenic purpura. *Journal of Thrombosis and Haemostasis* **8**: 631–40.

Scully M, Hunt BJ, Benjamin S *et al.* (2012) Guidelines on the diagnosis and management of thrombotic thrombocytopenic purpura and other thrombotic microangiopathies. *British Journal of Haematology* **158**: 323–35.

CHAPTER 44

Heritable thrombophilia

44

Trevor Baglin[1] and David Keeling[2]

[1]Cambridge University Hospitals NHS Trust, Addenbrookes Hospital, Cambridge, UK
[2]Oxford University Hospitals, Churchill Hospital, Oxford, UK.

Introduction

In most instances heritable thrombophilia is a late-onset genetic disease with low penetrance. There is no internationally accepted definition of thrombophilia. The term is sometimes used to describe disorders of haemostasis detected in the laboratory that appear to predispose to venous thrombosis. A clinical definition of heritable thrombophilia describes an inherited tendency for venous thrombosis, i.e. deep vein thrombosis (DVT) with or without associated pulmonary embolus (PE). The incidence of a first episode of venous thrombosis is 1.5 per 1000 person-years, with a per-person lifetime incidence of 5%. Venous thrombosis is a multicausal disease in which genetic and acquired risk factors interact synergistically. Understanding the genetic risk factors and gene–environment interactions is necessary to understand why an individual develops thrombosis at a specific point in time. Interaction occurs when risk factors combine to produce an effect that exceeds the sum of their separate effects. Familial thrombophilia is likely a multiple gene disorder and family members may have one, or more, identifiable defect(s), probably in combination with additional unknown defects. It is likely that this largely explains why a dichotomous testing strategy with a limited number of factors (e.g. presence or absence of factor (F) V Leiden) has limited predictive value, resulting in little if any clinical utility for most patients (clinical utility refers to the ability of a screening or diagnostic test to improve clinical outcome).

Deficiency of the natural anticoagulant antithrombin was the first reported inherited risk factor for venous thromboembolism. Since then deficiencies of the naturally occurring anticoagulants protein C and protein S have been linked with familial thrombosis. In recent years, several other potential thrombophilic risk factors have been investigated but only the *F5*G1691A (FVR506Q, FV Leiden) and the *F2*G20210A gene mutations have been shown to be unequivocally associated with an increased risk of thrombosis, i.e. odds ratio (OR) of 2 or greater from case–control studies. In the 1980s and 1990s thrombophilia testing became common in unselected patients and their relatives, despite the fact that there was no evidence that testing had clinical utility. It is now apparent that testing for heritable thrombophilia typically does not predict likelihood of recurrence in the majority of unselected patients with symptomatic venous thrombosis (Figure 44.1). Furthermore, testing for inherited thrombophilia did not reduce recurrence of venous thrombosis in a large cohort study. There is a low risk of thrombosis in affected asymptomatic relatives of patients with thrombophilic defects followed prospectively and the results of thrombophilia tests are frequently misinterpreted. Given these considerations the clinician must consider the potential clinical utility of testing for heritable thrombophilic defects in each individual patient or family affected by venous thrombosis.

Heritable thrombophilias associated with venous thrombosis

The heritable thrombophilias shown to be associated with at least a twofold increased risk of venous thrombosis are: (i) deficiencies of the natural anticoagulants antithrombin, protein C and protein S, due to mutations in the corresponding genes *SERPINC1*, *PROC* and *PROS*, and (ii) two common mutations in genes encoding procoagulant factors, namely *F5*G1691A (FVR506Q, FV) and *F2*G20210A (commonly referred to as the

Postgraduate Haematology, Seventh Edition. Edited by A Victor Hoffbrand, Douglas R Higgs, David M Keeling and Atul B Mehta.

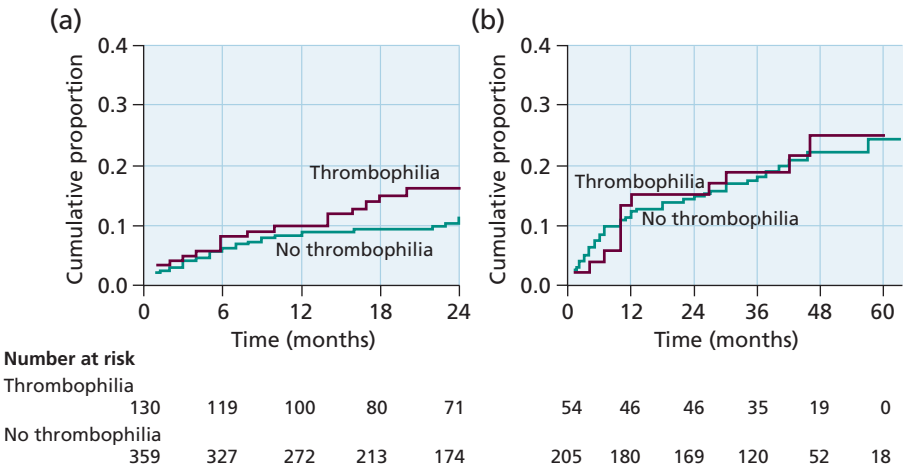


Figure 44.1 Cumulative recurrence of venous thrombosis in consecutive patients with and without heritable thrombophilia followed from the end of anticoagulation after a first episode of venous thrombosis in the first (a) and second (b) Cambridge cohort studies. The hazard ratio for recurrence in patients with thrombophilia compared with those without was 1.3 (95% CI 0.8–2.0) and 1.0 (95% CI 0.5–2.1), respectively. In the second cohort, patients with postoperative thrombosis were excluded, which may account for the apparent slightly higher recurrence rate. (For related publications on these cohorts, see Baglin *et al.* 2003 and Baglin T, Palmer CR, Luddington R, Baglin C (2008) Unprovoked recurrent venous thrombosis: prediction by D-dimer and clinical risk factors. *Journal of Thrombosis and Haemostasis* **6:** 577–82.)

prothrombin gene mutation). The causal association between these heritable thrombophilias and venous thrombosis has been confirmed by comparing the prevalence of defects in patients with venous thrombosis and controls. In such studies, measurement errors between cases and controls do not arise in relation to genotype and so the case–control study is the method of choice for measuring simple associations between genotype and disease. The prevalence of the heritable thrombophilias in cases and controls is shown in Table 44.1. The increased risk of venous thrombosis is often expressed as the relative risk (RR) (the probability of the disease in those with the risk factor divided by the probability of the disease in those without the risk factor) or the odds ratio (OR) (the odds of the disease in those with the risk factor divided by the odds of the disease in those without the risk factor). The RR for each defect is indicated in Table 44.1. Deficiencies of antithrombin, protein C and protein S are rare and so the estimates of risks are uncertain, but appear to be of the order of about a 10-fold increased risk of venous thrombosis. The FVR506Q and *F2*G20210A occur much more frequently and so the estimates of risk are more certain.

Gene–environment interaction is present when the effect of genotype on disease risk depends on the level of exposure to an environmental factor. Heritable thrombophilia is subject to a strong gene–environment interaction. For example, there is a synergistic increase in risk of venous thrombosis due to an interaction between the factor V Leiden mutation and oestrogen

Table 44.1 Prevalence of heritable thrombophilias shown to be associated with at least a twofold increased risk of venous thrombosis in patients and controls and relative risk of a first episode of venous thrombosis.

	General population	Consecutive patients with a first episode of venous thrombosis	Relative risk of venous thrombosis
Antithrombin	0.03%	1%	10–20
Protein C deficiency	0.3%	3%	10
Protein S deficiency	0.3%	3%	?–10
FVR506Q (FV Leiden)*	4%	15%	4
*F2G20210A**	2%	4%	2

*The prevalence of FVR506Q and *F2*G20210A is determined by ethnic origin as these mutations are only prevalent in whites.

exposure and obesity. Another example is observed in patients with antithrombin deficiency where venous thrombosis frequently occurs when an affected patient is exposed to an environmental risk. This was evident in a study of families with type 1 antithrombin deficiency, where the incidence of venous thrombosis was 20 times greater in affected family members, but was strongly dependent on acquired risks. In this study the annual incidence of venous thrombosis in affected family members in any year in which they were exposed to surgery, trauma, plaster cast, hospitalization or immobilization was 20% (provoked venous thrombosis), but in any year in which there was no exposure the incidence was only 0.3% (i.e. unprovoked venous thrombosis rate). In this study the unprovoked venous thrombosis rate is only slightly higher than the background incidence in the general population.

Antithrombin deficiency

The gene for antithrombin (*SERPINC1*) is located on the long arm of chromosome 1 and contains seven exons spanning 13.5 kb. Antithrombin (previously termed antithrombin III) is a glycoprotein (464 amino acids, 58 kDa, plasma concentration 2–3 μmol/L). Antithrombin is a serpin (serine protease inhibitor) and its primary targets are thrombin, factor FXa and FIXa. Its concentration in the circulation is higher than that of prothrombin, and as only a fraction of prothrombin is converted to thrombin during coagulation, antithrombin is potentially present in vast excess over the protease, although the site will determine the local concentration of the enzyme and inhibitor. The plasma half-life of antithrombin is about 90 hours.

The serpin mechanism was confirmed by successful resolution of a crystal structure of the final serpin–protease complex by Jim Huntington in 2000. In contrast to the classical lock-and-key mechanism of protease inhibition, serpins form a stable covalent complex in which there is a dramatic conformational change in both the inhibitor and the inhibited protein that alters the properties of each, resulting in rapid clearance from the circulation. Antithrombin consists of three β-sheets and nine α-helices organized into an upper β-barrel domain and a lower helical domain. These two domains are bridged by the main structural feature of a serpin, the central A sheet (shown in red in Figure 44.2). The flexible reactive centre loop (RCL, shown in yellow) contains the scissile bond formed by the P1 arginine and P1′ serine. Amino acids in the RCL are designated P1, P2 and so forth, moving towards the N-terminus of the peptide sequence and P1′, P2′ and so forth moving towards the C-terminus of the sequence. The length and flexibility of the loop renders it highly susceptible to proteolysis. Cleavage of the RCL by thrombin, FXa or FIXa produces a conformational change during which the RCL inserts as the central strand of the expanded central A sheet (Figure 44.2c). This results in translocation of the protease to the opposite pole of the serpin. The incorporation of the RCL as the central strand produces a hyperstable molecule such that the protease is crushed by intermolecular clashes with the hyperstable serpin. Distortion of the active site architecture disrupts the catalytic triad and results in disorganization of almost 40% of the protease architecture. The poor inhibitory activity of the circulating native form of antithrombin is conferred by intramolecular contacts that restrain the RCL and orientate the P1 arginine away from attacking proteases (Figure 44.2a). Activation causing release of the RCL is dependent on binding of glycosaminoglycans to the 'heparin-binding site' located in the region of the D-helix (cyan in Figure 44.2). An initial low-affinity interaction induces high-affinity binding (referred to as the induced fit mechanism) and this produces a conformational change in the antithrombin that releases the RCL for protease attack (Figure 44.2b). Activation increases the rate of protease inactivation 1000-fold. Successful inhibition of the protease depends on a race between complete proteolysis of the RCL with release of the regenerated protease from the cleaved loop (the substrate pathway) and incorporation of the cleaved loop into the central A-sheet with resultant capture and crushing of the tethered protease (the inhibitory pathway). The rate of inhibition relative to the substrate reaction is defined as the stoichiometry of inhibition, where a value of 1 represents a perfect inhibitor and a value of 2 would indicate that half the reactions led to cleaved antithrombin and a re-generated protease. Some mutations in the *SERPINC1* gene produce a dysfunctional antithrombin molecule in which the insertion of the RCL into the central A-sheet is retarded, turning the molecule into a substrate rather than an inhibitor. This is an example of how this highly regulated mechanism of thrombin or FXa inhibition renders antithrombin susceptible to dysfunction through genetic mutation.

Two laboratory phenotypes of heritable antithrombin deficiency are recognized. Type 1 is characterized by a quantitative reduction of antithrombin, with a parallel reduction in function (measured as inhibitory activity) and level of protein in the plasma (measured immunologically as the antigenic level). Type 2 deficiency is due to the production of a qualitatively abnormal antithrombin protein with dysfunction due to disturbance of the complex mechanism of protease inhibition by a mutation in *SERPINC1* gene. The functional activity is discrepantly low compared with the antigenic level. Type 2 deficiency is subclassified according to the nature of the functional deficit:

- Type 2 reactive site (RS): mutations alter the sequence of the RCL, reducing the ability to inhibit thrombin or FXa in the presence or absence of heparin in a laboratory assay.
- Type 2 heparin-binding site (HBS): mutations affect the ability of antithrombin to bind and be activated by glycosaminoglycans and therefore result in reduced ability to inhibit thrombin or FXa only in the presence of heparin in a laboratory assay. These mutations are typically located in the designated 'heparin-binding site' in the region of the D helix of the molecule.
- Type 2 pleiotropic (PE): a single mutation produces multiple effects on the structure–function relationship of the molecule,

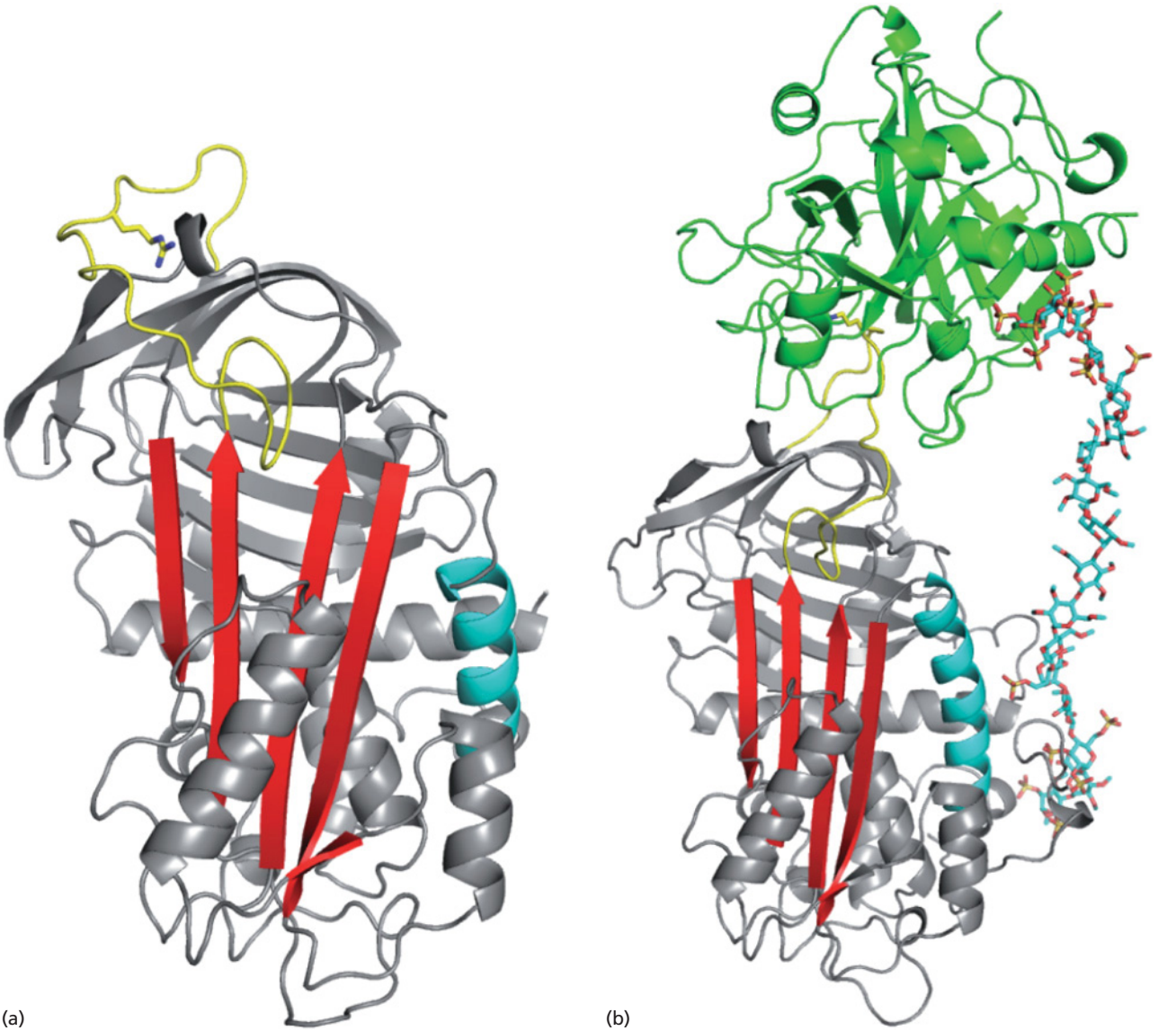
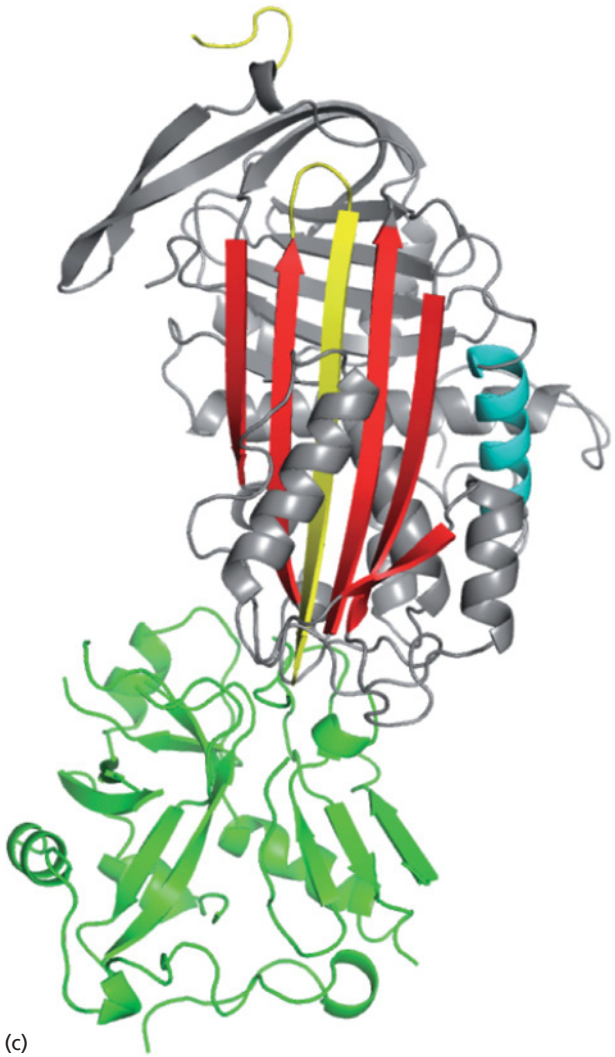


Figure 44.2 Model of thrombin destruction by antithrombin. Native antithrombin (a) circulates predominantly in an inert configuration with the reactive centre loop (RCL, shown in yellow) restrained on the C-sheet with the P1 arginine buried in a hydrophobic pocket and inaccessible to proteolytic attack. Binding of a glyscosaminoglycan such as heparin sulfate to helix D (shown in cyan) induces a conformational change in the molecule that results in closure of the central A-sheet (shown in red), release of the RCL and exposure of the P1 arginine for proteolytic attack by protease (shown in green) resulting in the formation of an encounter complex (b). A heparin chain is shown bridging antithrombin and thrombin. The heparin chain binds to the D-helix of antithrombin, inducing the conformational change and stabilizing the resultant encounter complex by binding to exosite II on thrombin. Cleavage of the RCL results in opening of the central A-sheet and incorporation of the cleaved loop as the central strand (c). This movement throws the protease to the opposite pole of the molecule where it is crushed by intermolecular clashes with antithrombin, which is now in a hyperstable conformation. Inhibition of the protease depends on incorporation of the cleaved loop into the central A-sheet with resultant capture and crushing of the tethered protease with distortion of the catalytic triad before completion of proteolysis of the RCL, which would release the protease. (Source: Dr Jim Huntington, Department of Haematology, University of Cambridge. Reproduced with permission.)



(c)

Figure 44.2 *(Continued)*

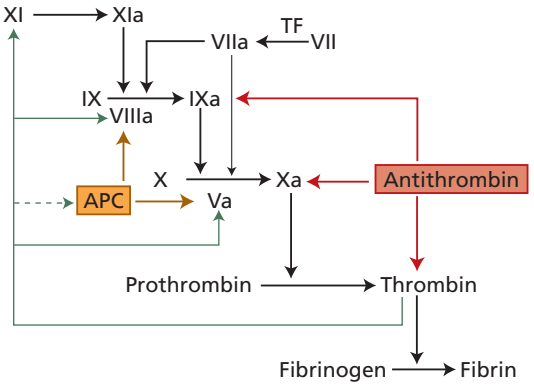


Figure 44.3 The thrombin explosion occurs in two phases: initiation and propagation. In the initiation phase, tissue factor (TF)–FVIIa generates small amounts of thrombin that enhances platelet activation (in conjunction with exposed subendothelial collagen at the site of vascular injury) and cleaves FV and FVIII to their active forms. The activation of FVa and FVIIIa amplifies thrombin generation and increasing concentrations of thrombin then lead to back-activation of FXI and the propagation phase. The activated forms of FV and FVIII are non-enzymatic cofactors that assemble and orientate the enzymes (FIXa and FXa) on the negatively charged surface of activated platelets. Antithrombin can neutralize the enzymatic coagulant factors FIXa, FXa and FIIa (thrombin). Thrombin bound to thrombomodulin cleaves protein C to its active form (APC), which proteolysis FVa and FVIIIa and dismantles the tenase (FVIIIa–FIXa) and prothrombinase (FVa–FIIa) complexes on phospholipid surfaces, which attenuates thrombin generation. Protein S is a cofactor for the APC-mediated inactivation of FVa and FVIIIa.

often associated with low plasma levels due to either effects on secretion or instability.

Approximately 100 point mutations (missense, nonsense or insertions or deletions causing frameshifts) and several whole or partial gene deletions have been identified as causes of type 1 deficiency. Numerous point mutations causing qualitative type 2 deficiency have been identified. Homozygous type 1 deficiency and probably homozygous type 2 RS mutations are incompatible with life. Homozygous type 2 HBS and PE mutations are associated with a lower risk of thrombosis and are compatible with life.

Protein C deficiency

The gene for protein C (*PROC*) is located on chromosome 2 and contains nine exons spanning 11 kb. Protein C is a vitamin-K-dependent glycoprotein synthesized in the liver (419 amino acids, 62 kDa, plasma concentration 65 nmol/L). Its concentration in plasma is about 100 times less than that of antithrombin. The plasma half-life of protein C is about 6 hours.

Protein C is the zymogen precursor of activated protein C (APC). Protein C is activated by thrombin bound to thrombomodulin on the endothelial surface. This is facilitated by the endothelial protein C receptor (EPCR). APC acts as a physiological anticoagulant by cleaving the cofactors FVa and FVIIIa and so thrombin generation (which is dependent on FVa and FVIIIa cofactor activity) is attenuated (Figure 44.3). Protein S is a cofactor for APC-mediated inactivation of FVa and FVIIIa.

Protein C deficiency is classified into type 1 and 2 defects on the basis of functional and antigenic assays. The relative risk of thrombosis in relation to type 1 and the various type 2 defects has not been characterized. Most heritable protein C deficiency is due to type 1 deficiency. The majority of type 1 defects are due to point mutations. Multiple type 2 defects have been reported involving the catalytic active site, the phospholipid-binding Gla domain, the propeptide cleavage activation site and sites of interaction with substrates or cofactors.

Protein S deficiency

The gene for protein S (*PROS*) is located on chromosome 3 and contains 15 exons spanning 80 kb. A pseudogene with approximately 95% homology is located adjacent to *PROS* and this can cause difficulty with identification of mutations within the *PROS* gene.

Protein S is a vitamin-K-dependent glycoprotein produced in the liver, endothelial cells and megakaryocytes (676 amino acids, 69 kDa, plasma concentration of free protein 145 nmol/L). Protein S is a non-enzymatic cofactor for APC-mediated inactivation of FVa and FVIIIa. Approximately 60% circulates bound to C4b-binding protein and is inactive. The remaining 40%, designated free protein S, is uncomplexed and is the active form. Free protein S increases the affinity of APC for negatively charged phospholipid surfaces on platelets or the endothelium and increases complex formation of APC with the activated forms of FV (FVa) and FVIII (FVIIIa). The binding of protein S by C4b-binding protein is very strong and free protein S is essentially the molar excess of protein S over C4b-binding protein. In addition to APC cofactor activity, protein S has an independent anticoagulant effect by acting as a cofactor for TFPI (tissue factor pathway inhibitor).

Protein S defects have been divided into three types. In type I deficiency, both total and free protein S antigen levels are low (and functional activity is low if measured). A type III deficiency characterized by a normal level of total protein S, but a reduced level of the free protein S has been described. This may be a phenotypic variation resulting from the same genetic mutations associated with type I deficiency (not recognized with the less sensitive assay for total protein S), but many patients with an apparent type III phenotype do not have heritable protein S deficiency and this may be related to the increase in C4b levels with age. Type II defects are characterized by reduced activity in the presence of normal total and free levels of protein S. Type II deficiency is difficult to diagnose because functional protein S assays are imprecise.

FVR506Q (FV)

The gene for FV is located on chromosome 1, near the *SERPINC1* gene, and contains 25 exons spanning 80 kb. FV (2196 amino acids, 330 kDa, plasma concentration of free protein 30 nmol/L) is the cofactor for activation of prothrombin by FXa in the prothrombinase complex (Figure 44.3). The plasma half-life is 15 hours. It has no cofactor activity until cleaved by thrombin or FXa at Arg709, Arg1018 and Arg1545. In 1993, Bjorn Dahlbach and colleagues in Malmo described the phenomenon of resistance to APC in unrelated patients with venous thrombosis. Patient plasma was resistant to the anticoagulant effect of a fixed amount of added exogenous APC. The effect was measured by dividing the activated partial thromboplastin time (APTT) in the presence of added APC by the APTT in the absence of added APC. APC resistance was defined as a ratio less than 2.0. Further work confirmed that increased APC resistance cosegregated with thrombosis in families with familial venous thromboembolism. The following year it was reported by a team from Leiden that the majority of patients with familial APC resistance had the same point mutation in the gene for FV, a guanine to adenine transition at nucleotide 1691 in exon 10 of the *F5* gene. This is known as the FV Leiden mutation (FVR506Q). Mutant FV Leiden has normal procoagulant activity, but substitution of glutamine for arginine at position 506 results in slower inactivation by APC. Prothrombinase activity is dependent on FVa maintaining the integrity of the enzymatic complex on a phospholipid surface. Thus thrombin generation is rate limited by FVa, and inactivation of FVa by APC attenuates thrombin generation. Inactivation of FVa occurs through cleavage of the protein at positions Arg306, Arg506 and Arg679 by APC. Initial cleavage at R506 is necessary for subsequent cleavage at R306 and R679. Mutation of these residues (*F5*G1691A, FVR506Q or FV Leiden; *F5*G1091C, FVA306T or FV Cambridge) results in resistance to inactivation by APC. APC resistance is observed in the laboratory as an impaired plasma anticoagulant response to APC added *in vitro*, usually with an APTT-based clotting test. However, nowadays the mutation is frequently detected by direct DNA analysis rather than by a clotting assay (see section on laboratory testing). The FVR506Q mutation is present in about 4% of the white population and about 15% of unselected consecutive patients with a first venous thrombosis. The prevalence is highest in northern Europeans. The mutation is infrequent in other populations. The high prevalence and founder effect suggests positive selection and this may relate to a favourable effect on embryo implantation and hence reproduction rather than a lower risk of fatal haemorrhage in females during childbirth, as originally thought.

APC also inactivates FVIIIa by cleavage at positions Arg336 and Arg562, but mutations affecting these sites in FVIII causing an increased risk of venous thrombosis have not been reported.

APC resistance in the absence of FVR506Q occurs in about 5% of patients with a history of venous thrombosis. This is often related to high FVIII levels or low FV levels, which may be found in association with the *F5*HR2 haplotype. There is evidence that APC resistance in the absence of FVR506Q is associated with venous thrombosis, but defining the cause of resistance in individual patients has no clinical utility. Acquired APC resistance is explained by increased FVIII levels, pregnancy or oestrogen exposure in some individuals.

*F2*G20210A

The gene for prothrombin is located on chromosome 2 and contains 14 exons spanning 21 kb. Prothrombin is a vitamin-K-dependent protein synthesized in the liver (579 amino acids, 72 kDa, plasma concentration 2 μmol/L). Prothrombin is the zymogen precursor of thrombin.

A single nucleotide change, guanine to adenine at position 20210 in the 3′ untranslated region of the prothrombin gene (*F2*G20210A), is associated with elevated plasma prothrombin levels and an increased risk of venous thrombosis. This mutation was identified by the Leiden team when investigating candidate genes in the Leiden Thrombophilia Study cohort. The prevalence of the *F2*G20210A mutation is about 2% of whites, with a higher prevalence in southern compared with northern Europeans. The mutation increases the plasma level of prothrombin by about 30%, but the mechanism has not been explained. No specific clotting test for the presence of the mutation has been described and diagnosis depends on detection of the genetic mutation by DNA analysis.

Genome-wide association studies (GWAS) and deep sequencing of candidate genes

Peter Reitsma and Frits Rosendaal from Leiden have detailed the history of unravelling the genetic architecture of venous thrombosis risk. Initially, extended pedigrees of families with venous thrombosis were studied and plasma levels of natural anticoagulants were measured. With the advent of DNA technology, genes encoding the natural anticoagulant proteins were cloned and sequenced and hundreds of private mutations were reported, often with mutations found only in one family. Subsequently, the common public mutations (*F5*G1691A and *F2*G20210A) were found in case–control studies. Currently, common genetic markers in every gene encoding a protein involved in haemostasis are being studied in order to determine whether these markers are associated with thrombosis risk. Two types of genetic mutation cause venous thrombosis: those that reduce function (e.g. a mutation causing antithrombin deficiency) and those that increase function (e.g. FVR506Q, which renders the mutant protein resistant to inactivation, resulting in increased thrombin generation). The private mutations in *SERPINC1*, *PROC* and *PROS* and public mutations in *F5* and *F2* account for about 50% of the thrombosis risk attributable to genetic mutation. There are two approaches to identifying the remaining 50% of genetic mutations. A linkage study requires extended pedigree analysis and is limited by heterogeneity in the genetic cause of thrombosis in different families. This dilutes the apparent effect of a genetic mutation when data from different pedigrees are combined. The alternative approach is to use genetic markers of common risk alleles in association studies of unrelated individuals using genome-wide genotyping of a large set of single-nucleotide polymorphisms (SNPs). This approach relies on the principle that common diseases are explained by common risk alleles. The weakness of this approach is that there is little evidence that this principle is true. With this approach there is also a high risk of misidentifying a risk allele (type I error) due to multiple testing and in order to offset this it is necessary to raise the significance level of testing, which then leads to decreased sensitivity (type II error). Reitsma and Rosendaal suggest that a deep re-sequencing project may identify genetic factors and allow powerful genetic risk profiling. This hypothesis is supported by an initial proof of principle study showing that common SNPs, which in isolation confer a very small risk of venous thrombosis (hazard ratio <1.5), interact synergistically to significantly increase the risk of venous thrombosis (hazard ratio >5.0 for four SNPs).

Other natural anticoagulants

Tissue factor pathway inhibitor (TFPI) is an inhibitor of FXa and tissue factor-bound FVIIa. TFPI knockout is embryonically lethal in mice. TFPI is predominantly extravascular and so measurement in plasma samples is not representative of total TFPI availability. In a few studies plasma TFPI levels were slightly lower in patients with venous thrombosis than in normal individuals and the low levels may be a mild risk factor for venous thrombosis. Protein Z is a vitamin-K-dependent protein that circulates in complex with PZI (protein-Z-dependent protease inhibitor) and catalyses the inhibition of FXa. A causal relationship between PZI and/or PZ deficiency and venous thrombosis has not been demonstrated. Heparin cofactor II, a serpin with a similar structure–function relationship to antithrombin, inhibits thrombin when activated by glycosaminoglycans. Heparin cofactor II knockout mice are healthy and do not develop spontaneous thrombosis, but do have a shorter time to thrombosis in an arterial injury model. A relationship between heparin cofactor II deficiency and thrombosis in humans is uncertain. A relationship between thrombosis and mutations affecting thrombomodulin or EPCR has not been established.

Other procoagulant factors

Increased FVIII levels are associated with an increased risk of both venous and arterial thrombosis, but a heritable basis for this, other than through ABO blood group effects on von Willebrand factor levels, has not been established. Elevated FIX and FXI levels are also associated with venous thrombosis risk, but a heritable basis for high levels associated with venous thrombosis is not established. There is equivocal evidence for a causal relationship between fibrinogen levels and venous thrombosis. Approximately 300 examples of heritable dysfibrinogenaemia have been reported. The majority are asymptomatic or associated with increased bleeding and very few examples are considered to be genuinely associated with an increased tendency to thrombosis, which may be arterial or venous or a combination. Dysfibrinogenaemia has been found in less than 1% of patients with a history of venous thromboembolism. It is associated with postpartum thrombosis and an increased risk of pregnancy loss. The identification and characterization of rare dysfibrinogenaemias associated with thrombosis is beyond the capability of most coagulation laboratories. Polymorphisms in the prothrombin gene have been described that may further increase

the risk of venous thrombosis associated with the *F2*G20210A mutation, but the effect is mild. It was previously thought that deficiency of FXII was a risk factor for venous thromboembolism, but subsequent investigation strongly indicates that this is unlikely. A polymorphism in the FXIII gene (FXIIIV341L) is significantly less common in patients with coronary heart disease than in control subjects, and a protective effect for venous thromboembolism has been reported.

Fibrinolysis

A causal relationship between levels of individual specific proteins involved in regulating fibrinolysis has not been established. However, a global measure of fibrinolytic potential as measured by a plasma-clot lysis assay was associated with an approximately twofold increased risk of venous thrombosis in patients with clot lysis times above the 90th percentile of controls in the Leiden Thrombophilia Study. Analysis of the large MEGA (Multiple Environmental and Genetic Assessment) study confirmed this finding and demonstrated that hypofibrinolysis in combination with established acquired and genetic risk factors, such as FVR506Q, had a synergistic effect on venous thrombosis risk. A genetic basis for hypofibrinolysis in these patients was not investigated. Genetic mutation in the thrombin-activatable fibrinolytic inhibitor (TAFI) gene has not been linked to thrombosis risk, although plasma TAFI levels appeared to be weakly linked to thrombosis risk in some studies.

Homocysteine

Hyperhomocysteinaemia may be caused by genetic abnormalities, but only the severe inherited abnormalities of homocysteine metabolism (homozygous cystathionine β-synthase deficiency and homozygous deficiency of methylenetetrahydrofolate reductase) result in congenital homocysteinuria associated with an increased risk of both arterial and venous thrombosis, as well as premature atherosclerosis and mental retardation, epilepsy, and skeletal and eye problems. Half of patients present with venous or arterial thrombosis before the age of 30 years (see also Chapter 5).

The thermolabile variant of methylenetetrahydrofolate reductase (MTHFR) due to a common genetic polymorphism (C677T) is not a risk factor for venous thrombosis. The gene frequency of the T allele is 30%, which means that 40% of the population is heterozygous and 10% homozygous for the T polymorphism. Although the TT genotype has been shown to be associated with raised plasma homocysteine concentrations in some studies, this is only evident in areas of the world with a low folate intake and an analysis of 23 studies gave the relative risk of vascular disease associated with the TT genotype as 1.12 (0.92–1.37). Large, randomized clinical trials (predominantly performed in regions with a good folate intake) have shown that vitamin supplementation to reduced homocysteine levels does not reduce cardiovascular risk. Genetic testing for MTHFR should not be part of a thrombophilia screen and has no clinical utility.

Treatment of patients with venous thrombosis and heritable thrombophilias

There is no evidence that heritable thrombophilia should influence the intensity of anticoagulation. In patients with antithrombin deficiency, heparin resistance is infrequent and recurrence or extension of thrombosis while on treatment is no greater than that observed in individuals without antithrombin deficiency. Warfarin-induced skin necrosis is extremely rare, even in patients with protein C or S deficiency, such that most individuals with protein C or S deficiency do not develop skin necrosis. The intensity of maintenance therapy with warfarin should not be influenced by laboratory evidence of inherited thrombophilia. There is no evidence that recurrence on treatment is more likely in patients with heritable thrombophilia.

Long-term prospective cohort outcome studies have shown that finding a heritable thrombophilia does not usefully predict recurrence in unselected patients after an episode of venous thrombosis. However, these studies were not powered to exclude an increased risk of recurrence specifically in relation to rare thrombophilias, such as antithrombin or protein C deficiency. Therefore, it remains uncertain if mutations in *SERPINC1*, *PROC* or *PROS* causing deficiency of the corresponding protein might predict a sufficiently high risk of thrombosis to justify long-term (lifelong) anticoagulation. Studies in thrombosis-prone families indicate a risk of recurrence that would justify long-term anticoagulation, but these studies may not be representative of risk associated with the *SERPINC1*, *PROC* and *PROS* mutations per se, but may be influenced by additional defects in the thrombosis-prone family members. Further studies are required to define the relative risk of recurrent events associated with the actual *SERPINC1*, *PROC* and *PROS* mutations.

Systematic reviews indicate a risk of 1.4 for recurrent venous thromboembolism in patients heterozygous for the FVR506Q mutation and 1.2–1.7 for the *F2*G20210A mutation. The authors concluded that the magnitude of the increase in risk was modest and by itself did not justify an extended duration of anticoagulation. An analysis of the MEGA study showed that testing for inherited thrombophilia did not reduce recurrence of venous thrombosis.

As a result of these observations, at present the following conclusions seem reasonable:

- Indiscriminate testing for heritable thrombophilias in unselected patients presenting with a first episode of venous thrombosis is not indicated.
- Decisions regarding duration of anticoagulation (lifelong or not) in unselected patients should be made with reference to whether a first episode of venous thrombosis was provoked,

other risk factors and risk of anticoagulant therapy-related bleeding, regardless of whether a heritable thrombophilia is known.

- Testing for heritable thrombophilias in selected patients, such as those with a strong family history of unprovoked recurrent thrombosis, may influence decisions regarding duration of anticoagulation. Unfortunately, in this regard, identifying patients for testing is not straightforward, as criteria for defining thrombosis-prone families have not been validated and the association between family history of thrombosis and detection of inherited thrombophilia is weak.

Case finding

Testing for heritable thrombophilias in patients presenting with venous thrombosis allows case finding of affected asymptomatic family members. The rationale is that this permits avoidance of environmental risks (e.g. use of the combined oral contraceptive pill [COCP] by females) or provides an opportunity for targeted thrombophylaxis at times of unavoidable high risk (such as surgery). This rationale assumes such action would not be taken for family members without thrombophilia. However, individual risk is affected by multiple genetic and environmental factors that will be different, even among first-degree relatives. Prospective cohort studies have determined the annual risk of venous thrombosis in asymptomatic family members identified after testing unselected patients presenting with venous thrombosis. The studies included more than 3500 patient-years of observation. The annual risk of venous thrombosis in asymptomatic family relatives of index patients was 0.1% for those with *F2*G20210A, 0.6% for those with FVR506Q, about 2% for protein C and protein S deficiency, and 4% for antithrombin deficiency. High-risk periods contribute to approximately half of all events in patients with FVR506Q and thrombophylaxis appears to reduce risk. In the EPCOT registry, patients were referred to specialist centres for thrombophilia testing if they had a personal or family history of venous thrombosis. The incidence of thrombosis on study entry was determined retrospectively in asymptomatic relatives. The risk of venous thrombosis was 16 times higher in affected relatives, with the greatest risk in relatives of patients with deficiency of a natural anticoagulant or multiple defects. The highest risk was in individuals with antithrombin deficiency (1.7% per year) or combined defects (1.6% per year). Targeted case finding of relatives with 'severe' or 'high-risk' thrombophilias, such as deficiency of antithrombin, protein C or protein S, has been suggested, although the clinical utility of such methods remains uncertain and contentious among experts.

Given uncertainty, some experts argue that it is reasonable to perform testing if it is anticipated that clinical management will be influenced, for example an intensified or extended period of prophylaxis during a high-risk period. If a family history suggests a high degree of genetic penetrance, then it might be reasonable to test a symptomatic patient and then their relatives with a view to enhanced prophylaxis at times of high risk in affected members, for example thrombophylaxis in pregnancy when there is a family history of pregnancy-associated thrombosis, or intensified or extended surgical thrombophylaxis when there is a history of thrombophylaxis failure in affected members. In all cases, the risks, benefits and limitations of testing should be discussed in the context of explained inheritance and disease risk. The importance of this is demonstrated by reported anxiety after testing positive and an over-estimated perception of risk. Simple methods for quantifying a positive family history do not discriminate patients with and without thrombophilia and therefore the decision to test for inherited thrombophilia cannot be accurately guided by the presence or absence of a family history.

Prevention of thrombosis associated with oestrogen-containing hormone preparations

In some women, heritable thrombophilia has already been established, while in others it is perceived that testing would enable informed decision-making regarding use of a COCP or hormone-replacement therapy (HRT). However, the absolute risk of thrombosis is low and the fact that venous thrombosis has a multiple genetic basis with incomplete penetrance makes counselling in relation to genetic testing uncertain. In many instances, an alternative effective contraceptive is acceptable and unlike the COCP the progesterone only pill is not associated with an increased risk of thrombosis. Similar principles apply to HRT, although the baseline risk is higher as the population is older. Rarely is there a therapeutic indication for HRT and in most instances there is only a weak indication. If HRT is considered essential, then a non-oral formulation is associated with a significantly lower risk of venous thrombosis.

A first-degree relative with a history of venous thrombosis is a relative contraindication to an oestrogen-containing hormonal preparation. The risk is dependent on the circumstances of thrombosis in the relative. For example, a history of an elderly relative who developed venous thrombosis as a complication of cancer is not a contraindication. In contrast, a relative with unprovoked venous thrombosis, or specifically a sibling developing venous thrombosis while taking a COC, should be considered a strong contraindication. In families with known heritable thrombophilias the risk of venous thrombosis can be increased in unaffected members as well as affected and so a negative thrombophilia result does not exclude an increased risk of venous thrombosis. Therefore, decisions regarding use of COCPs and whether thrombophilia testing is likely to be informative should be made with reference to individual clinical risk factors and the circumstances associated with venous thrombosis in the family.

Prevention of pregnancy-associated venous thrombosis

Pregnancy is associated with a fivefold to tenfold increased risk of venous thrombosis compared with non-pregnant women of comparable age, with an absolute risk of 1–2 per 1000 deliveries. The risk of thrombosis, compared with the general age-matched female population, is increased 100-fold in pregnancy in women with a previous thrombosis. Retrospective studies in women with previous venous thrombosis for whom detailed information on the type of thrombophilia was available indicate that the rate of recurrence is similar in women with and without thrombophilia, although studies excluded women with high-risk thrombophilias (anticoagulant deficiency, multiple defects). Thrombosis rarely occurs in women whose initial event is provoked. In general, the absolute risk of pregnancy-associated venous thrombosis in women with heritable thrombophilia with no previous history is small, but women with antithrombin deficiency or those homozygous for the FVR506Q or *F2*G20210A mutations or who are double heterozygotes should be regarded as being at higher risk. The number of women with these defects is very small.

In women with a previous history of venous thrombosis, the major factor in determining if prophylaxis should be given is whether prior venous thrombosis was provoked or not. If the episode was unprovoked, prophylaxis should be considered, and thrombophilia testing is not required if prophylaxis is given. In women with a first provoked event, the decision to test should be influenced by the strength of the provocation, for example thrombosis associated with major trauma and subsequent immobility would not be an indication for prophylaxis or testing. In women with a first-degree relative with thrombosis, the decision to test should be influenced by whether the event in the relative was unprovoked or provoked and the strength of the provocation. If the event in the first-degree relative was associated with pregnancy or COCP, then testing and finding thrombophilia should prompt consideration of prophylaxis, particularly if the symptomatic relative was known to have the same defect, especially deficiency of antithrombin or protein C. When testing in pregnancy is performed it is necessary to interpret the results with reference to the effect of pregnancy on the tests.

Pregnancy morbidity

There is evidence of an association between heritable thrombophilia and pregnancy morbidity, including early and late pregnancy loss, pre-eclampsia and intrauterine growth retardation. A recent study has shown that antepartum prophylactic dalteparin does not reduce the occurrence of venous thromboembolism, pregnancy loss, or placenta-mediated pregnancy complications in pregnant women with thrombophilia at high risk of these complications and is associated with an increased risk of minor bleeding.

Purpura fulminans

Purpura fulminans is a rare syndrome characterized by progressive haemorrhagic skin necrosis that occurs in neonates with congenital severe protein C deficiency at birth or in the first few days of life, and in association with infection in children and adults. The condition may occur in children without congenital anticoagulant deficiency following viral infections, with an onset within 10 days of infection. Acquired severe protein S deficiency has been reported in purpura fulminans following chickenpox infection. With bacterial infections, disseminated intravascular coagulation (DIC) is often present, for example meningococcal infection. In patients with very severe skin necrosis, testing for acquired protein C or S deficiency should be considered, as plasma exchange may be beneficial. Neonates homozygous for protein C or S deficiency may be born with skin necrosis or DIC.

Vitamin K antagonist-induced skin necrosis

Vitamin K antagonist-induced skin necrosis, such as that seen with warfarin, is extremely rare, even in patients with protein C or S deficiency, such that most individuals with protein C or S deficiency do not develop skin necrosis. However, a high proportion of patients with vitamin K antagonist-induced necrosis have heritable protein C deficiency. The condition has been described in patients with other heritable thrombophilias. Vitamin K antagonist-induced necrosis associated with heritable thrombophilias tends to involve a central distribution.

Thrombophilia and arterial thrombosis

The genetic risk factors for venous thrombosis and arterial atherothrombosis overlap, but the material contribution of each differs between the two diseases. Hypercoagulability resulting from variation in the genetic control of blood coagulation produces a greater material contribution to venous thrombosis. However, it is becoming increasingly recognized that both genetic and acquired risk factors are common denominators for venous and arterial thrombosis. In patients presenting with venous thrombosis before the age of 40 there appears to be an increased risk of acute myocardial infarction subsequently.

Evidence of an association between heritable thrombophilia and arterial thrombosis is limited to case reports and small studies. It is possible that heritable defects that result in increased coagulability increase the likelihood of atherothrombosis. However, the material contribution of heritable thrombophilia, as

compared with established cardiovascular risk factors, is not sufficient to change therapy for primary and secondary prevention. Therefore, testing young patients for heritable thrombophilia after an arterial occlusive event is unhelpful. Because there is no established causal relationship and as treatment and secondary prevention should be in relation to established cardiovascular risk factors, thrombophilia testing is not recommended. In patients with arterial thrombosis at a young age that cannot be explained by conventional cardiovascular risk factors, various causes should be considered, including antiphospholipid antibody syndrome, a cardiac or large-artery source of embolus, vasculitis, aneurysm, a collagen-vascular disorder causing arterial dissection.

Neonatal stroke (see also Chapter 50)

Heritable thrombophilic defects are associated with perinatal brain injury. The pathogenesis is thrombosis in either cerebral arteries or cerebral veins or sinuses. Secondary brain infarction results, with subsequent cystic change. Thrombophilia is not a significant risk factor for ischaemic stroke in adults, as it does not increase the risk of arterial thrombosis (but cerebral vein thrombosis can present as stroke). In the neonate the pathophysiology of stroke is quite different and thrombophilia may increase the risk of arterial thrombosis and stroke in the neonate or fetus. There are two potential explanations for this:

1 The circulation in the neonate is different. In the adult venous blood is 'filtered' by the pulmonary circulation. In the fetus and neonate the lungs are bypassed by the ductus arteriosus and so it is possible for clots arising in the venous system to bypass the lungs and occlude a cerebral artery.

2 The blood coagulation system of the neonate is different to that of the adult. It is possible that the coagulability of blood in fetal and neonatal arteries is influenced by thrombophilic defects. However, this remains to be investigated.

Testing for defects may therefore identify a material contributory factor, but does not typically inform management decisions. For example, anticoagulant therapy is not usually considered in children found to have a suffered a stroke in the neonatal period and there may be significant time before a neurological deficit is recognized or the cause of stroke determined. In practice, testing is sometimes performed in order to explain to parents why a stroke possibly occurred.

Laboratory methodology and testing strategy

The laboratory diagnosis of heritable thrombophilias is difficult as the tests are subject to considerable preanalytical variables. Low levels of antithrombin, protein C and protein S occur in a variety of circumstances and test results and the clinical implications of both positive and negative results are frequently misinterpreted.

A full blood count and platelet count are useful indicators of general health and will identify myeloproliferative disorders that increase thrombotic risk. The APTT may detect some lupus anticoagulants, but a sensitive assay should be used. The thrombin time will detect heparin contamination and the prothrombin time (PT) is useful for the interpretation of protein C and protein S levels. Testing is usually delayed until at least 1 month after completion of a course of anticoagulation with a vitamin K antagonist. Testing should be avoided during intercurrent illness and use of COCPs or HRT should be noted. Sometimes testing is performed during pregnancy and so interpretation of results must be made with reference to the effect of the pregnancy on results.

Functional assays should be used where accuracy and precision are acceptable. However, no single method will detect all defects. For example, a protein C chromogenic assay will not detect a dysfunctional protein C molecule with impaired phospholipid binding due to a mutation in the Gla domain. A clot-based protein C assay would be sensitive to this defect, but imprecision of the assay would result in reduced sensitivity and specificity for other defects compared with a chromogenic assay. Similarly, the performance of antithrombin assays will be influenced by heparin concentration, the preincubation time with heparin, use of FXa or bovine thrombin or human thrombin as a substrate. For example, an assay utilizing a short heparin incubation time will detect heparin binding-site defects, which may not be associated with an appreciably increased risk of venous thrombosis.

Even in families with characterized defects, a phenotypic assay may fail to accurately discriminate affected and non-affected individuals. True heritable deficiencies may not be detected and false-positive diagnoses are common, 1 in 40 will have a level below the lower limit of the reference range and few of these have inherited deficiency.

Preanalytical variables

Most assays are affected by clots in the sample. Antibody assays may be affected by atypical antibodies such as paraproteins, rheumatoid factor and heterophile antibodies. Chromogenic assays are subject to interference by lipaemia and haemolysis. Clot-based assays are subject to imprecision due to variation in the baseline clotting time and the levels of other factors.

Low levels of antithrombin, protein C or protein S may relate to age, sex, acquired illness or drug therapy and therefore interpretation requires knowledge of the patient's condition at the time of blood sampling. Interpretation of results in children must be with reference to the normal range for age.

Low levels of antithrombin, protein C or protein S suspected to be the result of heritable mutations should be confirmed on one or more separate samples. Demonstrating a low level in

other family members supports a diagnosis of heritable deficiency and characterization of the genetic mutation can confirm the diagnosis and give some indication of thrombosis risk, particularly in the case of antithrombin deficiency.

General recommendations for laboratory tests and interpretation

- Testing at the time of acute venous thrombosis is not indicated, as the utility and implications of testing need to be considered and the patient needs to be counselled before testing. As treatment of acute venous thrombosis is not influenced by test results, testing can be performed later.
- The PT should be measured to detect the effect of oral vitamin K antagonists, which will cause a reduction in protein C and S levels.
- Functional assays should be used to determine antithrombin and protein C levels.
- Chromogenic assays of protein C activity are less subject to interference than clotting assays and are preferable.
- Immunoreactive assays of free protein S antigen are preferable to functional assays. If a protein S activity assay is used in the initial screen, low results should be further investigated with an immunoreactive assay of free protein S.
- If an APC resistance assay is performed to detect FVR506Q, then the modified APC sensitivity test (predilution of the test sample in FV-deficient plasma), as opposed to the original APC sensitivity test, should be used as a phenotypic test for FV Leiden. If positive, the mutation should be confirmed by a direct genetic test. An APC resistance assay is unnecessary if a direct genetic test for FV Leiden is used initially.
- Repeat testing for identification of deficiency of antithrombin, protein C and protein S is indicated and a low level should be confirmed on one or more separate samples. Deficiency should not be diagnosed on a single abnormal result.
- Rigorous internal quality assurance and participation in accredited external quality assessment schemes are mandatory.
- Thrombophilia testing must be supervised by experienced laboratory staff and the clinical significance of the results must be interpreted by an experienced clinician who is aware of all relevant factors that may influence individual test results in each individual case.

Antithrombin assays

The activity of antithrombin can be measured as progressive activity towards thrombin in the absence of heparin (designated accordingly 'progressive activity'). More usually antithrombin activity against either thrombin or FXa is measured in the presence of heparin (designated 'heparin cofactor activity'). Activity assays typically use a chromogenic substrate in which a small synthetic peptide is linked to *p*-nitroaniline. Cleavage of the peptide releases the *p*-nitroaniline, which is yellow and can be measured quantitatively. Patient plasma is incubated with heparin and then an excess of thrombin or FXa is added. The heparin-activated antithrombin is complexed and neutralizes some of the thrombin or FXa. The remaining free protease quantified by cleavage of the appropriate chromogenic substrate is inversely proportional to the antithrombin level in patient plasma (Figure 44.4).

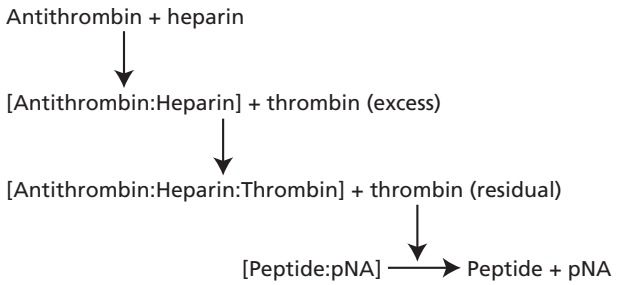


Figure 44.4 Chromogenic antithrombin heparin cofactor assay. Plasma (the source of antithrombin) is incubated with an excess of thrombin (human or bovine) in the presence of heparin. Thrombin is complexed and neutralized by heparin-activated antithrombin until all antithrombin is complexed. The residual thrombin is then detected by the chromogenic indicator. The concentration of antithrombin in the sample is inversely proportional to the residual thrombin concentration.

Interference by other proteases and inhibitors in the plasma sample is a potential problem. Interference by heparin cofactor II, another heparin-activated serpin, can be eliminated by using FXa or bovine thrombin as the target protease, as heparin cofactor II only inhibits human thrombin. The effect of other proteases in the patient plasma can be minimized by addition of a protease inhibitor such as aprotonin.

The total amount of antithrombin protein can be measured immunologically with antibodies, for example by ELISA. As antithrombin antigen levels may be normal or near normal in type II deficiency, immunological assays may fail to identify patients with these variants.

Antithrombin levels are slightly reduced in infants, during pregnancy, in women taking oestrogen preparations and in some patients during heparin therapy. Levels also decrease with age and acquired low levels occur in liver disease, DIC, nephrotic syndrome, during L-asparaginase therapy, and in protein malnutrition and cachexia.

Protein C assays

Protein C defects are divided into type 1, with a reduced amount of functionally normal protein, and type 2, in which there is reduced activity due to a dysfunctional protein. In this case there is discordance between the functional and antigenic levels. The laboratory diagnosis of protein C deficiency is based on a functional assay. Most commercially available functional

assays of protein C employ the specific activator Protac, a protein C activator derived from the venom of the southern copperhead snake (*Agkistrodon contortrix*). Chromogenic assays using a synthetic substrate for APC are simple to perform and can detect all type 1 defects and type 2 defects affecting the catalytic site. A chromogenic assay will not detect most other type 2 defects. Clotting assays utilizing the prolongation of either the PT or APTT after Protac activation have the capacity to detect other, but not all, type 2 defects, but are subject to imprecision due to variation in the baseline clotting time and the levels of other factors, for example if there is a short APTT due to a high FVIII level, or the presence of the factor V Leiden mutation which results in a falsely low protein C activity due to APC resistance, or in the presence of a lupus anticoagulant. The diagnosis of protein C deficiency is problematic because of the wide overlap in protein C activity between heterozygous carriers and unaffected individuals. Protein C levels are not affected by pregnancy or oestrogen exposure. Acquired low levels of protein C occur during anticoagulant therapy with oral vitamin K antagonists, vitamin K deficiency, DIC and liver disease.

Protein S assays

Protein S is usually quantified immunologically rather than measured functionally. Originally, free protein S was measured in plasma supernatant after precipitation of C4b-bound protein S by polyethylene glycol (PEG). Nowadays PEG precipitation is not required as monoclonal antibodies that detect only free protein are used. Functional protein S assays are imprecise and are not used in the majority of coagulation laboratories.

Protein S levels are significantly lower in females, so much so that different reference ranges are required for males and females and if this is not done there is a significant risk of a false-positive diagnosis of protein S deficiency in women. Protein S levels fall progressively during normal pregnancy and are reduced by oestrogens. Acquired low levels of protein S occur during anticoagulant therapy with oral vitamin K antagonists, vitamin K deficiency, DIC and liver disease.

FVR506Q and APC resistance

Nowadays, the FVR506Q mutation is usually detected by direct mutation analysis. APC resistance assays typically utilize the APTT. Samples are tested with and without an added fixed concentration of APC. The clotting times are expressed as a sensitivity ratio, i.e. the APTT (seconds) of patient plasma plus APC divided by the APTT (seconds) of patient plasma without added APC; the lower the ratio, the greater the APC resistance. Platelet contamination and activation must be avoided as this will lead to APC resistance. The ratio can be normalized by dividing the patient's sensitivity ratio by the sensitivity ratio of a pooled normal plasma. If normalization is used, FVR506Q carriers should be excluded from the normal plasma pool. The APC resistance assay is abnormal not only in the presence of FVR506Q, but with all other causes of heritable and acquired APC resistance and in individuals with a prolonged baseline APTT due to clotting factor deficiencies or anticoagulant therapy. It is therefore not specific for FVR506Q. Modification of the test by a 1 in 4 predilution of the test plasma in FV-deficient plasma increases the sensitivity and specificity of the assay for FVR506Q to almost 100% and this test is reliable if genetic analysis is not available.

*F2*G20210A

A laboratory clotting test with sensitivity and specificity for this mutation is not available and so detection is by direct mutation analysis.

Next-generation global thrombophilia tests

A test that predicts a high rate of recurrent venous thrombosis would be clinically useful. The clinical utility of such a test would be the identification of patients in whom the benefit of continued long-term anticoagulant therapy outweighed the risk. Patients with a first unprovoked venous thrombosis are at highest risk of recurrence, but it is arguable if the level of risk estimated from clinical risk factors alone justifies long-term anticoagulation. A new generation of thrombophilia tests that measure the degree of coagulability are now being investigated in clinical studies. These include biomarkers, such as D-dimer, which reflect the level of thrombin generation *in vivo*, and thrombin generation assays, which measure the capacity of a patient plasma sample to generate thrombin in response to a predefined coagulation trigger, for example low concentration tissue factor. It has now been demonstrated in several studies that the level of D-dimer correlates with risk of recurrent venous thrombosis and this test is being used increasingly to stratify patient risk and influence clinical decisions regarding duration of anticoagulant therapy after an episode of venous thrombosis.

Counselling and genetic testing

Clinical utility refers to the ability of a screening or diagnostic test to prevent or reduce adverse health outcomes, including mortality, morbidity and disability. A test does not have inherent clinical utility; rather it is the adoption of a therapeutic or preventive intervention on the basis of the test result that influences health outcomes. Clinical utility can more broadly refer to any use of a test result to inform clinical decision-making, and this may include information considered important by individuals and families.

A large proportion of genetic counselling focuses on presenting risk information. Making sense of information can be

difficult. Whenever a patient is offered an intervention, there is a trade-off between benefit and risk. The benefit and risk information must be presented in ways that are relevant and understandable. In addition to the quantitative information, the qualitative aspect of risk is equally important. A relative risk does not indicate how likely it is that an outcome will occur, only how much relatively more or less likely it is. It is the absolute risk that indicates how likely an event will be. For example, the relative risk of venous thrombosis is increased twofold to sixfold in a female aged between 20 and 40 years who takes a COCP, depending on her age and the COCP preparation. It is accepted that COCP exposure increases the risk of venous thrombosis and some women presented with this information might choose not to use a COCP as, despite its proven efficacy and attractions. When informed that the risk would be 35-fold higher than the population baseline risk if the woman had the factor V Leiden mutation and used a COCP, some women would again choose to avoid use of a COCP. However, in order to make an informed decision it is necessary for the woman to know how likely it is she would suffer venous thrombosis if she did not use the COCP and how much the absolute risk would change if she did use a COCP. The population baseline risk of thrombosis is about 1 per 10,000 per year under the age of 40. If a woman used a COCP it would increase about fourfold to about 4 per 10,000 per year. Clearly, the risk of venous thrombosis associated with COCP use is very low. For a woman with the factor V Leiden mutation, her baseline risk of thrombosis would be about 1 per 2000 per year under the age of 40, a risk that is increased about fivefold compared with a woman without this mutation. If this woman were to take a COCP, her absolute risk would increase from 1 per 2000 to 1 in 300, a risk that is increased about sevenfold for her but which is 35-fold greater than the population baseline risk. For many women an annual risk of 1 in 300 is acceptable, given the proven efficacy, ease of use and other obvious advantages of a COCP.

Selected bibliography

Baglin T, Luddington R, Brown K, Baglin C (2003) Incidence of recurrent venous thromboembolism in relation to clinical and thrombophilic risk factors: prospective cohort study. *Lancet* **362**: 523–6.

Baglin T, Gray E, Greaves M *et al.* (2010) Clinical guidelines for testing for heritable thrombophilia. *British Journal of Haematology* **149**: 209–20.

Baglin T (2010) Unravelling the thrombophilia paradox: from hypercoagulability to the prothrombotic state. *Journal of Thrombosis and Haemostasis* **8**: 228–33.

Bates SM, Greer IA, Middeldorp S; American College of Chest P (2012) VTE, thrombophilia, antithrombotic therapy, and pregnancy: Antithrombotic Therapy and Prevention of Thrombosis, 9th ed: American College of Chest Physicians Evidence-Based Clinical Practice Guidelines. *Chest* **141**: e691S-736S.

Bertina RM, Koeleman BP, Koster T *et al.* (1994) Mutation in blood coagulation factor V associated with resistance to activated protein C. *Nature* **369**: 64–7.

Coppens M, Reijnders JH, Middeldorp S, Doggen CJ, Rosendaal FR (2008) Testing for inherited thrombophilia does not reduce the recurrence of venous thrombosis. *Journal of Thrombosis and Haemostasis* **6**: 1474–7.

Dahlback B, Carlsson M, Svensson PJ (1993) Familial thrombophilia due to a previously unrecognized mechanism characterized by poor anticoagulant response to activated protein C: prediction of a cofactor to activated protein C. *Proceedings of the National Academy of Science USA* **90**: 1004–8.

Lijfering WM, Brouwer JL, Veeger NJ *et al.* (2009) Selective testing for thrombophilia in patients with first venous thrombosis: results from a retrospective family cohort study on absolute thrombotic risk for currently known thrombophilic defects in 2479 relatives. *Blood* **113**: 5314–22.

MacCallum P, Bowles L, Keeling D (2014) Diagnosis and management of heritable thrombophilias. *British Medical Journal* **349**, g4387.

Middeldorp S, van HylckamaVlieg A (2008) Does thrombophilia testing help in the clinical management of patients? *British Journal of Haematology* **143**: 321–35.

CHAPTER 45

Acquired venous thrombosis

45

Beverley J Hunt[1] and Henry G Watson[2]

[1]Guy's and St Thomas' NHS Foundation Trust, London, UK
[2]Department of Haematology, Aberdeen Royal Infirmary, Foresterhill Health Campus, Aberdeen, UK

Epidemiology of venous thromboembolism (VTE)

Deep vein thrombosis (DVT) is common, principally affecting the lower limbs. In western populations the annual incidence is around 1 per 1000 population. Incidence is age-dependent, being exceptional in childhood and increasing with age with an incidence of around 1 per 10,000 population per annum in young adults, and 1% per annum in the very elderly. The incidence of pulmonary embolism is lower, possibly around 1 per 3000 population annually, although subclinical pulmonary embolism occurs frequently in subjects with lower limb proximal DVT. In patients who die in hospital there is evidence of pulmonary embolism in up to 20% of cases and in many it contributes to the cause of death. In some cases pulmonary embolism presents without identifiable DVT, reflecting that 80% of DVTs are clinically silent.

Virchow's triad, described in 1859, postulates that venous thrombosis could result from changes in blood flow, hypercoagulability of the blood, or abnormalities of the blood vessel wall. Hypercoagulability as a result of inherited thrombophilia is discussed in Chapter 44. The main risk factors for VTE are shown in Table 45.1 and in many cases a combination of factors described by Virchow are present. In this chapter, the acquired conditions that predispose to venous thromboembolism (VTE) are considered, although in a minority of cases it is likely that the individual genotype influences the likelihood of thrombosis in particular circumstances. In addition to the factors listed in Table 45.1, there has been recent interest in an association between VTE and artherosclerotic disease and the metabolic syndrome.

Pregnancy and venous thromboembolism

Around 1 in 1000 pregnancies is complicated by DVT or pulmonary embolism. Widespread use of thromboprophylaxis in at-risk pregnancies has contributed to massive pulmonary embolism no longer being the leading causes of pregnancy-related death in the UK. However, it still remains a major cause of death in the Triennial Confidential Enquiry. Although the postpartum phase has been considered to be the time of highest risk, overall at least as many episodes of VTE occur antepartum, including some in the first trimester. There are notable clinical features in pregnancy-related DVT: around 70% of thromboses affect the deep veins of the left lower limb and involve the iliofemoral vein from the time of presentation; thrombosis restricted to the deep veins of the calf is uncommon in pregnancy, but is more common after delivery. The left-sided predominance most likely relates to the anatomical relationship between the iliac artery and vein on that side. Many affected women develop post-thrombotic syndrome.

The pathogenesis of pregnancy-related VTE is multifactorial, as in other situations. Pre-exisiting patient factors, such as obesity, along with specific pregnancy-related factors such as raised intra-abdominal pressure and vascular compression caused by the gravid uterus, the suppressed systemic fibrinolytic capacity partly due to the production of plasminogen activator inhibitor-2 by the placenta, the progressive increase in plasma concentrations of clotting factors from around 10 weeks' gestation, including factor (F)VIII and fibrinogen, and the fall in plasma concentration of the anticoagulant cofactor protein S may all contribute to thrombosis risk. Operative delivery,

Postgraduate Haematology, Seventh Edition. Edited by A Victor Hoffbrand, Douglas R Higgs, David M Keeling and Atul B Mehta.

Table 45.1 Risk factors for venous thromboembolism.

Inevitable/environmental
- Increasing age
- Pregnancy and puerperium
- Prolonged immobility, e.g. long-haul travel
- Obesity
- Trauma

Iatrogenic
- Surgery
- Indwelling venous devices
- Pharmacological
- Oestrogen related: combined oral contraceptive, hormone-replacement therapy, tamoxifen
- Ovarian stimulation regimens
- Chemotherapy and radiotherapy
- Heparins

Disease-related
- Antiphospholipid syndrome
- Cancer
- Myeloproliferative diseases
- Acute promyelocytic leukaemia
- Inflammatory states e.g. nephrotic syndrome, inflammatory bowel disease, connective tissue disorders
- Stroke
- Paroxysmal nocturnal haemoglobinuria
- Thrombotic thrombocytopenic purpura
- Sickle cell disease and other sickling syndromes
- Congestive cardiac failure

especially emergency caesarean section increases the risk further.

Although heritable thrombophilias contribute to the risk of pregnancy-related VTE, in around 70% of cases, no hereditary thrombophilia can be identified and in 25% there is also no obvious risk factor other than otherwise normal pregnancy.

Currently, diagnosis is dependent on imaging because the use of the Wells score and D-dimers has not been validated in pregnancy, and in late pregnancy D-dimers are often increased. It is useful to perform bilateral Doppler ultrasound of the legs initially in the diagnostic algorithm for pulmonary embolism, because if it is positive it negates the need for radiation exposure. The use of CT pulmonary angiography increases the risk of breast cancer in the mother by 10%. V/Q scanning delivers a lower dose of radiation to the mother but may carry a slightly increased risk of childhood cancer; in both situations the absolute risk is very small. V/Q SPECT has a similar sensitive and specificity to CTPA. Where feasible, women should be involved in the decision to undergo CTPA or V/Q scanning.

When VTE is diagnosed in pregnancy it is treated with low-molecular-weight heparin (LMWH) in doses based on the woman's booking or early pregnancy weight. There is insufficient evidence to recommend whether the dose of LMWH should be given once daily or in two divided doses. Treatment with therapeutic doses of subcutaneous LMWH should be employed during the remainder of the pregnancy and anticoagulation should continue for at least six weeks postnatally and until at least 3 months of treatment has been given in total. Once labour is established LMWH should be stopped; where delivery is planned, either by elective caesarean section or induction of labour, LMWH should be stopped 24 hours prior to planned delivery. Regional anaesthetic or analgesic techniques should not be undertaken until at least 24 hours after the last dose of therapeutic LMWH. LMWH should not be given for 4 hours after the use of spinal anaesthesia or after the epidural catheter removal, and the epidural cannula should not be removed within 10–12 hours of the most recent injection.

Immobility as a risk factor for venous thromboembolism

It is likely that venous stasis is significant in the pathogenesis of many cases of VTE. Sitting still for 90 min reduces the flow through the popliteal vein by 50%. Immobility is a major contributor to hospital-acquired VTE, and is a factor in DVT in paralysed limbs and associated with splinting, for example in a plaster cast.

There has been considerable interest in the relationship between VTE and the physically cramped conditions encountered during travel, especially long-haul aircraft journeys. The term 'economy class syndrome' has been coined to encompass the occurrence of VTE during and after air travel. The level of risk of VTE has been severely exaggerated, partly due to media attention. Epidemiological data suggest a relative risk of around twofold, which persists for a few weeks after the journey. A similar level of risk applies to other modes of powered travel, including by road. Observational data indicate that the incidence of fatal pulmonary embolism induced acutely by air travel is less than one in a million journeys. The duration of the flight is a factor, with some evidence that VTE risks begin to increase progressively with flight time of greater than 4 hours. The pathogenesis is multifactorial, and the elderly, travellers with recent surgical or medical admission and those with a previous history of VTE are more at risk. The contribution of the particular environmental conditions within a pressurised cabin, particularly hypobaric hypoxia, to hypercoagulability is disputed. At present it is reasonable to assume that any form of long-distance powered travel carries a low, but slightly increased, risk of VTE and to recommend simple precautionary measures such as maintenance of hydration and calf exercises and ambulation in order to maintain venous flow in the lower limbs. Emerging studies shows that prolonged sitting in office or at home especially on line gaming, is associated with increased risk of VTE and is known as 'e-thrombosis'.

Iatrogenic venous thromboembolism

Hospital acquired thrombosis

VTE is common in hospitalized subjects, and has led to the widespread adoption of physical and pharmacological methods of thromboprophylaxis which reduce risk by 50–70%. Hospital-acquired VTE also known as hospital-acquired thrombosis (HAT) may occur up to 90 days after admission, with the peak incidence of DVT in the first week and highest rate of PE in the second to third week after discharge. Immobility and the acute phase response seen in postoperative patients and acutely ill medical patients, are factors. The UK National Institute for Health and Care Excellence (NICE) has recommended that all adults admitted to hospital are assessed for their risk of VTE. Medical patients with risk factors (and this includes reduced mobility) and surgical patients are offered LMWH or fondaparinux, unless there is a contraindication. In addition, surgical patients receive mechanical prophylaxis with antiembolism stockings (thigh or knee length), foot impulse devices or intermittent pneumatic compression devices. The highest prevalence of VTE occurs after major orthopaedic surgery to the lower limb and dense stroke. In hip replacement and knee replacement surgery subclinical DVT is frequently detectable by imaging, although there is evidence that clinically significant VTE may be becoming less common due to improvements in surgical techniques and materials, as well as wider use of thromboprophylaxis. Extended prophylaxis for up to 28–35 days has been shown to reduce VTE in those undergoing hip replacement, knee replacement and abdominal or pelvic surgery for cancer. For those undergoing hip or knee replacement, several non-vitamin-K antagonist oral anticoagulants (NOACs) are licenced for thromboprophylaxis.

Indwelling venous devices

The use of indwelling venous catheters is common in several groups of patients, including those with malignancy, acute renal failure and those requiring total parenteral nutrition. The fact that there are not more thrombotic complications of indwelling lines is possibly surprising when one considers the inherent concept of a foreign surface being placed in a vein for a significant period of time. Nevertheless thrombosis of these lines is fairly common with a higher incidence observed in patients with large diameter lines and those with a tip positioned distal to the superior vena cava. A number of other factors contribute to the thrombosis risk, including infection, extrinsic compression, the use of certain types of chemotherapy and a previous history of VTE. A Cochrane review evaluated the efficacy of oral and parenteral anticoagulants in the prevention of central venous catheter-related thrombosis reported up to February 2010. Neither warfarin nor prophylactic-dose LMWH was associated with reduction in risk and neither should be routinely used for prophylaxis.

The presentation of line-related thrombosis is site specific. For lines in the upper limb area the presentation often includes swelling, pain and discolouration of the arm and possibly face. Most diagnoses are made by ultrasound, although contrast venography is still used in some centres. Treatment is very dependent on the perceived value of the line, stage of illness of the patient and likelihood of being able to replace the line elsewhere if needed. There are options, including local thrombolysis and anticoagulation with or without line removal.

Pharmaceuticals

Combined oral contraceptive

VTE associated with use of the combined oral contraceptive (COCP) is of particular importance because it affects healthy women in a young age group. The overall risk of VTE in a COCP user is around four- to fivefold higher than in a non-pill user. The risk is also increased by family history of thrombosis, obesity, thrombophilia and older age. Fortunately, the absolute risk remains acceptable because of the low background incidence of thrombosis in women of childbearing age. Combined pills induce a state of activated protein C resistance, and this appears to be more marked with the third-generation contraceptives than second-generation preparations. Most preparations in current use contain an equivalent dose of oestrogen, usually 30–35 μg of ethinylestradiol. However, the progestagen content varies. In second-generation pills this is levonorgestrel, whereas in third-generation products it is desogestrel, gestodene or norgestimate. There is convincing epidemiological evidence that third-generation pills carry a greater risk of clinical VTE. Progestagen-only oral contraceptives appear to carry a lower risk of VTE such that their use can be considered in women in whom the COCP is contraindicated.

Hormone-replacement therapy

Oral hormone-replacement therapy (HRT) is also associated with VTE risk. The relative thrombosis risk is equivalent to second-generation COCPs, but the absolute risk is greater due to the higher background incidence of VTE in older women. Observational data indicate a very high rate of VTE in women with a previous history of venous thrombosis who embark upon HRT subsequently. Such is the level of risk that HRT is generally contraindicated in such women. If there are overwhelming indications for HRT (usually disabling menopausal symptoms) some clinicians consider warfarin thromboprophylaxis alongside HRT in high-risk situations. The procoagulant changes induced by oral HRT are not apparent with transdermal preparations and cohort studies suggest transdermal HRT may not increase the risk of VTE and may be considered in those thought to be at increased risk.

Drugs used in the treatment of malignancy

There is evidence that chemoradiotherapy is associated with a risk of inducing venous thrombosis. The mechanisms underlying this are unknown and almost certainly multifactorial. There has been a recent massive increase in the number of drugs that are available for systemic anticancer treatment and so it is likely that new additions will be made to the list of drugs associated with development of venous thrombosis. There are, however, some well-documented agents that are worthy of a mention.

Asparaginase

This has been used in the treatment of acute lymphoblastic leukaemia in adults and children for many years. It is associated with thrombotic complications including DVT, PE and cerebral sinus thrombosis. These occur more commonly in adults than in children. Asparaginase use results in a coagulopathy that includes a reduction in levels of fibrinogen, antithrombin and also the other natural anticoagulants, which led to the use of antithrombin concentrates, although there is poor evidence to support this. The aetiology of the coagulopathy is uncertain. Current recommendations are to actively give thromboprophylaxis with LMWH and only treat the coagulopathy if the patient is bleeding.

Thalidomide and lenolidamide

These two iMIDs are commonly used in the treatment of myeloma. It is in this indication that they seem to be most associated with a prothrombotic risk, especially when they are combined with high-dose steroids or other types of chemotherapy, notably anthracyclines. Incidences of thrombosis as high as 25–30% in the first few months of treatment have prompted studies into effective thromboprophylaxis for these patients. It is now routine practice for these patients to receive thromboprophylaxis while receiving these drug combinations.

Tamoxifen

Tamoxifen has been shown in large clinical trials to be a useful adjuvant therapy for oestrogen-receptor-positive breast cancer. Unfortunately it is associated with an increased thrombosis risk, which is probably mediated by its mild oestrogenic activity. Women taking tamoxifen have around a 3.5-fold increased risk of thrombosis, which is present only for the first two years of treatment and which is more prevalent in older rather than younger women.

Women developing thrombosis on tamoxifen are likely to continue to benefit from ongoing oestrogen receptor blockade. Options for their ongoing treatment may include continuing anticoagulation along with tamoxifen or switching to anastrozole, an aromatase inhibitor that reduces oestradiol synthesis and does not appear to carry a thrombosis risk.

Heparin-induced thrombocytopenia

Heparin-induced thrombocytopenia (HIT) is an important diagnosis because failure to recognize the syndrome carries a high risk of morbidity and mortality. It is due to development of antibodies directed against a complex of heparin and platelet factor 4. It is more commonly seen with unfractionated heparin (UFH) than LMWH and in surgical rather than medical or obstetric patients. Typically, the platelet count begins to fall between 5 and 10 days after first exposure to heparin. In subjects with previous recent exposure it may occur earlier and very occasionally the thrombocytopenia develops after heparin has been discontinued. The condition may complicate heparin administered at any dose. In contrast to other drug-related immune thrombocytopenias, the degree of thrombocytopenia is rarely severe ($<20 \times 10^9$/L) and bleeding is not a common feature. There may be systemic symptoms, including fever and a local reaction at the heparin injection site, such as erythema or even skin necrosis after subcutaneous administration. Because of platelet activation there is often new thrombosis, and this may manifest before the platelet count reaches thrombocytopenic levels. When heparin has been administered to treat arterial disease the new thrombosis is most commonly arterial, and in treatment of DVT there is usually extension of the presenting thrombus, often with pulmonary embolism. Thrombotic stroke, myocardial infarction, visceral ischaemia and disseminated intravascular coagulation have all been reported.

Confirmatory laboratory tests are not entirely satisfactory. Immunoassays for heparin/platelet factor 4 are sensitive, but lack specificity because a significant proportion of subjects treated with heparin have positive antibody tests, but neither thrombocytopenia nor thrombosis. In contrast, bioassays are specific, but high sensitivity is difficult to achieve. Few centres provide a reliable bioassay for routine diagnosis.

The most practical approach to diagnosis outside the intensive care unit, where thrombocytopenia is common due to the other causes that may coexist with HIT, is the use of a pretest probability scoring system based on the timing of onset of thrombocytopenia, presence or absence of new thrombosis, presence or absence of other potential cause of thrombocytopenia, and the degree of thrombocytopenia (the four Ts score). If the probability score is high or intermediate, heparin should be stopped and an alternative non-cross reacting anticoagulant commenced while further investigation is carried out. An alternative anticoagulant such as argatroban, danaparoid or fondaparinux must be administered immediately as the risk of thrombosis persists for several days after heparin withdrawal. The NOACs have not been used, but are likely to be effective. Warfarin given alone is contraindicated as massive thrombosis has been described in this situation. When a high or intermediate score is combined with the results of an immunoassay including high-dose heparin neutralization, this results in improved sensitivity and specificity for the diagnosis of HIT. The antibody may cross-react with

low-molecular-weight heparins and they should be avoided. The pathogenetic antibodies tend to become unmeasurable in the ensuing few months, but further exposure to heparin should be avoided in the long term.

Antiphospholipid syndrome

Antiphospholipid syndrome (APS) is an important acquired thrombophilic condition because it is of high prevalence and is associated with considerable morbidity and mortality. The essential features are arterial or venous thrombosis or recurrent pregnancy loss or placental dysfunction occurring in a subject in whom laboratory tests for antiphospholipid antibody are persistently positive. Diagnostic criteria have been established in order to standardize diagnosis for the purposes of clinical studies (Table 45.2).

DVT, pulmonary embolism and ischaemic stroke are the principal thrombotic manifestations, but any vessel may be involved, including the microvasculature, for example renal thrombotic microangiopathy. A characteristic of APS in an individual patient is the tendency to recurrent problems in the same vascular bed. The full range of clinical features is not necessarily present. For example, women can have severe thrombotic events without pregnancy complications and those with pregnancy complications most commonly have no other manifestation. In addition to thrombosis and pregnancy failure, additional clinical and laboratory features are variably present in APS. These include mild thrombocytopenia and an unusual dermatological feature, livedo reticularis. Cardiac valvular abnormalities occur in up to 30% of patients, but are usually subclinical. Most commonly the mitral or aortic valve is affected, often with valve thickening or incompetence. However, haemodynamic changes as a consequence of valvular damage are rare. Exceptionally, the syndrome may manifest as widespread multiorgan microvascular occlusion, with multiorgan failure, so-called catastrophic APS.

APS may occur with another chronic systemic autoimmune disease, usually systemic lupus erythematosus (SLE), when the term 'secondary antiphospholipid syndrome' used to be used. In primary APS there is no evidence for another relevant underlying condition.

Table 45.2 Diagnostic criteria in APS.

*Clinical criteria**
• Thrombosis: arterial, venous or microvascular thrombosis in any tissue or organ
*Laboratory criteria**
• Antiphospholipid antibody
• IgG or IgM anticardiolipin antibodies at moderate or high concentration†; IgG or IgM anti-β_2-GPI >99th percentile *and/or*
• Lupus anticoagulant
Pregnancy complications
• Unexplained death of morphologically normal fetus at or beyond 10 weeks of gestation
• Three or more unexplained consecutive miscarriages before 10 weeks
• One or more premature births of a morphologically normal fetus before 34 weeks of gestation due to pre-eclampsia, eclampsia or severe placental insufficiency

*There must be at least one clinical and at least one laboratory criterion present. The laboratory test must be consistently positive on at least two occasions 12 weeks apart as transient antibodies may occur, for example in infection. Such antibodies are not usually associated with clinical events.
†Evidence of the definition of moderate/high concentration. In general, values of IgG anticardiolipin >40 GPLU are considered to be moderate-titre antibodies; >99th percentile has been employed also. The significance of low-titre antibodies is less clear.

Antiphospholipid antibodies

Antiphospholipid antibodies are a family of antibodies reactive with proteins that have the property of binding to negatively charged phospholipids. The most important is β_2-glycoprotein I (β_2-GPI), as pathogenic antibodies (i.e. those most strongly associated with clinical events) are most frequently reactive with this protein. β_2-GPI is a member of the complement control protein family. It has five domains. It is now known that antiphospholipid antibodies reactive with each of these five domains may occur, but it is only those that recognize a specific epitope on domain I that are pathogenic. Others with an affinity for negatively charged phospholipid have been implicated in the syndrome, especially anti-annexin V.

The lupus anticoagulant (LA) is an *in vitro* phenomenon in which the antiphospholipid antibody slows clot formation, thereby lengthening the clotting time. This is probably due to impairment of the assembly of the components of prothrombinase on phospholipid due to interference by the antibody (Figure 45.1). LA is due to antibodies reactive to β_2-GPI/phospholipid or to prothrombin/phospholipid. The β_2-GPI-dependent antibodies also bind in anticardiolipin assays, as the glycoprotein is present in test serum and often in assay reagents. LA due to prothrombin-reactive antibodies may be negative in anticardiolipin assays. Therefore some subjects with APS have LA and anticardiolipin, and some LA only. Others have anticardiolipin without LA, due to the presence of non-β_2-GPI and non-prothrombin-dependent antibody or possibly to relative insensitivity of the coagulation assays for LA. Because antiphospholipid antibodies are so heterogeneous, a comprehensive laboratory approach is essential for their reliable detection. In most laboratories, coagulation-based assays for LA and ELISAs employing cardiolipin and β_2-GPI are all required

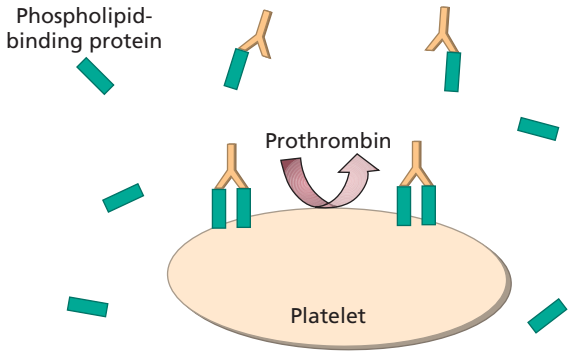


Figure 45.1 Antiphospholipid antibody binding: explanation for the lupus anticoagulant phenomenon.

to ensure appropriate diagnosis. The pathogenicity of IgA anticardiolipin antibodies is disputed and their detection is not utilized routinely in diagnosis.

Pathogenic mechanisms in APS

The pathogenesis of pregnancy failure and thrombosis in APS is not fully understood. Animal experiments support a pathogenic role for antibodies to β_2-GPI. Important candidate mechanisms for thrombosis are antibody-induced concentration of prothrombin on phospholipid surfaces *in vivo*, resulting in enhanced thrombin generation, interference with the activated protein C anticoagulant pathway, and increased monocyte and endothelial tissue factor expression. Increased platelet activation and inhibition of fibrinolysis by antiphospholipid antibodies have also been proposed. In relation to pregnancy failure, again numerous pathogenic mechanisms have been proposed. It is noteworthy that early miscarriage is a common feature and occurs before full placentation and is attributed to antiphospholipid antibodies causing inhibition of early trophoblast growth. There is an accumulating body of evidence for a role of complement activation in pregnancy loss in APS in a murine model, where heparin is effective in preventing embryo loss through inhibition of complement rather than via an anticoagulant mechanism. This raises the possibility that similar mechanisms may be relevant to human APS. Placental dysfunction is the hallmark of second- and third-trimester complications of intrauterine fetal death, intrauterine growth restriction, pre-eclampsia and placental abruption. The aetiology of the placental dysfunction has been related to placental infarction and an acute atherosis in the maternal spiral arteries, although these are not universal features. One interesting hypothesis for the cause of placental insufficiency is displacement of annexin V from trophoblast by antiphospholipid antibodies, with resultant acceleration of thrombin generation on the exposed negatively charged phospholipid.

Laboratory diagnosis of antiphospholipid antibodies

In diagnosing APS it is essential to consider that antiphospholipid antibodies are not specific to the syndrome. In addition to transient antibodies, which may, for example, be triggered by intercurrent infection, some chronic infections are associated with antiphospholipid antibodies, such as syphilis and hepatitis C. Antiphospholipid antibodies may also be detected incidentally in healthy subjects and they occur in relation to use of some drugs, particularly chlorpromazine. These drug-induced and infection-related antibodies do not usually appear to be associated with the clinical thrombotic manifestations of APS. In some cases they appear to be neither β_2-GPI nor prothrombin dependent.

The diagnosis of APS relies on the demonstration of the persistent presence of either LA by coagulation tests or of antiphospholipid antibodies by solid-phase immunoassays for anticardiolipin and/or β_2-GPI. Reliance on just one type of assay may lead to false-negative assessment of antiphospholipid antibodies. Overall, among the commonly applied assays, it appears that those for LA associate most strongly with clinical events. In relation to solid-phase assays, high-titre antibodies and IgG antibodies associate more strongly with clinical manifestations than do IgM and lower-titre antibodies.

Coagulation assays (lupus anticoagulant tests)

The LA assay is a double misnomer for it is neither a test for systemic lupus erythematosus nor for an *in vivo* anticoagulant. LA tests are indirect assays that rely on slowing of the clotting time of plasma through interference by antiphospholipid antibodies. The tests most frequently employed are the activated partial thromboplastin time (APTT), the dilute Russell's viper venom time (DRVVT) and, less frequently now, the kaolin clotting time (KCT).

Platelet activation causes exposure of negatively charged phospholipid at the cell surface and therefore contamination of test plasma with platelets must be minimized, as these will limit the sensitivity of tests, particularly when plasma must be stored frozen prior to testing. Platelet depletion is achieved most commonly by double centrifugation.

None of the above coagulation assays is specific for LA. Specificity and sensitivity are also reagent dependent. For example, some partial thromboplastin reagents are insensitive to LA. If the same reagent is employed for LA tests as in routine laboratory screening for coagulopathy, sensitivity to LA must be assured. Factors that lengthen or shorten clotting times (other than antiphospholipid antibodies) potentially interfere in LA tests. Examples are anticoagulant drugs and clotting factor deficiencies and inhibitors, which lengthen clotting times, and increased clotting factor levels, especially FVIII, which shortens the time to clotting in the APTT, for example, potentially masking the presence of LA.

In order to reduce the risks of false interpretation, in addition to prolongation of clotting time in a phospholipid-dependent coagulation test, the criteria for LA positivity also include: (i) evidence of an inhibitor demonstrated by mixing studies and (ii) confirmation of the phospholipid-dependent nature of the inhibitor. In principle, laboratory tests should employ a detection or screening stage (prolongation of the clotting time) and a confirmation stage showing: (i) failure of correction of the prolongation when normal plasma is added in order to exclude factor deficiency as the cause of the prolongation and (ii) that the prolongation is phospholipid dependent, for example by showing that addition of excess phospholipid corrects the clotting time. To achieve this, and because no LA test consistently shows 100% specificity and sensitivity, more than one test system should be used for detection of LA. The prothrombin time and thrombin time should also be performed as they are not usually affected by the presence of LA and the results assist in the interpretation of LA tests, for example when there is undisclosed anticoagulant therapy.

The APTT is commonly employed as the initial screening test for LA. Its specificity for inhibitor detection is improved by inclusion of a mixing study with platelet-free normal pooled plasma. When prolongation of the APTT is due to coagulation factor deficiency, the clotting time corrects when the test is repeated on an equal mixture of patient and normal plasma, whereas the prolongation above normal may persist with LA, consistent with its inhibitory activity. However, correction in a mixing study does not exclude LA, as a weak antibody is obviously diluted out by addition of normal plasma and this may be sufficient to abolish its effect. Very occasionally LA causes enhancement of the prolongation of the APTT when normal plasma is added. This phenomenon has been called the lupus cofactor effect, but is the exception rather than the rule. Inhibitors to clotting factors, usually FVIII, are associated with bleeding rather than thrombosis, but also cause prolongation of the clotting times which may not be corrected by addition of normal plasma. Typically, FVIII inhibitors are time-dependent, unlike LA. Diagnostic confusion may arise due to LA that prolongs the APTT, causing erroneously low results in coagulation factor assays based on the APTT. A normal APTT is insufficient to exclude LA and additional tests must be performed.

The DRVVT does not involve the clotting factors of the extrinsic system, unlike the APTT, nor FVII, unlike the prothrombin time. Any inhibition of coagulant-active phospholipid in the test by LA results in a prolonged DRVVT. However, as is the case with all LA tests, it is not specific. Deficiencies of clotting factors, for example FII and FX due to warfarin therapy, will also prolong the DRVVT. The specificity of the test is improved by repeating it in the presence of a high concentration of phospholipid, which should result in partial or complete correction of the prolonged clotting time if it is due to LA. This phospholipid is conveniently provided as platelet membranes in which negatively charged phospholipid is exposed by freezing and thawing. In the presence of LA, the ratio of test to normal plasma clotting time is often in excess of 1.2, and corrects to less than 1.2, or at least partially, in the platelet neutralization procedure. As with all coagulation tests, because of variations in reagents and techniques it is essential that laboratories derive local normal ranges using a large number of plasma samples from healthy volunteers. The LA is unreliable in patients receiving a direct oral anticoagulant with anti Xa activity.

In the KCT no additional phospholipid is employed. The test therefore resembles the APTT in that it involves the extrinsic and common pathways of coagulation, but the sensitivity to LA is enhanced because the small amount of phospholipid present is only that derived from residual platelets in the test sample and plasma lipids. The test is affected by clotting factor deficiencies and anticoagulants, but specificity can be improved by use of normal plasma mixing at more than one ratio to test plasma. LA is identified when the KCT fails to correct, even after relatively large proportions of normal plasma are added, whereas in factor deficiency the KCT is corrected with small amounts of normal plasma.

Alternative tests for LA may be employed. They include the tissue thromboplastin inhibition test and clotting tests that use venoms other than Russell's viper venom. An example is the Taipan venom: this test has been shown to have a strong positive predictive value in patients receiving warfarin. Many laboratories rely on commercial assay kits for LA testing. It is essential that steps are taken to ensure internal and external quality assurance.

Solid-phase assays (for anticardiolipin and anti-β_2-GPI antibodies)

Solid-phase assays for antiphospholipid antibodies, such as the anticardiolipin ELISA, allow rapid processing of numerous serum samples and the results are not affected by factor deficiency or the use of anticoagulants. The introduction of international standards allows the calculation of anticardiolipin results in IgG or IgM antiphospholipid units (GPLU and MPLU, respectively) related to a given concentration of affinity-purified anticardiolipin immunoglobulin. Despite this, there remains a lack of precision, and comparability between laboratories using different assays is not ensured. Clinicians should be aware of the performance of the assay in use.

The anticardiolipin assay is not necessarily detecting the same set of antiphospholipid antibodies seen in positive LA testing, due to the antibody heterogeneity referred to above. Furthermore, the clinical significance of low-titre anticardiolipin is doubtful. Thus, in cases where the anticardiolipin titre is less than 30 GPLU and tests for LA are negative, a diagnosis of APS is inconclusive. Under these circumstances it is important to consider other causes of thrombosis or pregnancy morbidity.

Specific assays for anti-β_2-GPI antibodies have been developed, and several commercial kits are available. Anti-β_2-GPI

antibody assays may show higher precision and better correlation with the thromboembolic complications in APS and SLE than assays for anticardiolipin, and are less likely to show transient positive results in association with infection. Antiprothrombin antibodies generally exhibit poor specificity for venous thrombosis and recurrent fetal loss, and may be found in patients with infection, and are not included in the consensus criteria for diagnosis of APS.

The prevalence of antiphospholipid antibodies in subjects with thrombosis varies with selection criteria for testing. Because the risk of recurrent thrombosis appears to be great, antiphospholipid antibodies should be sought in subjects with arterial, venous or microvascular thrombosis where no other cause is apparent. Examples are younger subjects with ischaemic stroke in the absence of cardiovascular disease and subjects with unprovoked VTE.

The prevalence of persistent antiphospholipid antibodies among women with recurrent first-trimester miscarriage is around 15%, although a proportion of these are low-titre antibodies. In women with recurrent miscarriage due to APS the prospective fetal loss rate may be as high as 90%. In contrast, the prevalence of positive tests for antiphospholipid antibodies in unselected women of childbearing age is around 3% and they are not sensitive predictors of poor pregnancy outcome in women with no history of pregnancy complications. Because miscarriage is a common phenomenon, screening for antiphospholipid antibodies is not indicated after a single event. Maternal antiphospholipid antibodies may be downregulated during pregnancy, so tests are best performed preconceptually when possible. A small proportion of women with antiphospholipid antibodies also have anti-Ro antibodies. Their detection is important as anti-Ro is associated with a 2% risk of complete heart block in the fetus and a 10% chance of neonatal lupus.

Management of APS

Thrombosis

There is wide variability in severity of prothrombotic states between individuals with APS. The management of patients with antiphospholipid antibodies and previous thrombosis remains contentious. Retrospective observational studies suggested that these patients should remain on indefinite oral anticoagulation, maintaining an International Normalized Ratio (INR) of 3–4. However, two subsequent prospective randomized studies both indicated that a lower target INR of 2–3 is effective in preventing recurrent thrombosis in the majority of patients. There is less certainty for those with arterial thrombosis as the number of patients was smaller; clinical observation suggests that there is a significant subpopulation of APS with small-vessel thrombosis, some evident as lacunar infarcts on magnetic resonance imaging, who appear to require a target INR of 3–4 to prevent recurrent cerebral thrombosis.

Immunosuppressive therapy is not indicated in primary APS other than in the very rare case where thrombosis recurs despite intensive anticoagulant therapy. Corticosteroids and other immunomodulatory therapies have been administered. An exception may be catastrophic APS when combination treatment with antithrombotics, corticosteroids and other immunomodulatory therapies such as rituximab is administered as a potentially life-saving emergency measure.

Pregnancy

The management of a pregnancy in a mother with a poor obstetric history due to APS is based on the use of anticoagulation with empirical doses of LMWH in combination with low-dose aspirin. Initial studies indicated that this approach increases the chance of a successful outcome of a healthy live birth from 30% to 70–80%, although the total number of cases in randomized studies is limited and trials have given conflicting results. Nevertheless, LMWH prophylaxis has become standard, although some clinicians believe that combination antithrombotic therapy as a first line should be reserved for women with a previous history of placental dysfunction, while those with recurrent first-trimester-loss miscarriage should be given supporting care as first line, with aspirin or aspirin/heparin reserved for those with further pregnancy failure.

Thromboprophylaxis for those with APS and a previous history of thrombosis is based again on use of LMWH, although there is no international consensus on dosing or the need to monitor therapy.

Venous thromboembolism and cancer

The association between cancer and venous thromboembolism is well established. Thrombotic events rank second only to the direct effects of cancer as a cause of death in cancer patients (Table 45.3). In addition to intrinsic biological features of the primary disease there are other factors, including surgery, chemotherapy, radiotherapy, hormonal therapies and the use of indwelling central lines that increase the risk of thrombosis in this group of patients. The development of venous thromboembolism in the context of malignancy is generally a poor prognostic feature often being indicative of advanced and aggressive malignancy. Several mechanisms may contribute to the propensity to develop venous thromboembolism in the context of active cancer; indeed abnormalities in most of the pathways related to normal haemostasis have been described. However, amongst the most compelling are the expression of tissue factor and cancer procoagulant on tumour cells and the prothrombotic properties of mucins secreted by adenocarcinomas. Of note there is evidence that tissue factor expression, which may result from proto-oncogene expression and tumour-suppressor gene inhibition, confers a proangiogenic state, which may enhance the aggressiveness and invasiveness of cancers. Several tumour types

Table 45.3 Some pathogenic factors for thrombosis in cancer patients.

Stasis
• Immobility/surgery
Treatment
• Surgery, chemoradiotherapy and adjuvant therapies
Vessel wall/endothelial perturbation
• Local tumour infiltration
• Central venous catheters
Hypercoagulability
• Infection
• Cytokine-related prothrombotic changes
• Tissue factor/cancer procoagulant expression on tumour cells
• Disseminated intravascular coagulation
• Increased platelet activation
• Tumour mucins
• Altered fibrinolysis

including lung, brain, pancreas, stomach, ovary, renal and lymphoma have the strongest reported association with thrombosis development. Whether this is related to the direct effects of thrombosis or reflects inherent tumour aggressiveness is not clear. It has also been observed that cancer patients with VTE have poorer outcomes in terms of the need for hospital admission, higher rates of major bleeding and higher rates of thrombosis recurrence on anticoagulant therapy, compared with patients who do not have cancer.

It is now usual practice for the majority of patients with active cancer who are admitted to hospital in the UK to receive thromboprophylaxis during their stay. The role of routine thromboprophylaxis in ambulatory cancer patients receiving chemotherapy is less clear. A Cochrane review of nine RCTs involving 3538 patients, which compared patients receiving thromboprophylaxis (eight LMWH and one warfarin) with control groups found a reduction in VTE risk without a significant increased risk of bleeding. However, the number needed to treat to prevent one episode of thrombosis was 60, suggesting that further stratification to identify those individuals at the highest risk was needed. Risk assessment scores such as that developed by Khorana, which identified patients with stomach, pancreas, lung, lymphoma, gynaecologic, bladder and testicular cancer combined with a platelet count $>350 \times 10^9$/L, haemoglobin <100 g/L, leucocyte count $>11 \times 10^9$/L and a body mass index of >35 kg/m^2 may be helpful.

The management of established venous thromboembolism in patients with active cancer differs from usual therapy. Three trials have shown that continuous treatment with LMWH at therapeutic doses compares favourably with the use of coumarins. Guidelines from various societies, including the BCSH and ISTH now recommend this. The reasons for this difference in efficacy of LMWH is not clear, but it has been suggested that better compliance with therapy, lower rates of drug interactions and less interference around times of poor or limited oral intake (when vitamin K intake is variable) may contribute. The management of patients with cancer with an ongoing indication for anticoagulation after 6 months is less clear and a variety of options, including LMWH, warfarin and NOACs are used. Although it has been suggested that heparins improve life expectancy through an anticancer effect, recent data do not support this suggestion.

Whether to screen patients over 40 years of age who present with a first unprovoked VTE for a possible underlying cancer is contentious. In the UK, NICE suggested we should consider screening with an abdominopelvic CT scan (and a mammogram for women), but a recent randomized trial has shown that routine screening with CT of the abdomen and pelvis did not provide a clinically significant benefit. Patients with bilateral DVTs, recurrences on anticoagulation or very high D-dimers are more likely to have an undiagnosed cancer.

Thrombotic risk in myeloproliferative disease (polycythaemia rubra vera and essential thrombocythaemia) (see also Chapter 26)

Untreated, polycythaemia vera (PV) affects life expectancy, with the majority dying of vascular occlusion, which may involve large vessels or the microvasculature. There is clearly significant overlap between PV and essential thrombocythaemia (ET), and thrombosis may involve both arterial and venous circulation. There does appear to be a predisposition towards thrombosis in the cerebral circulation, but events occur elsewhere and notably involve the splanchnic vessels. A well-described entity is the development of unprovoked intra-abdominal venous thrombosis in individuals who have the JAK-2 V617F mutation without the typical peripheral blood phenotype of MPD.

In PV there is good evidence that the incidence of vascular occlusion is positively related to the packed cell volume, with the lowest incidence in patients with good control of haematocrit. The role of thrombocytosis in risk of thrombosis in PV is controversial. The risk of thrombosis in PV increases with advancing age, a past history of thrombosis and presence of other risk factors for thrombosis. Thrombosis risk is reduced by control of haematocrit and antiplatelet therapy.

The risk of thrombosis in ET is determined by platelet count, age and the presence of other risk factors for thrombosis. Recent data indicate an increased risk of thrombosis in patients who are JAK-2 V617F positive and in those with elevated white cell counts.

The microvascular events and vasomotor manifestations of both PV and ET almost certainly relate to quantitative and qualitative changes in platelets, for they are not seen in other forms of polycythaemia or thrombocytosis. It is possible that similar

small-vessel occlusive and vasomotor changes seen in the feet and hands may also occur in other parts of the body. Erythromelalgia, a syndrome consisting of painful burning red extremities with normal peripheral pulses, is the characteristic vasomotor disturbance. Physical findings may be absent or there may be warmth, duskiness and mottled erythema of the involved areas. Livedo reticularis is occasionally found. In the digital vessels, usually of the toes, but occasionally the fingers, the development of thrombosis can lead to digital ischaemia and gangrene. In relation to the cerebral circulation, a range of symptoms and signs, including transient cerebral ischaemia, transient monocular blindness, migraine, headaches and seizures are seen. Both the cerebrovascular complications and erythromelalgia of myeloproliferative disease may respond promptly to low-dose aspirin and/or platelet cytoreduction.

Acute promyelocytic leukaemia

In comparison with the severe haemorrhagic presentation in the majority of cases of acute promyelocytic leukaemia (APL), acute arterial thrombosis is a rare presenting feature. Thrombosis is associated with the hypogranular variant of APL, which accounts for 25% of all cases (see Chapters 19 and 20) and has also been attributed to tissue factor expression on leukaemic cells. Vascular occlusion is more common when a high white cell count ($>150 \times 10^9$/L) predisposes to leucostasis. It is important to recognize this variant of APL because the use of all-*trans* retinoic acid may increase the risk of thrombosis.

Inflammation and thrombosis

Systemic inflammation is a potent prothrombotic stimulus (Table 45.4). Inflammation upregulates procoagulant factors, downregulates physiological anticoagulants, inhibits fibrinolytic activity and increases the platelet count. Inflammatory mediators further promote coagulation by causing endothelial cell activation. Thrombin, a key enzyme in coagulation, also has cytokine-like activities, for it augments leucocyte adhesion and can activate endothelium, leucocytes and platelets. Thrombin-mediated signalling through protease-activated receptor (PAR)-1 on vascular cells leads to recruitment of inflammatory cells including monocytes. This increases the expression of negatively charged phospholipids, such as phosphatidylserine, on the surface of the cells, promoting surface procoagulant activity.

The interleukin (IL)-6 family of molecules mediate the acute-phase response. The protein C anticoagulant pathway appears to be highly influenced by inflammation. Thrombomodulin and endothelial cell protein C receptor are both downregulated by inflammatory cytokines. There is inhibition of the promoter of the thrombomodulin gene, as well as active pinocytosis to remove existing surface molecules, while neutrophil elastase readily cleaves thrombomodulin from the endothelial cell surface. Moreover, thrombomodulin is very sensitive to oxidation of exposed methionine by oxidants produced by leucocytes.

Endothelial cell activation causes multiple prothrombotic changes, including downregulation of fibrinolysis by increased production of plasminogen activator inhibitor (PAI)-1.

As a consequence of these changes, conditions that provoke an inflammatory response are associated with increased risk of venous thrombosis. These include inflammatory bowel disease, Behçet disease, systemic tuberculosis, SLE and diabetes. Atherosclerosis can be considered a chronic inflammatory state and has recently been associated with an increased incidence of venous thrombosis and, more recently, the metabolic syndrome has been implicated. Furthermore, epidemiological studies indicate an excess of minor illness, including infections, in the weeks preceding episodes of VTE. The association between inflammation and thrombosis emphasizes the need for thromboprophylaxis in patients hospitalized on medical wards.

Table 45.4 Effects of inflammation on haemostasis.

Increased
• Tissue factor expression
• Surface procoagulant activity, negatively charged phospholipid
• Platelet reactivity
• Levels of fibrinogen and other coagulation proteins
Decreased
• Thrombomodulin expression
• Endothelial cell protein C receptor
• Half-life of activated protein C
• Protein Z
• Fibrinolytic activity due to increased PAI-1
• Endothelial gylcosaminoglycans
• Plasminogen activator inhibitor

Haematological prothrombotic states due to non-malignant diseases of the blood and bone marrow

Paroxysmal nocturnal haemoglobinuria (see Chapter 11)

Venous thrombosis occurs in up to 40% of patients with paroxysmal nocturnal haemoglobinuria (PNH) and represents a significant cause of morbidity and mortality. It may be the presenting feature or may predate the diagnosis. Thrombosis occurs at all sites, but there is an excess of events in the hepatic and portal veins, elsewhere in the splanchnic circulation and in the cerebral veins. Occasionally, painful discoloured skin lesions occur when the dermal veins are affected. These lesions rarely ulcerate. Occasionally, skin lesions can resemble purpura fulminans;

these can affect large areas of skin with necrosis and demarcation. PNH in pregnancy is associated with an increased risk of fetal loss (40%) as a result of thrombosis and haemorrhage.

Thrombosis at presentation is associated with a poorer prognosis with a 40% 4-year survival rate. The occurrence of thrombosis in patients with PNH is associated with a relative risk of death of up to 15-fold.

The pathogenesis of thrombosis in PNH remains uncertain. The effect of the PNH clone size is not entirely clear and although thrombosis has been linked to larger clone size, thrombosis is also markedly increased in patients with much smaller clones. Platelet activation is likely to be one of the main causes of thrombosis in this condition. Although platelets are capable of compensating for the decreased expression of decay accelerating factor (CD55) the absence of membrane inhibitor of reactive lysis (MIRL, CD59) renders them susceptible to complement-mediated lysis. Resultant platelet vesicles or micro particles are very procoagulant. The externalized phosphatidyl serine acts as a binding site for prothrombinase and tenase complexes. Fibrinolysis is also affected in PNH. Urokinase-type plasminogen activator receptor is also a GPI-bound protein that is absent from PNH cells. It binds urokinase to the cell surface and converts plasminogen to plasmin. As discussed in Chapter 11 the inhibition of complement (C5a) by eculizumab has a profound effect on thrombosis in PNH.

Thrombotic thrombocytopenic purpura

See Chapter 43.

Sickle cell disease (see Chapter 7)

In sickle cell disease there is microvascular and macrovascular occlusion with sickled cells, with resultant thrombosis distal to the sickling. Additional evidence for a prothrombotic state is enhanced thrombin generation, as shown by increased levels of prothrombin fragment 1+2 and thrombin–antithrombin complexes, in patients with sickle cell disease in their steady state when compared with age-matched controls. There is evidence that sickled erythrocytes adhere more readily to vascular endothelium and strongly accelerate coagulation due to abnormal exteriorization of procoagulant anionic membrane phospholipids. Comparison of the coagulation and fibrinolytic pathways in sickle cell patients with ethnically matched controls has shown increased levels of von Willebrand factor during sickling, although it is not clear whether this is specific or part of the acute-phase reaction.

Selected bibliography

Breen KA, Grimwade D, Hunt BJ (2012) The pathogenesis and management of the coagulopathy of acute promyelocytic leukaemia. *British Journal of Haematology* **156**(1): 24–36.

Carrier M, Lazo-Langner A, Shivakumar S *et al*. (2015) Screening for occult cancer in unprovoked venous thromboembolism. *New England Journal of Medicine*. Available at: http://www.nejm.org/doi/full/10.1056/NEJMoa1506623. Accessed on 9 August 2015.

Keeling D, Mackie I, Moore GW *et al*. (2012) Guidelines on the investigation and management of antiphospholipid syndrome. *British Journal of Haematology* **157**: 47–58.

Hill A, Kelly RJ, Hillmen P (2013) Thrombosis in paroxysmal nocturnal haemoglobinuria. *Blood* **121**:4985–96.

Venous thromboembolic diseases: the management of venous thromboembolic diseases and the role of thrombophilia testing. http://www.nice.org.uk/guidance/cg144 (accessed June 2015).

The acute management of thrombosis and embolism during pregnancy and puerperium (Green-top Guideline No. 37b) https://www.rcog.org.uk/en/guidelines-research-services/guidelines/gtg37b/ (accessed June 2015).

Truelove E, Fielding A, Hunt BJ (2013) The coagulopathy and thrombotic risk associated with L-aspararinase treatment in adults with acute lymphatic leukaemia. *Leukaemia* **27**: 553–9.

Watson HG, Davidson S, Keeling D (2012) The management of heparin-induced thrombocytopenia; second edition. *British Journal of Haematology* **15**: 528–40.

Watson HG, Keeling DM, Laffan M *et al*. (2015) Guidelines on cancer associated venous thrombosis. *British Journal of Haemotology* (In press).

CHAPTER 46

Antithrombotic agents

46

Trevor Baglin[1] and David Keeling[2]

[1]Cambridge University Hospitals NHS Trust, Addenbrookes Hospital, Cambridge, UK
[2]Oxford University Hospitals, Churchill Hospital, Oxford, UK.

Heparins

Unfractionated heparin (UFH) is a naturally occurring glycosaminoglycan produced by mast cells. Most of the heparin used worldwide is extracted from porcine intestine. Heparin consists of saccharide chains of alternating uronic acid and glucosamine residues with varying degrees of sulfation. The long chains of UFH have a mean molecular weight of approximately 15,000. For most clinical indications there has been a move from UFH to low-molecular-weight heparin (LMWH). LMWH is manufactured from UFH by controlled depolymerization of the chains using chemical (nitrous acid or alkaline hydrolysis) or enzymatic (heparinase) methods. The LMWHs currently available for clinical use, dalteprin, enoxaparin and tinzaparin, have a mean molecular weight between 3000 and 5000.

The ability of UFH and LMWH chains to act as anticoagulants depends on the presence of a specific pentasaccharide sequence, which binds with high affinity to antithrombin and produces a conformational change that releases the reactive centre loop (RCL), greatly potentiating its activity (see Chapter 44). This conformational change makes antithrombin a potent inhibitor of FXa. Potentiation of thrombin inhibition by the RCL also needs the heparin chain to bind to exosite II on thrombin, and this requires a minimum total chain length of 18 saccharides. For this reason, relative to UFH, LMWH inhibits FXa more than IIa (Figure 46.1).

Fondaparinux is a synthetic pentasaccharide and is not an animal product. The pentasaccharide results in antithrombin inhibiting Xa but not IIa.

Heparins are only active when administered parenterally. UFH can be given intravenously or subcutaneously, but is usually given by continuous intravenous infusion. Metabolism is by a saturable mechanism, involving binding to endothelial cells and clearance by the reticuloendothelial system, and a non-saturable mechanism involving mainly renal clearance and, at typical therapeutic doses, the half-life of intravenous UFH is 45–60 min. There is no evidence that heparin crosses the placenta. The APTT is used for routine monitoring of therapeutic doses of UFH and in the early studies a therapeutic range of an APTT ratio of 1.5–2.5 was established, which corresponded to a heparin level of 0.2 to 0.4 IU/mL by protamine titration or 0.3 to 0.7 IU/mL by anti-Xa assay. However, APTT methods vary markedly in their responsiveness to UFH so local calibration of the APTT assay should be employed in the construction of the local therapeutic range. Directly monitoring using an anti-Xa assay has some theoretical advantages being insensitive to baseline levels of FVIII and other clotting factors and the presence of lupus anticoagulant, and has greater precision. An inadequate APTT response in the first 24 hours may increase the risk of recurrence of thromboembolism, although this does not seem to be critical if the starting infusion rate is at least 1250 IU/h. A validated regimen is to give a bolus dose of 80 IU/kg and to start the infusion at 18 IU/kg/h, performing the first APTT/anti-Xa estimate, and if necessary dose adjustment, within 6 hour. If bleeding occurs, stopping an UFH infusion and general haemostatic measures are often sufficient. If required, protamine sulfate (1 mg per 80–100 units UFH) will reverse UFH; it should be given slower than 5 mg/minute and a maximum dose of 50 mg protamine is usually sufficient.

Inhibition of aldosterone secretion by heparin can result in hyperkalaemia; patients with diabetes mellitus, chronic renal failure, acidosis, raised plasma potassium or those taking

Postgraduate Haematology, Seventh Edition. Edited by A Victor Hoffbrand, Douglas R Higgs, David M Keeling and Atul B Mehta.

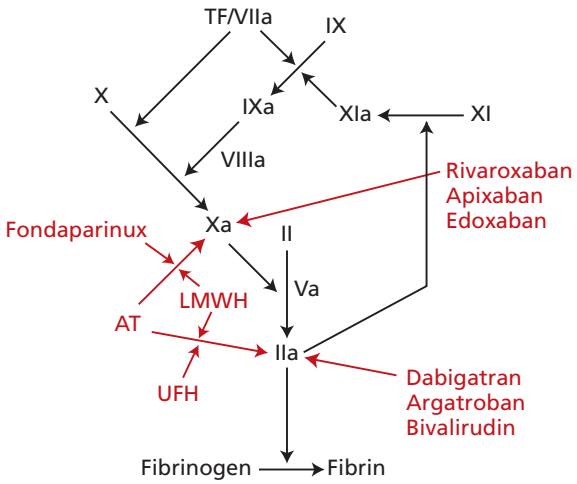


Figure 46.1 Inhibitors of coagulation. AT, antithrombin; UFH, unfractionated heparin; LMWH, low-molecular-weight-heparin.

potassium-sparing drugs seem to be more susceptible. A slight fall in platelet count (<30% compared with baseline) occurs in up to one-third of patients within the first four days of starting heparin. This reversible dose-dependent phenomenon is not associated with bleeding or thrombotic complications and does not require cessation of heparin. More marked falls in platelet count (>50% compared with baseline) after 5 days of treatment should raise the possibility of heparin-induced thrombocytopenia (HIT) (see Chapter 45). Prolonged use of UFH for months can lead to osteoporosis by increasing osteoclast activity and reducing osteoblast numbers.

LMWH has better bioavailability after subcutaneous injection and the anticoagulant effect is more predictable due to less binding to plasma proteins. This means that the dose can be calculated by body weight and be given subcutaneously without any monitoring or dose adjustment. Despite a half-life of 3–5 hours, once daily dosing is adequate for the treatment and prevention of venous thromboembolism (VTE). The actual dosage used differs slightly with the different LMWH and the manufacturers' recommendations should be followed, but a typical dose is 200 IU/kg subcutaneously once a day. Monitoring is rarely needed, but at therapeutic doses on a once a day regimen peak levels (3–4 hours post dose) are approximately 1.0–1.5 anti-Xa units/mL (measured by chromogenic assay). If bleeding occurs protamine sulfate can be given, but is likely to be less effective than for UFH. HIT and osteoporosis are less common with LMWH than with UFH.

Fondaparinux is given as a once a day subcutaneous injection and has a half-life of 17 hours. Clearance is exclusively renal, so dose adjustment in renal failure is required and it is not recommended in those with a creatinine clearance below 30 mL/min. Fondaparinux does not bind to platelet factor 4, so has no capacity to cause HIT.

Danaparoid

Danaparoid is a heparinoid composed of heparan sulfate, dermatan sulfate and chondroitin sulfate. It is mainly used in patients with HIT. It indirectly inhibits Xa and to a lesser degree thrombin. It has a predictable dose response and a long half-life of approximately 24 hours. Danaparoid does not prolong the PT and has a minimal effect on the APTT which cannot be used to monitor it. If monitoring is required, a specific anti-Xa assay calibrated for danaparoid should be used. Monitoring may be of value only in patients with severe renal impairment and body weight greater than 90 kg.

Direct thrombin inhibitors

Bivalirudin

Bivalirudin is a recombinant peptide which directly inhibits thrombin. Bivalirudin is cleared predominantly through proteolysis by thrombin (80%) and only 20% is excreted renally. The half-life is approximately 25 minutes and it is given by continuous intravenous infusion and normally monitored by the APTT. It is licensed for acute coronary syndrome patients undergoing percutaneous coronary intervention.

Argatroban

Argatroban is a reversible direct thrombin inhibitor. It is rapidly eliminated by hepatic cytochrome P450 and has a plasma half-life of approximately 50 minutes. It is licensed for anticoagulation in patients with heparin-induced thrombocytopenia and is very useful in those who also have renal impairment due to its hepatic metabolism. It is given as a continuous intravenous infusion monitored by the APTT.

Vitamin K antagonists (VKA)

Vitamin K is needed for the post-translational modification of the vitamin-K-dependant coagulation factors (II, VII, IX, X, protein C and protein S). Vitamin K is a cofactor for vitamin-K-dependent carboxylase, an enzyme that catalyses the carboxylation of glutamic acid (Glu) residues in these proteins, forming their Gla-domains (Chapter 36). The Gla domains of the vitamin-K-dependent coagulant and anticoagulant proteins allow calcium-dependent binding to negatively charged phospholipids which is necessary for the formation of coagulation complexes. Vitamin K antagonists, such as warfarin, acenocoumarol, phenindione, fluindione and phenprocoumon therefore inhibit synthesis of normal FII, FVII, FIX and FX and so result in anticoagulation. There are a number of common genetic polymorphisms of the CYP2C9 enzyme

Table 46.1 Many drugs interact with warfarin; some important interactions, where the evidence is strongest, are listed.

Potentiation	Inhibition
Amiodarone	Barbiturates
Cimetidine	Carbamazepine
Clarithromycin	Griseofulvin
Clofibrate	Rifampicin
Cotrimoxazole	Sucralfate
Erythromycin	
Fluconazole	
Glucosamine	
Metronidazole	
Miconazole	
Omeprazole	
Paracetamol	
Statins	
Sulfinpyrazone	

primarily responsible for warfarin metabolism. The most common polymorphisms, CYP2C9*2 and CYP2C9*3, are each seen in around 10% of whites, and are associated with a reduced warfarin requirement. There is, however, insufficient evidence to warrant genotype-guided dosing. Although it has been shown that genetic testing can predict the maintenance dose, for initiation, information from previous dosing rapidly becomes more important and response to a standard dosing algorithm can accurately predict maintenance dose without genotyping. The anticoagulant effect is also affected by dietary vitamin K intake and warfarin and has many drug interactions (Table 46.1). Due to these environmental and genetic factors the anticoagulant effect varies between individuals and over time and needs to be monitored by the International Normalized Ratio (INR). This is a manipulation of the prothrombin time (PT) to allow for the different sensitivities of various laboratory reagents to the warfarin-induced coagulopathy. The INR equals $(PT/MNPT)^{ISI}$, where MNPT is the (geometric) mean normal PT and ISI is the international sensitivity index of the thromboplastin used in the assay. For the treatment of deep vein thrombosis (DVT) and pulmonary embolism (PE), and stroke prevention in atrial fibrillation (AF) the target INR should be 2.5 (target range 2.0–3.0). A higher target INR is used for some mechanical heart valves (MHV).

VKAs take a number of days to become effective, so if immediate anticoagulation is needed, heparin is given initially. For warfarin induction there is no clear advantage for 10 mg or 5 mg loading doses other than in the elderly, where lower initiation doses or age-adjusted doses may be more appropriate, leading to fewer high INRs. The dosing algorithm used in Oxford is shown in Table 46.2

Table 46.2 A warfarin dosing algorithm starting with 5 mg doses.

	Day 3		Day 4	
Days 1 and 2	INR	Dose (mg)	INR	Dose (mg)
Give 5 mg each day if baseline INR ≤ 1.3	<1.5	10	<1.6	10
	1.5–2.0	5	1.6–1.7	7
	2.1–2.5	3	1.8–1.9	6
	2.6-3.0	1	2.0–2.3	5
	>3.0	0	2.4–2.7	4
			2.8–3.0	3
			3.1–3.5	2
			3.6–4.0	1
			>4.0	0

When a VKA is started, the vitamin-K-dependent factors fall according to their half-lives. Factor VII and protein C have the shortest half-lives, so that despite a prolongation of the INR due to factor VII deficiency, warfarin may initially be procoagulant. This is the mechanism for the rare problem of warfarin-induced skin necrosis, most often described in those with protein C deficiency. Other rare side-effects of vitamin K antagonists are purple toe syndrome, rash, hair loss and hepatitis. The purple toe syndrome is a very painful, burning, dark-blue discoloration of the toes and sides of the feet, it is hypothesized that this is possibly due to cholesterol embolization from atherosclerotic plaques that have become friable owing to reduced fibrin deposition or haemorrhage into the plaques several weeks after initiation of anticoagulation. Vitamin K antagonists also have a teratogenic effect, with skeletal malformations, optic atrophy and mental impairment, occurring in a small percentage of babies of exposed mothers.

For some invasive procedures such as joint injections, cataracts and certain endoscopic procedures, warfarin does not need to be stopped. If anticoagulation has to be stopped for surgery or an invasive procedure, it should be stopped for 5 days beforehand. Consideration then needs to be given to which patients require 'bridging' therapy with an alternative short-acting anticoagulant (usually LMWH). Patients with a VTE within the previous 3 months, patients with AF with a history of previous stroke or transient ischaemic attack (TIA) or multiple other risk factors, and patients with a mitral mechanical heart valve should be considered for bridging therapy, though a recent study has suggested that bridging is not beneficial even in AF patients with a history of previous stroke or TIA (Table 46.3). Warfarin can be resumed, at the normal maintenance dose, the evening of surgery or the next day if there is adequate haemostasis.

Emergency anticoagulation reversal in patients with major bleeding should be with 25–50 U/kg four-factor prothrombin

Table 46.3 Risk stratification for bridging therapy when warfarin is stopped for an invasive procedure.

	Bridging with treatment dose heparin not required	Bridging with treatment dose heparin considered
VTE	Last episode >3 months ago	Last episode within previous 3 months
AF	AF with no prior stroke/TIA and without multiple other risk factors	AF and previous stroke/TIA* or multiple other risk factors
MHV	Bileaflet aortic MHV with no other risk factors	Mitral MHV Non-bileaflet aortic MHV Bileaflet aortic MHV with other risk factors

*A recent study has suggested that bridging is not beneficial even in those with a history of previous stroke/TIA.
AF, arterial fibrillation; MHV, mechanical heart valves; VTE, venous thromboembolism.

complex concentrate (PCC) and 5 mg intravenous vitamin K. Anticoagulation reversal for non-major bleeding should be with 1–3 mg intravenous vitamin K, whilst patients with an INR >8.0, but who are not bleeding can receive 1–5 mg of oral vitamin K.

Non-vitamin K antagonist oral anticoagulants (NOACs)

Small molecules that selectively and specifically inhibit coagulation serine proteases have been developed for clinical use (Table 46.4). NOACs are more convenient than oral vitamin K antagonists (VKA) and are at least as safe and effective as warfarin. In some patients they will be superior or safer than VKA, but for the majority of patients with good anticoagulant control with VKAs the advantages of the new agents are primarily convenience and few drug interactions. Originally called new or novel oral anticoagulants (NOACs) they are now neither new nor novel. Consequently some experts have termed these drugs direct oral anticoagulants (DOACs), whilst others have retained the acronym NOAC with N representing 'Non-VKA'; we will adopt this latter approach.

Ximelegatran, a prodrug of melagatran, was the first oral direct inhibitor of thrombin and was an effective antithrombotic drug. However, liver toxicity with long-term dosing resulted in a failed approval and withdrawal of the drug. Currently NOACs are either inhibitors of thrombin (dabigatran) or inhibitors of Xa (rivaroxaban, apixaban, edoxaban) (Figure 46.1). All anticoagulant drugs inhibit coagulation by reducing thrombin generation, blocking thrombin activity, or both. The results of studies with thrombin and FXa inhibitors do not indicate that one target is better than the other. Direct comparisons could determine differences in efficacy and safety between NOACs, but are unlikely to be performed. Indirect comparisons between trials are problematic because of confounding due to differences in patients, interventions and outcomes, and as yet there is no convincing evidence that any particular NOAC is more effective or safer.

NOACs are prescribed at fixed dose without the need for monitoring or dose adjustment. A rapid onset of action and short half-life make initiation and interruption of anticoagulation considerably easier than with VKAs (Table 46.5). Specific antidotes to NOACs are not yet available for clinical use, but are in development as rapid reversal agents and will soon be available. As with all anticoagulants produced so far there is a correlation between intensity of anticoagulation and bleeding. Consequently, the need to consider the balance of benefit and risk in each individual patient is no less important than with VKA therapy.

NOACs offer an improvement over VKAs by virtue of:

- predictable dose responses and so no need for routine monitoring
- no food interactions
- limited drug interactions.

Safety and efficacy of NOACs was originally demonstrated in patients undergoing elective hip and knee replacement surgery as the endpoint (radiographically detectable venous thrombosis) occurs frequently within 14 days of surgery and therefore efficacy is demonstrable quickly in relatively small studies. Much larger studies conducted over several years were required to demonstrate efficacy in comparison to warfarin for stroke prevention in patients with atrial fibrillation and for treatment and secondary prevention of venous thromboembolism.

Poor compliance with VKAs is not necessarily an indication to switch to a NOAC, as the combination of short half-life and lack of monitoring with NOACs may make the balance of benefit and risk unfavourable in patients with poor compliance. However, a NOAC may be preferred if instability or low therapeutic-time-in-range (TTR) with a VKA is due to drug interactions. NOACs are not associated with heparin-induced thrombocytopenia (HIT), although they have not been investigated as a treatment option in patients who develop HIT, with or without associated thrombosis.

In clinical trials in patients with venous thromboembolism 5% to 7% of patients had cancer and NOACs were as effective as warfarin. However, in patients with solid tumours, treatment with VKA is inferior to therapeutic-dose low-molecular-weight heparin (LMWH) and so LMWHs remain the preferred treatment option in patients with cancer and thrombosis. A decision to use a NOAC can be made on an individual basis taking into account ease of use and quality of life.

All NOACs are excreted to a variable degree by the kidneys (Table 46.4). Dabigatran is contraindicated when the creatinine clearance is less than 30 mL/min but anti-Xa inhibitors can be

Table 46.4 Properties of NOACs. Low protein binding may allow removal of drug by dialysis. Factor Xa inhibitors contraindicated when CrCl <15 mL/min and dabigatran contraindicated when CrCl <30 mL/min. Apixaban appears to have least effect on PT and APTT.

NOAC	Target	Time to peak (h)	$t_{1/2}$ (h)	Dose* AF	Dose VTE	LMWH used for initial Tx of VTE	Renal excretion	Reduced dose*	Indications for reduced dose* in AF	Protein binding	Quantitative assay	Qualitative assay
Dabigatran	IIa	<4	12–17	150 mg bd	150 mg bd	Yes	80%	110 mg bd	Age >80; Concomitant use of verapamil; Consider if CrCl 30 to 50 mL/min	Low	Dilute plasma thrombin time, e.g. Hemoclot ECT	APTT
Rivaroxaban	Xa	<4	7–13	20 mg od	15 mg bd for 3 weeks then 20 mg od	No	33%	15 mg od	CrCl 15 to 49 mL/min	High	Calibrated anti-Xa	PT
Apixaban	Xa	<4	10–14	5 mg bd	10 mg bd for 7 days then 5 mg bd	No	25%	2.5 mg bd	CrCl 15 to 29 mL/min, or two of the following: • age 80 years or greater; • body weight 60 kg or less • serum creatinine of 133 µmol/L or more.	High	Calibrated anti-Xa	Possibly none
Edoxaban	Xa	<4	9–11	60 mg od	60 mg od	Yes	35%	30 mg od	CrCl 30 to 50 mL/min Body weight <60 kg	Low	Calibrated anti-Xa	PT

*All dosing regimens should be checked in current summary of product characteristics. CrCl, creatinine clearance; ECT, ecarin clotting time.

Table 46.5 Suggested times for stopping NOACs prior to elective surgery based on $t_{1/2}$ and renal function and suggested doses for restarting. Re-introduction should be delayed due to fast onset of anticoagulation and achievement of therapeutic anticoagulation within 4 hours of first therapeutic dose. Longer periods should be considered in elderly patients. Low-dose LMWH or prophylactic dose of NOAC may be considered for thromboprophylaxis when therapeutic dose NOAC is interrupted.

NOAC	Major surgery	Minor surgery
Dabigatran (CrCl >80 mL/min)	48	24
Dabigatran (CrCl 50–80 mL/min)	72	48
Rivaroxaban	48	24
Apixaban	48	24
Edoxaban	48	24

used, albeit at reduced dose, with a creatinine clearance down to 15 mL/min. Therefore, when there is moderate renal impairment a factor Xa inhibitor may be preferable. The estimated glomerular filtration rate (eGFR) may over-estimate renal function and so it is preferable to calculate the creatinine clearance when the eGFR is less than 50 mL/min. There is no liver toxicity associated with NOACs and they can be prescribed as long as there is no coagulopathy.

NOACs should not be prescribed in pregnancy or during breast-feeding and women of childbearing age should be warned of this before starting treatment.

NOACs should not be used for anticoagulation in patients with mechanical heart valve prostheses. Dabigatran has been compared to warfarin in one trial and despite being used at higher doses than those used in patients with atrial fibrillation it was not as effective as warfarin at preventing thromboembolism. As a result of the higher doses the incidence of major bleeding was more than doubled. The relative effectiveness of anti-Xa inhibitors in patients with mechanical heart valves is currently unknown.

Clinically significant interactions may occur with NOACs and drugs that affect P-glycoprotein (P-gp) and CYP3A4. The only drugs that interact with dabigatran are inhibitors or inducers of P-gp. CYP3A4 has almost no role in the metabolism of dabigatran, whereas it is involved in the metabolism of the factor Xa inhibitors. Unlike VKAs, NOACs are not metabolized by CYP2C9. Drugs that inhibit P-gp include:

- Azoleantimycotics – ketoconazole, itraconazole, voriconazol, posaconazole
- Immunosuppresants – ciclosporin, tacrolimus
- Amiodarone, verapamil, quinidine
- Rifampicin
- Antiepileptics – carbamazepine, phenytoin
- St John's wort.

Pharmacokinetics of factor Xa inhibitors are affected by drugs that affect P-gp and/or CYP3A4. CYP3A4 is responsible for oxidative metabolism of Xa inhibitors and drugs that act both as strong inhibitors of both CYP3A4 and of P-gp have been shown to increase plasma concentrations. These drugs include azole antimycotics and viral protease inhibitors used for treatment of HIV and hepatitis C. Coadministration of strong inducers of P-gp, such as phenytoin, carbamazepine, phenobarbitone and St. John's wort should be avoided.

Dabigatran

Dabigatran etexilate is an oral prodrug that is hydrolysed by carboxyl esterases to the active compound dabigatran after absorption. It is a direct specific competitive inhibitor of free and fibrin-bound thrombin which binds to the active site of thrombin with high affinity. Peak plasma levels are reached 2 to 3 hours after ingestion, although this may be delayed for up to 6 hours after the first postoperative dose. Renal excretion accounts for 80% of dabigatran clearance. With a creatinine clearance (CrCl) of 80 mL/min, the half-life is 13 hours and it increases to 27 hours when the CrCl is below 30 mL/min. Severe renal insufficiency, defined by a CrCl less than 30 mL/min, is a contraindication to dabigatran. Concomitant use of antiplatelet agents such as aspirin or clopidogrel should only take place after careful consideration of the risks and benefits and non-steroidal anti-inflammatory drugs should be avoided, as bleeding rates are 50% higher in patients receiving antiplatelet drugs with dabigatran. Dyspepsia occurs in 10% of patients, which may be due to tartaric acid in the capsule formulation.

The recommended dose is 150 mg bd with a dose reduction to 110 mg bd in patients over the age of 80 and those taking concomitant verapamil. In those with a CrCl of between 30 and 50 mL/min, a dose reduction should be considered. The lower dose can also be used when it is thought that the balance of thromboembolism and bleeding risk is shifted towards bleeding. It is recommended that the CrCl is calculated or the glomerular filtration rate is estimated (eGFR) before starting treatment and this is repeated at least annually or if there is a suspected deterioration in renal function.

Food has no significant effect on absorption of dabigatran etexilate. The capsules should not be opened as this increases absorption unpredictably by up to 75% and this can lead to high dabigatran levels. Only blister packs should be used, as the formulation loses potency after exposure and capsules should be discarded after 60 days of exposure.

Rivaroxaban

Rivaroxaban is a direct competitive inhibitor of factor Xa and so limits thrombin generation. It binds to the active site of factor Xa with high affinity. Absorption of drug is rapid with peak levels 2 to 3 hours after ingestion. Two-thirds of the drug is metabolized

in the liver with only one-third excreted by the kidneys. Food increases the absorption of rivaroxaban and it is recommended that the capsules should be taken with food. Due to a ceiling effect on absorption no increase in plasma levels are found with doses of rivaroxaban above 50 mg.

The recommended dose in patients with atrial fibrillation is 20 mg once daily, and for treatment of acute venous thrombosis 15 mg twice daily for three weeks followed by 20 mg once daily. No dose adjustment is required in the elderly. However, a dose reduction from 20 mg once daily to 15 mg once daily is recommended for patients with atrial fibrillation and a CrCl between 15 and 30 mL/min. Rivaroxaban is not recommended when the CrCl is less than 15 mL/min.

Apixaban

Apixaban is a direct competitive inhibitor of factor Xa. Apixaban is rapidly absorbed with maximum concentrations 3 to 4 hours after intake. Food does not affect absorption of apixaban. The initial treatment dose for patients with VTE is 10 mg bd for 7 days followed by 5 mg bd for at least three months. For those requiring long-term anticoagulation (>6 months) for secondary prevention of VTE the dose is reduced to 2.5 mg bd. The treatment dose for AF is 5 mg bd, reduced to 2.5 mg bd in patients with a CrCl of less than 30 mL/min or with two of the following:

- age 80 years or greater
- body weight 60 kg or less
- serum creatinine of 133 μmol/L or more.

Apixaban is not recommended in patients with a CrCl of 15 mL/min or less.

Edoxaban

Edoxaban is another factor Xa inhibitor which has completed Phase III trials in patients with atrial fibrillation and venous thromboembolism.

Measurement of anticoagulant effect of NOACs

In most circumstances, NOACs have predictable bioavailability and pharmacokinetic and pharmacodynamic profiles. However, there will be clinical circumstances when assessment of the anticoagulant effect of these drugs may be required, such as:

- before surgery or invasive procedure when a patient has taken a drug in the previous 24 hours (or longer if CrCl <50 mL/min on dabigatran)
- when a patient is bleeding
- when a patient has taken an overdose
- when a patient has developed renal failure or in patients with deteriorating renal function
- establishing the optimal dose in patients taking other drugs that are known to significantly affect pharmacokinetics
- establishing the optimal dose in patients at extremes of body weight
- when a patient has thrombosis on treatment to assess whether there is failure of therapy or lack of adherence (this may have limited application due to the short half-lives of NOACs in comparison to VKAs).

Ideally a quantitative assay should be used to measure the level of drug. For Xa inhibitors this will generally be with a calibrated anti-Xa assay. For thrombin inhibitors a modified dilute plasma thrombin time such as the Hemoclot® assay can be used. The Ecarin clotting time (ECT) can also be used.

The result of a qualitative test such as the prothrombin time (PT) or activated partial thromboplastin time (APTT) can indicate whether anticoagulation is approximately supratherapeutic, therapeutic or subtherapeutic, but cannot be used to determine the plasma concentration of the drug. Laboratories should be aware of the sensitivity of their own assays to each drug. This can be achieved by comparing the PT and APTT results for each NOAC against the measured level in a quantitative assay. If quantitative assays are not readily available, plasma samples on which the PT and APTT have been measured can be sent to a specialist coagulation laboratory using quantitative assays. The APTT using most reagents can be used for urgent determination of the relative intensity of anticoagulation due to dabigatran. The APTT cannot be used to determine the drug level. A normal thrombin time indicates a very low level of dabigatran. With an appropriate reagent the PT (or APTT with some reagents) can be used for the urgent determination of the relative intensity of anticoagulation due to rivaroxaban. It cannot be used to determine the drug level. PT and APTT reagents are generally insensitive to apixaban.

The interpretation of all test results is dependent on when the last dose of drug was taken, the dose, the anticipated half-life and factors that influence pharmacokinetics. It is likely that coagulation tests will be performed on patients taking anticoagulants as part of clinical assessment, e.g. admission to Accident and Emergency. The prothrombin time (PT), activated partial thromboplastin time (APTT), thrombin time (TT) and fibrinogen can be affected, and recognition of this is required for interpretation of results. NOACs do not interfere with the D-dimer assay, but D-dimer levels are lower in patients treated with anticoagulant drugs. Patients taking NOACs may have a prolonged APTT and/or PT (and a low fibrinogen with dabigatran with some reagents) and the results might wrongly be interpreted as suggesting disseminated intravascular coagulation (DIC). However, NOACs do not cause thrombocytopenia and the D-dimer level is likely to be low.

Interruption of anticoagulant treatment and switching with other anticoagulant drugs

Surgery may require interruption of anticoagulant therapy. Ideally spinal and epidural anaesthesia should not be performed

until no anticoagulant effect is detectable by a quantitative assay. In patients taking dabigatran, the standard thrombin time is very sensitive and if the thrombin time is normal dabigatran will likely be undetectable in a quantitative assay. After removal of an epidural catheter there should be an interval of 2 hours before the next dose of an NOAC. NOACs should be stopped 24 to 48 hours before surgery, depending on renal function (and so half-life) and whether the surgery carries a low or high risk of major bleeding. As NOACs have a rapid onset of anticoagulation, re-introduction at full dose after surgery should be delayed, for example for 48–72 hours, and then a lower dose or prophylactic LMWH can be considered, if haemostasis has been secured.

In case of suspected NOAC overdose, the anticoagulant effect should be measured. Oral activated charcoal may reduce absorption of NOACs if given in the first few hours after ingestion. As protein binding is low, dabigatran can be dialysed. This may be applicable when there is severe renal failure as the greatly prolonged half-life in the presence of renal failure can result in continued anticoagulation for several days.

When switching from an LMWH to an NOAC, the new drug can be started instead of the LMWH when the next dose is due. Likewise, when switching from an NOAC to subcutaneous LMWH, the latter can be started instead of the NOAC when the next dose is due.

When switching from a VKA to an NOAC the VKA should be stopped and the NOAC started when the INR is <2.0 or 2.5. When switching from an NOAC to a VKA the duration of drug overlap will be determined by the starting dose of VKA (loading or known maintenance dose). Measurement of the INR before the next dose of NAOC (i.e. at trough levels) can be used to indicate when the warfarin can be stopped (when the INR >2.0).

Management of bleeding patients treated with NOACs

The management of patients who are bleeding whilst taking NOACs is influenced by the severity of bleeding. The time of the last dose of NOAC should be determined and the half-life should be estimated from measurement of serum creatinine and calculation of the CrCl. The anticoagulant activity of the NOAC should be determined by the most appropriate laboratory assay. When bleeding is not severe it may only be necessary to stop the drug temporarily and determine the cause of bleeding and the material contribution, if any, of the NOAC. For more severe bleeding, treatment may include:

- mechanical compression (e.g. for epistaxis or superficial wounds)
- surgical haemostasis (sutures and cautery)
- fluid replacement
- correction of anaemia by transfusion of red cells
- correction of additional coagulopathy (e.g. dilutional coagulopathy) with platelet transfusion and appropriate blood products.

Fresh frozen plasma (FFP) does not reverse the anticoagulant effect of NOACs. The effect of desmopressin and tranexamic acid on bleeding due to NOACs is not known, but a general haemostatic effect independent of thrombin or factor Xa might be beneficial, and given its effect in traumatic bleeding the use of tranexamic acid is reasonable. The effect of prothrombin complex concentrate (PCC), activated PCC (aPCC) and recombinant factor VIIa (rVIIa) have not been studied in clinical trials in human patients with bleeding. Notably, prolongation of clotting times such as the PT and APTT due to anti-Xa inhibition can be partially reversed by PCC, whilst the effect of thrombin inhibition is not. In some animal experiments, bleeding in the presence of NOACs has been variably reduced by administration of PCC and aPCC with evidence of a dose response. Until specific antidotes are available (a)PCCs may be considered for patients with severe haemorrhage associated with NOACs and in such a cases a dose of 50 U/kg would be reasonable based on extrapolation from animal studies. Idarucizumab is humanized monoclonal mouse antibody with high dabigatran binding affinity which has been developed as a specific antidote to dabigatran. Andexanet is a truncated recombinant FXa that lacks the GLA domain and is catalytically inactive because of a mutation (S419A) in the catalytic triad, it avidly binds Xa inhibitors and has been developed as a specific antidote to the Xa inhibitors. Both should be available in 2016.

Antiplatelet drugs

Antiplatelet drugs reduce the risk of the combined outcome of stroke, myocardial infarction (MI) and ischaemic stroke. They reduce vascular death by about 15% and non-fatal vascular events by 30%. They also reduce the incidence of myocardial infarction in patients with unstable angina, reduce acute occlusion of coronary bypass grafts and improve walking distance and decrease vascular complications in patients with peripheral vascular disease.

Aspirin

Aspirin (acetylsalicylic acid) acetylates and thus inactivates cyclo-oxygenase (COX), the enzyme responsible for the conversion of arachidonic acid to prostaglandin H2 and subsequently the generation of thromboxane (TXA_2) and prostacyclin (PGI_2). As acetylation of COX is irreversible and the platelet is unable to synthesize new enzyme, COX activity is lost for the platelet lifetime (8–10 days). About 10% to 12% of circulating platelets are replaced every 24 hours; however, the recovery of platelet function is faster than that predicted by the rate of platelet

turnover due to the aspirin-treated platelets being able to respond to TXA_2 generated by new platelets. Consequently it is only necessary to stop aspirin for 5 to 7 days to restore normal platelet function.

Oral aspirin is rapidly absorbed in the stomach and upper intestine and peak levels occur 30 to 40 minutes after ingestion. A dose of 75–100 mg/day is below the dose required to reduce pain and inflammation, but abolishes synthesis of TXA_2 without significant impairment of prostacyclin formation. However, low-dose aspirin is not without risk: a proportion of peptic ulcer bleeds in people aged over 60 years occur from prophylactic low-dose aspirin. Genetic polymorphisms affecting COX-1 and other genes involved in thromboxane biosynthesis may cause aspirin resistance in terms of the ability of aspirin to reduce TXA_2 biosynthesis. However, the term *aspirin resistance* has been used to describe different phenomena. The critical factor is the ability of aspirin to protect patients from thrombosis. The fact that some patients suffer recurrent events despite long-term aspirin therapy should be considered as treatment failure rather than aspirin resistance. Given the multifactorial nature of thrombosis and the fact that platelet-dependent thrombosis may not be responsible for all vascular events, it is not surprising that only about 25% of vascular complications can be prevented by antiplatelet therapy.

There is evidence that aspirin reduces the incidence of venous thrombosis by about 30%. However, as this is significantly inferior to the efficacy of anticoagulant drugs, aspirin is not the preferred treatment option to reduce venous thrombosis for the majority of at-risk patients.

There is evidence of a decreased risk of pre-eclampsia, preterm birth and fetal or neonatal death in women given antiplatelet therapy.

Dipyridamole

Dipyridamole reversibly inhibits platelet phosphodiesterase, and consequently the cyclic AMP concentration is increased and platelet activity reduced. The absorption of dipyridamole is variable and can result in low bioavailability. However, a slow-release preparation, when used in conjunction with aspirin, has been shown to reduce the incidence of stroke in patients with prior stroke or (TIA). However, three times as many patients discontinue the combination as compared to aspirin alone, mainly due to headache.

Clopidogrel, prasugrel and ticagrelor

Clopidogrel is a thienopyridine derivative that inhibits ADP-dependent platelet aggregation. It is used to prevent myocardial infarction and stroke in high-risk patients and in combination with aspirin after initial placement of drug-eluting stents. Clopidogrel is inactive, but is metabolized by cytochrome P-450 CYP2C19 in the liver, converting it to its active form. The active metabolite has an elimination half-life of about 0.5 to 1.0 h, and acts by forming an irreversible disulfide bridge with the platelet ADP receptor. Consequently the duration of the antiplatelet effect is the duration of the platelet lifespan. An initial dose of 300 mg is followed by a daily dose of 75 mg. About 15% of people with variants in cytochrome CYP2C19 (CYP2C19) have lower levels of the active metabolite of clopidogrel, less inhibition of platelets, and a fourfold greater risk of treatment failure.

Prasugrel is a thienopyridine derivative like clopidogrel. It inhibits ADP-induced platelet aggregation more rapidly, more consistently and to a greater extent than standard-dose clopidogrel. A 60 mg loading dose is followed by a daily maintenance dose of 10 mg. It is also a prodrug that is rapidly metabolized to a pharmacologically active metabolite. Prasugrel is a more effective antiplatelet drug than clopidogrel, but is associated with a higher bleeding rate.

Ticagrelor has a binding site different from ADP and is a reversible allosteric antagonist. In addition, the drug does not need hepatic activation and so it may be better for patients with genetic variants affecting CYP2C19 activity. It has a faster onset and shorter duration of action than clopidogrel and so has to be taken twice daily instead of once a day, which may be a disadvantage in terms of compliance, but its effects are more quickly reversible and this can be advantageous if surgery is required or if bleeding occurs.

Abciximab, eptifabatide and tirofiban

The platelet glycoprotein IIb/IIIa complex is the predominant platelet integrin, restricted to megakaryocytes and platelets. It mediates platelet aggregation by the binding of fibrinogen and von Willebrand factor (vWF) and GPIIb/IIIa antagonists have been developed as antiplatelet agents. As blockers of the final common pathway of platelet aggregation (the binding of fibrinogen) they are more complete inhibitors of platelets than aspirin, clopidogrel or prasugrel, which act only on the cyclooxygenase or ADP pathways. GPIIb/IIIa antagonists also have an anticoagulant effect by reducing the availability of platelet membrane anionic phospholipid required for thrombin generation on the platelet surface. The inhibition of platelet function is dose dependent.

Abciximab is a monoclonal antibody Fab fragment that binds to the GPIIb/IIIa complex with a high affinity and slow dissociation rate. It is administered intravenously and is cleared rapidly from plasma with a $t^1/_2$ of 20 minutes. A 0.25 mg/kg bolus is followed by an infusion of 0.125 μg/kg/min for 12 hours. This produces immediate and sustained inhibition of platelet function, with platelet aggregation gradually returning to normal about 96 to 120 hours after discontinuation of the drug. The dose causes and maintains blockade of more than 80% of receptors, causing a greater than 80% reduction in aggregation. Abciximab is

used in patients undergoing percutaneous coronary intervention (angioplasty with or without stent placement) in whom it reduces the incidence of ischaemic complications with the procedure and reduces the need for repeated coronary artery revascularisation. It can be used in patients with renal failure and can be used in combination with heparin and an antiplatelet drug such as aspirin. It should not be used if a patient requires emergency surgery because of the sustained effect.

Platelet transfusion after cessation of abciximab is necessary for refractory or life-threatening bleeding. After transfusion, the antibody redistributes to the transfused platelets, reducing the mean level of receptor blockade and improving platelet function. Severe thrombocytopenia occurs in 1% of treated patients. Abciximab-induced thrombocytopenia is usually rapid, occurring hours after administration, but may occur for up to 2 weeks after treatment. Transfusing platelets is the only known treatment, but may have limited effectiveness as the drug binds to transfused platelets. Abciximab-associated EDTA-induced platelet clumping also occurs causing pseudothrombocytopenia. Therefore, a low platelet count after abciximab should be confirmed using a non-chelating anticoagulant in the sample tube.

Eptifibatide is a cyclic heptapeptide based upon the Lys-Gly-Asp sequence. Tirofiban is a non-peptide mimetic. Both are competitive inhibitors of the GPIIb/IIIa complex, with lower affinities and higher dissociation rates than abciximab and relatively short plasma $t^1/_2$ values (2–2.5 h). Platelet aggregation returns to normal from 30 min to 4 hours after discontinuation. Eptifibatide and tirofiban are effective in acute coronary syndromes.

Selected bibliography

Ageno W, Gallus AS, Wittkowsky A *et al.*; American College of Chest Physicians (2012) Oral anticoagulant therapy: Antithrombotic Therapy and Prevention of Thrombosis, 9th ed: American College of Chest Physicians Evidence-Based Clinical Practice Guidelines. *Chest*: **141**: e44S–88S.

Douketis JD, Spyropoulos AC, Kaatz S *et al.* (2015) Perioperative Bridging Anticoagulation in Patients with Atrial Fibrillation. *New England Journal of Medicine*. [Epub ahead of print]. Available at: http://www.ncbi.nlm.nih.gov/pubmed/26095867. Accessed on 5th August 2015.

Eikelboom JW, Hirsh J, Spencer FA, Baglin TP, Weitz JI (2012) Antiplatelet drugs: Antithrombotic Therapy and Prevention of Thrombosis, 9th ed: American College of Chest Physicians Evidence-Based Clinical Practice Guidelines. *Chest* **141**: e89S–119S.

Garcia DA, Baglin TP, Weitz JI, Samama MM and the American College of Chest Physicians (2012) Parenteral anticoagulants: Antithrombotic Therapy and Prevention of Thrombosis, 9th ed: American College of Chest Physicians Evidence-Based Clinical Practice Guidelines. *Chest* **141**: e24S–43S.

Keeling D, Baglin T, Tait C *et al.* (2011) Guidelines on oral anticoagulation with warfarin – fourth edition. *British Journal of Haematology* **154**: 311–24.

Kitchen S, Gray E, Mackie I *et al.* Committee, B (2014) Measurement of non-coumarin anticoagulants and their effects on tests of Haemostasis: Guidance from the British Committee for Standards in Haematology. *British Journal of Haematology* **166**: 830–41.

Makris M, Van Veen JJ, Tait CR, Mumford AD, Laffan M (2013) Guideline on the management of bleeding in patients on antithrombotic agents. *British Journal of Haematology* **160**: 35–46.

Ruff CT, Giugliano RP, Braunwald E *et al.* (2014) Comparison of the efficacy and safety of new oral anticoagulants with warfarin in patients with atrial fibrillation: a meta-analysis of randomised trials. *Lancet* **383**: 955–62.

Schulman S (2014) New oral anticoagulant agents: general features and outcomes in subsets of patients. *Thrombosis and Haemostasis* **111**: 575–82.

Weitz JI, Eikelboom JW, Samama MM and the American College of Chest Physicians (2012) New antithrombotic drugs: Antithrombotic Therapy and Prevention of Thrombosis, 9th ed: American College of Chest Physicians Evidence-Based Clinical Practice Guidelines. *Chest* **141**: e120S–151S.

CHAPTER 47

Management of venous thromboembolism

47

Trevor Baglin[1] and David Keeling[2]

[1]Cambridge University Hospitals NHS Trust, Addenbrookes Hospital, Cambridge, UK
[2]Oxford University Hospitals, Churchill Hospital, Oxford, UK.

Introduction

Venous thromboembolism (VTE), which comprises deep vein thrombosis (DVT) and pulmonary embolism (PE), has an incidence of about 1 per 1000 per annum, with two-thirds presenting as DVT and one-third as PE. The case fatality is approximately 5%, but there is also major morbidity caused by post-thrombotic syndrome (PTS) of the leg and chronic thromboembolic pulmonary hypertension (CTPH). Anticoagulants are highly effective in secondary prevention, but leave the patient at risk of bleeding. The risk of recurrence is considerable and patients need to be assessed to identify those who require long-term anticoagulation as secondary prevention.

Diagnosis of deep vein thrombosis

Before the introduction of compression ultrasound, contrast ascending venography was the standard investigation for suspected DVT of the lower limb. Venography can detect both proximal DVT (involving the trifurcation area, the popliteal vein or above) and isolated calf DVT, but it is invasive and can be painful. Venography was largely replaced with compression ultrasound, often only with scanning of the proximal veins, but with a repeat scan one week later in those without proximal DVT. More recently, clinical medicine has seen a revolution in the diagnosis of DVT with the use of an integrated system, which includes clinical assessment, laboratory testing (in the form of D-dimer analysis) and ultrasound.

Isolated calf DVT

DVT usually starts in the calf, but by the time symptoms develop most patients have thrombus in the proximal veins. If isolated calf DVT is left untreated, extension into the proximal veins has been reported in 3–18%, usually within one week. Isolated calf DVTs that do not extend rarely lead to clinically significant emboli and it is safe to withhold anticoagulant treatment from patients with clinically suspected DVT who have normal compression ultrasonography of the proximal veins at the time of presentation and one week later. This has led to two different ultrasound strategies for DVT diagnosis. Many deliberately only look at the proximal veins and then perform a repeat test one week later in selected patients with a negative first scan. An alternative strategy is to scan the whole leg (proximal and calf veins). Both strategies are acceptable and safe, and have been shown to be clinically equivalent in randomized trials. Rapid compression ultrasound restricted to the proximal veins is seen as simple, convenient and technically simple, but requires repeat testing one week later in 25–50% of patients. The whole leg examination offers a one-day answer in all, but is technically more difficult and risks unnecessary anticoagulation. If the whole leg is scanned and an isolated calf DVT found, it is usually treated, though in the absence of severe symptoms or risk factors

Postgraduate Haematology, Seventh Edition. Edited by A Victor Hoffbrand, Douglas R Higgs, David M Keeling and Atul B Mehta.

for extension, the American College of Chest Physicians suggest serial imaging instead of initial anticoagulation.

Algorithms for the diagnosis of a deep vein thrombosis

The objective diagnosis of DVT depends on diagnostic imaging. However, because of the increasing number of negative tests, strategies have been developed that can exclude the diagnosis in some patients without the need for imaging. These strategies rely on the use of information from the clinical history and examination to estimate a pretest probability (PTP) and assays to detect D-dimers. The commonest PTP probability assessment tool is the Wells score (see Table 47.1). An 'unlikely' PTP score is not on its own enough to rule out a DVT, but an unlikely score can be used to rule out the diagnosis if the D-dimer test is also negative.

Thrombus formation is followed by a fibrinolytic response, which releases fibrin degradation products (containing the D-dimer motif) into the circulation. It follows that absence of a rise in D-dimers implies that thrombosis is not occurring. The strategy for using D-dimer assays in the diagnosis of DVT is to employ a sensitive test with a high negative predictive value. The specificity of the assays for DVT is poor and false positive results are common. All patients who have a 'likely' PTP or who have a raised D-dimer need to proceed to diagnostic imaging. This is illustrated in Figure 47.1a and b. If only the proximal veins are scanned we now know that patients with an 'unlikely' PTP and those with a normal D-dimer do not need a second scan.

Table 47.1 Wells clinical score used for predicting the probability of DVT prior to further testing.

Risk factors, symptoms or signs	Points
Active cancer (patient receiving treatment for cancer within the previous 6 months or currently receiving palliative treatment)	1
Paralysis, paresis or recent plaster immobilizsation of the lower extremities	1
Recently bedridden for 3 days or more, or major surgery within previous 12 weeks	1
Localized tenderness along the distribution of the deep venous system	1
Entire leg swollen	1
Calf swelling at least 3 cm larger than the asymptomatic leg (measured 10 cm below tibial tuberosity)	1
Pitting oedema confined to the symptomatic leg	1
Collateral superficial veins (non-varicose)	1
Previously documented deep vein thrombosis*	1
Alternative diagnosis at least as likely as deep vein thrombosis	–2

*The authors give a point for previous DVT or PE.
Score ≥2 points, likely pretest high probability; score ≤1 point, unlikely pretest probability

Diagnosis of pulmonary embolism

Clinical presentation of PE depends on the size, location and number of emboli, and the patient's underlying cardiorespiratory reserve. The classic triad of chest pain, haemoptysis and dyspnoea is present in less than 20% of patients, but nearly all will have at least one of pleuritic chest pain, dyspnoea or tachypnoea.

Diagnostic imaging is necessary to make the diagnosis and CT pulmonary angiography (CTPA) and ventilation–perfusion (V/Q) lung scanning are both used. Perfusion scanning with technetium-99-labelled albumin macroaggregates is a sensitive, but not very specific, method. The specificity is improved by adding ventilation scanning in which a gaseous radionuclide such as xenon-133 or technetium-99 diethylenetriaminepentaacetic acid (DTPA) is inhaled in an aerosol form, looking for a V/Q mismatch (no perfusion in ventilated lung). However, V/Q scanning is often non-diagnostic, resulting in a non-informative 'intermediate probability' scan report. CTPA, which may also identify alternative diagnoses, is becoming more commonly used.

In order to reduce the need for diagnostic imaging a similar strategy to that used for DVT is applied. The initial step for patients presenting with a possible PE is also to assess them for their pretest probability using a scoring system such as the Wells PE score (Table 47.2). Patients scoring as 'unlikely' can have the diagnosis excluded without the need for diagnostic imaging if they also have a negative D-dimer. If, however, the D-dimer is positive, the patient should be sent for diagnostic imaging, as should all patients who have a 'likely' PTP (Figure 47.2).

Treatment of VTE

The majority of deaths (>90%) due to PE occur in untreated patients in whom the diagnosis is not made in life. Treatment with therapeutic dose anticoagulation reduces the mortality to approximately 2%, with half of deaths directly attributable to PE rather than comorbidity, such as cancer. Patients with symptomatic PE are five times more likely to die of fatal PE than patients with symptomatic DVT and patients with massive PE with systemic hypotension (high-risk PE) are 10 to 20 times more likely to die. Rapid commencement of anticoagulation, and thrombolysis in patients with high-risk PE, is therefore required to prevent death. In patients presenting with DVT without symptomatic PE death is unusual once treatment with anticoagulation has started.

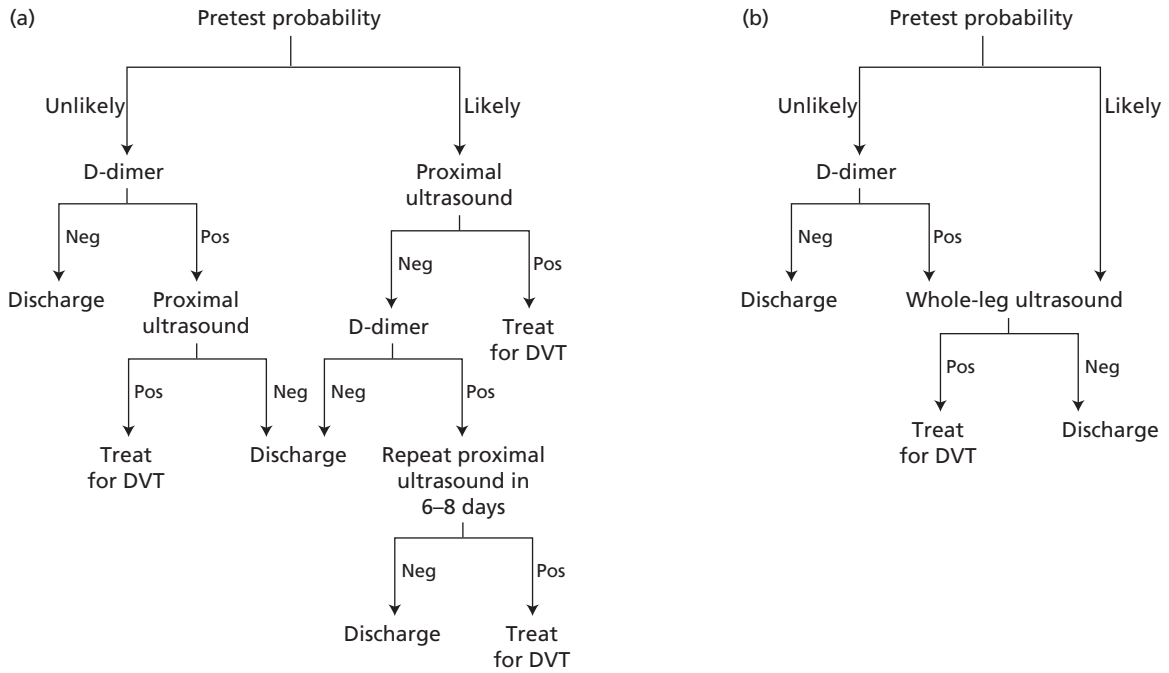


Figure 47.1 (a) Diagnostic algorithm for DVT with scanning of proximal veins only. (b) Diagnostic algorithm for DVT with scanning of whole leg.

Blood samples for haemoglobin, platelet count, PT and APTT serum creatinine and liver function should be taken. A group and save should also be ordered for patients who will be treated with thrombolysis. Thrombophilia testing is of no utility at this stage, as the results do not affect immediate patient management with any drug.

Initial anticoagulant therapy

Once a diagnosis of DVT or PE is made, or strongly suspected, a rapidly acting anticoagulant should be started immediately. Traditionally this has been heparin. A vitamin K antagonist (VKA) such as warfarin is usually started at the same time as heparin, or slightly later, but takes several days to produce an anticoagulant effect, hence the need for initial heparin in patients who are to receive warfarin. Nowadays a low-molecular-weight heparin (LMWH) rather than unfractionated heparin (UFH) is preferred. The LMWH dose is based on weight and the total daily dose is given as a single subcutaneous injection without the need for monitoring or dose adjustment. LMWH should continue for at least five days and until the International

Table 47.2 Wells clinical score used for predicting the probability of PE prior to further testing.

Risk factors, symptoms or signs	Points
Signs and symptoms of DVT	3
Alternative diagnoses less likely than PE	3
Heart rate >100	1.5
Immobility or surgery within 4 weeks	1.5
Previous DVT/PE	1.5
Haemoptysis	1
Active cancer	1

Score >4 points, likely pretest probability; score ≤4 points, unlikely pretest probability

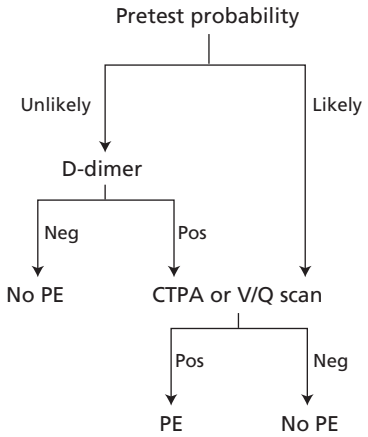


Figure 47.2 Diagnostic algorithm for PE.

Normalized Ratio (INR) has been greater than 2.0 for two consecutive days, when it is used in combination with warfarin. Unfractionated heparin (UFH) is now used infrequently, but is still used in some patients with massive (high-risk) pulmonary embolus.

When initiating rapid loading with warfarin, a loading nomogram should be used, as individual warfarin dose requirements vary considerably. The advantages of using a dosing nomogram are:

- A therapeutic intensity of anticoagulation is reached in the shortest possible time without an excessive risk of overanticoagulation.
- The daily maintenance warfarin dose, which will be required after the initial loading, is predictable. This reduces the risk of over-anticoagulation and hence bleeding following initial loading, and increases the likelihood of maintaining a therapeutic level of anticoagulation.

Alternatively, patients with acute DVT or PE can be treated with a non-VKA oral anticoagulant (NOAC). In clinical trials of rivaroxaban and apixaban, the NOAC was given at a higher dose initially without overlapping LMWH, whereas dabigatran and edoxaban were given with LMWH initially.

In patients with VTE who stop anticoagulant therapy after 3–6 months the incidence of subsequent fatal PE is approximately 0.25 per 100 patient years.

Thrombolytic therapy

Streptokinase and recombinant tissue plasminogen activator (rtPA) are widely available fibrinolytic agents. Streptokinase, a glycoprotein purified from the supernatant of β-haemolytic streptococci, in complex with plasminogen, cleaves other plasminogen molecules to plasmin. This indirect action results in complex pharmacokinetic characteristics. Furthermore, streptokinase is immunogenic, which may cause allergic reactions, as well as a reduced or absent response on repeated use or after a recent infection with streptococci. tPA has a direct enzymatic effect on plasminogen that is strongly enhanced by the presence of fibrin. Guidelines on the management of acute PE now recommend stratifying patients into high, intermediate and low risk (Table 47.3). Low-risk patients can be considered for home treatment. Patients with intermediate- and high-risk PE require hospitalization and patients with high-risk PE should be considered for immediate thrombolytic therapy as this may reduce mortality. Whilst life-saving in the acute phase the haemodynamic benefit of thrombolysis compared to heparin is lost after the first few days. At the present time, patients with intermediate risk are not given thrombolytic therapy, but they should be monitored, as 25% of patients deteriorate and may require urgent escalation to thrombolytic therapy. Escalation of patients to thrombolysis in clinical trials probably explains why there has been no difference in mortality in randomized trials of thrombolysis in intermediate-risk patients. In routine clinical practice the opportunity for escalation is lost if patients are not monitored, resulting in avoidable fatal PE.

Table 47.3 Stratifying pulmonary embolism by risk of death.

Haemodynamically unstable PE = high risk
Systolic blood pressure <90 mmHg or a pressure drop of ≥40 mmHg for >15 minutes if not caused by an arrhythmia, hypovolaemia or sepsis
About 5–10% of patients are in this high-risk group and have a risk of early death of >15%
Haemodynamically stable PE
Intermediate risk
These patients are haemodynamically stable, but with evidence of right heart strain or myocardial injury. Early mortality is 3–15%
Low risk
Haemodynamically stable without evidence of right heart strain or myocardial injury. Early mortality is <1%

Up to 5% of patients have massive PE with hypotension at presentation. Fatal PE occurs in >15% of patients presenting with high-risk (massive) PE, 3–15% of patients with intermediate-risk PE and in less than 1% of patients with low-risk PE or DVT alone. Right ventricular dysfunction (RVD) is present in up to 40% of patients with acute PE and normal blood pressure, and fatal PE is twice as likely in those with RVD. In patients with RVD, biomarkers such as troponin may identify those patients at greatest risk, but the clinical utility of these measurements is not yet known. Death due to acute PE results from an acute fall in cardiac output. In patients without pre-existing heart or lung disease the haemodynamic disturbance correlates with the extent of obstruction of the pulmonary circulation.

Contraindications to thrombolysis include active internal bleeding, a stroke within 2 months, and an intracranial process such as neoplasm or abscess. Relative contraindications include surgery within 10 days, uncontrolled hypertension and pregnancy. Haemorrhagic complications are higher in patients with a recent invasive procedure such as pulmonary angiogram or placement of an inferior vena cava (IVC) filter. There is a reported incidence of intracranial haemorrhage of approximately 2%, with higher rates in the elderly and those with poorly controlled hypertension. Up to 20% of patients suffer major bleeding. If heparin is being given it should be stopped. Heparin therapy is started 2 hours after administration of thrombolysis.

Thrombolytic therapy after DVT increases the rate of early re-canalization, but is associated with a higher risk of bleeding than with anticoagulant therapy. In clinical studies, venographic obstruction and PTS were less frequent in patients treated with streptokinase compared with anticoagulation alone, but PTS assessments in these studies were not performed using any validated instrument, and bleeding complications were much more

frequent in patients treated with streptokinase. Consequently, systemic streptokinase infusions are no longer used for DVT treatment. Systemic infusion of recombinant tissue plasminogen activator (rt-PA) has limited efficacy. Clinical studies suggest that the systemic administration route may not reliably achieve a therapeutic rt-PA concentration at its target sites within the thrombus. However, major bleeding is more frequent than with anticoagulant therapy. Catheter-directed intrathrombus thrombolysis (CDT) refers to the infusion of a fibrinolytic drug directly into the venous thrombus via a multisidehole catheter, which is embedded in the thrombus using imaging guidance. This is the preferred treatment currently if thrombolytic treatment is given. However, even with low-dose rt-PA (0.5 to 1.0 mg/h) major bleeding occurs in about 4% of patients. The long-term benefits remain unclear and so thrombolytic therapy is not recommended routinely in patients with DVT. Thrombolysis may be considered in individual patients with symptomatic iliofemoral DVT who have all of the following:

- symptoms of less than 14 days' duration;
- good functional status;
- a life expectancy of 1 year or more;
- a low risk of bleeding.

Long-term complications of VTE

Post-thrombotic syndrome (PTS)

Following acute DVT, re-canalization occurs slowly. Residual vein occlusion (RVO) is present in 50% of patients at 6 months and remains present in 25% after 3 years. Clot resolution is identical in patients with and without thrombophilia. Because of frequent incomplete clot resolution the diagnosis of recurrent ipsilateral DVT is problematic. Of particular clinical value in this regard is the demonstration that in patients with previous DVT, a negative high-sensitivity D-dimer assay result safely excludes recurrence. Post-thrombotic syndrome (PTS) is chronic pain and swelling, which occurs in up to one-third of patients who have a DVT; it is typically worse at the end of the day or with prolonged sitting or standing, and better after a night's rest and leg elevation. Patients may describe heaviness in the leg, cramps, pruritis and paraesthesia. There may be oedema, induration of the skin, reddish-brown hyperpigmentation and venous ectasia. PTS can occur early or, more unusually, have a latency of up to 10 years; the cumulative frequency has been estimated as 23% at 2 years and 28% at 5 years; in 5–10% PTS is classed as severe (this includes all those with ulcers).

A major risk factor for PTS is recurrent ipsilateral DVT. During anticoagulant treatment after DVT patients with prolonged periods of subtherapeutic anticoagulation are more likely to develop PTS. There is no evidence that prolonged anticoagulant therapy reduces the risk of PTS. This indicates that early suboptimal anticoagulation rather than duration of anticoagulation influences the development of PTS and that thrombin generation may contribute to ongoing valvular damage immediately after DVT.

Early walking after acute DVT and continued high levels of physical activity appear to reduce the likelihood of clot progression and reduce the risk of PTS. Initial randomized trials demonstrated a 50% reduction in PTS with compression stockings used for up to two years. However, the large SOX trial failed to show a benefit when compression stockings were used for 50% of the time compared to 'sham' stockings. Consequently, any benefit of stockings in reducing the incidence and severity of PTS after DVT is now contentious, although many patients use them for symptomatic relief early after acute DVT.

Chronic thromboembolic pulmonary hypertension (CTPH)

Following treatment of acute symptomatic PE with either thrombolysis or anticoagulation, 85% of patients still have partial occlusion of the pulmonary circulation after 1 week. At 6 months more than 50% of patients have persistent follow-up scan defects and complete resolution does not occur in the majority of patients by 1 year. Consequently, there is the difficulty of confirming or excluding recurrent acute PE in patients who do not have a follow up scan at completion of initial treatment. Appropriate follow-up after completion of anticoagulant therapy for evaluation of risk of chronic thromboembolic pulmonary hypertension (CTPH) remains to be defined. Less than 5% of patients with persistent occlusion of the pulmonary circulation after 6 months will develop CTPH. Recurrent and previous multiple PE are significant risk factors for CTPH, but heritable thrombophilic defects are not associated with an increased risk.

Risk of recurrence and duration of anticoagulant therapy

It is important to identify an underlying clinical condition that might have caused VTE, such as cancer, by careful history, examination and appropriate investigations. Overweight patients should be advised to lose weight and smokers advised to stop. The preferred anticoagulant treatment option for patients with active malignancy and those receiving treatment for cancer is continued LMWH.

If isolated calf DVT is diagnosed it is usually treated for 3 months but 6 weeks may be sufficient for isolated calf DVT following surgery. Patients with proximal DVT or PE should be treated for at least 3 months. The decision as to whether to continue anticoagulant treatment indefinitely after a first proximal DVT or PE is dependent on the risk of recurrence without continued treatment and the perceived risk of bleeding on anticoagulation. Long-term anticoagulation is usually given after

recurrent unprovoked VTE. For patients who receive long-term anticoagulation the risk–benefit ratio of continued treatment should be reassessed at regular intervals, e.g. annually or following an episode of severe over-anticoagulation (INR >8) or bleeding.

Recurrent DVT or PE is associated with a risk of death and a greatly increased likelihood of PTS and CTPH, respectively. Secondary prevention of VT significantly reduces the burden of disease, but recurrent VTE is only prevented for as long as anticoagulation is continued. Therefore, there is a need to improve risk stratification for both recurrent VTE and anticoagulant-related bleeding. Following an episode of venous thrombosis, patients are 40 times more likely to suffer an event in the future than patients who have not previously suffered from venous thrombosis. However, risk varies between patient groups. Patients with VTE provoked by surgery have an annualized recurrence rate of <1%. Patients with VTE provoked by a non-surgical factor have an annualized risk of 4% and patients with unprovoked VTE have an annualized event rate of 7%. Many experts consider an annual rate higher than 5% justification for long-term anticoagulant therapy, as the risk of future fatal PE is then likely greater than the risk of fatal bleeding on long-term anticoagulation.

Decisions regarding duration of anticoagulation should be made with reference to whether or not a first episode of VTE was provoked or unprovoked (Table 47.4). Current guidelines recommend that patients with provoked VTE receive anticoagulant therapy for 3 months. For patients with unprovoked VTE, a decision should be made after 3 months as to whether the patient should continue on anticoagulation or stop. In patients with an unprovoked first event, male sex, a high D-dimer on completion of initial anticoagulant therapy and possibly younger age predict higher rate of recurrence. Decision rules for stopping anticoagulation, which include testing for D-dimer, have been evaluated. Increased D-dimer levels following completion of anticoagulant therapy after DVT are associated with residual vein obstruction (RVO) and likelihood of recurrence is not determined by RVO when adjusted for D-dimer level. Testing for heritable thrombophilia has little, if any, clinical utility in deciding whether a patient should remain on anticoagulant therapy or not, and routine testing in order to determine the duration of anticoagulant therapy is not recommended. If it is planned to stop anticoagulation treatment in a patient who has had unprovoked proximal DVT or PE, then testing for antiphospholipid antibodies should be offered, as a positive result would argue for reconsidering long-term treatment.

A further consideration when deciding whether a patient should continue anticoagulation or not after an initial three months of treatment is whether the first presentation was with symptomatic PE or DVT without symptoms of PE. Whilst the risk of recurrence is the same, patients who present with an initial PE are three times as likely to suffer from PE if they suffer a recurrence, as compared to those whose first event is a DVT.

Other treatments

Vena caval filters

Vena caval filters are prosthetic devices that are deployed to prevent PE in patients at high risk, but in whom standard anticoagulant therapy is contraindicated, or in selected patients with PE despite therapeutic anticoagulation. Contraindications to anticoagulation insertion might be:

- central nervous system haemorrhage;
- overt gastrointestinal bleeding;
- retroperitoneal haemorrhage;
- massive haemoptysis;
- stroke in the last 14 days;
- central nervous system trauma;
- severe thrombocytopenia.

Filter placement is associated with potential complications. Because anticoagulation is usually contraindicated at the time of filter insertion there is a risk of thrombosis at the vascular access site. However, most of these are asymptomatic. There are no data on the long-term consequences of this complication. Mechanical problems such as caval penetration, filter migration, tilting and filter fracture are infrequent asymptomatic events with symptomatic events occurring in about 1% of patients.

Table 47.4 Suggested duration of anticoagulation.

3 months	3 months then consider for long-term	Usually long-term
• First proximal DVT or PE associated with a transient risk factor* • First unprovoked isolated calf DVT (consider 6 weeks if post surgery)	• First unprovoked proximal DVT or PE**	• Recurrent unprovoked DVT or PE • Proximal DVT or PE with a significant continuing risk factor (e.g. active cancer, antiphospholipid antibodies) and no other cause

*Transient risk factors: surgery, significant trauma e.g. fracture, plaster cast; combined pill/hormone replacement therapy; pregnancy/puerperium
**A stronger case for continuing in: males; those with raised D-dimers after completing anticoagulation; those with PE

With or without subsequent anticoagulation there is a significant incidence of subsequent thrombosis in the vena cava or leg veins resulting in PTS in approximately 20%. This is not significantly different to the incidence of PTS following DVT treated with anticoagulant therapy, but due to incomplete follow up of IVC filter patients this figure may be an underestimate. Filters are associated with a lower incidence of early PE, but there is a higher incidence of recurrent DVT and the thrombus frequently involves the inferior vena cava. Case-series have not demonstrated a benefit from introducing anticoagulation for the sole purpose of preventing filter-related thrombotic events and anticoagulant therapy for patients with IVC filters should be guided by an assessment of the patient's risk of recurrent VTE and major bleeding, and not the presence of the filter alone.

Thrombectomy

Thrombectomy is rarely performed in patients with massive PE, but is an option in cardiothoracic centres in patients with life-threatening PE. Thrombectomy is not recommended in patients with DVT. There are insufficient data to evaluate the long-term benefit of surgical removal of the thrombus in patients.

Venous thrombosis in unusual sites

Cerebral vein thrombosis

Cerebral venous sinus thrombosis (CVST) most often affects young adults, with approximately three-quarters of patients being female. Obstruction of cerebral veins causes cerebral oedema and venous infarction, while occlusion of venous sinuses causes intracranial hypertension. It can present with recent unusual headache or stroke-like symptoms. The most sensitive diagnostic test is magnetic resonance venography. If MRI is not available then high-resolution CT is useful but can be normal initially. Recognized underlying causes include infection, myeloproliferative disorders and leukaemia (especially with asparaginase treatment) and use of oestrogen-containing oral contraceptives. Mortality in the first month is approximately 5%, but 80% of surviving patients recover completely or have only a mild functional or cognitive deficit. Anticoagulant therapy was associated with a pooled relative risk of death of 0.33 (0.08 to 1.21) and of death or dependency of 0.46 (0.16 to 1.31), based on two small studies and it is recommended that patients with CVST without contraindications to anticoagulant therapy should be treated with therapeutic-dose heparin. The optimal duration of anticoagulation is unknown, but it is suggested that a minimum of 3 months treatment is given. Long-term anticoagulation has been suggested for patients with 'severe' thrombophilia and in patients with unprovoked CVST with incomplete re-canalization, but with no evidence.

Retinal vein thrombosis

Retinal vein occlusion (RVO) is predominantly a disorder of the elderly. The prognosis depends on the development of macular oedema and neovascularization and standard therapy has been laser photocoagulation. The role of anticoagulant therapy is unclear. For branch RVO, antiplatelet or anticoagulation therapy is not recommended. Patients with acute central RVO are considered for a short course of anticoagulation with LMWH or with low-dose tPA, but routine therapy with warfarin or antiplatelet agents is not recommended.

Upper limb DVT

Upper extremity DVT may involve the axillary, subclavian and brachial veins and may be associated with thoracic outlet syndrome, or with strenuous effort, or with placement of a central venous catheter. Compression ultrasound is an acceptable alternative to venography in the investigation of suspected upper limb DVT. PE occurs in up to 30% of cases and PTS of the arm is a risk. Recurrence is less common than in lower limb DVT. The majority of patients are managed with anticoagulation for three months.

Intra-abdominal vein thrombosis

Portal vein thrombosis

The portal vein is formed from the superior mesenteric and splenic veins. The most common underlying aetiology of portal vein thrombosis is cirrhosis; other causes are intra-abdominal infection, inflammation or malignancy and blunt trauma or surgery. Myeloproliferative disorders account for up to a quarter of cases and all patients should be assessed for the JAK2 V617F mutation. Paroxysmal nocturnal haemoglobinuria is another important cause of intra-abdominal thrombosis. Presentation can be acute, with abdominal pain, fever and nausea, or chronic, with symptoms of portal hypertension. The diagnosis can be made by CT, MRI or Doppler ultrasound. In patients with cirrhosis, the risk of anticoagulation-related bleeding is high, and anticoagulation may be regarded as too high risk. In acute portal vein thrombosis without cirrhosis, anticoagulation is usually given. There is no evidence to guide duration of anticoagulation. Although it has been suggested that anticoagulation should be given long term in the presence of heritable thrombophilia, or if there is a failure of re-canalization, this is not based on any evidence.

Hepatic vein thrombosis

The usual presentation is with abdominal pain, ascites and hepatomegaly (Budd–Chiari syndrome). Many patients have a myeloproliferative disorder and a large proportion of patients are JAK2 positive. The diagnosis can be made with Doppler ultrasound, CT or MRI. Most patients are anticoagulated and

in one-third this is followed by transjugular intrahepatic portal systemic shunting (TIPSS).

Mesenteric vein thrombosis

The most common causes are cancer, intra-abdominal infection, cirrhosis and surgery. Presentation is with abdominal pain and diarrhoea or peritonitis. Diagnosis is with CT or MRI. Surgery is required in patients with peritonitis, but otherwise anticoagulation is the treatment of choice.

Splenic vein thrombosis

Isolated splenic vein thrombosis is rare, and pancreatic disease is the most common aetiology. It can present with variceal bleeding and splenomegaly.

Superficial thrombophlebitis

Superficial venous thrombosis (SVT) of the lower limb, involving long or short saphenous veins or their branches, has the same risk factors as DVT. SVT was considered to be a benign and self-limiting condition, however, it is now appreciated that a significant proportion of those presenting with SVT will have concomitant DVT or PE, or are at significant risk of developing DVT or PE. SVT within 3 cm of the sapheno-femoral junction has such a high risk of progression to DVT that such patients are given therapeutic anticoagulation as for DVT. In other cases non-steroidal anti-inflammatory drugs (NSAIDs) and heparin have been used as treatment. A Cochrane review concluded an intermediate dose of LMWH might be appropriate and for SVT $\geq$5 cm in length this or a prophylactic dose of fondaparinux (2.5 mg daily) for 45 days should be used, the latter reducing symptomatic DVT or PE by 82%.

Selected bibliography

Anderson FJ, Wheeler H, Goldberg R *et al.* (1991) A population-based perspective of the hospital incidence and case-fatality rates of deep vein thrombosis and pulmonary embolism. The Worcester DVT Study. *Archives of Internal Medicine* **151**: 933–8.

Bates SM, Jaeschke R, Stevens SM *et al.*; American College of Chest Physiscians (2012) Diagnosis of DVT: Antithrombotic Therapy and Prevention of Thrombosis, 9th ed: American College of Chest Physicians Evidence-Based Clinical Practice Guidelines. *Chest* **141**: e351S–418S.

Bernardi E, Camporese G, Buller HR *et al.* (2008) Serial 2-point ultrasonography plus D-dimer vs whole-leg color-coded Doppler ultrasonography for diagnosing suspected symptomatic deep vein thrombosis: a randomized controlled trial. *Journal of the American Medical Association* **30**: 1653–9.

Iorio A, Kearon C, Filippucci E *et al.* (2010) Risk of recurrence after a first episode of symptomatic venous thromboembolism provoked by a transient risk factor: a systematic review. *Archives of Internal Medicine* **170**: 1710–16.

Kearon C, Akl EA, Comerota AJ *et al.* (2012) Antithrombotic Therapy for VTE Disease: Antithrombotic Therapy and Prevention of Thrombosis, 9th ed: American College of Chest Physicians Evidence-Based Clinical Practice Guidelines. *Chest* **141**: e419S–94S.

Kearon C, Akl EA (2014) Duration of anticoagulant therapy for deep vein thrombosis and pulmonary embolism. *Blood* **123**: 1794–1801.

NICE (2012) Venous thromboembolic diseases: the management of venous thromboembolic diseases and the role of thrombophilia testing. http://guidance.nice.org.uk/cg144 (accessed June 2015).

Scott G, Mahdi AJ, Alikhan R (2015) Superficial vein thrombosis: a current approach to management. *Br J Haematol* **168**: 639–45.

Tait C, Baglin T, Watson H *et al.*; British Committee for Standards in Haematology (2012) Guidelines on the investigation and management of venous thrombosis at unusual sites. *British Journal of Haematology* **159**: 28–38.

Tosetto A, Iorio A, Marcucci M *et al.* (2012) Predicting disease recurrence in patients with previous unprovoked venous thromboembolism: a proposed prediction score (DASH). *Journal of Thrombosis and Haemostasis* **10**: 1019–25.

Verhovsek M, Douketis JD, Yi Q *et al.* (2008) Systematic review: D-dimer to predict recurrent disease after stopping anticoagulant therapy for unprovoked venous thromboembolism. *Annals of Internal Medicine* **149**: 481–90, W494.

Wells PS, Anderson DR, Rodger M *et al.* (2003) Evaluation of D-dimer in the diagnosis of suspected deep-vein thrombosis. *New England Journal of Medicine* **349**: 1227–35.

Wells PS, Anderson DR, Rodger M *et al.* (2000) Derivation of a simple clinical model to categorize patients probability of pulmonary embolism: increasing the models utility with the SimpliRED D-dimer. *Thrombosis and Haemostasis* **83**: 416–420.

CHAPTER 48

Haematological aspects of systemic disease

48

A Victor Hoffbrand[1] and Atul B Mehta[1,2]

[1]University College London, London, UK
[2]Royal Free Hospital, London, UK

Anaemia of chronic disease

The anaemia of chronic disease (ACD) is a common normochromic or mildly hypochromic anaemia that occurs in patients with a systemic disease (Table 48.1). It is characterized by a reduced serum iron and iron-binding capacity, and normal or raised serum ferritin with adequate iron stores (Table 48.2). It is not due to marrow replacement by tumour, bleeding, haemolysis or haematinic deficiency, although these often complicate it.

Pathogenesis

Changes in iron metabolism

An important factor in the pathogenesis of chronic anaemia in the setting of inflammation, infection and malignancy is an increased plasma level of hepcidin. Hepcidin binds to ferroportin, accelerates its degradation and so inhibits iron absorption and export from macrophages (Chapter 3). The body utilizes this mechanism to reduce the supply of iron to microorganisms, but it also results from sequestration of iron within macrophages and reduction of iron absorption, in lowering plasma iron and so reducing supply of iron to developing erythrocytes. A fall in serum transferrin and a rise in serum ferritin also occur as part of the acute-phase response. The fall in serum iron can occur as early as 24 hours after the onset of a systemic illness, and will persist during the course of a prolonged illness. The cytokines IL-1,IL-6, TNF and interferon-α or -β released directly from macrophages or indirectly together with interferon-γ from stromal cells or T lymphocytes stimulate hepatocyte synthesis of hepcidin.

Anaemia usually reduces hepcidin secretion. In ACD, this effect is clearly abrogated by the effect of inflammation or malignancy increasing its secretion. Increased lactoferrin, occurring in response to inflammation and mediated by cytokines, competes with transferrin for iron and forms a complex that is taken up by macrophages in the liver and spleen. Increased intracellular apoferritin synthesis occurs in response to inflammation and malignancy and this too will bind iron. Both of these mechanisms reduce the amount of iron available for binding to plasma transferrin. In the mouse there is evidence that erythroferrone synthesized by erythroblasts (Chapter 3) lowers hepcidin levels during recovery from infection, thus enabling macrophages to release iron to transferrin.

Inhibition of erythropoiesis

This results from increased levels of inflammatory cytokines, including interleukin (IL)-1, IL-6, tumour necrosis factor (TNF) and transforming growth factor (TGF)-β. These interact with accessory marrow stromal cells and with the erythroid progenitors themselves to reduce their sensitivity to erythropoietin. In this way, the marrow is attempting to aid the recruitment of pluripotent stem cells to produce white blood cells in order to combat infection/inflammation and possibly malignancy. These inflammatory cytokines have also been shown to reduce the circulating level of erythropoietin. However, the deficiency in erythropoietin is relative. Although the plasma erythropoietin level in ACD remains inversely correlated with the haemoglobin, compared with patients with other types of anaemia and normal renal function, it is inappropriately low. TNF and IL-1 have been shown in experimental systems to reduce erythropoietin

Postgraduate Haematology, Seventh Edition. Edited by A Victor Hoffbrand, Douglas R Higgs, David M Keeling and Atul B Mehta.

Table 48.1 Conditions associated with anaemia of chronic disorders.

Chronic infections
Especially osteomyelitis, bacterial endocarditis, tuberculosis, abscesses, bronchiectasis, chronic urinary tract infections, osteomyelitis, HIV
Other chronic inflammatory disorders
Rheumatoid arthritis, juvenile rheumatoid arthritis, polymyalgia rheumatica, systemic lupus erythematosus, scleroderma, inflammatory bowel diseases, thrombophlebitis, severe trauma
Malignant diseases
Carcinoma (especially metastatic or associated with infection), lymphoma, myeloma
Others
Congestive heart failure, ischaemic heart disease, AIDS

production by cultured hepatoma cells. In addition, IL-1α inhibits erythropoietin production by isolated serum-free perfused kidneys. A mild decrease in red cell lifespan occurs in ACD, but is at a level that could be compensated by a normal marrow. TNF has been shown to increase apoptosis of erythroid cells.

Red cell survival is shortened in ACD due to the changes in the environment rather than in the red cells themselves. The mechanism is not certain, but raised TNF and IL-1 levels have been implicated. As the anaemia of chronic disorders responds to erythropoietin therapy it appears that inhibition of erythropoiesis and inadequate erythropoietin secretion are more dominant causes of the anaemia than the changes in iron metabolism.

Treatment

The severity of the anaemia correlates with the activity and severity of the underlying chronic disease. Successful therapy of this leads to a reduction in the levels of the mediator cytokines, increased erythropoietin production and reduced inhibition of erythropoiesis. Correction of the anaemia may take weeks or months. Pharmacological doses of recombinant erythropoietin have been used successfully to improve anaemia in patients with rheumatoid arthritis, cancer and myeloma.

Table 48.2 Laboratory features of anaemia of chronic disease.

Haemoglobin	Not less than 90 g/L
Mean corpuscular volume	Normal or mildly reduced (usually 77–82 fL)
Mean corpuscular haemoglobin	Usually normal; occasionally reduced
Serum iron	Reduced
Total iron-binding capacity (transferrin)	Reduced
Transferrin saturation	Mildly reduced
Serum ferritin	Normal or increased
Serum and urine hepcidin	Raised
C-reactive protein	Usually raised
Erythrocyte sedimentation rate	Usually raised

Malignancy

Anaemia

Anaemia is the most frequent haematological abnormality in cancer patients and may be due to many causes (Table 48.3). ACD (see above) will affect almost all cancer patients at some stage of their illness. The degree of anaemia reflects the extent of the malignancy and may be worsened by the myelotoxic effects of chemotherapy. Plasma erythropoietin levels tend to be inappropriately low and therapy with recombinant erythropoietin can reduce transfusion requirements by improving the haemoglobin level in cancer patients undergoing chemotherapy.

Haemolysis

Warm antibody autoimmune haemolytic anaemia (AIHA) is most frequently found in association with the following malignant diseases: chronic lymphocytic leukaemia, Hodgkin lymphoma and non-Hodgkin lymphoma. However, it has also been reported in association with solid tumours (e.g. carcinoma of the ovary). Cold antibody AIHA is less common, but occurs in

Table 48.3 Causes of anaemia in cancer patients.

Type of anaemia	Associations
Anaemia of chronic disease	All neoplasms
Blood loss	Gastrointestinal, gynaecological neoplasms
Haemolysis	
Immune	CLL, lymphoma, ovarian carcinoma, others
Non-immune fragmentation syndrome	Mucin-secreting adenocarcinomas
Haemolysis: secondary to drugs	Mitomycin, ciclosporin, cisplatin
Pure red cell aplasia	Thymoma
Megaloblastic	Chemotherapy, folate or cobalamin deficiency (gastric carcinoma)
Leucoerythroblastic	Metastatic disease in bone marrow
Marrow hypoplasia	Chemotherapy/radiotherapy
Myelodysplasia	Chemotherapy/radiotherapy

chronic cold agglutinin disease, lymphoplasmacytic lymphoma and myeloma.

Microangiopathic haemolytic anaemia (MAHA) with intravascular haemolysis may occur in association with disseminated carcinoma. An abrupt onset of anaemia and thrombocytopenia often occurs, with a leucoerythroblastic blood picture, reticulocytosis and red cell fragmentation. Renal failure may occur as a complication. Mucin-secreting adenocarcinomas especially gastric, breast and lung cancer are the most common causes; in about one-third of patients with MAHA, it is the presenting feature of the tumour. Abnormal blood vessels may be within the tumour itself (or within metastatic tumour thrombi, especially in the lungs) or fibrin deposition may occur at other sites because of disseminated intravascular coagulation (DIC). Widely disseminated disease with bone marrow infiltration is almost always present and the outlook is poor. A syndrome resembling MAHA and idiopathic haemolytic–uraemic syndrome (HUS) has been reported with a number of chemotherapeutic agents and following allogeneic bone marrow transplantation. Principal among the drugs is mitomycin, although cisplatin, carboplatin, bleomycin and ciclosporin have also been reported to be responsible. Immune complex deposition has been implicated in the pathogenesis.

Red cell aplasia

Acquired pure red cell aplasia is associated with a thymoma in approximately 50% of patients, although it complicates only approximately 5% of thymomas. Antibodies to erythroid precursors have been demonstrated in some patients, and removal of the thymoma (which is usually benign) leads to resolution of the anaemia in about half of those affected. Immunosuppressive therapy with cyclophosphamide, ciclosporin, steroids or plasma exchange may be helpful in patients who relapse. Red cell aplasia may also occur in patients with chronic lymphocytic leukaemia (CLL) or non-Hodgkin lymphoma and with large granular lymphocytic (LGL) leukaemia or as part of general marrow aplasia due to chemotherapy or radiotherapy.

Leucoerythroblastic anaemia

The blood film (Figure 48.1) shows the presence of erythroblasts and granulocyte precursors (e.g. myelocytes and myeloblasts). It is frequent when there is marrow infiltration by tumour. This disturbs the marrow microvasculature and allows early release of the precursors. Marrow infiltration is most commonly observed in breast (Figure 48.2), prostate and haematological malignancies, but also in tumours of lung, thyroid, kidney, gastrointestinal tract and melanoma. It can also occur as a reflection of active bone marrow response to peripheral consumption (acute haemolysis, DIC, septicaemia, hypersplenism) or of extramedullary haemopoiesis (e.g. myelofibrosis or megaloblastic anaemia).

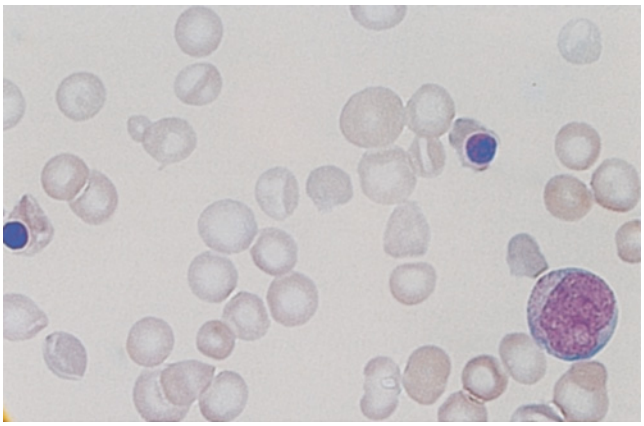


Figure 48.1 Nucleated red cells and an immature myeloid precursor in the peripheral blood film of a patient with a leucoerythroblastic anaemia.

Other causes of anaemia

Megaloblastic and dyserythropoietic anaemias are most commonly due to chemotherapy-induced disturbance of DNA synthesis within the bone marrow. Folate deficiency may also occur in patients with a poor diet and widespread disease. Cobalamin (vitamin B_{12}) deficiency due to underlying pernicious anaemia may complicate cancer of the stomach. Both chemotherapy and non-ionizing radiotherapy may lead to the development of myelodysplastic syndrome (MDS), which may progress to acute myeloid leukaemia (AML). Alkylating agents, especially melphalan and chlorambucil, nitrosoureas and epipodophyllotoxins in particular have been implicated, and there is evidence of a synergistic effect of these agents with small chronic doses of radiotherapy. The principal categories of patients affected are

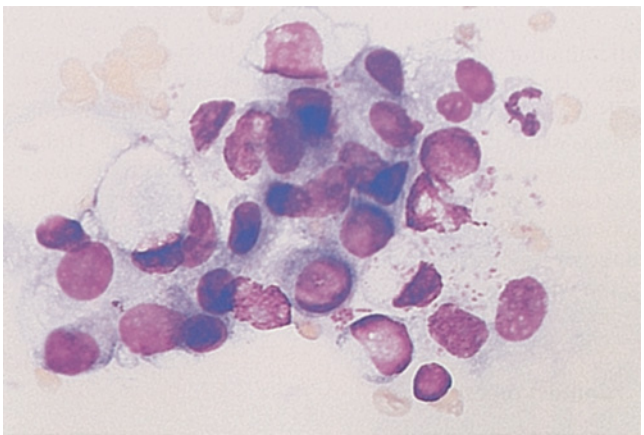


Figure 48.2 Bone marrow aspirate showing infiltration by metastatic breast carcinoma.

those who have received therapy for a haematological malignancy, but patients treated for non-haematological malignancies (especially ovarian and gastrointestinal carcinoma) are also susceptible. The median latency prior to onset of MDS is 2–3 years and AML usually supervenes 6 months to 2 years later.

Treatment

Erythropoietin (EPO) is effective in ameliorating anaemia in a proportion of patients with malignant disease, for example myeloma, lymphoproliferative disorders, MDS and carcinomas. It is considered if the haemoglobin is less than 100 g/L or in symptomatic patients with haemoglobin of 100–120 g/L. Erythropoietin is not justified if the haemoglobin is above 120 g/L because of the risk of thrombosis. Patients most likely to respond are those who have low pretreatment EPO levels (<100 mU/mL), well-preserved marrow function and normal/low levels of serum ferritin (<400 ng/mL). A number of studies have demonstrated that EPO therapy can lead to a reduced need for blood transfusion, better quality of life for patients and possibly improved overall outcome of anticancer therapy. The expression of both EPO and its receptor occurs in a range of human tumours, including breast, prostate, colon, ovary, uterine, and head and neck squamous tumours. Caution is therefore needed in the use of EPO to treat cancer-related and cancer-treatment-related anaemia.

Polycythaemia

This is a rare complication of non-haematological malignancy. It usually arises through elaboration by tumour cells of erythropoietin and erythropoietin-like peptides. The tumours most commonly associated are renal cell carcinoma and hepatoma; others include uterine myoma, androgen-secreting ovarian tumours, phaeochromocytoma and cerebellar haemangioblastoma. Non-malignant conditions affecting these organs (e.g. renal cysts, viral hepatitis) may also rarely be associated.

White cells (Table 48.4)

Neutrophilia

A raised neutrophil count is a frequent manifestation of non-haematological malignancies. In part, the response is due to inflammation induced by the tumour. Interaction of tumour cells with host T lymphocytes and mononuclear phagocytic cells leads to the production of a range of cytokines, which induce white cell proliferation and differentiation. Tumour cells may also secrete specific agents that directly stimulate reactive proliferation. Cancer patients also frequently have opportunistic infections, may bleed and are typically on a range of medications, all of which may influence the level of the neutrophil count.

Table 48.4 White cell changes in malignancy.

Neutrophils increased	Most, especially renal, Hodgkin lymphoma
Neutrophils decreased	Bone marrow infiltration Hypersplenism Chemotherapy/radiotherapy Large granular lymphocytic leukaemia
Basophils increased	Myeloproliferative disorders
Eosinophils increased	Hodgkin lymphoma T cell lymphomas Metastatic adenocarcinoma Other tumours (e.g. lung) Drug allergy Opportunistic infection
Monocytes increased	Carcinoma Hodgkin lymphoma
Monocytes decreased	Treatment induced
Lymphocytes increased	Lymphoid malignancies Postsplenectomy Opportunistic infection
Lymphocytes decreased	Treatment induced Radiotherapy/chemotherapy Lymphoma Opportunistic infection

Neutropenia

This is most frequently due to chemotherapy or radiotherapy, but may also occur with widespread marrow infiltration, for example by lymphoma or due to LGL leukaemia.

Platelets

Thrombocytopenia (Table 48.5)

This may arise through decreased production or accelerated peripheral destruction and/or hypersplenism. The former may result from extensive marrow infiltration or chemotherapy or

Table 48.5 Thrombocytopenia in patients with malignant disease.

Decreased production
Chemotherapy/radiotherapy
Marrow infiltration
Accelerated destruction
Hypersplenism
DIC
Drug-induced HUS
Autoimmune thrombocytopenia

Abbreviations in this and subsequent Tables: see text

radiotherapy. Hypersplenism is usually due to splenic infiltration by a haematological malignancy (e.g. lymphoma, CLL), but is rarely due to obstruction of the splenic or portal vein by hepatic and pancreatic malignancies. Increased destruction occurs with DIC. Immune thrombocytopenia may occur in association with haematological malignancy (CLL, lymphoma, MDS) and rarely may complicate solid tumours.

Thrombocytosis

A raised platelet count is frequently seen as a reactive phenomenon in patients with malignancy, but is usually less than 1000×10^9/L and is rarely of clinical significance. Iron deficiency may also cause a rise in platelets.

Platelet function abnormalities

Impaired platelet function leading to excessive bleeding is primarily seen in myeloproliferative disorders (e.g. essential thrombocythaemia) and MDS. Paraproteins, especially IgM in Waldenström macroglobulinaemia and IgG in myeloma, are frequent causes of impaired platelet aggregation and adhesion. Increased platelet aggregation occurs in cancer patients when tumour cells release adenosine diphosphate (ADP), prostaglandins or thrombin. Increased platelet adhesiveness has also been reported, but is probably secondary to other coagulation changes.

Coagulation (Table 48.6)

A wide range of coagulation changes predispose to either haemorrhage or thrombosis. The distinction between activation of the coagulation part of the haemostatic pathway (e.g. in DIC), which leads physiologically to activation of fibrinolysis (secondary fibrinolysis), and primary or direct activation of fibrinolysis may be difficult. These patients may also have other general medical problems that affect coagulation, including infection, surgery, chemotherapy or radiotherapy and will typically be on a range of medications.

Table 48.6 Coagulation changes in malignancy.

Bleeding tendency
DIC, acute or chronic
Primary fibrinolysis
Acquired platelet function defect
Thrombocytopenia
Circulating anticoagulants/inhibitors
Treatment-related bleeding disorders
Thrombotic tendency
Venous stasis: bed rest, venous compression/invasion by tumours
Increased coagulation factors: FI, FV, FVII, FVIII, FIX, FXI
Decreased inhibitors of coagulation: low antithrombin, proteins C and S
Direct activation of coagulation by tumour cells: FVII, FX
Indirect activation: via trypsin release, mucin secretion, monocytes or endothelial damage
Increased platelet aggregability and adhesiveness
Thrombocytosis

Disseminated intravascular coagulation (DIC)

Chronic or compensated DIC is probably underdiagnosed in cancer patients. It occurs particularly in those with gastrointestinal, lung, pancreatic and breast neoplasms. Thrombosis, including migratory thrombophlebitis (Trousseau syndrome) and non-bacterial thrombotic endocarditis, is a frequent manifestation of DIC. In contrast, acute or uncompensated DIC is uncommon with solid tumours, but occurs frequently with acute promyelocytic leukaemia (APL) and is associated with excessive bleeding. The triggering event in APL is release from the malignant promyelocytes of procoagulants and proteases, which directly activate both coagulation and fibrinolysis (Chapters 41 and 46). Tumour cells may activate coagulation by release of tissue factor (TF). Direct activation of FX through the action of a cancer procoagulant has been reported in lung, kidney, colon and breast cancer. The sialic acid moiety of secreted mucin can directly activate FX, whereas the systemic release of trypsin in pancreatic tumours can also activate coagulation. Tumour cells may also activate the monocyte/macrophage system to produce procoagulant materials including TF and FX activators. Deep vein thrombosis is frequent in patients with malignant diseases and is discussed in Chapter 46.

Primary fibrinolysis

This is less common as a cause of increased bleeding than DIC, but can occur, for example, in patients with prostatic cancer who undergo surgery. The release of proteases from leukaemic cells in both APL and monocytic leukaemia may induce fibrinolysis. Platelet counts tend to be higher than those seen in DIC, and fibrinogen levels are low, with raised fibrin degradation products (FDPs).

Acquired circulating anticoagulants (Table 48.7)

The most frequent is an acquired von Willebrand syndrome (both type 1 and type 2 disease) in patients with a paraprotein or a B-lymphoid malignancy. A number of different mechanisms may operate. The paraprotein may be directed against an epitope within the FVIII–von Willebrand factor (VWF) molecule and inactivate it, or reduce its plasma half-life. Alternatively, immune complexes may form that bind non-specifically to FVIII–VWF and accelerate its clearance, or the malignant lymphoid cells may adsorb FVIII–VWF onto their surface. Paraproteins may also interfere with cross-linking of fibrin.

Table 48.7 Circulating anticoagulants in malignancy.

Factor inhibitors
Factor V
Factor VII
Factor VIII
von Willebrand factor inhibitors
Paraproteinaemic disorders
Lymphoma
Myeloproliferative disorders
Chronic lymphocytic leukaemia
Others
Heparin-like anticoagulants: dysproteinaemias

Treatment-induced bleeding disorders

Thrombocytopenia and MAHA may occur as a result of therapy (see further on). Cancer patients with poor nutrition or who receive long-term antibiotics may develop vitamin K deficiency. L-asparaginase induces defective hepatic protein synthesis and can lead not only to impaired production of coagulation factors, but also to low levels of antithrombin, plasminogen and proteins S and C, and so give rise to thrombosis, most seriously of the cerebral veins. Mithramycin, which is used in the treatment of malignant hypercalcaemia, causes thrombocytopenia as well as platelet function defects, coagulation factor deficiencies and increased fibrinolytic activity.

Connective tissue disorders (Table 48.8)

Table 48.8 Haematological changes in connective tissue disorders.

Anaemia
ACD
Iron deficiency (drug-induced blood loss)
Folate deficiency
Sideroblastic anaemia
Pure red cell aplasia (PRCA), especially systemic lupus erythematosus (SLE)
Haemolytic anaemia: immune (especially SLE)/non-immune
White cells
Neutropenia (e.g. Felty syndrome)
Neutrophilia
Eosinophilia (e.g. Churg–Strauss syndrome, polyarteritis nodosa)
Platelets
Thrombocytopenia: immune/non-immune
Platelet dysfunction
Thrombotic thrombocytopenic purpura (TTP)
Thrombocytosis
Pancytopenia
Systemic lupus erythematosus (SLE)
Coagulation
Lupus anticoagulant
Specific factor deficiencies
DIC
Others
Myelofibrosis
Drug-related changes (e.g. aplastic anaemia due to gold, phenylbutazone; PRCA due to penicillamine)
Cryoglobulinaemia
Amyloidosis

Anaemia

ACD is the most common haematological abnormality seen in patients with rheumatoid arthritis (RA). Iron deficiency may coexist, particularly in patients taking non-steroidal anti-inflammatory agents. Folate deficiency may occur with severe disease and poor diet. Warm-type AIHA with IgG and complement on the red cell surface is most frequently seen in systemic lupus erythematosus (SLE), although it can occur in the other connective tissue disorders, notably RA and mixed connective tissue disorders. Ring sideroblasts in the marrow have been reported in both SLE and RA, but MDS must be excluded. Pure red cell aplasia and dyserythropoietic anaemia with ineffective erythropoiesis are rare complications of SLE. Haemolysis can also occur as part of thrombotic thrombocytopenic purpura (TTP), complicating SLE.

White cells

The inflammatory process in connective tissue disorders can lead to a neutrophilia. Neutropenia is a feature of Felty's syndrome, which is associated with splenomegaly in patients with RA. The pathogenesis involves increased margination of neutrophils, sequestration of neutrophils within the enlarged spleen, and immune-complex-mediated and humoral inhibition of granulopoiesis in the marrow. Antibodies to mature neutrophils have also been detected in SLE. Lymphopenia occurs in both SLE and RA, and may be a measure of disease activity. Eosinophilia may be seen in SLE, RA, polyarteritis nodosa and Churg–Strauss syndrome. The pathogenesis is unknown, but presumably involves release of cytokines by T lymphocytes. Functional defects in neutrophil and lymphocyte function have been reported in SLE and RA.

Platelets

Immune thrombocytopenia is a common manifestation of SLE and also occurs in mixed connective tissue disorders, scleroderma, RA and dermatomyositis. Autoantibodies to platelets may also impair platelet function. TTP is an association of SLE. Thrombocytosis is a non-specific reaction to inflammation and tissue damage in connective tissue disorders.

Coagulation

A wide diversity of coagulation changes including DIC may occur in patients with connective tissue disorders. In part, this may be due to liver and renal disease or to drug therapy. DIC has been reported in SLE patients who have high levels of circulating immune complexes and resulting angiopathy. The lupus anticoagulant (see Chapter 46) occurs as a complication in about 10% of patients with SLE and is associated with a thrombotic tendency, thrombocytopenia, recurrent miscarriages and pulmonary hypertension. Specific coagulation factor inhibitors encountered in patients with connective tissue disorders (especially SLE) include antibodies to VWF and to FVIII, FVII and fibrinogen.

Other changes

RA is one of the commonest causes of amyloidosis. An increased incidence of haematological malignancies (principally Hodgkin and non-Hodgkin lymphomas and B-lymphoproliferative disorders, including paraproteinaemias) occurs in SLE, RA and, particularly, Sjögren syndrome. A wide range of haematological abnormalities also results from immunosuppressive therapy in these patients.

Renal disease (Table 48.9)

Anaemia

In acute renal failure, anaemia is commonly due to the drug or condition causing the renal failure, for example haemolysis due to sepsis or TTP. In chronic renal failure, anaemia is the most important haematological abnormality and its management has been revolutionized by recombinant human EPO. Patients with acute or chronic renal failure develop a normochromic normocytic anaemia, with the presence of echinocytes (burr cells) in the blood film. The reticulocyte count is normal or slightly low, and the bone marrow shows normoblastic erythropoiesis without the erythroid hyperplasia expected at that level of anaemia. Patients who have undergone nephrectomy tend to be more severely anaemic than patients with polycystic disease. Reduced EPO levels occur in renal failure and this is the dominant cause of anaemia.

An increase in serum creatinine above 133 μmol/L is associated with the loss of the normal inverse linear relation between plasma EPO and haemoglobin concentration, but there is no direct correlation between reduction in glomerular filtration rate and impairment of renal EPO production. Circulating inhibitors of erythropoiesis have also been demonstrated, but are not clinically significant. Red cell survival is diminished in renal failure, but this is also a minor factor. Iron deficiency can arise through blood loss (exacerbated by haemodialysis).

Table 48.9 Haematological changes in renal disease.

Anaemia
Failure of EPO production
Raised hepcidin
Haemolysis: Hus, TTP
Iron deficiency
Folate deficiency
Hyperparathyroidism
Aluminium toxicity
Polycythaemia
Renal cell carcinoma
Other renal diseases (e.g. cysts, hydronephrosis, nephritic syndrome, renal transplantation)
Thrombocytopenia
HUS
TTPDIC
Platelet function abnormalities
Abnormal aggregation to ADP, adrenaline, collagen
Decreased platelet adhesiveness
Reduced platelet factor 3 availability
Acquired storage pool defect
Abnormal prostaglandin metabolism
↑ Prostacylin
Defective platelet cyclooxygenase?
Coagulation
Hypocoagulability
↓ FXII, FXI, prothrombin
↓ FXII or inhibition
Hypercoagulopathy
↓ Protein C
↓ Antithrombin
↓ Fibrinolysis

Elevated levels of hepcidin in plasma may be due to reduced excretion of this polypeptide and so impair iron release from macrophages. Folate deficiency arises in dialysed patients but is now prevented by prophylactic folic acid therapy. Renal failure is associated with elevated levels of 2,3-diphosphoglycerate and a right shift of the haemoglobin–oxygen dissociation curve.

Recombinant EPO therapy can fully correct anaemia in renal failure. It can be administered intravenously, subcutaneously or intraperitoneally. Depending on the preparation used it is given once, twice or three times a week or every fortnight. Anaemia is corrected up to a level of 100–120 g/L at a rate of 10 g/L per month. Subclinical iron deficiency, detected by serum ferritin, percentage hypochromic red cells or reticulocyte haemoglobin concentration (Chapter 3), and impaired mobilization of storage iron are often present, so concomitant iron therapy is usually required. This can often be easily accomplished by the administration of intravenous iron (Chapter 3). An impaired

response to EPO should prompt a suspicion of iron, cobalamin or folate deficiency, haemolysis, infection, occult malignancy, aluminium toxicity or secondary hyperparathyroidism. Angiotensin-converting enzyme inhibitors have also been associated with anaemia in this setting. Hypertension occurs in about one-third of patients treated with recombinant erythropoietin and is dose dependent; the risk of thrombosis of an arteriovenous fistula is also increased.

The optimum dose of EPO is one that restores the haemoglobin level to the normal or near-normal range and improves symptoms without increasing the risk of thrombosis; for most patients, this is approximately 120–125 g/L and should not exceed 140 g/L.

Polycythaemia

Secondary and inappropriate polycythaemia may result from either ectopic EPO production by renal tumours or regional renal hypoxia (in benign disease and following renal transplantation). Up to 5% of patients with renal cell carcinoma have paraneoplastic polycythaemia.

Haemostatic abnormalities

Abnormal platelet function is probably due to the accumulation of toxic metabolites (e.g. guanidinosuccinic and phenolic acids). DDAVP (1-deamino-8-D-arginine vasopressin) therapy, which leads to the appearance of large multimers of VWF, can reduce the tendency to bleed in anaemic patients. Dysfibrinogenaemia has been reported rarely, whereas FDPs are often elevated and may prolong the thrombin time. Hypercoagulopathy with a predisposition to thrombosis can also occur, especially after recombinant EPO therapy. Haemodialysis with heparin anticoagulation can cause platelet activation. Fibrinolytic activity, antithrombin and protein C are all reduced in renal failure, and FV, FVII, FVIII:C and FX are increased. Thrombosis (particularly of the renal vein) is a particular feature of the nephrotic syndrome. Platelet hyperaggregability with increased plasma β-thromboglobulin is described and hypoalbuminaemia may enhance the synthesis of prostaglandins involved in platelet activation.

Endocrine disease (Table 48.10)

Both hyperthyroidism and hypothyroidism are associated with mild anaemia, which is usually normochromic and normocytic, but may be macrocytic in hypothyroidism. A raised MCV can occur without anaemia in hypothyroidism, and a low MCV has been described in thyrotoxicosis. Thyroid hormones stimulate erythropoiesis, and tissue oxygen demands are increased in hyperthyroidism, whereas in hypothyroidism oxygen utilization is reduced. However, plasma volume is also increased and part of the anaemia in hypothyroidism is dilutional and/or due to defective iron utilization. Coexistent deficiencies of iron (due to menorrhagia or achlorhydria, both more frequent in hypothyroidism), folate and cobalamin must be excluded.

There is an increased incidence of pernicious anaemia in patients with hypothyroidism, hypoadrenalism and hypoparathyroidism. Antithyroid drugs (carbimazole, methimazole and propylthiouracil) can cause aplastic anaemia or agranulocytosis. Anaemia in patients with diabetes mellitus is usually due to complications of diabetes, although hyperglycaemia itself may lead to glycosylation of the renal EPO receptor, shortened red cell lifespan and decreased erythrocyte deformability. Polycythaemia (usually pseudo) can also occur with endocrine diseases e.g. Cushings syndrome, phaeochromocytoma. In anterior pituitary disease, androgen deficiency and adrenal insufficiency, a normochromic normocytic anaemia is common. Changes in leucocyte number and function are rarely of clinical significance, although many have been reported. Chemotaxis, phagocytosis and intracellular killing may all be disturbed in diabetes mellitus. Coagulation changes may

Table 48.10 Haematological changes in endocrine disease.

Red cells
Anaemia
Thyrotoxicosis (normochromic, normocytic or microcytic)
Hypothyroidism (normochromic, normocytic, occasionally macrocytic)
Diabetes mellitus (usually when complicated by infection, cardiac disease, renal failure, enteropathy)
Hyperparathyroidism (normochromic, normocytic)
Hypoadrenalism (normochromic, normocytic)
Hypogonadism (normochromic, normocytic)
Hypopituitarism (normochromic, normocytic)
Polycythaemia (pseudo)
Phaeochromocytoma
Cushing syndrome
White cells
Neutrophil leucocytosis: Cushing syndrome, phaeochromocytoma
Lymphocytosis: hyperthyroidism
Neutropenia: antithyroid drugs
Impaired neutrophil function: diabetes mellitus
Platelets
Abnormal function: diabetes mellitus: hyperthyroidism
Coagulopathy
Diabetes mellitus: ↑ platelet aggregability, ↓ prostacyclin, ↑ FVIII, ↓ antithrombin
Oestrogen therapy: ↑ FVIII, ↑ VWF
Cushing syndrome: ↑ FII, FIV, FIX, FXI, FXII

contribute to a mild bleeding diathesis in hypothyroidism and to the thrombotic predisposition in diabetes mellitus.

Anorexia nervosa is usually accompanied by anaemia and a moderate fall in white cells and platelets is frequent. The bone marrow is hypocellular with a gelatinous deposit and necrosis may be present. The same changes also occur in starvation and are reversed by feeding.

Liver disease (Table 48.11)

Liver disease causes a greater range of haematological change than does disease in any other organ, other than the bone marrow. The liver is an important source of EPO in the fetus, and serves as a haemopoietic organ *in utero*; extramedullary haemopoiesis occurs within the adult liver only in pathological states (e.g. primary myelofibrosis, severe haemolysis or megaloblastic anaemia). The liver is the major source of thrombopoietin in adults.

Anaemia

Anaemia occurs in up to 75% of patients with chronic liver disease. Portal hypertension often results in splenomegaly, which may cause haemodilution and pooling of red cells. Haemorrhage is a frequent complication, often due to oesophageal varices, and the red cell lifespan is shortened, even in uncomplicated liver disease. Mechanisms that operate in ACD may be relevant. Macrocytosis occurs in approximately two-thirds of patients and is particularly frequent in alcoholics, in whom reversible sideroblastic change may also occur. Target cells occur as the surface area of the cell increases, due to increased membrane lipid content without an increase in volume. Echinocytosis is fairly common because of binding of the red cell membrane by abnormal high-density lipoproteins. In contrast, true acanthocytes are uncommon. They are a characteristic finding in 'spur-cell anaemia' (non-immune haemolytic anaemia in patients with alcoholic cirrhosis). Zieve syndrome, comprising haemolytic anaemia with hypertriglyceridaemia in patients with alcoholic liver disease, is also rare. Haemolysis due to the direct toxicity of copper ions on red cells is characteristically an early presentation of Wilson disease. Reduction of hepatocyte glucose-6-phosphate dehydrogenase (G6PD) levels in G6PD-deficient individuals, and in neonates, may exacerbate and prolong hyperbilirubinaemia with haemolysis. Viral hepatitis, including hepatitis A, B and C, but most frequently hepatitis viruses yet to be characterized, may lead to a transient and mild pancytopenia or to severe aplastic anaemia (Chapter 11).

Platelets and haemostasis

These are discussed in Chapter 41.

Table 48.11 Haematological changes in liver disease.

Red cells
Anaemia
ACD
Folate deficiency
Iron deficiency (blood loss)
Aplastic anaemia (viral hepatitis, rare)
Sideroblastic anaemia (alcohol)
Hypersplenism
Microangiopathy/ DIC (rare)
Autoimmune (rare)
Zieve syndrome (rare)
Polycythaemia
Hepatocellular carcinoma (rare)
Infectious hepatitis (rare)
White cells
Neutrophilia
Infection
Haemorrhage
Malignancy
Haemolysis
Neutrophil function
Impaired chemotaxis (?due to lowered complement levels)
Neutropenia
Hypersplenism
Eosinophilia
Parasitic infestation
Chronic active hepatitis (rare)
Platelets
Thrombocytopenia
Hypersplenism, hepatic sequestration
DIC
Autoimmune (e.g. associated with viral hepatitis, primary biliary cirrhosis)
Post liver transplantation
Thrombocytosis
Hepatoma (rare)
Impaired platelet function
Inhibitory factors (including high-density lipoprotein and apolipoprotein E)
Other
Benign monoclonal gammopathy (biliary + other cirrhosis)
Cryoglobulinaemia (hepatitis B, hepatitis C, alcohol)

Liver transplantation

Orthotopic liver transplantation (OLT) is increasingly used for end-stage liver disease. Thrombocytopenia is frequently present prior to transplantation. The count tends to fall postoperatively,

Table 48.12 Haematological changes in viral infection.

Red cells
Anaemia
Autoimmune
- Measles
- Epstein-Barr virus (EBV)
- Hepatitis
- Cytomegalovirus (CMV)
- Human immunodeficiency virus (HIV)

Others including herpesviruses, varicella, influenza
Non-immune
- Microangiopathic haemolytic anaemia

Reduced red cell production
- Marrow hypoplasia
- EBV (especially in X-linked lymphoproliferative syndrome)
- Hepatitis viruses
- HIV
- CMV (especially after renal or bone marrow transplantation)
- Others (rare) include togaviruses epidemic haemorrhagic fevers, dengue

Red cell aplasia
- Parvovirus B19, especially with haemolytic anaemia

White cells
Neutrophilia
- Especially HIV, influenza, hepatitis, rubella, adenoviruses, measles, mumps, CMV and EBV as part of nearly all viral infections

Neutropenia
- Aplasia (see above)
- Complicating myalgic encephalitis
- Enteroviruses, EBV

Lymphocytosis
- Wide variety, especially early in course of infection

Malignant transformation
- HTLV-I, EBV, HIV

Platelets
Thrombocytosis (e.g. Kawasaki)
Thrombocytopenia
- Often history of viral prodrome in childhood immune thrombocytopenic purpura
- Autoimmune: EBV, hepatitis, rubella, CMV, HIV
- ↓ Production: aplasia (see above), measles, dengue, CMV, others
- ↑ Consumption: disseminated intravascular coagulation (DIC)/haemolytic–uraemic syndrome (see below)

Coagulation changes
DIC, especially varicella, vaccinia, rubella, arbovirus with/without microangiopathy, epidemic haemorrhagic fevers
Haemolytic-uraemic syndrome: coxsackievirus, mumps, echoviruses
Haemophagocytosis
Herpesviruses, adenoviruses, CMV

despite platelet transfusions and this may be due to platelet sequestration in the transplanted liver. Immune thrombocytopenia has also been reported after OLT. Antibody-mediated haemolysis occurs in recipients of ABO-incompatible grafts, but the engrafted liver can also produce mild haemolysis due to anti-recipient ABO antibody. This is a form of humoral graft-versus-host disease (GVHD), but T cell-mediated GVHD has also been reported after OLT. Although aplastic anaemia is a rare complication of viral hepatitis, it may occur in as many as 30% of patients transplanted for fulminant non-A, non-B viral hepatitis.

Congestive Heart Failure (CHF)

Between 30 and 50% of subjects with CHF are anaemic. The causes are multiple, including chronic renal failure, elevated cytokines causing the ACD, use of ACE inhibitors which may inhibit erythropoiesis, haemodilution, diabetes and iron deficiency due to associated gastrointestinal disorders. In some studies intravenous iron has improved haemoglobin levels, left ventricular ejection fraction and quality of life, and lowered plasma CRP and BNP levels.

Infections (Tables 48.12 and 48.13)

Infection may produce a tremendous variety of haematological changes. Many of these are covered in other sections of this book.

Viruses

Anaemia

Haemolytic anaemia due to red cell autoantibody production, usually of the warm type, may occur, although cold antibody syndromes occur in measles, influenza, infectious mononucleosis and mumps. Paroxysmal cold haemoglobinuria is rare and occurs in children due to Donath–Landsteiner IgG anti-P antibodies (Chapter 9). Non-immune MAHA may be associated with TTP or DIC, which may be the result of viral

Table 48.13 Haematological changes in bacterial, fungal and protozoal infections.

Anaemia
ACD
Haemolytic
Immune: *Mycoplasma*, malaria, syphilis (PCH), listeriosis
Non-immune: *Clostridium perfringens* (toxin related), *Bartonella bacilliformis* (Oroya fever)
Malaria/trypanosomiasis with microangiopathy/DIC, septicaemia
HUS: verotoxin-producing *Escherichia coli* and *Streptococcus pneumoniae*
Dilutional
Splenomegaly (e.g. malaria, schistosomiasis)
Blood loss
Helicobacter pylori, *Ankylostoma*
White cells
Neutrophilia
Virtually any bacterial/fungal infection
Neutropenia
Salmonella, *Rickettsia*, brucellosis, pertussis, disseminated tuberculosis (TB)
Overwhelming septicaemia
Neutrophil function defects
Rare (e.g. *Bacteroides*, endocarditis)
Lymphocytosis
Whooping cough (*Bordetella pertussis*), *Rickettsia*
Lymphopenia
TB, acute bacterial infections, brucellosis
Eosinophilia
Aspergillosis, coccidioidomycosis, *Chlamydia*, streptococcal infections, *Ancylostoma*
Eosinopenia
Common in acute *Bacteroides* infections
Monocytosis
Subacute/chronic infections (e.g. disseminated TB, listeriosis)
Pancytopenia
Bone marrow suppression (e.g. disseminated TB, listeriosis)
Haemophagocytosis: septicaemia
Peripheral destruction (e.g. DIC)

PCH, paroxysmal cold haemoglobinuria.

infection. Transient red cell aplasia occurs with parvovirus B19 infection in patients with haemolytic anaemias ('aplastic crisis'). This virus may also cause erythema infectiosum or fifth disease, in children. It invades and destroys red cell progenitors and the aplasia is terminated when neutralizing IgM and IgG antibodies develop. If the virus attacks pregnant women, it may cross the placenta and cause spontaneous abortion or hydrops fetalis. Intravenous immunoglobulin has been used for severe cases (e.g. in pregnancy, HIV infection and after stem cell transplantation). Bone marrow aspirate shows characteristic giant erythroblasts.

Pancytopenia occurs in virus-associated bone marrow aplasia, for example with hepatitis viruses. The presence of viruses, either within lymphocytes or on their cell surface, may lead to production of a range of cytokines (including TNF, IFN-α and IFN-γ), which inhibit haemopoietic cell proliferation *in vitro* and *in vivo*. This may cause a substantial reduction in erythropoiesis and is presumably the mechanism underlying neutropenia and lymphopenia in viral infections. In infectious mononucleosis and other viral infections such as viral hepatitis, the virus infects B lymphocytes and the activated lymphocytes seen on the blood film are a reactive population of T cells.

Platelets

Thrombocytopenia may occur due to multiple mechanisms. Children with autoimmune thrombocytopenic purpura frequently give a history of a preceding viral illness, and autoantibody production is well described in infectious mononucleosis, rubella and cytomegalovirus infections. Reduction of bone marrow thrombopoiesis is frequently subclinical, but it is particularly important in virus-associated aplasia and dengue fever. Thrombocytosis can also occur in response to viral infections.

Bacterial, fungal and protozoal infections

Anaemia

ACD can occur in acute infections, overwhelming septicaemia and chronic or suppurative infection. Haemolytic anaemia is less common, but can occur through both immune (e.g. cold antibodies with anti-I specificity in *Mycoplasma* infection) and non-immune mechanisms. Direct red cell invasion frequently results in severe haemolysis in infections caused by *Bartonella baciliformis*, with elements of intravascular haemolysis (due to increased red cell fragility) and extravascular haemolysis. *Clostridium perfringens* produces an α-toxin (a lecithinase) and a θ-toxin, and *Staphylococcus aureus* an α-toxin, which act as haemolysins to cause severe intravascular haemolysis. DIC and MAHA can occur in any severe bacterial, fungal or protozoal infection. HUS is associated with a range of bacterial infections, including *Salmonella*, *Shigella* and *Campylobacter* spp., but most frequently with verotoxin-producing strains of *Escherichia coli*. Hookworm and schistosomiasis are two of the most frequent causes worldwide of iron deficiency anaemia, due to intestinal haemorrhage and in the case of schistosomiasis, haematuria (see Chapter 3).

White cells

Neutrophilia is the most common manifestation. The neutrophil response shows great individual variation, with no clear relationship to the severity of the infection. The term *leukaemoid reaction* is used to describe marked leucocytosis ($>50 \times 10^9$/L),

with circulating immature forms occurring in patients with non-leukaemic conditions, typically severe infection or haemolysis or with generalized malignancy. Such reactions are more common in children. Features that distinguish such a reactive leucocytosis from chronic myeloid leukaemia include the presence of toxic granulation, Döhle bodies, and the lack of twin peaks of neutrophils and myelocytes. Neutropenia can also occur with virtually any bacterial infection, although it has been most frequently noted with *Salmonella*, *Rickettsia* and *Brucella*. Defects of neutrophil function may also occur.

Platelets

Thrombocytosis is frequent in patients with chronic infections, and during the convalescent phase of acute infections. Thrombocytopenia also occurs during severe bacterial or fungal infection, particularly where there is bloodstream invasion or in intensive care patients. Certain rickettsial infections (e.g. Rocky Mountain spotted fever) are almost always associated with thrombocytopenia. Accelerated platelet destruction is the most frequent mechanism and can arise through DIC or microangiopathy with platelet attachment to damaged endothelium. Immune destruction can also occur. Decreased platelet production is a less common mechanism, but may occur (e.g. in disseminated tuberculosis). The inflammatory and procoagulant responses to infections are related. TNF-α, IL-1α and IL-6 may activate coagulation and inhibit fibrinolysis, whereas thrombin may stimulate inflammatory pathways. In severe infection, the end result may be endovascular injury, DIC, multiorgan failure and death.

Haemostasis

DIC occurs frequently and may dominate the clinical picture in certain infections (e.g. bacterial meningitis). The acute-phase response that accompanies severe infection can lead to a rise in a range of coagulation factors, which may contribute to thrombosis. Suppurative thrombophlebitis, particularly in association with indwelling catheters, can occur in relation to both Gram-positive and Gram-negative infections. In patients with systemic inflammation and organ failure due to acute infection, plasma protein C levels are reduced. Activated protein C promotes fibrinolysis and inhibits thrombosis and inflammation and reduces circulating levels of D-dimers and IL6.

Malaria (see also Chapter 49)

Anaemia is most marked with *Plasmodium falciparum*, which invades erythrocytes of all ages (*P. vivax* and *P. ovale* invade only reticulocytes, *P. malariae* only mature cells) and can give parasitaemia levels as high as 50%. Cellular disruption and haemoglobin digestion lead directly to haemolysis. Parasitized cells have an increased osmotic facility and lose deformability; they thereby become sequestered and destroyed within the spleen, which often becomes massively enlarged. Non-parasitized cells may then become sequestered within the spleen and a raised plasma volume contributes to the anaemia. In addition, malarial antigens may attach to non-parasitized red cells to give rise to a positive direct antiglobulin test and haemolysis via a complement-mediated immune response. Acute intravascular haemolysis with haemoglobinuria, often leading to renal failure ('blackwater fever'), occurs rarely in *P. falciparum* infection.

An inadequate bone marrow response to anaemia is seen with relative reticulocytopenia at times of active infection, with some recovery after effective therapy. TNF levels are typically elevated and ACD occurs. Leucocyte numbers may be slightly increased or normal, but leucopoenia as a result of splenomegaly and impaired marrow function is characteristic. Eosinophilia is variable. Thrombocytopenia is seen in nearly 70% of *P. falciparum* infections and has multifactorial aetiology. Autoimmune mechanisms may operate as for red cells, splenic sequestration is a contributory factor, DIC (either acute, as in blackwater fever, or low grade and chronic) is common, and ADP release from damaged red cells may lead to platelet activation and consumption.

Haemophagocytic lymphohistiocytosis (haemophagocytic syndrome) (Table 48.14; see also Chapters 14)

This may occur in association with a wide range of systemic illness including malignancies (e.g. lymphoma), infections (particularly viral, e.g. Epstein–Barr virus, HIV) and autoimmune diseases. It typically manifests as fever, splenomegaly and jaundice with cytopenias. It is particularly common in patients who are immunosuppressed or who are acutely ill (e.g. septicaemic). A familial form is recognized (see Chapter 14). The macrophage activation syndrome is a variant form of haemophagocytic lymphohistiocytosis that occurs in children with juvenile RA. Pancytopenia is usual, although cytopenias affecting an individual cell lineage also occur, and coagulopathy due to associated DIC is frequently present. Abnormal liver function commonly coexists. The serum ferritin is usually markedly elevated. The bone marrow (Figure 48.3) shows the presence of increased numbers of histiocytes displaying haemophagocytosis. Myelofibrosis and/or marrow hypocellularity are present in a minority

Table 48.14 Conditions associated with reactive haemophagocytosis.

Infection
Viral (e.g. herpesviruses, adenoviruses, cytomegalovirus)
Bacterial, especially tuberculosis
Fungal
Tumours
Haematological
Others
Drugs
Phenytoin

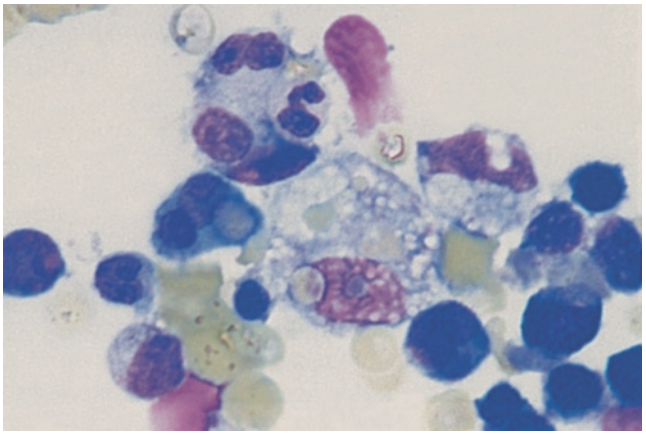


Figure 48.3 Bone marrow aspirate showing active haemophagocytosis, which in this patient antedated the development of high-grade non-Hodgkin lymphoma by 6 months.

of cases. The underlying mechanisms are poorly understood. Excessive production of cytokines (e.g. IFN-γ and TNF-α) by dysregulated T cells has been demonstrated. In macrophage activation syndrome, increased levels of IL-6, IL-8, IL-12, IL-18 and macrophage inflammatory protein (MIP)-1α have also been reported. Treatment should be directed at the underlying disease process. Possible precipitating or complicating infection must be treated after appropriate cultures have been taken. Tuberculosis should be excluded. Immunosuppressive therapy (e.g. methylprednisolone) is appropriate if the condition occurs in the setting of an autoimmune condition. Anti-TNF antibody therapy has been successfully used in selected cases. Epstein–Barr virus-associated haemophagocytic lymphohistiocytosis may be indistinguishable from T-cell lymphoma and responds to chemotherapy, such as etoposide-containing regimens. The condition is usually of brief duration until recovery or, sometimes, death occurs.

Haematological aspects of pregnancy (Table 48.15)

Pregnancy poses a major physiological challenge to the human body and a number of haematological changes accompany it.

Anaemia

Maternal plasma volume increases by approximately 50% during the first and second trimesters of pregnancy, whereas the corresponding increase in red cell mass is only 20–30%, resulting in a dilutional fall in haemoglobin concentration. Anaemia is defined as a haemoglobin <110 g/L in the first trimester, <105 g/L in the second or third trimesters and <100 g/L postpartum. The increase in maternal red cell mass, transfer of iron to the fetus (which takes place largely in the third trimester) and blood loss during labour together impose a requirement of about 800 mg of iron, so that iron deficiency frequently arises in mothers with normal or reduced iron stores. Folate requirements are also raised during pregnancy (Chapter 5). Routine supplementation is particularly advised before and during early pregnancy to prevent both megaloblastic anaemia and neural tube defects in the fetus. There is a fall in serum cobalamin to below normal in up to 30% of pregnancies which corrects postpartum. A physiological rise in MCV of 5–10 fL occurs during normal pregnancy. AIHA occurring during pregnancy is typically severe and refractory to therapy. The infant may have transient haemolysis for one or two months. An idiopathic direct antiglobulin negative haemolytic anaemia rarely occurs in the third trimester which is severe, may respond to corticosteroids or immunoglobulin and recovers spontaneously after the pregnancy. A microangiopathic anaemia is rare and associated with the HELLP syndrome (see below). Aplastic anaemia, pure red cell aplasia and sideroblastic anaemia have all been rarely described to have an onset in pregnancy with recovery spontaneously postpartum in a substantial proportion.

Table 48.15 Haematological changes during pregnancy.

Anaemia
Dilutional
Iron deficiency
Folate deficiency
Aplastic anaemia
White cells
Neutrophil leucocytosis
Platelets
Thrombocytopenia
Immune thrombocytopenic purpura
Eclampsia
HUS
TTPHELLP syndrome (haemolysis, elevated liver enzymes, low platelet count)
DIC
Drug induced
Coagulation
Coagulation factors: vitamin K-dependent factors FII, FVII, FIX, FXI, FX increased, FVIII increased, von Willebrand factor increased, fibrinogen increased
Coagulation inhibitors: protein C increased or no change, protein S decreased, antithrombin decreased or no change
Fibrinolytic activity decreased
DIC due to: abruptio placentae, intrauterine fetal death, amniotic fluid embolism, obstetric sepsis, eclampsia

White cells

A mild neutrophil leucocytosis with a left shift and occasional Döhle bodies occur during normal pregnancy.

Platelets

The normal range for the platelet count ($140–400 \times 10^9$/L) does not alter during pregnancy; thrombocytopenia occurring during pregnancy requires evaluation. Gestational thrombocytopenia complicates 8–10% of pregnancies and is characterized by mild thrombocytopenia occurring for the first time during pregnancy (platelets $80–150 \times 10^9$/L) and is usually not associated with neonatal thrombocytopenia or significant bleeding in the mother. Maternal immune thrombocytopenic purpura may antedate or present during pregnancy; it is often associated with increased levels of platelet-associated IgG, although this is a non-specific finding and the presence of serum platelet autoantibodies to platelet glycoproteins (GP)IIb/IIIa or GPIb/IX is more specific. The management of immune thrombocytopenic purpura is discussed in Chapter 49. Thrombocytopenia is regularly seen in pre-eclampsia. The mechanism is unknown, but increased aggregation is suggested, as low-dose aspirin therapy may increase the platelet count. TTP may occur at any time during pregnancy, but typically before 24 weeks; the use of fresh-frozen plasma and plasma exchange has been shown to improve fetal outcome. HUS typically occurs within 48 hours of delivery in an otherwise normal pregnancy.

The potentially fatal syndrome of haemolysis, elevated liver enzymes and low platelets (HELLP) occurs in up to 10% of pregnancies complicated by eclampsia. The existence of coagulation abnormalities with red cell fragmentation suggests that microangiopathy, DIC and endothelial damage all have a role in its pathogenesis. Fetal and maternal outcomes are characteristically poor.

Basophilic stippling, crenated red cells and large platelets are characteristic peripheral blood findings in acute fatty liver of pregnancy.

Coagulation changes

Normal pregnancy is associated with a range of alterations to haemostatic components (see Table 48.15), which combine to give an increased risk of haemorrhage, thrombosis and DIC, occurring in up to 40% of patients with abruptio placentae, leading to haemorrhage and shock. Amniotic fluid embolism typically occurs during the course of a difficult delivery in a multiparous woman and rapidly leads to a picture of chronic low-grade DIC, with onset over a period of 1–2 weeks. Venous stasis resulting from the gravid uterus combines with the coagulation changes to make pregnancy a hypercoagulable state; operative delivery imposes an additional risk. Postpartum DVT is frequent and prophylactic low-molecular-weight heparin is given to those at high risk, e.g. with a previous DVT.

Anaemia in the elderly

The WHO defines anaemia as a haemoglobin <130 g/L in men and <120 g/L in women. Other criteria are given by other authors. Using the WHO criteria, the incidence of anaemia is substantial in the elderly, e.g. >25% in men and >20% in women >85 years old. The incidence is higher in blacks, increases with age and predicts for shorter survival. The main causes are ACD, iron deficiency and renal disease, but about a third are unexplained. There is also an increased incidence of arterial and venous thrombosis with increasing age. This is partly due to increase in plasma levels of some of the clotting factors and reduced fibrinolysis, as well as to atheromatous plaques. Elderly subjects tend to be more sensitive than younger patients to anticoagulants and need careful monitoring to avoid haemorrhage.

Haematological aspects of HIV infection (Table 48.16)

HIV infection and its treatment cause a range of haematological effects. The newer generation of treatments with highly active antiretroviral therapy (HAART) is changing the natural

Table 48.16 Haematological complications of HIV.

Pancytopenia
Ineffective haemopoiesis, MDS
Anaemia
Drugs: anti-HIV, supportive care (e.g. ganciclovir)
Nutritional
Bleeding/phlebotomy
Anaemia of chronic disease
Parvovirus
Thrombocytopenia
Reduced platelet survival: antibodies, infection, splenomegaly, fever
Reduced production: megakaryocyte differentiation reduced, CD34 reduction
Neutropenia: antibody mediated, drug induced
TTP/HUS/microangiopathy: usually treatment induced
Coagulation abnormalities
Thrombosis
Antiphospholipid antibodies plus anticardiolipin antibodies
Protein S deficiency
Malignancy
Lymphoma: non-Hodgkin/Burkitt-like, primary CNS, Hodgkin, others
Myeloma
Acute myeloid leukaemia

history of AIDS and increasingly turning it into a chronic disease. HAART is also less toxic, particularly to the marrow, than the drugs it has replaced.

Pathophysiology

HIV infection suppresses haemopoiesis through the action of cytokines (e.g. IFN-γ, TNF-α). These are elaborated by activated lymphocytes in response to infection and have been shown to induce progenitor cell apoptosis and reduced growth of short-term marrow cultures. Direct infection of CD34 cells by HIV also occurs and cross-culture experiments have demonstrated that HIV infection impairs proliferation of CD34 cells from AIDS patients more than its effect on bone marrow stromal cells. Morphological evidence of ineffective myelopoiesis is seen in the marrow, with myelodysplastic changes in all three lineages. An increased number of plasma cells is also frequent.

Anaemia

This is the most common abnormality, occurring in up to 80% of patients; it may be due to a range of factors (Table 48.16). HIV infection itself is a prominent cause of ACD. Infiltration of the marrow by tumour, such as Hodgkin or non-Hodgkin lymphoma, is much more frequent among HIV-positive subjects than normal. There is also a relative reduction in EPO levels. Treatment of HIV infection by HAART leads to an improvement in anaemia.

Other common causes include bleeding and infection. Certain infections, such as *Mycobacterium avium* and parvovirus B19, may directly involve the marrow. Nutritional anaemia frequently arises from a defective diet, and gastrointestinal disease (e.g. due to infection or drug toxicity) is common. Vitamin B_{12} deficiency is seen in up to one-third of AIDS subjects and iron and folate deficiencies are also common. Although treatment-induced anaemia is less common in the HAART era, treatment is a contributory factor in more than 50% of cases.

White cells

Leucopenia is frequently seen and predominantly due to lymphopenia. Neutropenia also occurs and is multifactorial. Granulocyte and monocyte production are reduced in HIV infection; drug-induced changes are common and autoimmune neutropenia also occurs. Granulocyte colony-stimulating factor has been successfully used to improve the neutrophil count in infected HIV-positive subjects. Defects in neutrophil and monocyte function have also been demonstrated.

Platelets

Thrombocytopenia is seen in up to 50% of HIV-infected patients. It is usually due to increased platelet destruction. This is often antibody mediated (immune thrombocytopenia; see Chapter 42). HIV-specific antibodies have been shown to share a common epitope with antibodies against platelet GPIIb/IIIa. Non-specific absorption of immune complexes onto platelets also occurs and predisposes to immune thrombocytopenia. Reduced platelet production is common in HIV-positive subjects and direct infection of megakaryocytes by HIV has been described. Morphologically abnormal megakaryocytes are seen in the marrow. Other causes of thrombocytopenia include infection and microangiopathy. Both HUS and TTP are well described in the setting of HIV infection, particularly in the era before HAART. Treatment of thrombocytopenia depends on the cause. Treatment of HIV infection using HAART frequently leads to improvement in the platelet count. Immune thrombocytopenia in HIV-infected subjects is treated in the same way as in the non-HIV population. However, immunosuppression carries particular risks, and steroids should be used with extreme caution. Splenectomy is often effective, and anti-CD20 monoclonal antibody therapy may also be used.

Coagulation

Coagulation abnormalities can occur in the setting of infection and acute illness. Circulating coagulation inhibitors also occur. Protein S deficiency may predispose to thrombosis. DIC is common in the setting of infection and HIV infection.

Other changes

Lymphoma is the commonest tumour in HIV-positive subjects (Chapter 34). Other haematological malignancies also occur with increased frequency, for example myeloma and AML. HIV-positive subjects who develop these malignancies often display a prolonged premalignant phase, with smouldering myeloma and myelodysplasia frequently reported.

Selected bibliography

Camaschella C (2013) Iron and hepcidin: a story of recycling and balance. *Hematology*, ASH Education Program 2013: 1–8.

KDOQ1 Clinical practice guidelines and clinical practice recommendations for anemia in chronic kidney disease (2006) *American Journal of Kidney Diseases* **47**: S11–145.

Okonko DO, Van Veldhuisen DJ, Poole-Wilson PA *et al.* (2005) Anaemia of chronic disease in chronic heart failure: the emerging evidence. *European Heart Journal* **26**: 2213–14.

Ponikowski P, van Veldhuisen DJ, Comin-Colet J *et al.* (2015) Beneficial effects of long-term intravenous iron therapy with ferric carboxymaltose in patients with symptomatic heart failure and iron deficiency. *European Heart Journal* **35**: 657–68.

Redig AJ, Berliner N (2013) Pathogenesis and clinical implications of HIV-related anemia in 2013. *Hematology*, ASH Education Program 2013: 377–81.

Rizzo JD, Brouwers M, Hurley P *et al.* (2010) American Society of Hematology/American Society of Clinical Oncology clinical practice guidelines update on the use of epoietin and darbepoietinin adult patients with cancer. *Blood* **116**: 4045–59.

Sloand E (2005) Hematologic complications of HIV infection. *AIDS Reviews* **7**: 187–96.

Stancu S, Stanciu A, Zugravu A *et al.* (2010) Bone marrow iron, iron indices, and the response to intravenous iron in patients with non-dialysis dependent CKD. *American Journal of Kidney Diseases* **55**: 639–47.

Weiss G, Goodnough L (2005) Anemia of chronic disease. *New England Journal of Medicine* **352**: 1011–23.

CHAPTER 49

Haematological aspects of tropical diseases

49

Imelda Bates[1] and Ivy Ekem[2]

[1]Liverpool School of Tropical Medicine, Liverpool, UK
[2]University of Ghana Medical School, Accra, Ghana

Introduction

The chapter has been divided into two sections, covering tropical diseases in which organisms can be visualized in the blood or bone marrow and those that cause secondary haematological abnormalities. Discussing the haematological aspects of tropical diseases is very important as global travel has expanded and haematologists are increasingly involved with patients who have travelled to, or are from, tropical regions.

In 2010, 3% of the world's population lived outside their country of origin. The majority were economic migrants, often moving from tropical countries to temperate areas. The latest UK Office for National Statistics indicate that 624,000 people immigrated to the UK in the year ending September 2014, a statistically significant increase from 530,000 in the previous 12 months. Rapid increases in worldwide travel mean that haematologists need not only to know about the tropical diseases that can cause haematological abnormalities, but also be able to take relevant travel histories from patients and appropriately interpret haematological results obtained from investigations. They also need to be aware of ethnic variations in reference ranges to avoid wasting resources on unnecessary investigations and to avoid causing undue anxiety for the patient.

Ethnic variations in reference ranges

The white blood cell count and relative and absolute neutrophil counts are lower in people of African and Middle Eastern descent than in Caucasians. After the age of 1 year, Africans have lower counts than West Indians or African Americans (Table 49.1). The difference is due to a greater number of neutrophils in the storage pool of the non-Caucasian populations. Stimulation of a neutrophil response in these ethnic groups leads to rises in the neutrophil count to the same level as white populations, irrespective of the baseline level. Indian, Chinese and Southeast Asian populations have the same white blood cell and neutrophil counts as northern Europeans. The platelet counts in healthy West Indians and Africans may be 10–20% lower than in Europeans living in the same environment. It is not clear whether there are true ethnic variations in eosinophil counts, but counts of up to 2×10^9/L have been described in healthy blood donors in Africa. Interpretation of variations in blood counts should therefore take account of the individual's ethnic background. It must be borne in mind that 'normal' values derived from populations with high background prevalence of ill health may include individuals with asymptomatic conditions.

Tropical diseases with organisms in peripheral blood or bone marrow

Malaria

Epidemiology and biology

Until 2004, only four species of plasmodia were known to cause malaria in humans: *Plasmodium (P) falciparum*, *P. vivax*, *P. ovale* and *P. malariae*; since then there have been increasing numbers of cases of infection by a fifth plasmodium, i.e.

Postgraduate Haematology, Seventh Edition. Edited by A Victor Hoffbrand, Douglas R Higgs, David M Keeling and Atul B Mehta.

Table 49.1 Automated white cell (WBC) and neutrophil counts in adults of different ethnic origins (95% ranges).

	Male		Female	
	WBC ($\times 10^9$/L)	Neutrophil count ($\times 10^9$/L)	WBC ($\times 10^9$/L)	Neutrophil count ($\times 10^9$/L)
Caucasian	3.7–9.5	1.7–6.1	3.9–11.1	1.7–7.5
Afro-Caribbean	3.1–9.4	1.2–5.6	3.2–10.6	1.3–7.1
African	2.8–7.2	0.9–4.2	3.0–7.4	1.3–3.7

P. knowlesi initially reported in South-east Asia. *P. falciparum* remains by far the most dangerous and causes the greatest mortality and morbidity (see Figure 49.1).

Each year for the past 10 years, over 1500 cases of malaria have been imported into the UK. Over 80% of *P. falciparum* infections in the UK are acquired in sub-Saharan Africa, while 85% of *P. vivax* infections are acquired in South Asia, especially India and Pakistan. In 2013, 1501 cases of malaria were reported in the UK, of which 1192 (79%) were due to *P. falciparum* and 179 (12%) to *P. vivax*; the rest were due to *P. ovale and P. malariae*, There were seven deaths. In 2006 a case of *P. knowlesi* was reported in the UK. Malaria rates were highest in immigrants returning from visits to their countries of origin and accounted for one-third of all reports. It has also been reported in travellers who have transited through airports in malarious areas. The majority of travellers had not taken malaria prevention tablets.

Malaria is transmitted by the bite of an infected female *Anopheles* mosquito. The infecting agent is the spindle-shaped sporozoite and thousands of these may be injected by a single bite. Infrequently, transmission may also occur through blood transfusion, bone marrow transplantation and transplacentally (0.5% of all UK cases). There have also been reports of malaria transmission in aircraft or near airports in temperate zones due to infected mosquitoes being brought to non-malarious areas.

P. knowlesi is transmitted by *Anopheles latens*. Its absence in Africa may partly be due to the absence of its reservoir hosts, long-tailed and pig-tailed macaques, in Africa. *P. knowlesi* is the most common cause of malaria in childhood in the Kudat district of Sabah, Malaysia.

Within a few hours of an infected bite, the sporozoites enter the hepatocytes, where they divide (Figure 49.2). Rupture of the hepatocyte releases the parasites into the blood, where they attach to red cell membranes using specific receptors. In the red cells, asexual replication of the trophozoites (ring forms) occurs, giving rise to erythrocyte schizonts. The schizonts mature into merozoites and are released into the circulation to re-infect other red cells. The periodicity of this release varies with the species and is responsible for the classical cyclical nature of malaria fevers. Relapses, which can occur months or years after the primary illness, are characteristic of infection with *P. vivax* and *P. ovale*, and are due to maturation of persistent hypnozoites in the liver.

A few of the trophozoites develop into male and female banana-shaped gametocytes and are taken up by the mosquito at a blood meal. Inside the midgut of the mosquito, they undergo sexual reproduction and sporozoites migrate to the salivary glands, ready to infect another host when the mosquito bites. Unlike the schizonts of *P. vivax*, *P. ovale P. malariae*

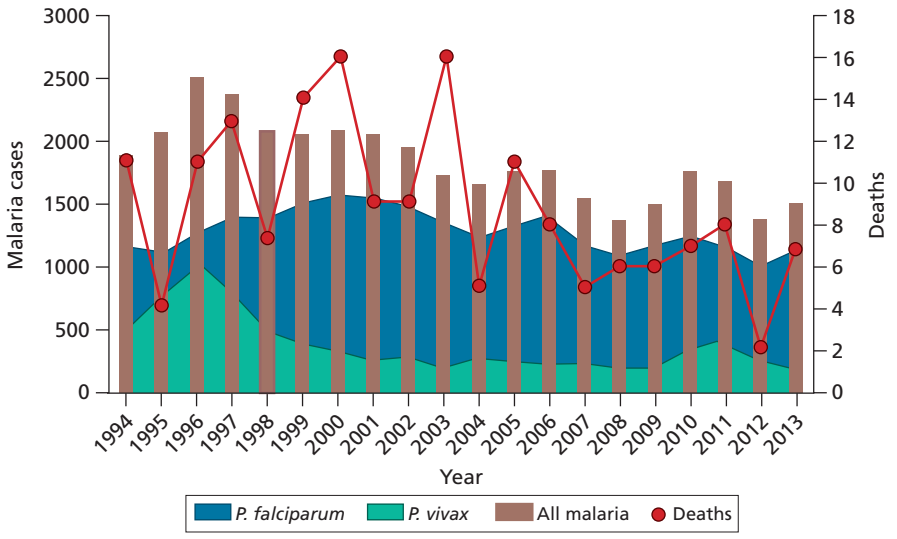


Figure 49.1 Imported malaria cases and deaths, United Kingdom: 1994–2013 (Data from the PHE Malaria Reference Laboratory).

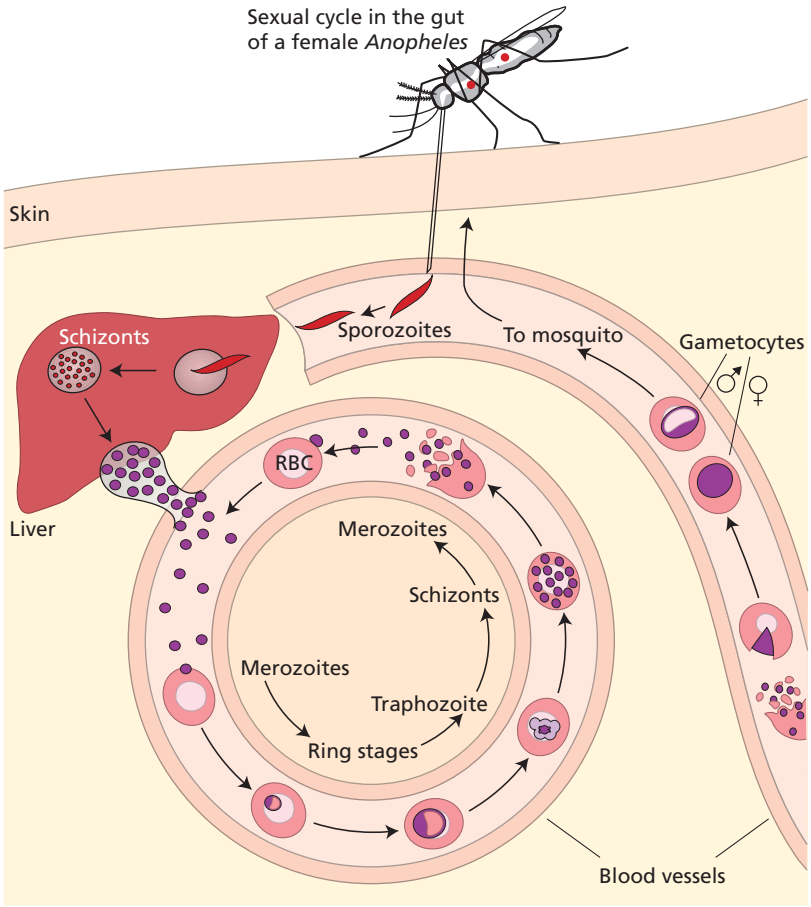


Figure 49.2 Life cycle of the malaria parasite.

and *P. knowlesi*, those of *P. falciparum* are not commonly seen in the peripheral blood of the human host. This is because *P. falciparum*-infected cells have surface cytoadherence molecules that enable them to be sequestered in the deep tissues. Therefore *P. falciparum* schizonts only appear in the blood in very severe infections or in splenectomized patients. Sequestration is responsible for some of the severe clinical consequences of *P. falciparum* malaria, such as cerebral malaria.

Clinical features

The time between the infected bite and the appearance of clinical symptoms and parasites in the peripheral blood varies between species. It is 7–30 days (mean 10 days) in *P. falciparum*, but can be months, or even years, with other species, particularly *P. vivax* and *P. ovale*, because of their hypnozoite stage.

All five *Plasmodium* species produce factors that cause release of tissue cytokines, especially from leucocytes. These cytokines produce fever and contribute to anaemia through marrow suppression. Splenomegaly is a common feature of acute malaria and mild jaundice may also occur secondary to haemolysis. Other clinical features vary with different species. Maximal immunity to malaria takes around 10 years to develop and is lost over the course of 1–5 years if the individual leaves a malarious area and is no longer exposed to infections.

Plasmodium falciparum

Plasmodium falciparum is the only species associated with complicated and severe disease (Figure 49.3) though recently *Plasmodium vivax* has also been associated with severe and fatal disease. In *P. falciparum* infections, death may occur after a single exposure to malaria, particularly in those with no immunity, such as non-immune travellers or young children in endemic countries. Recurrent attacks with different strains lead to the development of clinical immunity, but not necessarily to complete clearance of parasites. Parasites may therefore be detected in a high proportion of clinically asymptomatic adults in endemic areas. In addition to fever with rigors, nausea, and hot and cold phases, *P. falciparum* infection can also present with diarrhoea and cough (Table 49.2). Serious complications include severe anaemia, cerebral involvement and failure of major organs such as kidneys and liver.

During pregnancy, immunity to malaria is reduced and parasite density increases. Even when parasites cannot be visualized in the peripheral blood, they may be sequestered in

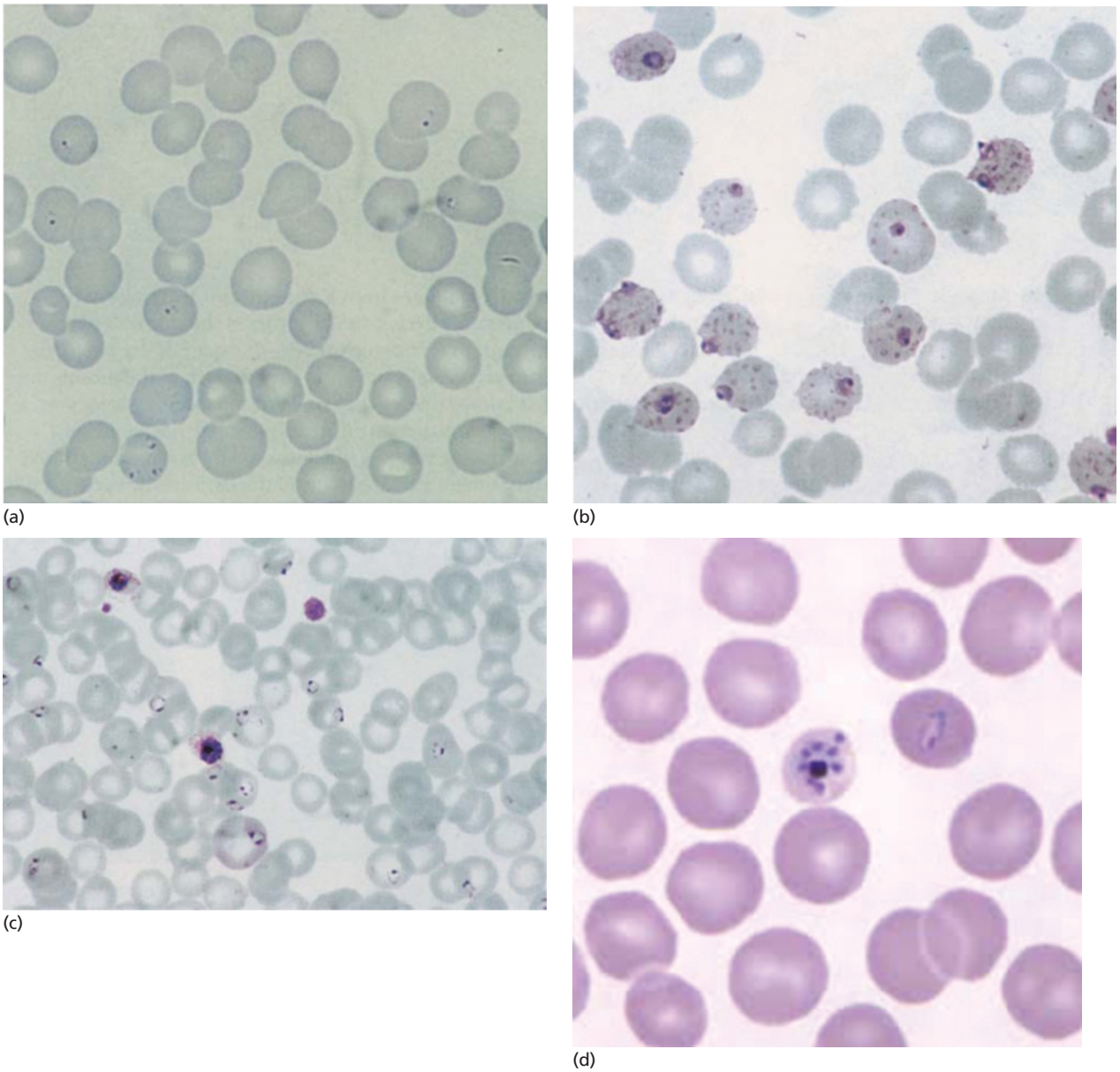


Figure 49.3 Stages in the life cycle of *Plasmodium falciparum* in Giemsa-stained thin films; the cells are not enlarged or decolorized: (a) delicate early ring forms, (b) ring forms with prominent Maurer's clefts and (c) ring forms and early and late schizonts (schizonts are not commonly seen in the peripheral blood). (d) Ring form and schizont of *Plasmodium knowlesi.* (Part (d) Source: © GEFOR 2012 (www.gefor.4t.com). Reproduced with permission of Galería de imágenes y recursos.)

the placenta and compromise fetal development. Malaria is an important cause of low birth weight in neonates and anaemia in pregnant women, both of which have a detrimental effect on the later development of the infant.

Plasmodium malariae

The incubation period of *P. malariae* may be several weeks. It is associated with recurrent fever, anaemia and enlargement of the liver and spleen. Without treatment or following incomplete treatment of the primary infection, re-crudescences may occur with decreasing severity over many years. Clinical symptoms of malaria have been reported up to 30 years after the initial infection.

Plasmodium vivax and *Plasmodium ovale*

These species cause a similar clinical picture with bouts of fever occurring periodically up to 5 years after the initial infection. These are relapses, due to reinvasion of red cells by

Table 49.2 Clinical features of *P. falciparum* malaria infection.

Target organ	Clinical features	Potential common misdiagnosis
Gastrointestinal	Diarrhoea, vomiting	Traveller's diarrhoea
Respiratory	Cough, pulmonary oedema	Pneumonia, cardiac failure
Neurological	Delirium, coma, convulsions, focal neurological signs	Encephalitis, meningoencephalitis
Renal	Oliguria, haemoglobinuria	Nephritis
Hepatic	Jaundice, hypoglycaemia	Hepatitis
Haematological	Anaemia, splenomegaly	Viral infection, lymphoma

hypnozoites. The trigger for hypnozoites to reactivate after dormancy is unknown.

Plasmodium knowlesi

The incubation period of *P. knowlesi* infection is about 11 days. It causes severe, and occasionally fatal, malaria. This is because the parasite invades red cells of all ages and has a short life cycle of 24 hours.

Haematological abnormalities

Normochromic normocytic anaemia is common in malaria, particularly in children, but the degree and rapidity of onset are variable. The haemoglobin may fall by up to 20 g/L each day. In malaria-endemic regions, chronic anaemia due to nutritional deficiencies, intestinal helminths, HIV infection and haemoglobinopathies may be compounded by anaemia due to malaria. In chronically anaemic patients, the oxygen dissociation curve is shifted to the right so they are better able to tolerate further falls in haemoglobin. The clinical effects of anaemia in malaria are therefore due to a combination of the degree and rate of fall of haemoglobin.

Anaemia due to malaria has multiple aetiologies: red cells containing parasites are removed from the circulation by the reticuloendothelial system; there is accelerated destruction of non-parasitized cells and dyserythropoiesis in the bone marrow. Both parasitized and non-parasitized red cells lose deformability and the high shear rates in the spleen enhance their removal. In acute malaria, reticulocyte response is suppressed and erythropoietin levels are elevated, although this may be less than expected for the degree of anaemia.

Uncommon complications of malaria that can exacerbate the anaemia are hyper-reactive malarial splenomegaly (HMS) and 'blackwater fever'. HMS is characterized by massive splenomegaly with hypersplenism (see later) and occurs as a result of a disordered immune response to malaria. Blackwater fever is associated with severe intravascular haemolysis with haemoglobinuria, and can lead to acute renal failure. It has been associated with antimalarial drugs, particularly quinine, and may be associated with glucose-6-phosphate dehydrogenase (G6PD) deficiency.

The high prevalence of severe anaemia in areas with intense malaria transmission has generally been ascribed to malaria, but studies in Malawian children have suggested that the anaemia is multifactorial, with bacteraemia, malaria, hookworm, HIV and deficiencies of G6PD, vitamin A and vitamin B_{12}, but not iron and folate deficiency, being important.

Case fatality rates of children with severe anaemia in Africa are 9–18% and mortality rises steeply at haemoglobin concentrations of less than 40 g/L. Severe decompensated malarial anaemia can be accompanied by hypovolaemia and acidosis and therefore requires careful rehydration (not with rapid fluid boluses which may be harmful) and blood transfusion. With the risks associated with transfusions, it is important to prevent and adequately treat the milder forms of anaemia so that transfusions can be avoided. Antimalarial prophylaxis, prompt treatment of malaria and avoidance of mosquito bites are valuable in reducing malaria infections and anaemia in vulnerable groups, such as pregnant women and young children.

The white cell count in malaria is usually normal, but may be raised in severe disease. Other white cell changes that have been described in malaria include a leucoerythroblastic response, monocytosis, eosinopenia and a reactive eosinophilia during the recovery phase. Neutrophil activation may be apparent in severe malaria. Thrombocytopenia due to increased splenic clearance, and thus increased platelet turnover and raised thrombopoietin levels is usually mild with counts around 100×10^9/L, but can be marked in *P. knowlesi* infection. Pancytopenia without massive splenomegaly has also been described.

The bone marrow of patients with acute malaria due to any of the species shows prominent dyserythropoiesis. This may persist for weeks after the acute infection and is caused by intramedullary cytokines produced by the infection. Erythrophagocytosis and macrophages containing malaria pigment are frequently seen in marrow samples.

Although malaria is associated with thrombocytopenia and activation of the coagulation cascade and fibrinolytic system, bleeding and haemorrhage are uncommon, even though the prothrombin and partial thromboplastin times may be prolonged. Disseminated intravascular coagulation is not thought to be important in the pathogenesis of severe malaria. Fibrinogen levels are often increased and there is rapid fibrinogen turnover with consumption of antithrombin and factor XIII resulting in increased fibrin degradation products. Microparticle formation from platelets, red cells and macrophages also causes widespread activation of blood coagulation. Malaria also results in increased levels of circulating active von Willebrand

factor. Haematological indicators of a poor prognosis in severe malaria include:

- leucocytosis $>12 \times 10^9$/L;
- severe anaemia (packed cell volume <15%);
- thrombocytopenia $<50 \times 10^9$/L;
- prolonged prothrombin time (>16 s);
- prolonged partial thromboplastin time (>40 s);
- reduced fibrinogen (<2 g/L);
- hyperparasitaemia >100,000/μL (high mortality >500,000/μL);
- >20% of parasites are pigment-containing trophozoites and schizonts;
- >5% of neutrophils contain visible malaria pigment.

Genetic haematological protection mechanisms

Plasmdium vivax and *P. knowlesi* need Duffy blood group antigen as a receptor to enter red cells. This antigen is absent in at least two-thirds of all Africans who consequently have a natural resistance to infection with *P. vivax and P. knowlesi*. Hb AS has a substantial (up to 10 times) protective effect against severe malaria and a similar, but less marked protection is associated with other genetic red cell abnormalities such as G6PD deficiency, thalassaemia trait and Hb C trait.

Diagnosis

Microscopy

Direct visualization of parasites by light microscopy using a combination of thick and thin blood films is the gold-standard diagnostic technique for malaria (Table 49.3 and Figure 49.3). A Romanowsky stain (e.g. Field, Giemsa, Leishman) pH 7.2 is used so that the parasite cytoplasm stains blue and the nuclear chromatin red. A thick blood film examined by two observers, each viewing a minimum of 200 high-power fields, should be used as the first screening tool as it allows larger volumes of blood to be examined than the thin film. However, the parasites appear distorted due to the process of lysing the red cells so this method cannot be used for parasite morphology and speciation. A thin blood film should be performed on any sample that yields a positive or uncertain result. This allows visualization of undistorted parasites and of the size and shape of the red cells, so speciation and quantification (parasites/μL) can be carried out. However, the thin film has low sensitivity because of the small amount of blood that can be examined.

The disadvantages of basing a diagnosis of malaria on blood film examination include the following:

- An initial negative film does not exclude malaria: at least three films taken during episodes of fever should be examined in the absence of antimalarial drugs to confirm a negative blood film.
- A positive film does not prove that symptoms are due to malaria: asymptomatic parasitaemia is common in adults from endemic areas.
- *P. malariae* and *P. knowlesi* cannot be distinguished morphologically.
- Parasites, particularly *P. falciparum* gametocytes, may be washed off the slide during staining and bulk staining may result in transfer of parasites between slides.
- Parasite density does not necessarily correlate with disease severity, although heavy parasitaemia (>5% of red cells infected) indicates a poor prognosis.

Malaria pigment may persist in phagocytic cells for several weeks after an acute attack and may be helpful in retrospective diagnosis of malaria. Automated haematology analysers may produce an abnormal pattern on the white cell differential count histogram. Debris below the white cell threshold may be due to malaria parasites and manual examination of blood films is indicated if this pattern is flagged up by the analyser.

Antigen detection

Rapid diagnostic tests for malaria are based on detection of the malaria antigen histidine-rich protein (HRP)2 or parasite lactate dehydrogenase (pLDH). They have been incorporated into immunochromatographic antigen-capture kits for rapid diagnosis. The sensitivity of these dipstick strip tests approaches that of thick film microscopy (i.e. 0.002% parasitaemia equivalent to 100–200 parasites/μL of blood). HRP2 protein may remain positive for 14 days after successful treatment and false positives due to rheumatoid factor have been reported. pLDH is only produced by viable parasites so it becomes negative 2–3 days after successful treatment. These tests should not replace microscopy, but are useful in on-call or emergency situations or when no experienced microscopist is available.

Antibody detection

Malarial antibodies can remain in the blood after the eradication of parasites, so their detection is not useful for diagnosis in the acute attack. The main uses of malarial antibody detection are for excluding malaria as a cause of recurrent fever, for population surveys and as a screening test for blood donors in non-endemic areas.

DNA-based methods

DNA probes have been developed for malaria diagnosis, but their use is generally restricted to research and epidemiological surveys. Although the current prevalence of *P. knowlesi* infections is relatively low, it may be misdiagnosed as *P. malariae*, especially when microscopy is used. *P. knowlesi* can only be accurately distinguished from *P. malariae* using PCR assay and/or molecular characterization.

Haematological implications of treatment for falciparum malaria

Current widely recommended first-line treatment for malaria is with combination therapy which includes artemisinin.

Pyrimethamine is used in combination with a long-acting sulfonamide, such as sulfadoxine (as in Fansidar), a dihydrofolate reductase inhibitor, which may therefore induce megaloblastic

Table 49.3 Differentiating epidemiological and stained thin blood film features of the five *Plasmodium* species affecting humans compared to *Babesia* species.

	P. falciparum	***P. knowlesi***	***P. vivax***	***P. ovale***	***P. malariae***	***Babesia* sp.**
Main site of distribution, proportion in the endemic area (%)	Global, most common in Africa, 80–90 (Africa)	South-east Asia,1–60	Mostly in Asia, 50–80	Africa, 5–8	Global, 0.5–3	
Life-threatening	Yes	Yes	Yes	No	No	
Parasitemia	Can be high	Can be high	<2%	<2%	<2%	
Treatment	Combinations with arthemisinine derivatives: atovaquone plus proguanil; mefloquine;quinine +/– doxycycline	Probably all the medications listed for the other plasmodia	Chloroquine followed by primaquine	Chloroquine followed by primaquine	Chloroquine	
Appearance of infected red blood cells (size and shape)	Normal size and shape	Normal size and shape although some have irregular edges	1.5–2 times larger than normal; shape normal or oval	As for *P. vivax*, but some have irregular edges	Normal shape; size normal or smaller	Both normal
Red cells with multiple parasites/cell	Common	Common	Occasional	As for *P. vivax*	Rare	Common
Stages present in peripheral blood	Rings and gametocytes; occasionally schizonts	All stages: late stages similar to *P. malariae*	All stages	As for *P. vivax*	All stages	Only rings and rare pear-shaped forms ('Maltese cross'); no gametocytes
Ring form (young trophozoite)	Delicate small ring; scant cytoplasm; sometimes at the edge of red cell ('accolé form')	Similar to *P. falcipraum*	Ring one-third to half of the diameter of cell; heavy chromatin dot	As for *P. vivax*	Ring one-third of the diameter of cell; heavy chromatin dot; vacuole sometimes 'filled in'	Resembles ring of *P. falciparum*; look for pear-shaped structure
Gametocyte	'Crescent' or 'sausage' shape is characteristic	Similar to *P. malariae*; comprise 1–3% of infected erythrocytes	Round or oval	Round or oval (smaller than *P. vivax*)	Round or oval; dark coarse pigment	No gametocyte

anaemia and other cytopenias. The sulfur component of these combinations may also cause methaemoglobinaemia.

Dapsone acts by inhibiting the synthesis of dihydrofolic acid and is used as part of a fixed combination with proguanil or chlorproguanil. It may be associated with haemolytic anaemia, methaemoglobinaemia and eosinophilia.

Primaquine, an inhibitor of protein synthesis, is active against the hypnozoites of *P. vivax* and the gametocytes of *P. falciparum*. It causes oxidant haemolysis in patients with G6PD deficiency and, rarely, methaemoglobinaemia.

Quinine is usually reserved for life-threatening infections, especially in endemic countries. It acts by disrupting the food vacuole of the parasite. Rarely, it is associated with immune thrombocytopenia and severe intravascular haemolysis.

Mefloquine, halofantrine and artemisinin-related compounds do not commonly cause significant haematological side-effects. A combination of artemether and lumefantrine (Coartem or Riamet) is commonly used as first-line treatment in endemic areas.

Some antibacterial drugs, such as doxycycline, clindamycin, trimethoprim and sulfonamides, have also been used for their antimalarial effect and may be associated with haematological side-effects.

For malaria due to *P. vivax*, *P. ovale, P. malariae and P. knowlesi*, chloroquine is still widely used for treatment, as resistance is generally low; primaquine is added to prevent relapses in vivax, ovale and knowlesi malaria.

Warfarin and malaria chemoprophylaxis

Travellers on warfarin should begin malaria chemoprophylaxis one week (2–3 for mefloquine) before travelling and the INR should be stabilized before travel.

Babesiosis in the differential diagnosis of malaria

Babesiosis is not a tropical disease, but is briefly described here as it can be confused clinically and haematologically with malaria. Babesiosis is endemic in the north-eastern and upper midwestern regions of the USA and is found sporadically in other parts of the USA, Europe, Asia, Africa and South America. It is primarily a disease of animals and rarely infects humans. It is due to a protozoan parasite transmitted by the bite of the ixodid tick. Following the bite, the organisms penetrate red cells, where they take on an oval, round or pear shape and multiply by budding. The erythrocytic ring forms of *Babesia microti* and *B. divergens* may be confused with malaria *P. falciparum* rings, but they do not produce pigment or cause alterations in red cell morphology. A minority of organisms take on a folded shape and are thought to be gametocytes.

Babesia bovis, *B. microti* and *B. divergens* are responsible for the majority of human infections. The consequences of babesiosis can range from asymptomatic infection to severe and occasionally fatal outcomes. Most of the cases reported from Europe have been due to *B. divergens* and occurred in patients without a functioning spleen. In North America almost all the cases have been due to *B. microti*, which usually produces a subclinical infection. Specific laboratory diagnosis of babesial infection is made by morphological examination of Giemsa-stained blood smears, serology and amplification of babesial DNA using polymerase chain reaction.

Filariasis

Epidemiology and biology

There are two groups of human filariasis: those that occur in the blood (lymphatic filariasis) and those that occur in the skin (onchocerciasis). Only lymphatic filariasis is considered in this chapter, as it is associated with detectable organisms in the peripheral blood.

Three species of filarial worms cause lymphatic filariasis in humans and are relevant for haematologists, *Wuchereria bancrofti, Brugia timori* and *Brugia malayi*. They have different geographical distributions, with *W. bancrofti* being the most widespread. More than 90% of infections due to *W. bancrofti* are found in Asia, although it also occurs in Africa, America and the Pacific islands. Filariasis due to *B. malayi* occurs in China, Indo-China, Thailand, Malaysia, Indonesia, the Philippines and south-west India, *Brugia timori* has predominantly been reported in Indonesia (Figure 49.4).

The worms are 4 mm (male) to 10 mm (female) in length and can live for over 10 years in the lymphatics. Microfilariae, which are 250–300 μm long, are produced by the female worm and released into the blood after 3–8 months, where they may live for up to 1 year. Microfilariae densities can reach 10,000/mL but are usually much lower. They exhibit daily periodicity in the blood and this timing is designed to match the biting habits of their mosquito vectors, culicine and anopheline mosquitoes. The microfilariae develop into infective larvae in the mosquito and pass into the proboscis, ready to be transmitted into a new host.

Clinical presentation

There is wide variation in the presenting features of lymphatic filariasis, which may occur 6 months or more after the infective bite. The symptoms and signs are due to lymphangitis. There are recurrent bouts of fever with heat, redness and pain over lymphatic vessels. In fair-skinned people, the lymphangitis can be seen to spread distally (i.e. the opposite direction to septic lymphangitis). In *W. bancrofti* infection, these repeated episodes of inflammation eventually result in the typical chronic picture of filariasis, including hydrocele, lymphoedema and elephantiasis, chyluria and tropical pulmonary eosinophilia. The clinical picture in *B. malayi* infection is similar, but it does not cause hydrocele or chylous urine. *B. timori* causes lymphodema characteristically limited to the leg below the knee.

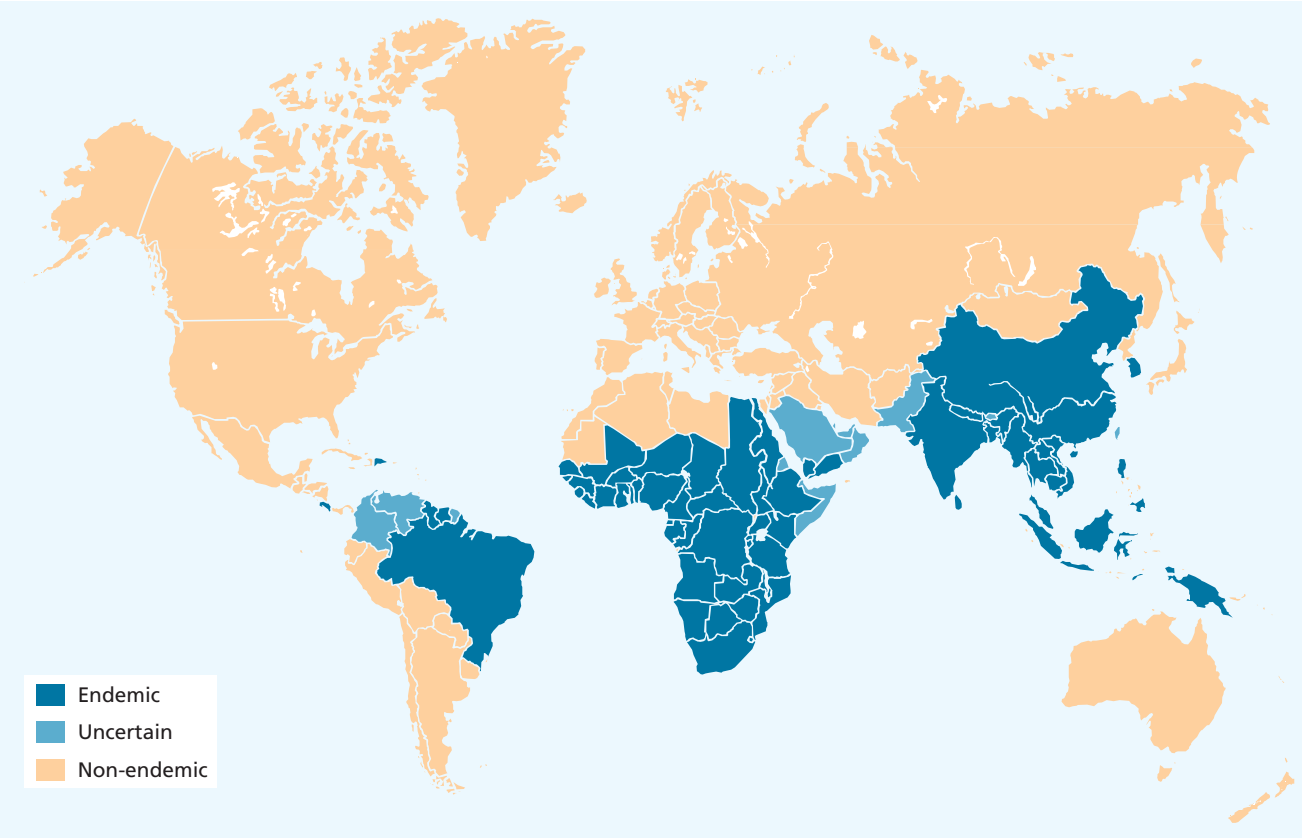


Figure 49.4 Global distribution of lymphatic filariasis.

Other filariae with blood-inhabiting larvae

Loa loa

This occurs in the rain-forest belt of Africa, especially West Africa. The adult worms migrate through the subcutaneous tissues, including the conjunctiva, and occasionally can be seen passing across the eye.

Mansonella perstans

This is a common infection in Africa. It can cause angioedema, pruritus, abdominal pain and eosinophilia. These organisms often coexist in the blood with *W. bancrofti*, but can be distinguished by their smaller size and absence of a sheath.

Mansonella ozzardi

This is probably non-pathogenic and occurs in the West Indies and South America. The adult worm lives in the body cavities and mesentery, rarely causing any symptoms.

Haematological abnormalities in filariasis

Eosinophilia is the major and most frequent haematological abnormality produced by lymphatic filariasis. Tropical pulmonary eosinophilia is an unusual complication of filariasis and is due to an immunological hyper-responsiveness to microfilariae in the lungs. It is more common in men than women. Although microfilariae are absent from the blood in this syndrome, they may be seen in lung biopsies and adult worms can be visualized in lymphatics on ultrasound. There is an extreme eosinophilia, with eosinophil counts of greater than 10×10^9/L; the level of eosinophilia is not related to the severity of symptoms. A study of filariasis in India in 2013 found that, in addition to eosinophilia, there was leucocytosis, neutrophila, thrombocytosis, raised erythrocyte sedimentation rate and raised mean corpuscular volume. Haemoglobin, total red cell count, lymphocytes, basophils, monocytes, mean corpuscular haemoglobin and mean corpuscular haemoglobin concentration were reduced. In tropical pulmonary eosinophilia, diethylcarbamazine treatment reduces the eosinophil count and produces resolution of symptoms. This rapid response to treatment distinguishes filariasis from other causes of marked eosinophilia associated with helminths such as *Ascaris*, *Strongyloides*, *Schistosoma* subsp. *trichinosis* and *Toxocara*.

Diagnosis of filariasis

The adult worms residing in the lymphatics are inaccessible, so diagnosis is based on finding microfilariae in the peripheral blood. The level of filaraemia is inversely related to the clinical

signs because much of the damage is due to immunological responses to the microfilariae rather than to the organisms themselves. Furthermore, the presence of microfilariae does not necessarily mean that they are causing clinical problems and, conversely, a lack of microfilariae in the blood does not exclude a diagnosis of filariasis. The peripheral blood findings must therefore be assessed in the context of the clinical picture.

To optimize the chances of finding scant microfilariae in the blood, the sample should be taken at the appropriate time for the expected peak concentration of microfilariae (i.e. around midnight or midday for nocturnally and diurnally periodic forms, respectively). There are many techniques for demonstrating microfilariae in the laboratory. The simplest method is a wet preparation of fresh blood. Microfilariae will survive in venous blood collected into EDTA for 2 days at room temperature. Motile microfilariae can be seen on a slide under low power and can be counted in a counting chamber. Numbers of microfilariae may be low, requiring concentration techniques such as blood filtering.

For species identification, thick and thin blood films should be stained with Giemsa or haematoxylin and the microfilariae differentiated according to the pattern of their sheaths, nuclei distribution and size (Figure 49.5). The edges of the film should be examined carefully as microfilariae tend to be concentrated at the periphery and are easily missed unless the whole film is scanned at low magnification.

Detection of circulating antigen by enzyme-linked immunosorbent assay (ELISA) or immunochromatography (ICT) has replaced microscopy for the diagnosis of bancroftian, but not brugian, filariasis. An antigen ICT card test is available for the detection of *W. bancrofti*, which does not react with other filariae and is highly sensitive (100%) and specific (92%). Filarial DNA can be detected by PCR, and ultrasound scans can help identify adult worms within the lymphatic system. Serological tests are not very helpful as most individuals from endemic areas have antibodies to crude filarial antigens and there is cross-reactivity with other filariae and nematodes.

Haematological implications of treatment for filariasis

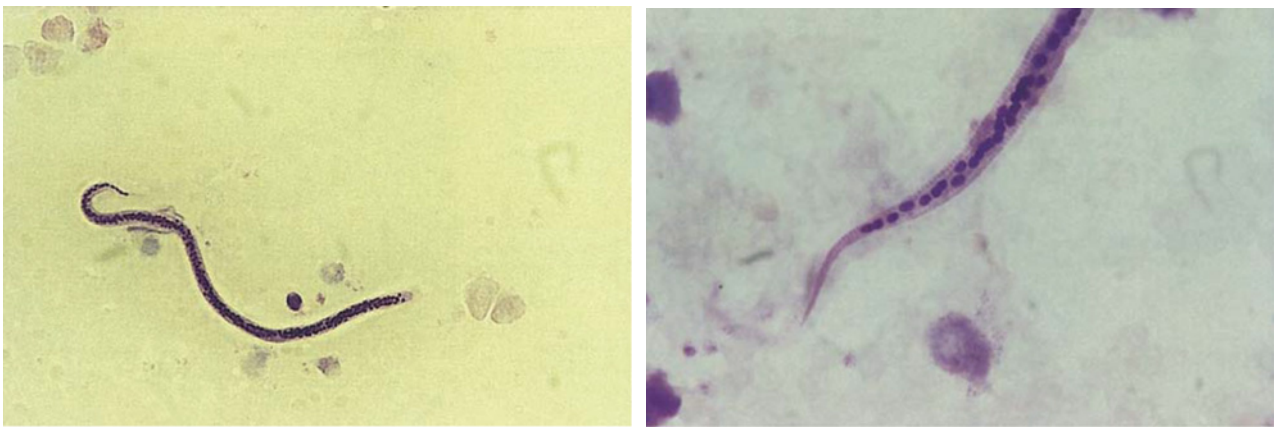
Oral diethylcarbamazine is the drug of choice in all forms of lymphatic filariasis, including subclinical infection. Alternative treatments include combinations of albendazole and ivermectin. None of these drugs has common, serious haematological side-effects. Depletion of *Wolbachia* endobacteria, a symbiont of *Onchocerca*, by tetracycline antibiotics leads to long-lasting sterility of adult female worms.

African trypanosomiasis (sleeping sickness)

Epidemiology and biology

African sleeping sickness is caused by the haemoflagellate protozoa *Trypanosoma brucei gambiense* in West and Central Africa, and *T. brucei rhodesiense* in eastern Africa (Figure 49.6). These parasites are fusiform in shape, 12–35 μm long and morphologically indistinguishable from each other. The disease is transmitted by the bite of the tsetse fly, which is only found in Africa. The trypanosomes multiply by fission in the vicinity of the infected bite and are then disseminated by the bloodstream. Congenital transmission has also been described.

The distribution of African sleeping sickness is determined by the ecological limits of the tsetse fly vector and lies in the region between Senegal and Somalia (Latitude 14° north) and the Kalahari and Namibian deserts (Latitude 20° south). Sleeping sickness is an important public health problem in Central African Republic, Chad, Democratic Republic of Congo, Côte d'Ivoire, Guinea, Malawi, Uganda and United Republic of Tanzania. During recent epidemic periods, the prevalence reached 50% in



(a) (b)

Figure 49.5 Microfilariae of *W. bancrofti* in thick film: (a) microfiliaria showing the negative impression of the sheath (× 365); (b) tail of the microfilaria showing that the nuclei do not extend into the tail (× 912).

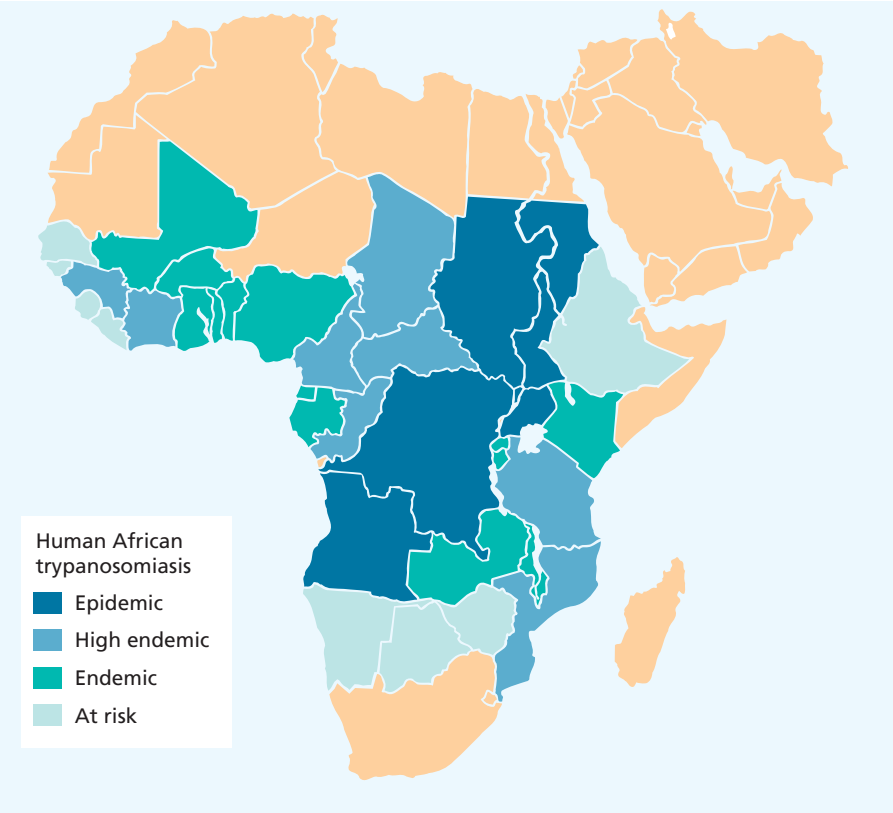


Figure 49.6 Global distribution of African trypanosomiasis.

several villages in the Democratic Republic of Congo, Angola and southern Sudan. Improved surveillance has consistently reduced the number of new cases over the last two decades. For the first time in 50 years, the number of reported cases dropped below 10,000 (9878) in 2009 and the decline has continued; 3796 cases were recorded in 2015, the lowest since data collection started 75 years ago.

Clinical features

The bite of a tsetse fly is very painful and causes a small indurated lesion that may persist for some days. The local multiplication of the trypanosomes may cause a marked inflammatory reaction (a chancre) that regresses after 2–3 weeks. Entry of the trypanosomes into the bloodstream is associated with fever, which tends to be less marked in West African trypanosomiasis than in the East African variety. East African trypanosomiasis is primarily a disease of cattle and only enters human hosts by accident. It is therefore less well tolerated than West African sleeping sickness, having a more aggressive course and intense symptoms.

The early stages of sleeping sickness can be associated with prominent lymphadenopathy, particularly of the posterior cervical nodes, and mild splenomegaly. These features may be suggestive of infectious mononucleosis, tuberculous lymphadenitis or a lymphoproliferative disorder. Severe anaemia, haemorrhages and petechiae may occur at this stage.

Both types of African sleeping sickness cause a protracted febrile illness which, despite the name, is not always associated with drowsiness. Death is inevitable if the disease is left untreated. As the disease progresses, parasitaemia decreases, trypanosomes invade the central nervous system (CNS) and neurological disturbances due to inflammatory chronic meningoencephalitis supervene. In West African trypanosomiasis, the disease runs its course over several years, but in East African trypanosomiasis infection CNS involvement may occur within weeks.

Haematological abnormalities

The aetiology of the anaemia in sleeping sickness is multifactorial, but primarily due to phagocytic removal of immune complex-coated red cells from the circulation. Trypanosomes liberate haemolytic factors that contribute to this process, and increases in plasma volume cause a dilutional anaemia. There is a failure to incorporate iron into red cell precursors and the resulting dyserythropoiesis means that the bone marrow is unable to compensate for the fall in haemoglobin. There may be a moderate leucocytosis with increased monocytes, lymphocytes and

plasma cells. Mott morular cells have also been described in sleeping sickness. The bone marrow is hypercellular, with fat cell atrophy, focal loss of haematopoietic cells and deposition of extracellular gelatinous substances, which histochemically are mucopolysaccharides, rich in hyaluronic acid.

As the disease advances, a bleeding tendency may develop due to thrombocytopenia, vascular injury and coagulopathy. Platelet dysfunction has also been described and is manifest as clumping and abnormal aggregation responses. DIC with raised FDPs may occur in the later stages. Although some of these haematological changes can be linked to the non-specific polyclonal activation of B cells, overall the underlying mechanisms are not well understood.

Diagnosis

Trypanosomes can be seen on stained thin blood films, but the number of trypanosomes in the circulation can vary considerably and is often low, so concentration techniques are usually required. Quantitative buffy coat method is the technique of choice for diagnosis of African sleeping sickness. This involves concentrating the trypanosomes at the plasma–platelet interface in a special microhaematocrit tube using differential centrifugation. Parasites are identified by labelling with the fluorescent marker acridine orange.

Wet preparations of fluid aspirated from the lymph nodes, bone marrow or cerebrospinal fluid (CSF) may also reveal live motile organisms. This technique is more likely to be productive in the case of infection with *T. brucei rhodesiense* than *T. brucei gambiense*. The organisms are fragile, so care must be taken not to damage them when making the smears. The highly specific and sensitive serological card agglutination test for trypanosomiasis (CATT) may be used in conjunction with a direct visualization method. If these tests are positive, then CSF examination is mandatory to determine the stage of the illness.

Haematological implications of treatment

Pentamidine and suramin are the drugs of choice for the early stages of West and East African trypanosomiasis, respectively. They have a cure rate of around 90%, but are only able to achieve modest CSF concentrations so they cannot be used for later stages of the disease. The most common haematological side-effects of pentamidine are leucopenia, thrombocytopenia and anaemia. Suramin has serious side-effects, including haemolytic anaemia and bone marrow toxicity.

Melarsoprol, an arsenic-based compound, has been the drug of choice for late-stage sleeping sickness, but is highly toxic, with a mortality of 4–12%. Its main adverse effect is a fatal encephalopathic syndrome; haematological toxicity is not a problem. Eflornithine is expensive, but is of benefit in late-stage, particularly West African, disease; 25–50% of patients treated with this drug exhibit bone marrow toxicity with pancytopenia.

American trypanosomiasis (Chagas disease)

Epidemiology and biology

Chagas disease is caused by *T. cruzi*, which is transmitted by triatomine bugs that infest poor-quality housing. It can also be transmitted through blood transfusions and congenitally. It was restricted to a region in the Americas between Argentina and the southern states of the USA; however, because of increasing population mobility between Latin America and the rest of the world it is now detected in other parts of the United States of America, Canada, many European and some Western Pacific countries.

Clinical features

The incubation period is usually 7-14 days but may exceed 40 days if transmission was through blood transfusion. In the acute phase, swelling at the site of entry of the organism (chagoma), may be accompanied by fever, hepatosplenomegaly and lymphadenopathy. The trypanosomes multiply intracellularly in muscle tissue, particularly the heart, colon and oesophagus. Once infection has occurred, the organisms will be present for life unless treatment is given. The chronic phase of the disease is associated with heart disease in 30% of cases, manifesting as arrhythmias and cardiomegaly. A small proportion of patients have involvement of the gastrointestinal tract and other hollow organs, resulting in loss of peristalsis, organomegaly and organ failure. Asymptomatic infection is common and poses a problem for blood transfusion services, so some countries routinely screen blood for American trypanosomiasis. Many cases do not become symptomatic until the chronic stage, which can occur 5 to 40 years after infection.

Diagnosis

Microscopy of blood smears can be helpful, but only in the acute phase of infection when parasites are circulating in blood. Sensitive serological tests based on enzyme immunoassay or immunofluorescence are therefore more commonly used as the primary diagnostic tool. Diagnosis is generally made based on at least two different serologic tests. PCR may also be useful.

Haematological implications of treatment

Treatment of *T. cruzi* infection is with nifurtimox or benznidazole and if started early, the disease is curable. Major haematological side-effects are uncommon with either drug, although agranulocytosis has been reported with benznidazole.

Leishmaniasis

Epidemiology and biology

Visceral and cutaneous leishmaniases are caused by protozoan flagellates that are transmitted through the infective bite of a phlebotomine sandfly. Leishmaniasis is second only to malaria as the most common parasitic cause of mortality. Following an infected bite, parasites spread from the inoculation site to

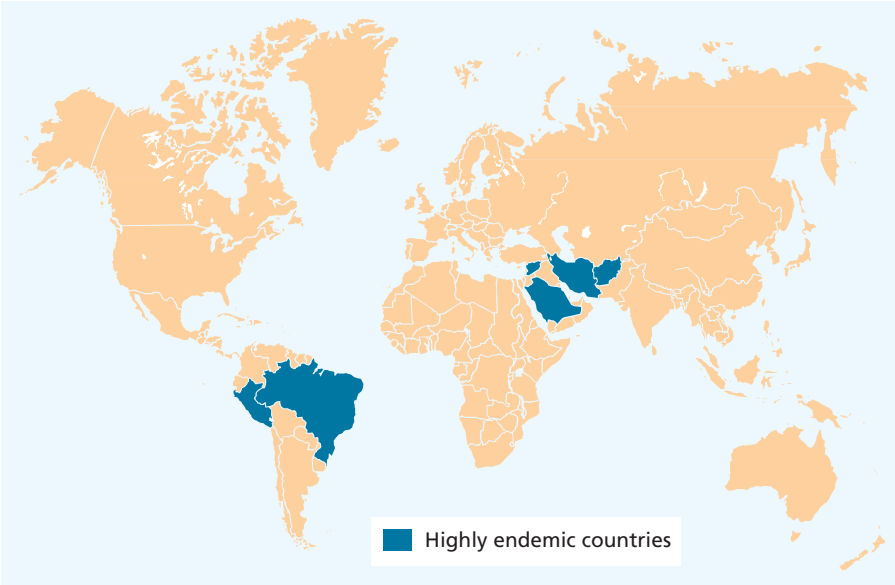


Figure 49.7 Global distribution of visceral leishmaniasis.

the mononuclear phagocytic system. Only the visceral form (kala-azar) is associated with organisms in haemopoietic tissues and thus is considered here. Visceral leishmaniasis is due to the species *Leishmania donovani* and *L. infantum*. In 2013, it was endemic in 75 countries throughout the world, with extension limits from 45° N to 32° S. Over 90% of cases occur in Bangladesh, Brazil, Ethiopia, India, South Sudan and Sudan (Figure 49.7). A concomitant HIV infection increases the risk of developing active visceral leishmaniasis by between 100 and 2320 times. In southern Europe, up to 70% of cases of visceral leishmaniasis in adults are associated with HIV infection.

Clinical features

This depends on both the genotypic potential of the parasite and the immunological response of the patient. Incubation period varies from days to years but is generally 2–6 months. Onset can be sudden with high fever, or gradual with intermittent fever. Diarrhoea, joint pain, weight loss and bleeding gums occur in the acute phase. This is followed by progressive muscle wasting, protuberant abdomen, fever, weight loss, anaemia and hepatosplenomegaly. The splenomegaly appears early and worsens with the duration of the disease, so that eventually it may reach into the left hypochondrium. In immunocompromised patients, such as transplant recipients and those with advanced HIV disease, kala-azar behaves like an opportunistic infection.

Haematological abnormalities

Normochromic normocytic anaemia is a frequent feature, and haemoglobin levels of 70–100 g/L are common. The massive splenomegaly results in hypersplenism with consequent pancytopenia. Liver dysfunction with jaundice, ascites and deranged coagulation may occur in the late stages and has a poor prognosis. The bleeding tendency may be exacerbated by thrombocytopenia. In all patients with unexplained splenomegaly, pancytopenia or fever, a high index of suspicion of leishmaniasis needs to be maintained to prevent fatalities.

Diagnosis

Definitive diagnosis is based on detection of the parasites, or their DNA, in smears of bone marrow, splenic aspirate or fluid aspirated from enlarged lymph nodes.

On microscopy *Leishmania* are usually seen as intracellular amastigotes in mononuclear cells, but can also be seen extracellularly (Figure 49.8). They are 2–6 μm in diameter and

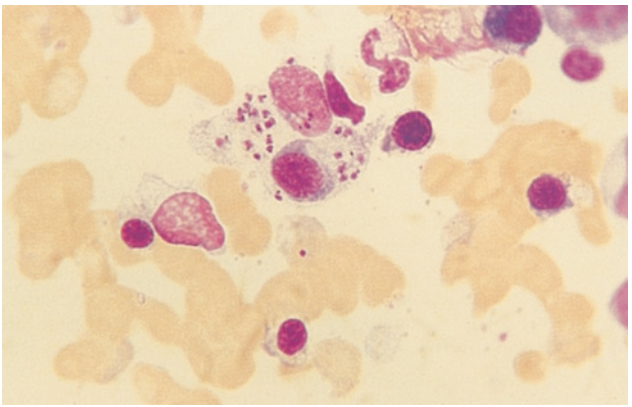


Figure 49.8 Bone marrow aspirate of leishmaniasis infection showing a macrophage containing numerous organisms. The presence of both a nucleus and a small paranuclear kinetoplast gives the organisms their characteristic 'double-dot' appearance (MCG × 940).

contain a nucleus lying close to the rod-shaped kinetoplast, and an internal flagellum. Using a Romanowsky stain, the nucleus and kinetoplast stain purple and can be clearly distinguished. Amastigotes can be seen in both bone marrow aspirates and in trephine-impression smears. They are rarely seen in peripheral blood and then only in buffy coat preparations.

Microscopy is less sensitive than molecular diagnosis, particularly when there is coinfection with HIV. PCR can be performed on lesion aspirate, marrow, blood and biopsy material. The indirect fluorescent antibody tests ELISA and DAT are useful for detecting antibodies to visceral leishmaniasis, but results may be inconclusive in immunosuppressed patients.

Haematological implications of treatment

For decades, the treatment of leishmaniasis has been based on pentavalent antimonials, although their mode of action is still unclear. Sodium stibogluconate is the most commonly used. It can be associated with worsening anaemia and thrombocytopenia, although its most detrimental effects are on cardiac function. Resistance levels to antimonials are high in countries like India, Bangladesh, Brazil and Sudan, so other options including amphotericin, paromomycin and miltefosine need to be considered. HIV-coinfected patients do not respond well to antimonials, so for these individuals amphotericin is the drug of choice. Amphotericin commonly causes normocytic normochromic anaemia, but its most serious toxicity is related to renal, cardiac, neurological and hepatic dysfunction. Side-effects can be reduced by using the liposomal preparation. Immunocompromised patients may require prophylaxis to prevent relapses.

Tropical diseases associated with changes in FBC and / or coagulation

Hookworm infection

Hookworms (*Ancylostoma duodenale*) and *Necator americanus*) are common in the tropics and subtropics and affect 600 million people worldwide. They cause significant iron deficiency anaemia in all age groups, including 30–50% of anaemia in pregnancy. The hookworm eggs are generally transmitted between humans through contact with faeces. Most infections are asymptomatic, but chronic blood loss through the gut eventually leads to severe iron deficiency anaemia, and even cardiac failure. Other symptoms include itch, rash and abdominal discomfort. Eosinophilia is common and characteristic eggs can be seen on faecal microscopy. Albendazole or mebendazole are the usual treatments and iron deficiency should also be treated. Transfusion is rarely necessary, and may be harmful since the anaemia develops very slowly and cardiac function may be compromised in severe anaemia.

Schistosomiases (Bilharzia)

Schistosomiasis is a common intravascular infection caused by trematode worms acquired through contact with contaminated water. It occurs across Africa, Asia and South America. *Schistosoma. mansoni, intercalatum, japonica* and *mekongi* all cause intestinal schistosomiasis, whilst *S. haematobium* causes urinary schistosomiasis. There is an acute stage (Katayama syndrome), and chronic and advanced stages. Schistosomiasis is the second most common cause of iron deficiency worldwide after hookworm infection. It is due to blood loss from the urinary and gastrointestinal tracts and may be severe. Eosinophilia occurs in 80% of cases. Anaemia of chronic disease, massive splenomegaly with hypersplenism, hepatomegaly and generalized lymphadenopathy can all occur. Iron deficiency should be treated and schistosomiasis treatment is with praziquantel, which has no significant haematological complications.

Viral haemorrhagic fevers

These are caused by arenaviruses, filoviruses, bunyaviruses and flaviviruses and are classified according to their reservoir hosts and their primary means of transmission. They are divided into:

- rodent-associated viruses (e.g. Lassa fever, hantaviruses);
- arthropod-borne viruses (e.g. dengue, yellow fever and Chikungunya viruses);
- unknown vectors or hosts (e.g. Marburg, Ebola).

They often occur in epidemics, have human-to-human transmission and, in non-tropical settings, may only be suspected if a relevant travel history is elicited from the patient. Dengue and yellow fever are becoming increasingly important imported infections.

Dengue

The four types of dengue virus belong to a group of flaviviruses and are transmitted by *Aedes* mosquitoes. Dengue is endemic in tropical areas of Asia, Africa, South America and the Caribbean and is particularly virulent in Southeast Asia, including Thailand and Vietnam. There has been a resurgence of the disease as a result of urbanization, poverty and increasing travel.

Primary infection in young children is usually asymptomatic. Older children and adults develop acute fever, headache and myalgia ('breakbone fever'). Leucopenia may accompany this stage of the illness. Severe complications may arise in those who have had previous infection. These include hypotensive shock, bone marrow hypoplasia with neutropenia and abnormal megakaryopoiesis, leading to severe thrombocytopenia and spontaneous bleeding.

Yellow fever

Yellow fever virus is transmitted by *Aedes* mosquitoes and exists throughout equatorial Africa, northern and central southern America. Ninety percent of cases and deaths each year occur in

Africa, and the number of cases has been increasing over the past two decades possibly due to a decline in population immunity.

The virus invades hepatocytes, causing hepatocellular dysfunction. After an incubation period of 3–4 days, fever, myalgia and back pain may be followed by jaundice, bleeding and, in the most severe cases, renal failure. Haematological changes include leucopenia with relative neutropenia, thrombocytopenia as part of a consumptive coagulopathy, initial haemoconcentration, and subsequent haemorrhage and haemodilution. Coagulation abnormalities include reduced fibrinogen and clotting factors II, V, VII, VIII, IX and X, with increased PT and APTT, as well as the presence of fibrin degradation products. In 85% of cases there is spontaneous improvement, with the disappearance of symptoms in 3–4 days. The remaining 15% enter a toxic phase which has a 50% mortality.

Diagnosis can be difficult, as yellow fever (preventable by vaccination) can be confused with several febrile and haemorrhagic illnesses. Antibodies can be detected in the serum and treatment is symptomatic.

Lassa fever, Ebola virus and Marburg virus

These are endemic in equatorial Africa and are important because they cause potentially fatal infections and have the ability to spread from person to person. Only about 10% of infected individuals become ill. Of these 1–2% develop fatal disease. The clinical features of these three haemorrhagic fevers are similar and are characterized by headache, fever and oesophagitis. Spontaneous bleeding occurs in 25% of hospitalized patients and is thought to be due to increased vessel permeability and abnormal platelet function. Death is due to multi-organ failure and shock. Case reports suggest that treatment with ribavirin may be helpful and a promising vaccine against Ebola virus has recently been tested in clinical trials in west Africa.

Non-specific haematological abnormalities associated with tropical diseases

Anaemia

Anaemia of chronic disease is a common and non-specific finding in many types of tropical infections. Some infections are responsible for specific types of anaemia. For example, hookworm, which is common in farmers who work bare-foot on infected soil, and schistosoma infections are the two most common causes of iron deficiency anaemia worldwide. For both conditions, a good clinical history and a thorough examination of a stool (hookworm and schistosomiasis) or urine sample (schistosomiasis) is important if the diagnosis is not to be missed. Secondary megaloblastic anaemia can result from consumption of vitamin B_{12} in infections with the tapeworm *Diphyllobothrium latum*. Intraerythrocytic parasites, such as those that cause malaria and babesiosis, may be associated with significant intravascular haemolysis.

White cell abnormalities

Severe infections, particularly those due to bacteria, may cause a neutrophilia and a leukaemoid reaction with circulating myeloid precursors and neutrophils with toxic granulation, vacuolation and Döhle body formation. Lymphocytosis with neutropenia, splenomegaly, nose bleeds, rash and neurological complications can occur in rickettsial infections (e.g. typhus, Q fever, trench fever). Some, such as trench fever, may persist for many years and are transmissible in blood transfusions. The absence of neutrophilia in the presence of marked fever is a particular feature of typhoid. Lymphocytosis or monocytosis may also be present in typhoid; other clinical features include hepatosplenomegaly and, in severe disease, haemorrhage from ileal ulcers exacerbated by DIC. Helminths and other predominantly tropical organisms that invade tissues may be associated with a significant eosinophilia ($>0.5 \times 10^9$/L). Such diseases include loaiasis, lymphatic filariasis, schistosomiasis, trichinosis, toxocariasis, strongyloidiasis, hydatid disease, oriental liver flukes and guinea worm.

Platelet abnormalities

Thrombocytosis may occur as a non-specific change associated with inflammatory conditions, including severe infections. Thrombocytopenia is usually secondary to, for example, hypersplenism (e.g. portal hypertension associated with schistosomiasis, hyper-reactive malarial splenomegaly or disseminated intravascular coagulation). Thrombocytopenia can occur in malaria, typically due to *P. falciparum*, either as a consequence of direct invasion of platelets by the parasite or by secondary mechanisms.

Hypersplenism

Hypersplenism is a clinical syndrome characterized by splenomegaly, with a reduction in the number of one or more types of blood cells. There is reduced survival of platelets and red cells and a normocellular or hypercellular bone marrow. Hypersplenism can result from any cause of splenomegaly, but it is most commonly associated with splenomegaly secondary to portal hypertension and haematological disorders. Tropical infections associated with massive splenomegaly (Figure 49.9) include hyper-reactive malarial splenomegaly (formerly tropical splenomegaly syndrome), visceral leishmaniasis, schistosomiasis and trypanosomiasis. In the tropics, portal hypertension associated with schistosomiasis is probably the commonest cause of hypersplenism.

The cytopenias in hypersplenism are due to a combination of sequestration and haemodilution and the degree of plasma

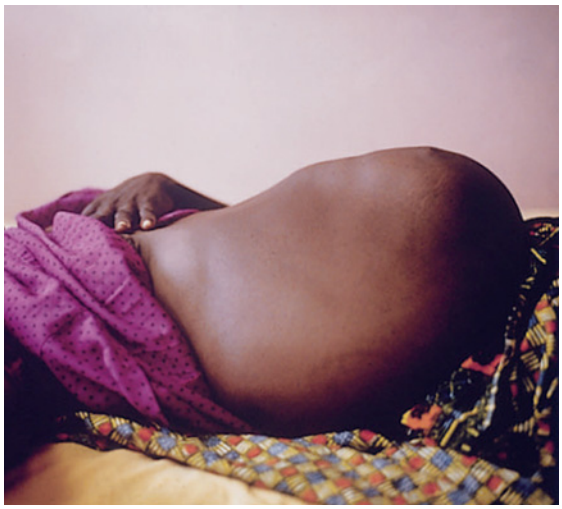


Figure 49.9 Patient with massive splenomegaly due to hyper-reactive malarial splenomegaly.

volume expansion is roughly proportional to the size of the spleen. The thrombocytopenia and neutropenia are rarely severe enough to cause clinical problems. The thrombocytopenia is primarily due to pooling in the spleen which, when massively enlarged, can hold up to 90% of the platelet mass. Neutropenia is the result of increased marginalization of granulocytes. Treatment of the underlying disorder generally leads to regression of the splenomegaly, with resolution of the haematological abnormalities. Occasionally, splenectomy is necessary, but this should only be undertaken after a careful assessment of the risks and benefits, and adherence to recommendations (e.g. vaccinations, antibiotic prophylaxis) regarding infection prevention.

Selected bibliography

Bain BJ (2015) *Blood Cells: A Practical Guide*, 5th edn. Wiley-Blackwell, Oxford.

Calis JC, Phiri KS, Faragher EB et al. (2008) Severe anaemia in Malawian children. *New England Journal of Medicine* **358**: 888–99.

Cook GC, Zumla A (eds) (2014) *Manson's Tropical Diseases*, 23nd edn. Saunders/Elsevier, Edinburgh. http://www.malariajournal.com/content/pdf/1475-2875-8-73.pdf

Migration: A world on the move. April 2013. Commission on Population and Development (a UNFPA report).

Migration Statistics Quarterly Report, February 2015. Office for National Statistics.

Antinori S, Galimberti L, Milazzo L, Corbellino M (2013) *Plasmodium knowlesi*: The emerging zoonotic malaria parasite. *Acta Tropica* **125**: 191– 201.

Trypanosomiasis, Human African sleeping sickness. Fact sheet N°259. Updated March 2014.

CHAPTER 50

Neonatal haematology

50

Irene Roberts[1] and Subarna Chakravorty[2]

[1]Department of Paediatrics and Weatherall Institute of Molecular Medicine, University of Oxford, Oxford, UK,
[2]Department of Paediatrics, Imperial College Healthcare and Imperial College London, London, UK

Developmental haemopoiesis

In humans, haemopoiesis begins in the yolk sac in the third week of gestation and moves sequentially to the aorto–gonado–mesonephros (AGM) by 5 weeks' gestation, the liver by 6–8 weeks and the bone marrow around the 11th week of gestation. The AGM involutes early in the first trimester and the liver is the principal site of haemopoiesis until the end of the third trimester. Fetal haemopoietic stem cells (HSCs) have different molecular and functional characteristics compared to adult HSCs. These differences ensure that fetal HSCs meet the specific needs of the fetus for different types of mature blood cells. The predominant lineage during fetal life is erythropoiesis, but platelets and all types of leucocyte found in adult blood are also seen from as early as 4–5 weeks' gestation.

There are a number of differences between erythropoiesis in neonates and adults: red cell morphology is distinctive with large numbers of crenated red cells, particularly in preterm neonates (Figure 50.1); red cell lifespan is reduced (35–50 days in preterm infants, 60–70 days in term infants); susceptibility to oxidant-induced injury is increased because of differences in the glycolytic and pentose phosphate pathways; the erythropoietin response to anaemia is blunted and specific embryonic and fetal globin chains are synthesized (Table 50.1). The first globin chain produced is ε-globin, followed by α- and γ-globin chains. Haemoglobin F (HbF, $\alpha_2\gamma_2$) is produced from 4–5 weeks' gestation and is the predominant haemoglobin until after birth. Adult haemoglobin (HbA, $\alpha_2\beta_2$) is produced from 6–8 weeks' gestation, but remains at low levels until after birth. In term babies, the average HbF at birth is 70–80%, the HbA is 25–30%, there are small amounts of HbA_2 and sometimes a trace of Hb Barts (γ_4).

Immediately after birth, rates of haemoglobin synthesis and red cell production fall in response to the sudden increase in tissue oxygenation at birth. In term babies, the haemoglobin reaches a mean of 130–140 g/L at the age of 4 weeks and 95–110 g/L at 7–9 weeks of age. Studies of well-preterm infants show a steeper fall in haemoglobin, reaching a mean of 65–90 g/L at 4–8 weeks of age. The reticulocyte count falls soon after birth as erythropoiesis is suppressed and increases back to normal values at 6–8 weeks of age. The blood volume at birth varies with gestational age and the timing of clamping of the cord. In term infants, the average blood volume is 80 mL/kg and in preterm infants 106 mL/kg (range 85–143 mL/kg). Term and preterm babies have adequate stores of iron, folic acid and vitamin B_{12} at birth. However, stores of both iron and folic acid are lower in preterm infants and are depleted more quickly, leading to deficiency after 2–4 months if the recommended daily intakes are not maintained.

Neonatal anaemia

Definition and pathophysiology

Any neonate with a haemoglobin of less than 130 g/L at birth should be considered anaemic and may require investigation (Figure 50.2). However, it is important to be aware that the haemoglobin concentration is affected by the site of sampling (it is up to 40 g/L lower in venous samples than in heel-prick samples in the first few days of life) and the timing of the clamping

Postgraduate Haematology, Seventh Edition. Edited by A Victor Hoffbrand, Douglas R Higgs, David M Keeling and Atul B Mehta.

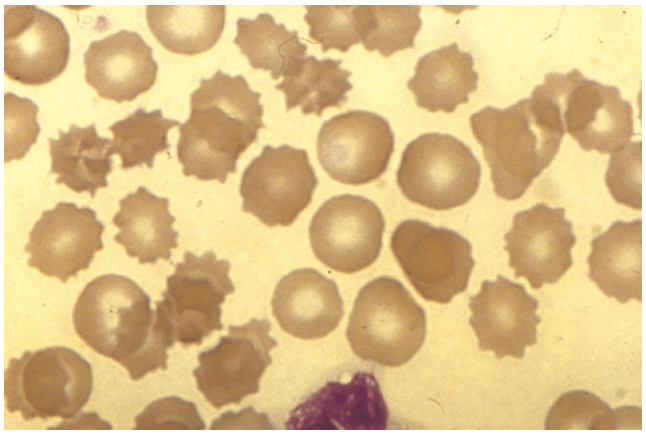


Figure 50.1 Typical erythrocyte morphology in a preterm neonate. Crenated red cells are a normal feature of the blood film of preterm neonates during the first few weeks of life. This film is from a neonate born at 26 weeks' gestation and shows the number of crenated cells present in neonates under 28 weeks' gestation. The numbers of these cells is inversely proportional to gestational age at birth.

of the cord (around 30 g/L higher after late clamping). The clinical significance of neonatal anaemia depends on whether the baby is able to maintain adequate tissue oxygenation. This in turn depends on the position of the haemoglobin–oxygen dissociation curve, which is principally determined by the concentrations of HbF and 2,3-diphosphoglycerate (2,3-DPG); a high HbF and low 2,3-DPG both shift the curve to the left, i.e. the affinity of haemoglobin for oxygen is increased and less oxygen is released to the tissues. This is the situation just after birth and may be more of a problem for very preterm babies as their HbF levels are greater than 90%. Over the first few months of life, 2,3-DPG levels rise and HbF levels fall so that the haemoglobin–oxygen dissociation curve gradually shifts to the right, the oxygen affinity of haemoglobin falls and oxygen delivery to the tissues increases, ameliorating the effects of the falling haemoglobin.

Table 50.1 Composition of haemoglobins in the human embryo, fetus and neonate (see also Chapter 6).

	Globin chains		
Haemoglobin	**α-Globin gene cluster***	**β-Globin gene cluster***	**Gestation**
Embryonic			
Hb Gower 1	ξ_2	ε_2	From 3–4 weeks
Hb Gower 2	α_2	ε_2	
Hb Portland	ξ_2	γ_2	From 4 weeks
Fetal			
HbF	α_2	γ_2	From 4 weeks
Adult			
HbA	α_2	β_2	From 6–8 weeks
HbA_2	α_2	δ_2	From 30 weeks

*The α-globin gene cluster is situated on chromosome 16 and the β-globin gene cluster on chromosome 11. Note that fetuses and neonates with α-thalassaemia major, who are unable to synthesize α-globin chains, will have Hb Portland as well as Hb Barts (β_4), detectable by haemoglobin electrophoresis or HPLC.

Causes of neonatal anaemia

Anaemia may be caused by reduced red cell production, increased red cell destruction (haemolysis) or blood loss (Table 50.2).

Neonatal anaemia due to reduced red cell production

The main diagnostic clues to reduced red cell production are the combination of a low reticulocyte count ($<20 \times 10^9$/L) with a negative direct antiglobulin test (Coombs test). The most common causes are congenital parvovirus B19 infection and genetic red cell aplasias, particularly Diamond–Blackfan anaemia (DBA).

Parvovirus B19 and fetal/neonatal anaemia

Maternal infection with parvovirus B19 can cause severe fetal anaemia and in 9% of cases leads to intrauterine death. The baby has marked reticulocytopenia (often $<10 \times 10^9$/L) and thrombocytopenia may also occur. The diagnosis is made by maternal serology and demonstration of B19 DNA in the fetus or neonate by dot-blot hybridization or polymerase chain reaction (PCR) of peripheral blood (bone marrow aspiration for morphology and parvovirus B19 PCR may be necessary in difficult cases). Severe cases require intrauterine transfusion, but have a good long-term outcome if they survive to delivery.

Genetic red cell aplasia

Apart from DBA, the genetic causes of congenital red cell aplasia are extremely rare. They include congenital dyserythropoietic anaemia (CDA) and Pearson syndrome; haematological evidence of the other inherited bone marrow failure syndromes, such as Fanconi anaemia, are almost never present at birth. These genetic causes of congenital red cell aplasia can normally be distinguished from each other through a combination of their distinctive bone marrow morphology, family history and the presence, or not, of additional congenital anomalies.

DBA (see also Chapter 10)

The incidence of DBA is estimated at five to seven cases per million live births. DBA has a clear family history in 20% of cases (autosomal dominant or recessive) and appears to be sporadic in the remaining 80%. Presentation in the neonatal period

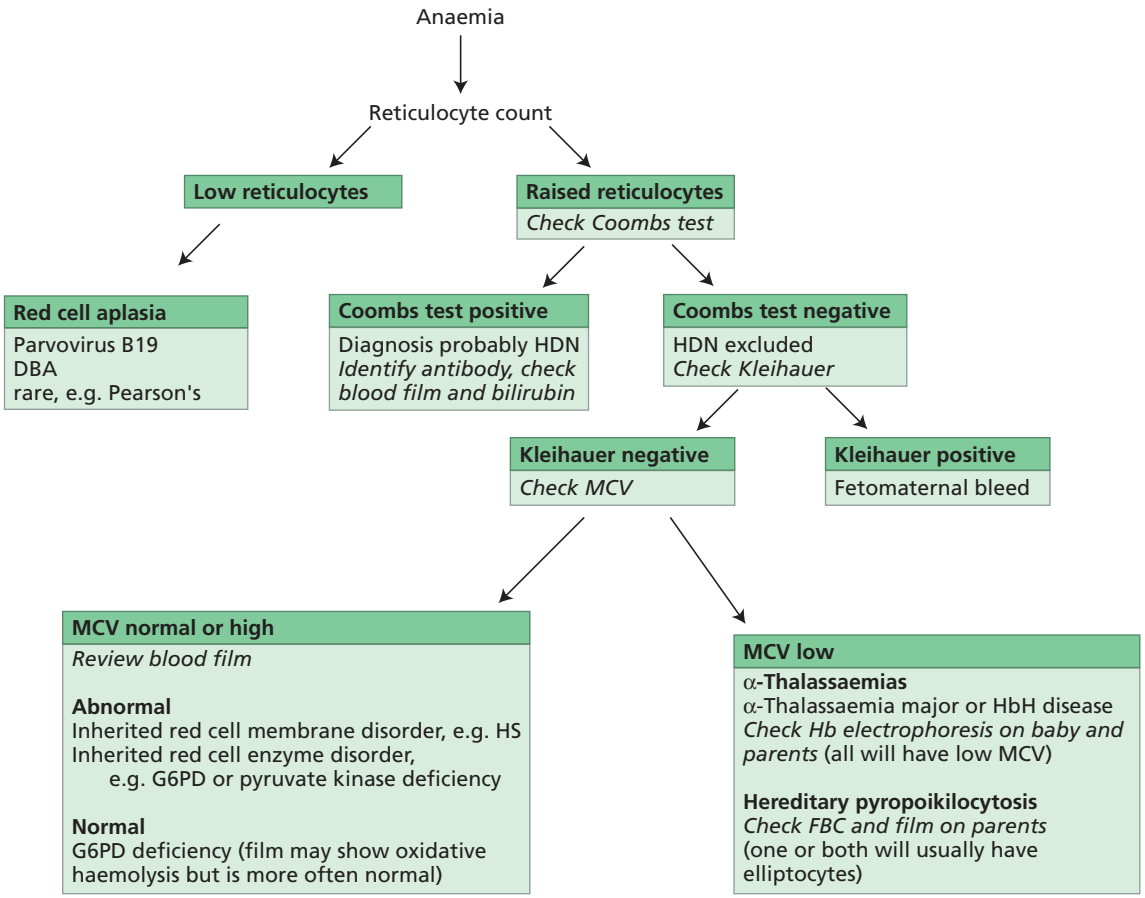


Figure 50.2 A diagnostic algorithm for neonatal anaemia. The most useful screening tests for investigating unexplained neonatal anaemia are the reticulocyte count, the Coombs test and the mean corpuscular volume (MCV) of the red cells. DBA, Diamond–Blackfan anaemia; G6PD, glucose-6-phosphate dehydrogenase; HDN, haemolytic disease of the newborn; HS, hereditary spherocytosis.

or first few months of life with increasing anaemia from birth is typical, but more severe cases manifest as second-trimester anaemia or hydrops fetalis. Around 40% of infants have associated congenital anomalies, particularly craniofacial dysmorphism, neck anomalies and thumb malformations similar to those seen in Fanconi anaemia. In neonates, the blood film shows normochromic red cells with an absence of polychromasia and nucleated red cells despite severe anaemia (Figure 50.3). Reticulocytopenia is usually severe, but automated reticulocyte counts of 20–30 $\times$ 10^9/L are sometimes seen. Erythroid precursors are very low or absent in the bone marrow aspirate. These features are diagnostic of DBA if parvovirus infection is excluded. Raised red cell levels of adenosine deaminase (ADA) in the patient and/or parents may be useful for confirming the diagnosis, although normal red cell ADA levels do not exclude DBA. As discussed in Chapter 10, recent studies indicate that most cases of DBA are due to defects in the synthesis and function of structural ribosomal proteins (RP). At present mutations in 12 genes have been identified and account for about 50% of DBA cases, *RPS19* mutations being the commonest (25% of all DBA). In just under 50% of patients, the molecular basis for the disease has proved elusive, even after screening all 80 ribosomal protein genes. However, recently mutations in the key haemopoietic transcription factor gene *GATA1* have been identified in a small proportion of DBA patients. In the neonatal period, the only treatment of DBA is red cell transfusion, although steroids are used in older infants and children (see Chapter 10).

Pearson syndrome (see also Chapter 3)

This rare condition nearly always presents within a few days or weeks of birth. Pearson syndrome is caused by mutations or deletions in mitochondrial DNA and presents with normochromic anaemia, neutropenia, thrombocytopenia and failure to thrive. Most affected neonates have low birthweight, and metabolic acidosis, exocrine pancreatic deficiency, and abnormal liver and renal function are common. The most useful haematological clue to the diagnosis is the highly characteristic vacuolation of early erythroid cells on the marrow aspirate (Figure 50.4). Interestingly, a recent report identified large

Table 50.2 Common of neonatal anaemia.

Reduced red cell production
• Diamond–Blackfan anaemia
• Congenital viral infections, e.g. parvovirus, cytomegalovirus
• Congenital dyserythropoietic anaemia
• Pearson syndrome
Increased red cell destruction (haemolysis)
• Alloimmune: haemolytic disease of the newborn (Rh, ABO, Kell, other)
• Red cell membrane disorders, e.g. hereditary spherocytosis
• Red cell enzyme deficiencies, e.g. pyruvate kinase deficiency
• Some haemoglobinopathies, e.g. α-thalassaemia major, HbH disease
Blood loss
• Occult haemorrhage before or around birth, e.g. twin–twin, fetomaternal
• Internal haemorrhage, e.g. intracranial, cephalhaematoma
• Iatrogenic, due to frequent blood sampling
Anaemia of prematurity
• Impaired red cell production plus reduced red cell lifespan

mitochondrial gene deletions in a small subgroup of DBA patients. When these cases were re-evaluated they were found to be cases of Pearson syndrome which had been misdiagnosed as DBA. Unfortunately, the prognosis for children with Pearson syndrome is very poor, with few surviving beyond the second year of life.

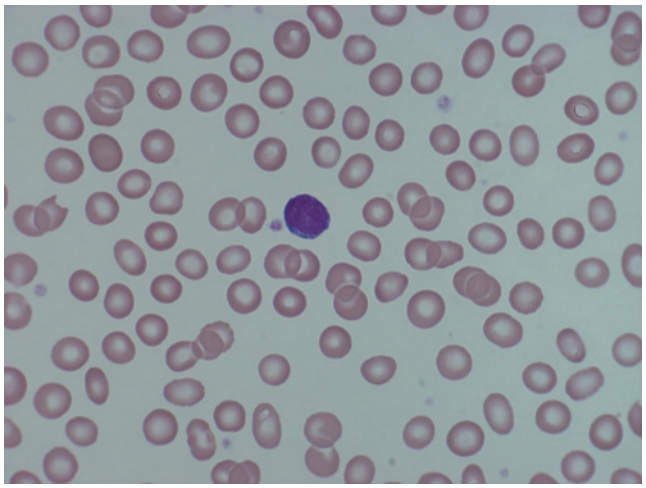


Figure 50.3 Blood film from a neonate with Diamond–Blackfan anaemia. This baby presented with fetal anaemia at 20 weeks' gestation and received intrauterine transfusion. At birth, the baby had normochromic anaemia and the blood film showed a complete absence of polychromasia and nucleated red cells, despite a haemoglobin of 70 g/L.

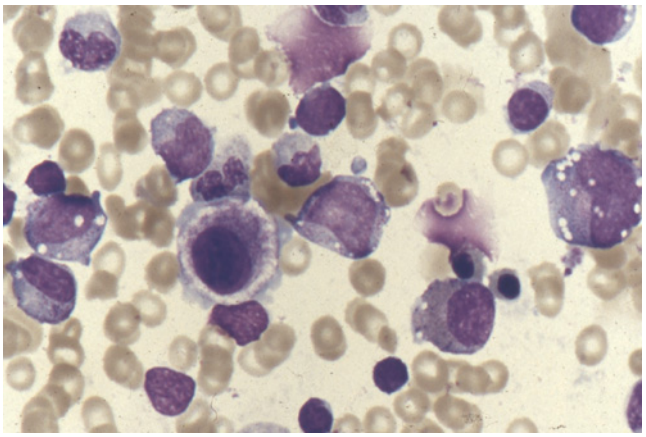


Figure 50.4 Pearson syndrome. Bone marrow aspirate from a neonate with Pearson syndrome showing typical vacuolation of erythroblasts and a dysplastic megakaryocyte.

CDA (see also Chapter 10)

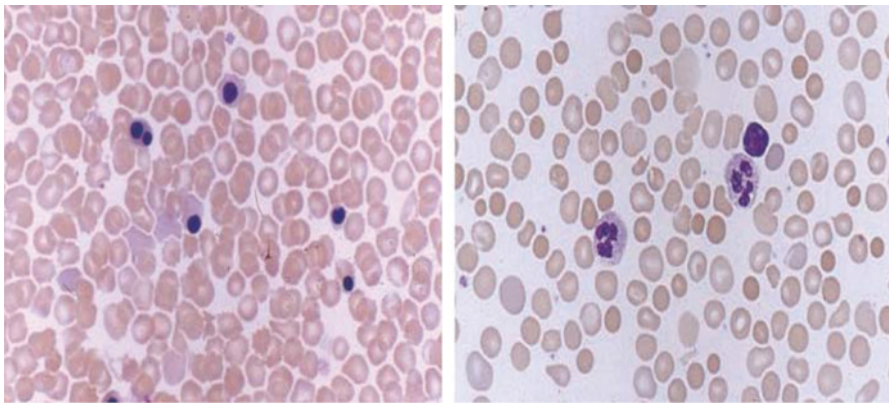
The clinical presentation of CDA is extremely variable. While the most severe forms may present as hydrops fetalis in association with fetal anaemia, milder forms of the disease may not present until adulthood. As discussed in Chapter 10, CDA is conventionally classified into three main types (I, II and III) with additional variant types of CDA. Most case reports of CDA presenting in the neonatal period are Type I or CDA variants. CDA type I is caused by mutations in the *CDAN-1* gene or, occasionally, in the *C15ORF41* gene. Type II CDA is due to mutations in the *SEC23B* gene and CDA type III, which is very rare, is due to mutations in the *KIF23* gene, while some variant forms of CDA are due to mutations in *LPIN2* or in the transcription factor genes *GATA1* or *KLF1*.

Neonatal anaemia due to increased red cell destruction (haemolysis)

The main diagnostic clues suggesting a haemolytic anaemia are increased numbers of reticulocytes and/or circulating nucleated red cells (Figure 50.5), unconjugated hyperbilirubinaemia, a positive Coombs test (if immune) and characteristic changes in the morphology of the red cells on a blood film (e.g. hereditary spherocytosis; Figure 50.5). The main cause of immune haemolytic anaemia is haemolytic disease of the newborn. The main causes of non-immune neonatal haemolysis are red cell membrane disorders, red cell enzymopathies and, occasionally, haemoglobinopathies (Table 50.2).

Immune haemolysis, including haemolytic disease of the newborn (see also Chapter 13)

The principal alloantibodies causing haemolytic disease of the newborn are those against Rh antigens (anti-D, anti-c and anti-E), anti-Kell, anti-Kidd (J^k), anti-Duffy (F^y) and antibodies



(a) (b)

Figure 50.5 Haemolytic disease of the newborn. (a) Blood film from a baby with rhesus haemolytic disease of the newborn due to anti-D showing polychromasia and large numbers of nucleated red cells, but relatively few spherocytes. (b) Blood film from a neonate with ABO haemolytic disease of the newborn due to anti-A, showing very large numbers of spherocytes, polychromasia and no nucleated red cells.

of the MNS blood group system, including anti-U. Anti-D is the most frequent alloantibody causing significant haemolytic anaemia, affecting 1 in 1200 pregnancies. Anti-Kell antibodies are less common, but can cause severe fetal and neonatal anaemia as they inhibit erythropoiesis as well as causing haemolysis. Most babies with haemolytic disease of the newborn present with jaundice and/or anaemia; in severe cases there is hepatosplenomegaly and/or skin deposits due to extramedullary haemopoiesis. ABO haemolytic disease occurs only in offspring of women of blood group O and is confined to the 1% of such women who have high-titre IgG antibodies. Haemolysis due to anti-A is more common (1 in 150 births) than that due to anti-B. In contrast with anti-Rh antibodies, both anti-A and anti-B usually cause hyperbilirubinaemia without significant anaemia; however, hydrops has occasionally been described. The blood film in ABO haemolytic disease characteristically shows very large numbers of spherocytes with little or no increase in nucleated red cells (see Figure 50.5a); this contrasts with rhesus haemolytic disease of the newborn, in which there are few spherocytes and large numbers of circulating nucleated red cells (see Figure 50.5b).

Management of haemolytic disease of the newborn: All neonates at risk should have cord blood taken for measurement of haemoglobin, bilirubin and a Coombs test and should remain in hospital, or under close outpatient supervision, until hyperbilirubinaemia and/or anaemia have been properly managed. Rhesus-alloimmunized infants with haemolysis should receive phototherapy from birth, as the bilirubin can rise steeply; this prevents the need for exchange transfusion in some infants. In haemolytic disease due to anti-Kell, anaemia is usually more prominent than jaundice and minimal phototherapy may be necessary despite severe anaemia. ABO haemolytic disease of the newborn usually just requires phototherapy, as anaemia is uncommon.

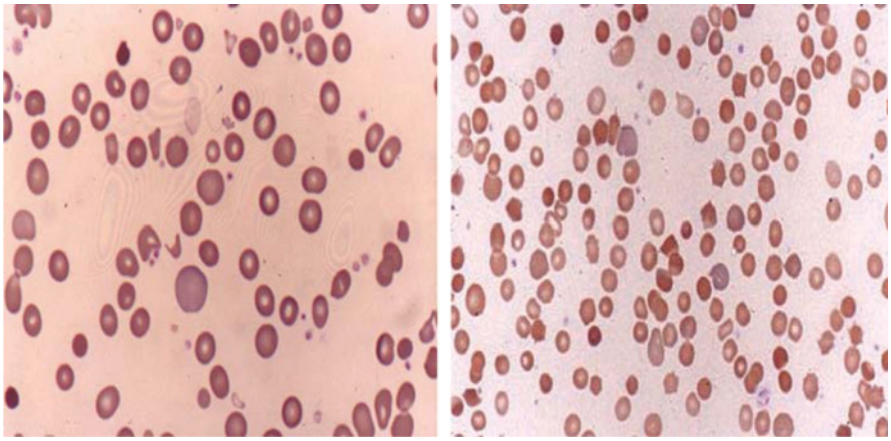
The indications for exchange transfusion in haemolytic disease of the newborn are:

- severe anaemia (haemoglobin <100 g/L at birth) and/or
- severe or rapidly increasing hyperbilirubinaemia.

The British Committee for Standards in Haematology (BCSH) have published useful guidelines for neonatal exchange transfusion. Intravenous immunoglobulin has been used to reduce the need for exchange transfusion, but the value of this approach remains to be proven in clinical trials. Neonates with haemolytic disease due to anti-Rh antibodies may develop 'late' anaemia at a few weeks of age, requiring 'top-up' transfusion; irradiated blood must be used for infants previously receiving intrauterine transfusion to prevent the risk of transfusion-associated graft-versus-host disease. Recombinant erythropoietin sometimes prevents the need for top-up transfusion for late anaemia but is not effective when haemolysis is brisk. Folic acid (500 μg/kg daily) should be given to all babies with ongoing haemolysis until they reach 3 months of age.

Neonatal haemolytic anaemia due to red cell membrane disorders (see also Chapter 8)

In neonates, these disorders usually present with jaundice, with or without anaemia, and abnormal red cell morphology on the blood film. Many cases of hereditary elliptocytosis present as an incidental finding in a neonate being investigated for a non-haematological disorder; anaemia is rare in heterozygotes and jaundice is often minimal. Most cases of hereditary spherocytosis do have jaundice in the neonatal period and moderate anaemia is fairly common. On the other hand, severe transfusion-dependent neonatal haemolytic anaemia due to an inherited red cell membrane disorder is very uncommon and typically due to hereditary pyropoikilocytosis which can be readily identified because of the characteristic low mean corpuscular volume (MCV) of 50–60 fL and bizarre red cell morphology. The main clues are a family history and an abnormal blood film, as red cell membrane disorders can nearly always be recognized by the characteristic shape of the red cells (Figure 50.6a). Identification of the exact type of membrane abnormality may require specialized investigations. The osmotic fragility test is of limited value in neonates and recent data indicate that the dye binding test is more useful. If the clinical phenotype is severe and family history or family studies are unhelpful, further



(a) (b)

Figure 50.6 Red cell membrane disorders. (a) Blood film from a baby with hereditary pyropoikilocytosis showing microspherocytes, red cell fragments and polychromasia. (b) Blood film from a neonate with hereditary spherocytosis showing large numbers of spherocytes and polychromasia.

investigation is useful to clarify the diagnosis and pattern of inheritance. The 'gold standard' investigation has been red cell membrane electrophoresis to look for deficiency or imbalance in the different red cell membrane proteins and hence the likely causative gene. Since this technique requires specialist expertise, is expensive and must be carried out on pretransfusion blood samples to minimize diagnostic confusion due to transfused cells, there is increasing use of molecular methods using a panel of known red cell membrane genes.

Hereditary spherocytosis

This is the most common of the red cell membrane disorders to cause symptomatic anaemia, affecting 1 in 5000 live births to parents of northern European extraction. Hereditary spherocytosis is caused by mutations in the genes for spectrin, ankyrin, protein 4.1 or protein 3 and is usually inherited in an autosomal dominant fashion. The blood film shows spherocytes and typically has an identical appearance to that of ABO haemolytic disease (Figure 50.6b). Neonatal anaemia due to hereditary spherocytosis is usually moderate (haemoglobin 70–100 g/L); it is not uncommon for affected neonates to require one or two transfusions during the neonatal period before a transfusion-free plateau haemoglobin of 80–100 g/L is achieved after a few months.

Hereditary elliptocytosis

This disorder is also caused by mutations in the spectrin, ankyrin or protein 4.1 genes, but which have a different functional consequence to the mutations associated with hereditary spherocytosis. In neonates, hereditary elliptocytosis usually has no clinical manifestations apart from elliptocytes on the blood film. However, neonates who are homozygous or compound heterozygous for hereditary elliptocytosis mutations have severe haemolytic anaemia. The most common form is *hereditary pyropoikilocytosis*, which causes severe, persistent, transfusion-dependent haemolytic anaemia presenting in the neonatal period. The diagnosis is usually easily made from the low MCV and blood films of the baby (which show bizarre fragmented red cells and microspherocytes; see Figure 50.6a). Examination of the parental blood films is very helpful since one or both parents typically has common hereditary elliptocytosis, although silent spectrin mutations and de novo mutations may also occur and produce a severe phenotype in the baby despite apparently normal parental blood films. The treatment of neonates with hereditary pyropoikilocytosis is red cell transfusion. Transfusions should continue on a regular basis until the child is old enough to undergo splenectomy, to which there is an excellent response, with all patients rendered transfusion-independent.

Neonatal haemolysis due to red cell enzymopathies (see also Chapter 8)

The commonest inherited red cell enzymopathies presenting in the neonatal period are glucose-6-phosphate dehydrogenase (G6PD) deficiency and pyruvate kinase deficiency.

G6PD deficiency

The prevalence of G6PD deficiency is particularly high in individuals from Central Africa (20%) and the Mediterranean (10%), but is also seen in the Indian subcontinent, the Far East and the Middle East. In neonates, G6PD deficiency nearly always presents with jaundice within the first few days of life; the vast majority of affected neonates are boys, as the *G6PD* gene is on the X chromosome. Neonatal jaundice due to G6PD deficiency is often severe, whereas anaemia is uncommon. The blood film is usually completely normal, but in the small number of neonates with anaemia due to G6PD deficiency, the blood film shows typical changes of oxidative haemolysis. The diagnosis is made by assaying G6PD on a peripheral blood sample. The pathogenesis of the jaundice is unclear, as most babies with G6PD deficiency have no evidence of haemolysis. Management of neonatal G6PD deficiency requires close monitoring of the bilirubin to prevent kernicterus, particularly when interactions with other risk factors for neonatal hyperbilirubinaemia are present, such as

Gilbert syndrome, haemolytic disease of the newborn or hereditary spherocytosis, and also the counselling of the babies' parents regarding which medicines, chemicals and foods may precipitate haemolysis. G6PD deficiency does not cause chronic haemolysis (except in a few extremely rare G6PD variants) and therefore folic acid supplements are not indicated.

Pyruvate kinase deficiency

Although pyruvate kinase deficiency is the second most common inherited red cell enzymopathy in neonates, the estimated prevalence is only 50 cases/million and therefore even large neonatal centres will only identify 1 case every 5 years. Pyruvate kinase deficiency is transmitted in an autosomal recessive fashion. It is clinically heterogeneous, varying from anaemia severe enough to cause hydrops fetalis to a mild unconjugated hyperbilirubinaemia. In severe cases, the jaundice has a rapid onset within 24 hours of birth and exchange transfusion may be required. The diagnosis is made by measuring pretransfusion red cell pyruvate kinase activity; in mild cases enzyme activity may be relatively modestly reduced making the diagnosis difficult and it is often useful to assay levels in the parents for confirmation. The blood film is sometimes distinctive but more often shows non-specific changes of non-spherocytic haemolysis. Management in the neonatal period depends on the severity of the jaundice and anaemia. Some, but not all, children with pyruvate kinase deficiency are transfusion dependent and folic acid supplements should be given to prevent deficiency due to chronic haemolysis.

Other red cell enzymopathies

The other red cell enzymopathies are rare. The most important in the neonatal period is *triosephosphate isomerase deficiency*, as one-third of patients present with neonatal haemolytic anaemia, often many months before the devastating neurological features of this disorder become apparent. Another cause of acute haemolysis confined to the neonatal period is *infantile pyknocytosis*. This poorly understood condition, which is probably due to transient glutathione peroxidase deficiency, usually presents at 1–6 weeks of age. Infantile pyknocytosis is more common in preterm neonates and when it occurs in twins, both are usually affected. The anaemia associated with infantile pyknocytosis is often moderately severe, requiring one to two red cell transfusions before gradually resolving by the age of 2–3 months. The blood film typically shows changes of oxidative haemolysis during the neonatal period, but returns to normal as the haemolysis resolves (Figure 50.7) Measurement of glutathione peroxidase levels in affected neonates may be useful (parental levels are normal) but the diagnosis is usually easily made from the blood film. There are no reports of recurrence of haemolysis outwith the neonatal period in babies with infantile pyknocytosis and care should be taken to exclude other causes of haemolytic anaemia in any cases where the haemolysis does not resolve, or recurs, after the age of 3 months.

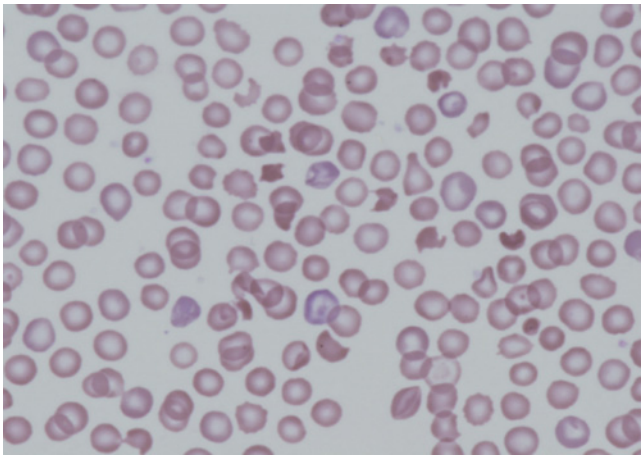


Figure 50.7 Infantile pyknocytosis. Blood film from a neonate with infantile pyknocytosis and a low level of erythrocyte glutathione peroxidase showing pyknocytes, fragmented red cells, polychromasia and occasional spherocytes.

Neonatal haemolysis due to haemoglobinopathies

The only haemoglobinopathy that typically presents in the neonatal period is α-thalassaemia major, as all four α-globin genes are deleted. Severe cases of HbH disease may also present with severe fetal or neonatal anaemia, although this is uncommon. Occasional non-thalassaemic structural α- and γ-globin gene mutations, which are clinically completely silent in adults and children, cause transient neonatal haemolytic anaemia in the neonate because they are unstable (e.g. Hb Hasharon, Hb Poole), whereas the major β-globin haemoglobinopathies (sickle cell disease and β-thalassaemia major) rarely manifest clinically in neonates.

α-Thalassaemia major

This predominantly affects families of Southeast Asian, Mediterranean or Middle Eastern origin, and presents with second-trimester fetal anaemia or hydrops fetalis, which is almost always fatal within hours of delivery. The only long-term survivors of α-thalassaemia major are those who have received intrauterine transfusions followed by regular postnatal transfusions and/or a bone marrow transplant. There is also a high incidence of hypospadias and limb defects in survivors, and others have severe neurological problems. If intrauterine transfusions are delayed until anaemia is severe, neonatal pulmonary hypoplasia is a cause of early mortality. The diagnosis of α-thalassaemia major should be suspected in any case of severe fetal anaemia that presents in the second trimester, and any case of hydrops fetalis with severe anaemia in which the parents come from high-prevalence areas, particularly Southeast Asia. Checking the blood counts of the parents will immediately identify whether they are at risk of having a child with α-thalassaemia major: both

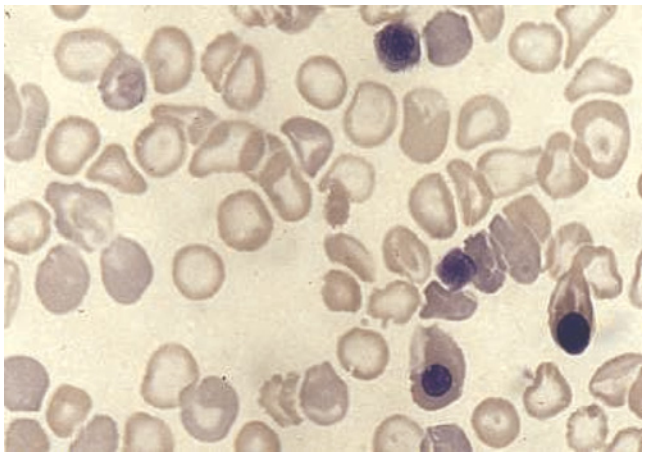


Figure 50.8 α-Thalassaemia major. Blood film from a neonate with α-thalassaemia major, born at 28 weeks' gestation, showing severe hypochromia, microcytosis, target cells, polychromasia and nucleated red cells.

parents will have hypochromic microcytic red cell indices, with MCV usually below 74 fL and mean corpuscular haemoglobin (MCH) usually less than 24 pg. The diagnosis of α-thalassaemia major is confirmed by haemoglobin electrophoresis or HPLC (which shows Hb Barts, Hb Portland and sometimes HbH; HbF and HbA are absent); the blood film shows hypochromic microcytic red cells with vast numbers of circulating nucleated red cells (Figure 50.8).

Neonatal anaemia due to blood loss

Blood loss as a cause of neonatal anaemia may be very obvious (e.g. a cephalhaematoma or rupture of the cord) or the blood loss may be concealed and easy to miss unless specifically sought (e.g. fetomaternal bleeds). Usually, the most serious blood loss occurs prior to delivery and important causes of this are twin–to-twin transfusion and fetomaternal blood loss.

Twin–to-twin transfusion

This occurs in monochorionic twins with monochorial placentas where there are two amniotic sacs (diamniotic). Twin-to-twin transfusion syndrome carries a mortality of up to 80% without treatment and 15–40% with treatment. In most cases bleeding is chronic, which usually results in a marked difference in birth weight between the twins. The donor twin is smaller, pale and lethargic and may have overt cardiac failure. By contrast, the recipient twin is often plethoric, with hyperviscosity, hyperbilirubinaemia, polycythaemia and, in the most severe cases, disseminated intravascular coagulation (DIC). Where twins survive to delivery, the difference in Hb between donor and recipient twin where there has been chronic blood loss is usually <50 g/L. Bleeding may also occur acutely during labour, with signs similar to those of fetomaternal haemorrhage.

Fetomaternal haemorrhage

Most significant fetomaternal bleeds occur during the third trimester, either spontaneously or secondary to trauma and involve very small quantities of blood (0.5 mL or less). Much more of a clinical problem are acute fetomaternal bleeds where there is loss of more than 20% of the blood volume. This may cause intrauterine death or the baby may be born with hydrops or circulatory shock. In some cases the baby appears well, but pale, at birth and collapses a few hours later. Diagnostic clues are anaemia at birth in an otherwise well baby with no or minimal jaundice. The most useful diagnostic tests are a Coombs test to exclude immune haemolysis, a reticulocyte count to exclude red cell aplasia, a Kleihauer test on maternal blood to quantify the number of HbF-containing fetal red cells in the maternal circulation and a blood film (Figure 50.9). Blood loss around the time of delivery is usually due to obstetric complications (e.g. placenta praevia, placental abruption); in these circumstances, the babies are often extremely ill, with circulatory shock, anaemia worsening rapidly after birth, large numbers of circulating nucleated red cells and DIC.

Anaemia of prematurity

The normal fall in haemoglobin in preterm neonates has been termed 'physiological anaemia of prematurity', as it does not appear to be associated with any abnormalities in the baby. The pathogenesis is not fully elucidated, but contributory factors include the reduced lifespan of fetal erythrocytes, the relatively low erythropoietin concentration, the rapid growth rate and iatrogenic blood loss. Routine supplementation of preterm neonates with folic acid and iron means that nutritional deficiency rarely plays a role. The diagnosis is usually straightforward: a well-preterm baby with a slowly falling haemoglobin, unremarkable blood film showing normochromic/normocytic red cells, slightly low reticulocytes (20×10^9/L) and no nucleated red cells.

Management of anaemia of prematurity and the role of erythropoietin

The severity of anaemia of prematurity and thereby the need for red cell transfusion can be reduced by: (i) limiting iatrogenic blood loss by appropriate use of blood tests, (ii) administering iron and folate supplements to all preterm infants (iron 3 mg/kg daily from 4–6 weeks of age or iron-fortified formula with 0.5–0.9 mg/dL of iron and folic acid 50 μg daily or 500 μg once weekly) and (iii) judicious use of erythropoietin. The many controlled trials of recombinant erythropoietin for prevention of neonatal anaemia have been reviewed extensively. Erythropoietin reduces the number of transfusions in relatively well infants with low transfusion requirements, but not in sick preterm infants who have a need for frequent phlebotomy and multiple transfusions. The main roles for recombinant erythropoietin in neonates are in ameliorating the anaemia in infants

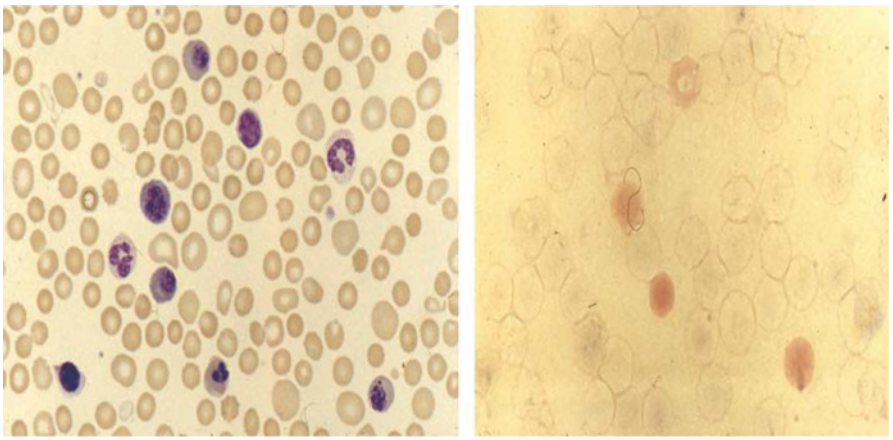

(a) (b)

Figure 50.9 Fetomaternal haemorrhage. (a) Blood film from a neonate with a haemoglobin of 50 g/L at birth due to a large fetomaternal haemorrhage; the main features are the marked polychromasia, large numbers of nucleated red cells and normal red cell morphology. (b) Kleihauer test showing HbA-containing maternal 'ghost' cells and pink-staining HbF-containing fetal cells.

who have received intrauterine transfusions for alloantibody-mediated anaemia and in preterm babies of Jehovah's Witnesses. The usual dose is 300 μg/kg of epoetin-β by subcutaneous injection three times per week, starting in the first week of life, together with oral iron supplements.

Indications for red cell transfusion in neonatal anaemia

The BCSH recently revised their guidelines for transfusion of fetuses, neonates and older children; the guidelines contain recommendations about the products and indications for red cell transfusion in neonates. These are consensus guidelines and need to be adapted for use in each individual neonatal intensive care unit, depending on the case mix of babies, as adherence to strict neonatal transfusion guidelines reduces both the number of transfusions and donor exposure.

A simple diagnostic approach to neonatal anaemia

Red cell disorders associated with neonatal or fetal anaemia present in three main ways: with a low haemoglobin, with jaundice due to haemolysis or with hydrops. A diagnostic algorithm to help identify which of these causes is most likely, which can be excluded and what further investigations are most appropriate is shown in Figure 50.2.

Neonatal polycythaemia

For both term and preterm infants, polycythaemia can be defined as a central venous haematocrit of greater than 0.65, as there is an exponential rise in blood viscosity above this level. However, even at haematocrits greater than 0.70, only a minority of neonates have clinical signs of hyperviscosity, such as lethargy, hypotonia, hyperbilirubinaemia and hypoglycaemia.

Table 50.3 Causes of neonatal polycythaemia.

- Intrauterine growth restriction
- Maternal hypertension
- Maternal diabetes
- Chromosomal disorders: trisomy 21, 18 or 13
- Twin–twin transfusion
- Delayed clamping of the cord
- Endocrine disorders: thyrotoxicosis, congenital adrenal hyperplasia

Causes of polycythaemia are shown in Table 50.3. Treatment is controversial and is not necessary in infants with very minor symptoms (e.g. borderline hypoglycaemia or poor peripheral perfusion). Infants with neurological signs and a haematocrit greater than 0.65 should have a partial exchange transfusion (using a crystalloid solution such as normal saline) to reduce the haematocrit to 0.55.

White cell disorders

Normal values

In the neonate, normal values for leucocytes, particularly neutrophils, are affected by a number of factors including gestational age, postnatal age, antenatal history, perinatal history and ethnic origin. Neutrophil counts in healthy babies increase for the first 12 hours then fall to a nadir at 4 days of age. The neutrophil count is higher in capillary samples and after vigorous crying; it is lower in neonates of African origin. Healthy preterm babies often have circulating myeloblasts and lymphoblasts, although these usually form less than 5% of the white cell differential count.

Neutropenia

A pragmatic approach is to consider a neutrophil count at birth of less than 2×10^9/L as abnormal and worth monitoring, and a neutrophil count during the first month of life of less than 0.7×10^9/L as significant enough to merit further investigation.

Causes of neutropenia

The commonest cause of neutropenia at birth in preterm neonates is reduced neutrophil production, secondary to intrauterine growth restriction or maternal hypertension. Most affected neonates also have thrombocytopenia and increased erythropoiesis (polycythaemia and/or increased circulating nucleated red cells). The underlying mechanism for these haematological abnormalities is chronic fetal hypoxia. The neutropenia resolves spontaneously usually within a few days of birth and does not persist beyond the first 2 weeks of life. The commonest cause of neutropenia in term infants is bacterial or viral infection. Infection-associated neutropenia is also self-limiting and therefore persistent neutropenia in a term or preterm baby should always be investigated. Other important causes of neutropenia are alloimmune neutropenia and severe congenital neutropenia (SCN) due to failure of neutrophil production (e.g. Kostmann syndrome), both of which predispose to severe neonatal infection.

Alloimmune neutropenia

This is the neutrophil equivalent of haemolytic disease of the newborn and alloimmune thrombocytopenia. Alloimmune neutropenia occurs when fetal neutrophils express paternally derived neutrophil-specific antigens absent on maternal neutrophils and against which the mother produces IgG neutrophil alloantibodies. The causative antibodies are usually anti-NA1 or -NA2. Severe cases present in the first few days of life with fever and infections of the respiratory tract, urinary tract and skin, particularly due to *Staphylococcus aureus*, and the mainstay of treatment is antibiotics. The diagnosis is made by demonstrating antineutrophil antibodies in the mother and baby, which react against paternal, but not maternal, neutrophil antigens. The neutropenia is self-limiting, usually resolving within 1–2 months. Since alloimmune neonatal neutropenia in everyday clinical practice is very uncommon, and yet case series show that it affects 3% of all deliveries, it is likely that most milder cases are missed as they do not present with clinical problems and routine full blood counts are not performed on well babies.

Congenital and inherited neonatal neutropenias (see also Chapter 14)

Although rare, it is important to identify SCN in the neonatal period, since such babies are at risk of life-threatening infection from birth. Investigations are indicated where neutropenia persists beyond the first 2 weeks of life, despite resolution of infection, if there is a relevant family history or consanguinity, if the neutropenia is severe ($< 0.5 \times 10^9$/L) or if the baby has typical dysmorphic features (e.g. thumb/radial abnormalities in Fanconi anaemia). SCN is the most likely cause in the neonatal period. Infants usually present with severe infections and a marked neutropenia ($<0.2 \times 10^9$/L), often with a compensatory monocytosis. The diagnosis is made by the severity of the neutropenia, the bone marrow appearance (arrest of differentiation at the myelocyte/promyelocyte stage) and the absence of antineutrophil antibodies. The inheritance of SCN can be autosomal recessive or dominant or X-linked. The molecular basis of most cases is increasingly being identified both through classical linkage studies and gene discovery approaches. The causative genes are important for normal neutrophil differentiation, survival and function. Mutations in *ELA2* (neutrophil elastase) are the most common cause of SCN, but other causes include mutations in *GFI1* (growth factor-independent protein 1), *WAS* (Wiskott–Aldrich syndrome protein, WASp), *JAGN1* and several genes affecting the endosomal–lysosomal system, such as *AP3B1, LAMTOR2, VPS13B* and *VPS45*. Recently, homozygous mutations in the *HAX1* gene were found to explain the SCN in the original Kostmann family, but this appears to be a rare cause of the disease.

Congenital leukaemias

Congenital leukaemia is rare, affecting ~5 neonates/per million live births. The most common form of congenital leukaemia is the transient leukaemia seen in neonates with Down syndrome, up to 10% of which will be affected. In children without Down syndrome, the majority of congenital leukaemias are associated with rearrangement of the MLL gene and manifest as acute lymphoblastic leukaemia (ALL) in 60% of cases and acute myeloid leukaemia (AML, typically monoblastic) in the remaining 40%.

Leukaemia in neonates with Down syndrome

Around 5–10% of neonates with Down syndrome have a transient megakaryoblastic leukaemia, also known as transient abnormal myelopoiesis (TAM) or transient myeloproliferative disorder (TMD) (see also Chapter 16), characterized by leucocytosis and circulating blast cells (Figure 50.10). TAM develops in fetal life and nearly always presents in the first week of life. Most cases spontaneously resolve within 2–3 months and no treatment is indicated. In severe cases (10–20%) neonates develop progressive disease within days or weeks of birth due to blast cell infiltration of the liver in association with liver fibrosis. In such cases pleural/pericardial effusions, ascites, hepatomegaly, liver dysfunction and coagulopathy are common and treatment with low-dose cytosine arabinoside is often successful. TAM is unique to neonates with Down syndrome and is now known to be caused by the effects of acquired mutations in the haemopoietic transcription factor gene *GATA1* in fetal liver

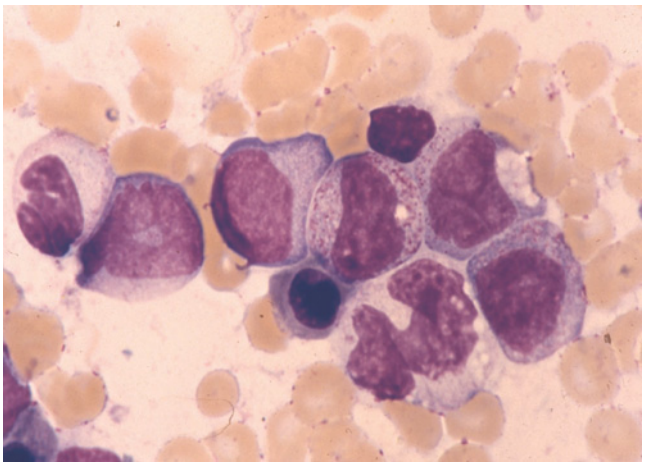


Figure 50.10 Transient leukaemia in a neonate with Down syndrome. Leucoerythoblastic blood film showing increased numbers of blast cells, which spontaneously returned to normal by 2 months of age.

haemopoietic cells. Acquired *GATA1* mutations are now known to affect almost 30% of all neonates with Down syndrome, but only half of these will manifest with clinical or haematological signs; the remaining cases are clinically silent because the mutant *GATA1* clones are very small. In most cases, the mutant *GATA1* clones disappear over the first few months of life, as all signs of TAM disappear. In up to 10% of neonates with Down syndrome who have TAM or silent *GATA1* mutations, an acute megakaryoblastic leukaemia (DS-AMKL) develops before the age of 5 years, usually between 12 and 18 months of age. In these cases, the TAM usually appears to resolve, but this is followed several months later by myelodysplasia, which subsequently evolves to acute leukaemia. DS-AMKL, unlike TAM, is not transient and must be treated with chemotherapy when the outcome is usually very good, with 80–90% long-term survival. Since the same *GATA1* mutations are present both in TAM and DS-AMKL, this indicates that the leukaemia arises from a mutant *GATA1* clone present at birth. Recently, mutations in the cohesin complex genes have been shown to be the most common mechanism responsible for transformation of TAM to DS-AMKL.

Leukaemia in children without Down syndrome

Affected babies present, often within hours of birth, with signs of anaemia, thrombocytopenia and/or skin lesions caused by leukaemic infiltration. Congenital leukaemia with nodular skin deposits of leukaemic cells is one cause of the classical 'blueberry muffin' rash in a newborn baby. Some cases present *in utero* with intrauterine death or hydrops. CNS leukaemia in a neonate may present as a bulging fontanelle. The blood film and bone marrow aspirate show large numbers of primitive blast cells. The most common cytogenetic abnormalities are translocations involving 11q23 resulting in rearrangement of the MLL gene and t(1;22), which is associated with the *OTT–MAL* fusion gene and an aggressive form of acute megakaryoblastic leukaemia. The prognosis is extremely poor (~20% long-term survival); few are cured by chemotherapy and bone marrow transplantation may be the best option.

Haemostasis and thrombosis in the newborn

Bleeding and thrombotic problems are relatively common in neonates, particularly in those who are preterm and/or sick. The vast majority of bleeding problems are acquired and secondary to perinatal complications, including perinatal asphyxia and severe infection. Inherited bleeding disorders, with the exception of haemophilia A and B, are rare in the newborn. By contrast, there has been a marked increase in the identification of inherited thrombophilias in neonates and the most challenging aspect of this has been understanding their significance, both for management of the neonate and for genetic counselling.

Developmental haemostasis

Coagulation proteins are present at measurable levels from the 10th week of gestation and gradually rise during fetal life. They do not cross the placenta, or do so in very small amounts, and therefore need to be independently synthesized by the fetus. 'Normal values' in the neonate vary not only with gestational age, but also with postnatal age. Data for babies at less than 30 weeks' gestation derive from fetoscopy samples: levels of the vitamin-K-dependent factors (FII, FVII, FIX and FX) and of FXI and FXII are all low (<40% of adult values) and remain so during the first month of life. In contrast, even in preterm babies (>30 weeks gestation at birth), levels of FV, FXIII and fibrinogen are normal at birth and levels of FVIII and von Willebrand factor (VWF) are normal or increased.

Platelet counts at birth in term and preterm neonates are within the normal adult range. However, increasing evidence points to significant differences in platelet production during fetal life. Fetal megakaryocytes are smaller and of lower ploidy despite being more proliferative than adult megakaryocytes and are more sensitive to thrombopoietin. Interestingly, in response to consumptive thrombocytopenia, neonates increase the number of megakaryocytes rather than megakaryocyte size consistent with fundamental differences in regulation of platelet production in the fetus and newborn compared to later in life. These differences may help to explain why many studies have found impaired *in vitro* function of neonatal platelets. The most consistent abnormalities are reduced aggregation in response to adrenaline, ADP and thrombin. Their significance in clinical practice is still not clear, as the bleeding time is normal in term and preterm infants (≤135 s).

Screening tests for bleeding disorders

Nearly all significant bleeding disorders in neonates can be identified using simple screening tests; exceptions are FXIII deficiency and platelet function disorders. It is often helpful to test both parents for coagulation abnormalities to help identify inherited disorders as there is considerable overlap between coagulation factor levels in inherited or acquired deficiency states and neonatal normal ranges, particularly in preterm neonates. In addition, it is advisable to repeat investigations for the levels of coagulation factors at the age of 6–12 months to confirm (or exclude) inherited deficiencies. The most useful screening tests in neonates are the prothrombin time (PT), activated partial thromboplastin time (APTT), thrombin time (TT), fibrinogen and platelet count. Bleeding times are generally unhelpful in neonates and investigation of platelet function abnormalities is often deferred until a few months of age when platelet aggregometry and PFA-100 become practical.

Inherited coagulation disorders (see also Chapters 38 and 39)

The commonest inherited disorders presenting in the neonatal period are FVIII deficiency (haemophilia A), which has a frequency of 1 in 5000 male births, and FIX deficiency (haemophilia B), which occurs in 1 in 30,000 male births.

Factor VIII deficiency

Almost 40% of patients with inherited FVIII deficiency present in the neonatal period. The clinical signs include intracranial haemorrhage, cephalohaematomas and bleeding after circumcision or from venous or arterial puncture sites. As in adults, the diagnosis is made by finding an isolated prolonged APTT and reduced FVIII clotting activity. Acute management of the bleeding neonate with haemophilia requires intravenous administration of recombinant FVIII (50–100 units/kg twice daily) to achieve FVIII levels of 1.0 unit/mL. As the half-life of FVIII is shorter than in adults, more frequent dosing or a continuous FVIII infusion may be required. For neonates with intracranial bleeding, treatment with FVIII should continue for at least 2 weeks. Fibrin glue may be useful in circumcision-associated bleeds. The optimal mode of delivery for carriers known to be carrying an affected child or at risk of carrying an affected child, remains controversial. However, there is general agreement that instrumental delivery, fetal blood sampling and use of fetal scalp electrodes should be avoided. Late amniocentesis is sometimes used to identify unaffected fetuses where delivery can safely proceed without intervention. The role of prophylactic FVIII administration to haemophiliac newborns following difficult delivery to reduce the risk of intracranial bleeding (1–4%) is controversial, but is increasingly used. Prophylactic FVIII should also be used for a newborn haemophiliac when a previous sibling has had a major intracranial bleed.

Factor IX deficiency

Deficiency of FIX is clinically indistinguishable from FVIII deficiency. As FIX levels are also low in liver disease and vitamin K deficiency, it is important to recheck FIX levels at 6 weeks and 6 months of age if the diagnosis is in doubt. Neonates with bleeding are treated with recombinant FIX concentrate (100 units/kg IV once daily, monitored to achieve a FIX level of 1.0 IU/mL).

von Willebrand disease in neonates

The two forms of von Willebrand disease (VWD) that can present in neonates are type 2B vWD and type 3 VWD. Type 2B disease presents with thrombocytopenia and bleeding and is very uncommon. Type 3 vWD has a clinical phenotype similar to haemophilia, as levels of both VWF and FVIII are low. The inheritance in type 3 vWD is autosomal recessive. The diagnosis is made by measuring VWF, FVIII and the pattern of VWF multimers. At present, type 3 vWD is treated with a high purity VWF:FVIII concentrate (such as Wilate®).

Factor XIII deficiency

This rare autosomal recessive disorder usually presents with delayed bleeding from the umbilical cord during the first 3 weeks of life. The diagnosis is made by measuring clot solubility in 5 mol/L urea solution as a screening test, followed by a specific FXIII assay. Molecular tests for the common FXIII mutations are also available. The routine diagnostic coagulation screen is normal in Factor XIII deficiency. Bleeding is treated with FXIII concentrate.

Acquired disorders of coagulation

Causes of acquired coagulopathy in neonates include vitamin K deficiency, DIC, liver disease, metabolic disorders, extracorporeal membrane oxygenation (ECMO) and giant haemangioendotheliomas (Kasabach–Merritt syndrome).

Vitamin K deficiency

Levels of vitamin-K-dependent procoagulant factors (FII, FVII, FIX and FX), protein C and protein S are low at birth because of poor placental transfer of vitamin K, low vitamin K stores at birth, low vitamin K in breast milk and the lack of bacterial vitamin K synthesis in the sterile neonatal gut. Vitamin K deficiency can lead to haemorrhagic disease of the newborn, also known as vitamin K deficiency bleeding (VKDB). Early VKDB presents in the first 24 hours of life with severe gastrointestinal and intracranial haemorrhage. It is usually due to maternal medication that interferes with vitamin K, for example, anticonvulsants (phenobarbital, phenytoin), antituberculous therapy or oral anticoagulants. Classical VKDB presents at 2–7 days in 0.25–1.7% of babies who have not received prophylactic vitamin K at birth, especially if breast-fed or with poor oral intake. Late VKDB occurs between 2 and 8 weeks after birth and presents with

sudden intracranial haemorrhage in an otherwise well breast-fed term baby or in babies with liver disease.

The diagnosis of VKDB is usually made by an isolated prolonged PT, which corrects following vitamin K administration. If doubt remains, assays of the inactive form of FII (decarboxyprothrombin; PIVKA II) can be used to confirm the diagnosis. Treatment of VKDB is vitamin K 1 mg intravenously or subcutaneously with fresh-frozen plasma only in severe haemorrhage. Classic and late VKDB can be prevented by a single intramuscular dose of vitamin K or, slightly less effectively, by weekly 1-mg oral doses of vitamin K over the first 12 weeks of life (in exclusively breast-fed infants neither single oral dose vitamin K nor daily very low-dose vitamin K 25 μg are effective in preventing classic VKDB). Some studies, but not others, suggest a link between intramuscular vitamin K at birth and later childhood malignancies. In healthy babies, the choice of which route of administration is increasingly being left to parents. However, infants of mothers taking drugs that inhibit vitamin K (for example, warfarin) are at risk of early VKDB and these babies should receive intramuscular vitamin K 1 mg as soon as possible after birth.

Disseminated intravascular coagulation

The main triggers of DIC in neonates are severe hypoxia and/or acidosis and sepsis. The main causes are DIC secondary to acute perinatal hypoxia (for example in association with placental abruption or a large, acute fetomaternal haemorrhage), septicaemia and necrotizing enterocolitis. DIC may also be seen secondary to perinatal Coxsackie virus infection. Neonates with DIC are nearly always very sick and the clinical signs include generalized bleeding, pulmonary haemorrhage and oozing from venepuncture sites. The usual pattern of coagulation abnormalities in neonatal DIC is prolongation of the PT, APTT and TT, together with low platelets and fibrinogen. D-dimers are increased, but are not specific and can be found in healthy neonates with no evidence of coagulopathy. The most important aspect of management of DIC is treatment of the underlying cause. Blood product replacement is indicated for clinical bleeding, aiming to maintain the platelet count above 30×10^9/L and fibrinogen greater than 1 g/L.

Neonatal thrombocytopenia

Thrombocytopenia occurs in 1–5% of all neonates and in up to 50% of neonates who are preterm and sick. It usually presents in one of two clinical patterns: early onset thrombocytopenia (within 72 hours of birth) and late onset thrombocytopenia (after 72 hours of life) (Table 50.4). The most frequent causes of early thrombocytopenia in preterm infants are intrauterine growth restriction and maternal hypertension or diabetes; the most frequent causes in term infants are neonatal alloimmune thrombocytopenia (NAITP) and thrombocytopenia secondary to maternal immune thrombocytopenic purpura (ITP). The most common causes of late thrombocytopenia are sepsis and necrotizing enterocolitis.

Table 50.4 Causes of neonatal thrombocytopenia.

Early onset (<72 hours)
- Placental insufficiency (PET, IUGR, diabetes)
- Neonatal alloimmune thrombocytopenia
- Birth asphyxia
- Perinatal infection (group B *Streptococcus*, *Escherichia coli*, *Listeria*)
- Congenital infection (CMV, toxoplasmosis, rubella, Coxsackie)
- Maternal autoimmune (ITP, SLE)
- Severe rhesus HDN
- Thrombosis (renal vein, aortic)
- Aneuploidy (trisomy 21, 18, 13)
- Congenital/inherited (TAR, Wiskott–Aldrich)

Late onset (>72 hours)
- Late-onset sepsis and necrotizing enterocolitis
- Congenital infection (CMV, toxoplasmosis, rubella)
- Maternal autoimmune (ITP, SLE)
- Congenital/inherited (TAR, Wiskott–Aldrich)

CMV, cytomegalovirus; HDN, haemolytic disease of the newborn; ITP, idiopathic thrombocytopenic purpura; IUGR, intrauterine growth restriction; PET, pre-eclampsia; SLE, systemic lupus erythematosus; TAR, thrombocytopenia with absent radii.

Neonatal alloimmune thrombocytopenia (NAITP)

NAITP is the platelet equivalent of haemolytic disease of the newborn and affects around 1 in 1000 pregnancies. It is frequently severe (platelet count $<30 \times 10^9$/L) and occurs in the first pregnancy in almost 50% of cases. Thrombocytopenia results from transplacental passage of maternal platelet-specific antibodies to human platelet antigens (HPA), which the mother lacks, but which the fetus inherits from the father. In 80% of cases, these are anti-HPA-1a antibodies and in 10–15% anti-HPA-5b; occasional cases are due to anti-HPA-3a. HLA-DRB3*0101-positive women are 140 times more likely to make anti-HPA-1a than HLA-DRB3*0101-negative women. Intracranial haemorrhage occurs in 10% of cases, with long-term neurodevelopmental sequelae in 20% of survivors. The diagnosis is made by demonstrating platelet antigen incompatibility between mother and baby and anti-HPA antibodies in the mother. Where possible the neonate and both parents should have platelet genotyping performed for the most common HPA alloantigens (HPA-1a, -2, -3, -5b and -15). Where a clinical diagnosis of NAITP is made and no antibodies or maternal–neonatal platelet incompatibility to these five HPA antigens has been excluded, it is not usually possible to identify the causal platelet antigen.

Very few of these unexplained cases are due to antibodies against minor HPA (such as HPA-6w and HPA-9w) and, therefore, most groups currently do not routinely screen for these very low frequency HPA antigens, except in selected severe cases.

Management of NAITP

Management depends on the gestation and postnatal age of the baby, the history of previously affected siblings, the platelet count and the extent of any bleeding. All neonates suspected to have NAITP should have a cranial ultrasound scan to exclude the presence of an ICH. Since the platelet count usually falls over the first 4–7 days of life, all thrombocytopenic neonates with NAITP should be monitored until there is a sustained rise in their platelet count into the normal range. Platelet transfusion is not indicated in a term neonate with platelet counts $>30 \times 10^9$/L who has no evidence of bleeding. Transfusion with HPA-compatible (usually HPA1a, 5b-negative) platelets is recommended for:

- platelet counts below 30×10^9/L in term neonates who have no evidence of bleeding;
- platelet counts below 50×10^9/L and an ICH and/or a previously affected sibling with an ICH and/or birthweight < 1000g.

Neonates who have received intrauterine platelet transfusions should be given irradiated platelets to prevent transfusion-associated graft-versus-host disease. In some cases, thrombocytopenia may persist for up to 8–12 weeks. In such cases intravenous IgG (total dose 2 g/kg, over 2–5 days) may be useful to reduce the need for repeated platelet transfusions. Antenatal management of NAITP is controversial and depends on the severity of previously affected neonates, the HPA genotype of the father and on the HPA antibody specificity. Most centres have abandoned the invasive approach, except in very high-risk cases, as this relies on fetal blood sampling, plus fetal transfusion with HPA-compatible platelets, which carries a risk of preterm delivery and fetal death. Instead, most centres now use a non-invasive approach relying on maternal intravenous IgG therapy during the second half of the pregnancy.

Neonatal autoimmune thrombocytopenia (see also Chapter 42)

The majority of cases are secondary to transplacental passage of maternal platelet autoantibodies in maternal ITP, which affects 1–2 in 1000 pregnancies. Most of the remaining cases are secondary to maternal systemic lupus erythematosus (SLE) although neonatal Kawasaki disease has also been reported. Around 10% of infants develop thrombocytopenia, of which around half will develop platelets $<20 \times 10^9$/L. The risk of major haemorrhage, including ICH, is estimated at ~1%. In babies with severe thrombocytopenia, intravenous IgG (400 mg/kg/day given over 2–4 hours for 5 days or 1 g/kg/day for 2 days, total dose 2 g/kg) is usually effective. In some cases neonatal thrombocytopenia is prolonged and may last up to 12 weeks before spontaneously resolving.

Indications for platelet transfusion

A number of countries have published consensus guidelines to help decide the indications for platelet transfusion in term and preterm neonates. In general, platelet transfusion is not considered necessary in well babies if the platelet count is above 20–30×10^9/L. Most groups agree that the threshold for transfusing sick babies, particularly preterm babies in the first few weeks of life, should be higher ($30–50 \times 10^9$/L).

Neonatal thrombosis: physiology and developmental aspects

Neonates have an increased risk of thrombosis (2.4 per 1000 hospital admissions) compared with older infants and children. This is largely due to the physiologically low levels of many of the inhibitors of coagulation and the frequent use of indwelling vascular catheters. Concentrations of antithrombin, heparin cofactor II and protein C at birth are 30–50% of adult values; protein S levels are also low, but overall protein S activity is normal as it exists mainly in its free active form due to the virtual absence of its binding protein (C4b-BP) in neonates. Levels of plasminogen at birth are only 50% of adult values so neonates have a reduced ability to generate plasmin in response to fibrinolytic agents.

Screening tests for thrombophilia in neonates

The only inherited deficiencies for which there is a proven role in neonatal thrombosis are deficiencies in proteins C and S, which cause purpura fulminans, and deficiency of antithrombin, although presentation in the neonatal period is extremely rare. FV Leiden and the prothrombin 20210A promoter mutation (prothrombin^{20210A}) have not yet been reported to cause neonatal thrombotic problems in isolation. The role of thrombophilia screening in neonates is controversial, since the significance of any findings remains very difficult to assess, particularly since the risk of recurrence of thrombosis after an episode in the neonatal period is very low (~3%). Guidelines from the Haemostasis and Thrombosis Task Force of the British Committee for Standards in Haematology (BCSH) state that congenital thrombophilia should be considered and screened for in any child with:

- clinically significant thrombosis, including spontaneous thrombotic events, unanticipated or extensive venous thrombosis, ischaemic skin lesions or purpura fulminans and
- a positive family history of neonatal purpura fulminans

The screening tests that should be performed in all suspected cases of thrombophilia include protein C activity, protein S, antithrombin, FV Leiden and prothrombin^{20210A}. In addition, babies with thrombosis born to mothers with SLE and/or antiphospholipid syndrome should be tested for lupus anticoagulant, as antiphospholipid antibodies may cross the placenta and are a rare cause of neonatal thrombosis in such babies. The

relevance of serum liporotein a and the MTHFR genotype to neonatal management is unclear at present.

Inherited thrombotic disorders in neonates

Proteins C and S deficiency

Protein C deficiency occurs in 1 in 160,000–360,000 births. Affected babies usually present within hours or days of birth with purpura fulminans, in which there is DIC and rapidly progressive life-threatening haemorrhagic necrosis due to dermal vessel thrombosis or cerebral, renal vein or ophthalmic thrombosis. The diagnosis of protein C deficiency is made by the clinical picture and undetectable levels of protein C (<0.01 U/mL) in the patient, together with heterozygote levels in the parents. Protein-C-deficient heterozygotes rarely present as neonates and have low protein C levels, which may overlap with the lower limit of normal in the first few months of life, delaying diagnosis until 6 months or later. Treatment is with protein C concentrate (40 U/kg, aiming to maintain a plasma level >0.25 U/mL). Protein S deficiency presents with identical features; levels of protein S are undetectable (<0.01 U/mL) and treatment is with fresh-frozen plasma (10–20 mL/kg) to maintain a plasma protein S in excess of 0.25 U/mL. The rarity of thrombosis in neonates with protein S deficiency may in part reflect the relatively high level of active protein S because of the undetectable levels of C4b-binding protein.

Antithrombin deficiency

Homozygous antithrombin deficiency usually presents later in childhood, but neonatal deep venous thrombosis and inferior vena cava thrombosis have been reported. Heterozygous antithrombin deficiency is more common (1 in 2000–5000 births); it usually presents in the second decade of life, but neonatal presentation with aortic thrombosis, myocardial infarction and cerebral dural sinus thrombosis may occur.

Acquired thrombotic problems in neonates

The most common risk factors for neonatal thrombosis are the presence of an intravascular catheter and shock, secondary to sepsis, hypoxaemia or hypovolaemia. Maternal factors, such as maternal thrombophilia, alcohol and cocaine use and pre-eclampsia, may also increase the risk of neonatal thrombosis. Thrombosis particularly affects the renal veins, the aorta, aortic arch or cerebral vessels. Catheter-related thrombosis causes more than 80% of venous thromboses and more than 90% of arterial thromboses. The diagnosis is made by Doppler ultrasound or contrast angiography. Treatment of catheter-related thrombosis depends on the severity and extent of thrombosis. The first step is prompt removal, where possible. If signs progress despite catheter removal, unfractionated heparin or low-molecular-weight heparin should be started using a dosing regimen adapted for neonates. Thrombolytic therapy with urokinase or tissue plasminogen activator has been used successfully for catheter-related thrombosis in neonates, but experience in preterm neonates is very limited. It is important to maintain fibrinogen at levels less than 1 g/L and the platelet count greater than 50×10^9/L in neonates receiving thrombolytic therapy, and heparin (starting dose 28 U/kg per hour) is often given to maintain vessel patency after thrombolytic therapy, although there is no evidence that this is beneficial. Excellent guidelines for antithrombotic therapy in neonates are regularly updated by the American College of Chest Physicians.

Selected bibliography

Bertrand G, Kaplan C (2014) How do we treat fetal and neonatal alloimmune thrombocytopenia. *Transfusion* **54**: 1698–703.

Bhat R, Monagle P (2012). The preterm infant with thrombosis. *Archives of Disease in Childhood (Fetal and Neonatal Edition)* **97**: F423–8.

Gibson BE, Todd A, Roberts I *et al.*; British Committee for Standards in Haematology Transfusion Task Force (2004) Transfusion guidelines for neonates and older children. *British Journal of Haematology* **124**: 433-53(and www.bcshguidelines.com/, accessed June 2015)

Chakravorty S, Roberts I (2012) How I manage neonatal thrombocytopenia. *British Journal of Haematology* **156**: 155–62.

Chalmers E, Williams M, Brennand J *et al.*, on behalf of the Paediatric Working Party of the United Kingdom Haemophilia Doctors' Organization (2011) Guideline on the management of haemophilia in the fetus and neonate. *British Journal of Haematology* **154**: 208–15.

Roy A, Roberts I, Vyas P (2012) Biology and management of transient abnormal myelopoiesis (TAM) in children with Down syndrome. *Seminars in Fetal and Neonatal Medicine* **17**: 196–201.

Shearer MJ (2009) Vitamin K deficiency bleeding (VKDB) in early infancy. *Blood Reviews* **23**: 49–59.

Veldman A, Fischer D, Nold MF, Wong FY (2010) Disseminated intravascular coagulation in term and preterm neonates. *Seminars in Thrombosis and Hemostasis* **36**: 419–28.

von Lindern JS, Lopriore E (2014) Management and prevention of neonatal anemia: current evidence and guidelines. *Expert Review of Hematology* **7**: 195–202.

Will A (2015) Neonatal haemostasis and the management of neonatal thrombosis. *British Journal of Haematology* **169**: 324–32.

CHAPTER 51

WHO Classification: Tumours of the Haematopoietic and Lymphoid Tissues (2008)

51

Myeloproliferative neoplasms

Chronic myelogenous leukaemia, *BCR–ABL1* positive
Chronic neutrophilic leukaemia
Polycythemia vera
Primary myelofibrosis
Essential thrombocythaemia
Chronic eosinophilic leukaemia, not otherwise specified
Mastocytosis
  Cutaneous mastocytosis
  Systemic mastocytosis
  Mast cell leukaemia
  Mast cell sarcoma
  Extracutaneous mastocytoma
Myeloproliferative neoplasms, unclassifiable
Myeloid and lymphoid neoplasms associated with eosinophilia and abnormalities of *PDGFRA*, *PDGFRB* or *FGFR1*
  Myeloid and lymphoid neoplasms associated with *PDGFRA* rearrangement
  Myeloid neoplasms associated with *PDGFRB* rearrangement
  Myeloid and lymphoid neoplasms with *FGFR1* abnormalities

Myelodysplastic/myeloproliferative neoplasms

Chronic myelomonocytic leukaemia
Atypical chronic myeloid leukaemia, *BCR–ABL1* negative
Juvenile myelomonocytic leukaemia
Myelodysplastic/myeloproliferative neoplasms, unclassifiable
  Refractory anaemia with ringed sideroblasts (RARS) associated with marked thrombocytosis*

Myelodysplastic syndromes

Refractory cytopenia with unilineage dysplasia
  Refractory anaemia
  Refractory neutropenia
  Refractory thrombocytopenia
Refractory anaemia with ring sideroblasts
Refractory cytopenia with multilineage dysplasia
Refractory anaemia with excess blasts
Myelodysplastic syndromes associated with isolated del(5q)
Myelodysplastic syndromes, unclassifiable
Myelodysplastic syndromes in children

Acute myeloid leukaemia

Acute myeloid leukaemia (AML) with recurrent genetic abnormalities
  AML with t(8;21)(q22;q22), *RUNX1–RUNX1T1*
  AML with inv(16)(p13.1q22) or t(16;16)(p13.1;q22), *CBFB–MYH11*
  Acute promyelocytic leukaemia with t(15;17)(q22;q11–12), *PML–RARA*
  AML with t(9;11)(p22;q23), *MLLT3–MLL*
  AML with t(6;9)(p23;q34), *DEK–NUP214*
  AML with inv(3)(q21q26.2) or t(3;3)(q21;q26.2), *RPN1–EVI1*
  AML (megakaryoblastic) with t(1;22)(p13;q13), *RBM15–MKL1*
  AML with mutated *NPM1**
  AML with mutated *CEBPA**

Postgraduate Haematology, Seventh Edition. Edited by A Victor Hoffbrand, Douglas R Higgs, David M Keeling and Atul B Mehta.

Acute myeloid leukaemia with myelodysplasia-related changes
Therapy-related myeloid neoplasms
Acute myeloid leukaemia, not otherwise categorized
  AML with minimal differentiation
  AML without maturation
  AML with maturation
Acute myelomonocytic leukaemia
Acute monoblastic and monocytic leukaemia
Acute erythroid leukaemia
  Acute erythroid leukaemia, erythroid/myeloid
  Acute pure erythroid leukaemia
Acute megakaryoblastic leukaemia
Acute basophilic leukaemia
Acute panmyelosis with myelofibrosis
Myeloid sarcoma
Myeloid proliferations related to Down syndrome
  Transient abnormal myelopoiesis
  Acute myeloid leukaemia associated with Down syndrome
Blastic plasmacytoid dendritic cell neoplasm
Acute leukaemias of ambiguous lineage
  Acute undifferentiated leukaemia
  Acute biphenotypic leukaemia

Precursor lymphoid neoplasms

B lymphoblastic leukaemia/lymphoma
B lymphoblastic leukaemia/lymphoma, not otherwise specified
B lymphoblastic leukaemia/lymphoma with recurrent cytogenetic/molecular genetic abnormalities
  B lymphoblastic leukaemia/lymphoma with t(9:22)(q34;q11.2), *BCR–ABL1*
  B lymphoblastic leukaemia/lymphoma with t(11q23), *MLL* rearranged
  B lymphoblastic leukaemia/lymphoma with t(12;21)(p13;q22), *TEL–AML1* (*ETV6–RUNX1*)
  B lymphoblastic leukaemia/lymphoma with hyperdiploidy
  B lymphoblastic leukaemia/lymphoma with hypodiploidy (hypodiploid ALL)
  B lymphoblastic leukaemia/lymphoma with t(5;14)(q31;q32), *IL3–IGH*
  B lymphoblastic leukaemia/lymphoma with t(1;19)(q23;p13.3), *E2A–PBX1* (*TCF3–PBX1*)
T-lymphoblastic leukaemia/lymphoma

Mature B-cell neoplasms

Chronic lymphocytic leukaemia/small lymphocytic lymphoma
B-cell prolymphocytic leukaemia
Splenic marginal zone lymphoma
Hairy-cell leukaemia
Splenic lymphoma/leukaemia, unclassifiable
  Splenic diffuse red pulp small B-cell lymphoma*
  Hairy-cell leukaemia variant*
Lymphoplasmacytic lymphoma
  Waldenström macroglobulinaemia
Heavy chain diseases
  Alpha heavy chain disease
  Gamma heavy chain disease
  Mu heavy chain disease
Plasma cell myeloma
Solitary plasmacytoma of bone
Extraosseous plasmacytoma
Extranodal marginal zone B-cell lymphoma of mucosa-associated lymphoid tissue (MALT lymphoma)
Nodal marginal zone B-cell lymphoma
  Paediatric type nodal marginal zone lymphoma
Follicular lymphoma
  Paediatric type follicular lymphoma
Primary cutaneous follicle centre lymphoma
Mantle cell lymphoma
Diffuse large B-cell lymphoma (DLBCL), not otherwise specified
  T-cell/histiocyte-rich large B-cell lymphoma
  DLBCL associated with chronic inflammation
  EBV-positive DLBCL of the elderly
Lymphomatoid granulomatosis
Primary mediastinal (thymic) large B-cell lymphoma
Intravascular large B-cell lymphoma
Primary cutaneous DLBCL, leg type
ALK-positive large B-cell lymphoma
Plasmablastic lymphoma
Primary effusion lymphoma
Large B-cell lymphoma arising in HHV8-associated multicentric Castleman disease
Burkitt lymphoma
B-cell lymphoma, unclassifiable, with features intermediate between DLBCL and Burkitt lymphoma
B-cell lymphoma, unclassifiable, with features intermediate between DLBCL and classical Hodgkin lymphoma

Mature T-cell and NK-cell neoplasms

T-cell prolymphocytic leukaemia
T-cell large granular lymphocytic leukaemia
Aggressive NK-cell leukaemia
Systemic EBV-positive T-cell lymphoproliferative disease of childhood (associated with chronic active EBV infection)
Hydroa vaccineforme-like lymphoma
Adult T-cell leukaemia/lymphoma
Extranodal NK/T-cell lymphoma, nasal type
Enteropathy-associated T-cell lymphoma

Hepatosplenic T-cell lymphoma
Subcutaneous panniculitis-like T-cell lymphoma
Mycosis fungoides
Sézary syndrome
Primary cutaneous anaplastic large-cell lymphoma
Primary cutaneous aggressive epidermotropic CD8-positive cytotoxic T-cell lymphoma*
Primary cutaneous γδ T-cell lymphoma
Primary cutaneous small/medium CD4-positive T-cell lymphoma*
Peripheral T-cell lymphoma, not otherwise specified
Angioimmunoblastic T-cell lymphoma
Anaplastic large-cell lymphoma, ALK positive
Anaplastic large-cell lymphoma, ALK negative*

Hodgkin lymphoma

Nodular lymphocyte-predominant Hodgkin lymphoma
Classical Hodgkin lymphoma
  Nodular sclerosis classical Hodgkin lymphoma
  Lymphocyte-rich classical Hodgkin lymphoma
  Mixed cellularity classical Hodgkin lymphoma
  Lymphocyte depleted classical Hodgkin lymphoma

Histiocytic and dendritic cell neoplasms

Histiocytic sarcoma
Langerhans cell histiocytosis
Langerhans cell sarcoma
Interdigitating dendritic cell sarcoma
Follicular dendritic cell sarcoma
Dendritic cell tumour, not otherwise specified
Indeterminate dendritic cell tumour
Fibroblastic reticular cell tumour

Post-transplant lymphoproliferative disorders

Early lesions
  Reactive plasmacytic hyperplasia
  Infectious mononucleosis-like
Polymorphic post-transplant lymphoproliferative disorder
Monomorphic post-transplant lymphoproliferative disorder (B- and T/NK-cell types)
Classical Hodgkin's lymphoma-type post-transplant lymphoproliferative disorder

* These represent provisional entities or provisional subtypes of other neoplasms. They are provisional because there are insufficient data to support their being a definite entity, significant controversies about their defining features and/or uncertainty about whether they are unique or closely related to other definite entities. Further classify according to the lymphoma they resemble.

Index

Page numbers in *italics* denote figures, those in **bold** denote tables.

Postgraduate Haematology, Seventh Edition. Edited by A Victor Hoffbrand, Douglas R Higgs, David M Keeling and Atul B Mehta.

WILEY END USER LICENSE AGREEMENT

Go to www.wiley.com/go/eula to access Wiley’s ebook EULA.