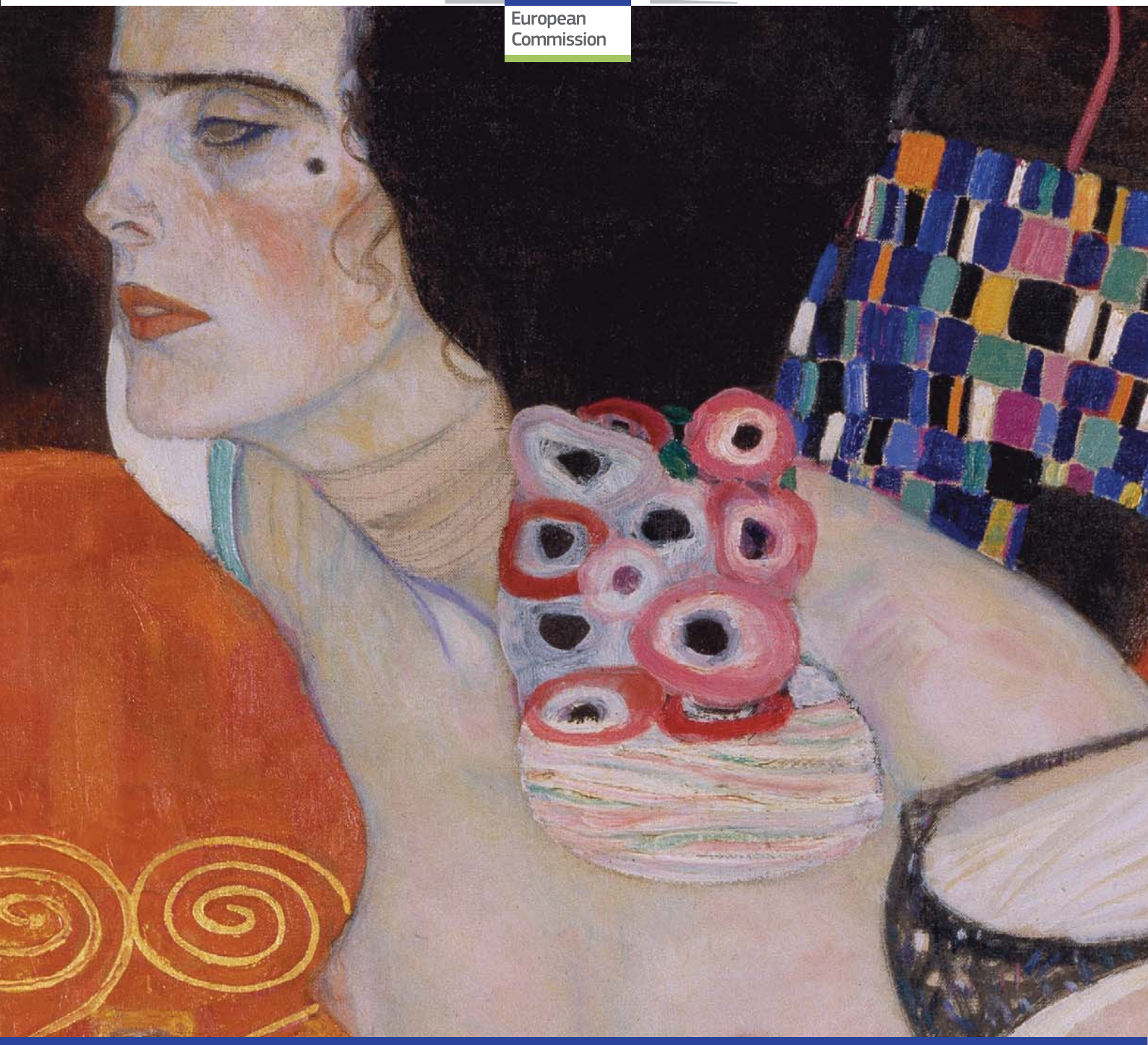
European Commission



European guidelines for quality assurance in breast cancer screening and diagnosis

Fourth edition - Supplements

Health and Consumers

Financial support was provided by the European Union Public Health Programme (Project no. 2006322, European Cooperation on Development and Implementation of Cancer Screening and Prevention Guidelines [ECCG]).

Special thanks are due to the IARC staff in the Quality Assurance, Screening and Communication Groups who provided technical and editorial assistance.

Further information on the European Union Public Health Programme is available at: http://ec.europa.eu/health/index_en.htm

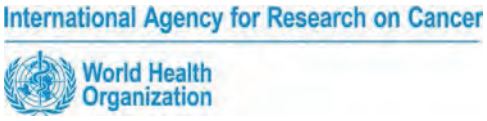
Cataloguing information can be found at the end of this publication.

Luxembourg: Office for Official Publications of the European Union, 2013

ISBN 978-92-79-32970-8

Printed in Belgium

Printed on white chlorine-free paper



European guidelines for quality assurance in breast cancer screening and diagnosis

Fourth Edition

Supplements

Editors
N. Perry
M. Broeders
C. de Wolf
S. Törnberg
R. Holland
L. von Karsa

N. Perry
The London Breast Institute
Princess Grace Hospital
42 Nottingham Place
London W1U 5NY / United Kingdom

M. Broeders
National Expert and Training Centre for Breast Cancer Screening
EUREF Office
P.O. Box 6873
6503 GJ Nijmegen / The Netherlands

C. de Wolf
Swiss Cancer Screening Federation
Effingerstrasse 40
Postfach 8219
CH-3001 Bern / Switzerland

S. Törnberg
Department of Cancer Screening
Regional Cancer Centre
Karolinska University Hospital
S-17176 Stockholm / Sweden

R. Holland
National Expert and Training Centre for Breast Cancer Screening
EUREF Office
P.O. Box 6873
6503 GJ Nijmegen / The Netherlands

L. von Karsa
Quality Assurance Group
Section of Early Detection and Prevention
International Agency for Research on Cancer
150 cours Albert Thomas
F-69372 Lyon Cedex 08 / France

Address for correspondence
Dr Lawrence von Karsa
Quality Assurance Group
International Agency for Research on Cancer
150 cours Albert Thomas
F-69372 Lyon cedex 08
France

Authors, contributors and editors[1]

Isabel Amendoeira, Serviço de Anatomia Patológica H.S. João and IPATIMUP
Porto, Portugal

Ahti Anttila, Finnish Cancer Registry
Helsinki, Finland

Jean-Pierre Bellocq, University Hospital Strasbourg
Strasbourg, France

Simonetta Bianchi, University of Florence
Florence, Italy

Magdalena Bielska-Lasota, National Institute of Public Health – NIH
Warsaw, Poland

Werner Boecker, University of Münster
Münster, Germany

Bettina Borisch, University of Geneva
Geneva, Switzerland

Hilde Bosmans[2], University Hospitals Leuven
Leuven, Belgium

Mireille Broeders, National Expert and Training Centre for Breast Cancer Screening
Nijmegen, The Netherlands

Birgitte Bruun Rasmussen, Herlev Hospital
Herlev, Denmark

[1] Authors, contributors and editors serve in their individual capacities as experts and not as representatives of their government or any organization with which they are affiliated. Affiliations are provided for identification purposes only. Each expert was asked to disclose pertinent research, employment, and financial interests. Current financial interests and research and employment interests during the past 4 years or anticipated in the future are identified here.
Minor pertinent interests are not listed and include stock valued at no more than US$ 10 000 overall, grants that provide no more than 5% of the research budget of the expert's organization and that do not support the expert's research or position, and consulting or speaking on matters not before a court or government agency that does not exceed 2% of total professional time or compensation. All grants that support the expert's research or position and all consulting or speaking on behalf of an interested party on matters before a court or government agency are listed as significant pertinent interests.

[2] H. Bosmans and the University of Leuven where she works hold shares in *qaelum*, a company established in October 2011 after her work on the digital mammography supplement was complete. The company provides software for quality control in diagnostic radiology and screening. She serves as president of the board and with her spouse has a 16% interest in the company. The value of their shares is estimated at less than 135 000 €.

Grace Callagy, National University of Ireland
Galway, Ireland

Ewa Chmielik, Maria Sklodowska-Curie Memorial Cancer Center
Gliwice, Poland

Alicia Cordoba, Hospital de Navarra
Pamplona, Spain

Gábor Cserni, University of Szeged, Szeged, Hungary &
Bács-Kiskun County Teaching Hospital, Kecskemét, Hungary

David R. Dance, Royal Surrey County Hospital
Guildford, United Kingdom

Peter B. Dean[3], International Agency for Research on Cancer, Lyon, France &
University of Turku, Turku, Finland

Chris de Wolf[4], Swiss Cancer Screening Federation
Bern, Switzerland

Thomas Decker, Dietrich Bonhoeffer Medical Center
Neubrandenburg, Germany

James DeGaetano, Mater dei Hospital
Valletta, Malta

Maria Drijkoningen, University Hospitals Leuven
Leuven, Belgium

Ian Ellis, University of Nottingham
Nottingham, United Kingdom

Daniel Faverly, CMP pathology laboratory and CCR Community Reference Center for cancer screening, Brussels, Belgium

Maria Pia Foschini, University of Bologna
Bologna, Italy

[3] P. B. Dean is employed part-time as a breast cancer screening radiologist for the Finnish national programme.

[4] C. de Wolf is a self-employed consultant to a non-profit organization working in the public interest (Swiss Cancer Screening Federation) and to various public entities in Switzerland (breast screening programmes in the cantons of Bern, Fribourg and Thurgau) that are engaged in the development, implementation and evaluation of population-based breast screening programmes. All income derived from these sources is paid to him through the Agency for Development and Evaluation of Health Policy, a company registered in Geneva, Switzerland with a value of 20 000 CHF, of which he is the sole proprietor and employee.

Snjezana Frkovic-Grazio, University Medical Center Ljubljana
Ljubljana, Slovenia

Dorthe A. Grabau[5], Skaane University Hospital
Lund, Sweden

Patrice Heid, Association ARCADES
Marseille, France

Paivi Heikkila, University of Helsinki and HUSLAB
Helsinki, Finland

Roland Holland, National Expert and Training Centre for Breast Cancer Screening
Nijmegen, The Netherlands

Efrosini Iacovou, Nicosia General Hospital
Nicosia, Cyprus

Jurgen Jacobs, Qaelum
Leuven, Belgium

Jocelyne Jacquemier, Association ARCADES
Marseille, France

Handan Kaya, Marmara University School of Medicine
Istanbul, Turkey

Janina Kulka, Semmelweis University
Budapest, Hungary

Manuela Lacerda, Instituto Português de Oncologia
Coimbra, Portugal

Barbara Lazzari, Azienda USL 3
Pistoia, Italy

Hans Lelivelt, National Expert and Training Centre For Breast Cancer Screening
Nijmegen, The Netherlands

Inta Liepniece-Karele, Riga Eastern Clinical University Hospital
Riga, Latvia

[5] D. Grabau currently collaborates part-time with researchers at the University of Lund on studies of HER2 in breast cancer patients that have been partially financed by Roche. The total funding provided by Roche to Lund University for this research was less than 50 000 € and ended in 2010.

Elsebeth Lynge[6], University of Copenhagen
Copenhagen, Denmark

Nicholas Marshall, University Hospital Leuven
Leuven, Belgium

Jose-Maria Martinez-Penuela, Hospital de Navarra
Pamplona, Spain

Nicholas Perry[7], The London Breast Institute
London, United Kingdom

Maja Primic Zakelj, Institute of Oncology
Ljubljana, Slovenia

Cecily Quinn, St. Vincent's University Hospital
Dublin, Ireland

Fritz Rank[†], Copenhagen University Hospital
Copenhagen, Denmark

Peter Regitnig, Medical University of Graz
Graz, Austria

Angelika Reiner, Danube Hospital
Vienna, Austria

Anna Sapino, University of Torino
Torino, Italy

Stephan Schopphoven, Referenzzentrum Mammographie Südwest
Marburg, Germany

Nereo Segnan[8], CPO Piemonte and City of Health and Science of Turin
Turin, Italy

Eero Suonio, International Agency for Research on Cancer
Lyon, France

[6] E. Lynge is undertaking a comparative study of new-generation HPV tests, involving collaboration with Roche Diagnostics A/S, Genomica S.A.U., Qiagen Gaithersburg Ltd., and GenProbe Inc., and has served as an unpaid advisor GenProbe and Norchip.

[7] N. Perry is employed part-time as director of the London Breast Institute at the Princess Grace Hospital, in London, UK, a privately owned and operated health care facility.

[8] N. Segnan has received consultancy fees of less than 1 500 € from Roche Diagnostics Ltd for his participation at an Advisory Board Meeting on CRC screening.

Martin Thijssen, National Expert and Training Centre For Breast Cancer Screening
Nijmegen, The Netherlands

Sven Törnberg, Karolinska University Hospital
Stockholm, Sweden

Tibor Tot, Central Hospital
Falun, Sweden

Johannes J.M. van Delden, Utrecht University
Utrecht, The Netherlands

Paul J. Van Diest, University Medical Center
Utrecht, The Netherlands

Ruben Van Engen, National Expert and Training Centre For Breast Cancer Screening
Nijmegen, The Netherlands

Chantal Van Ongeval, University Hospital Leuven
Leuven, Belgium

Zsuzsanna Varga, University Hospital Zurich
Zurich, Switzerland

Lawrence von Karsa, International Agency for Research on Cancer
Lyon, France

Clive Wells, University College London
London, United Kingdom

Jelle Wesseling, The Netherlands Cancer Institute
Amsterdam, The Netherlands

Kenneth C. Young, Royal Surrey County Hospital
Guildford, United Kingdom

Vassiliki Zolota, University of Patras
Patras, Greece

Enrique Zozaya-Alvarez, Hospital de Navarra
Pamplona, Spain

Preface

The completion of the Supplements to the Fourth Edition of the European Guidelines for quality assurance in breast cancer screening and diagnosis is a further milestone towards high quality cancer screening and diagnosis.

Cancers may be cured, or the prospects of cure greatly increased, if they are detected early. Well organized, nationwide screening programmes therefore have a huge potential to reduce cancer mortality. There have been considerable achievements in cancer screening, as a result of coordinated work at EU level to support Member States in the implementation of national cancer screening programmes.

This publication is a good illustration of the role the EU can play to improve the health of its citizens. The Supplements to the Fourth Edition of the European Guidelines for quality assurance in breast cancer screening and diagnosis take into account recent developments and will improve quality assurance before the publication of the next edition. They will help to give women a better chance of receiving timely treatment as a result of early diagnosis of breast cancer.

Cancer has been a priority issue for EU public health policy for more than 25 years and will remain high on the European agenda. At this juncture, I would like to thank all partners and contributors involved in this project for their valuable contribution to this publication which moves us an important step further down the road towards the common goal to improve public health throughout the European Union.

John-F. Ryan
Acting Director,
Public Health, DG Health and Consumers,
European Commission
Brussels

Foreword

The fourth edition of the European guidelines for quality assurance in breast cancer screening and diagnosis was published by the European Commission in 2006. The present supplements to the fourth edition have been produced by the same groups of experts originally established under the Europe Against Cancer programme that have developed and updated the guidelines since the early 1990s. Over the years, the scope and the depth of the multidisciplinary guidelines have expanded, and recommendations and protocols have been updated to keep pace with developments in the field. Many of the experts who contributed to the early editions of the guidelines are no longer active or have moved on to other endeavours, and in presenting the current supplements we wish to thank and pay tribute to the dedication of all current and previous contributors.

The present supplements are appearing at a time of transition. The European Parliament has called for European Union (EU)-wide accreditation of breast centres to ensure that all women in the EU have access to high-quality breast services. The European Commission has allocated funding to develop and pilot a voluntary accreditation scheme and to prepare the new, fifth edition of the guidelines that will be needed for the accreditation project. The Joint Research Centre of the European Commission has been commissioned to coordinate these key efforts.

The present supplements lay a cornerstone for a new, completely revised fifth edition of the guidelines. Their availability should not delay the important efforts at the European level to produce the fifth edition. The editors would also like to point out that for the foreseeable future dedicated, sustainable support is also required for the pan-European exchange of experience and collaboration in training, monitoring and evaluation, and other areas of quality assurance initiated in the European Breast Cancer Screening Network established under the former Europe Against Cancer programme. Such collaborative efforts will be essential to ensure that in the future breast screening and diagnostic services fulfilling the European standards are available to all women in the EU.

Nicholas Perry
The London Breast Institute
Princess Grace Hospital
London

Sven Törnberg
Department of Cancer Screening
Regional Cancer Centre
Karolinska University Hospital I
Stockholm

Mireille Broeders
National Expert and Training Centre For Breast Cancer Screening / EUREF Office
Nijmegen

Roland Holland
National Expert and Training Centre for Breast Cancer Screening / EUREF Office
Nijmegen

Chris de Wolf
Swiss Cancer Screening Federation
Fribourg

Lawrence von Karsa
Quality Assurance Group
Section of Early Detection and Prevention
International Agency for Research on Cancer
Lyon

Table of contents

S2 Pathology update

Executive summary

Authors
N. Perry
M. Broeders
C. de Wolf
S. Törnberg
R. Holland
L. von Karsa

Authors
N. Perry, United Kingdom
M. Broeders, The Netherlands
C. de Wolf, Switzerland
S. Törnberg, Sweden
R. Holland, The Netherlands
L. von Karsa, IARC

Declarations of interest
Interests of N. Perry and C. de Wolf are reported on pages VII and V respectively.

Disclaimer
The views expressed in this document are those of the authors and do not necessarily reflect the official position of the European Commission. Neither the European Commission nor any other organization, nor any person acting alone or on behalf of others can be held responsible for any use that may be made of the information in this document.

Acknowledgements
Financial support was provided by the European Union Public Health Programme (Project no. 2006322, European Cooperation on Development and Implementation of Cancer Screening and Prevention Guidelines [ECCG]).

Please cite this publication as follows
Perry N, Broeders M, de Wolf C, Törnberg S, Holland R, von Karsa L (2013). Executive summary. In: *European guidelines for quality assurance in breast cancer screening and diagnosis. Fourth edition*, *Supplements*. Perry N, Broeders M, de Wolf C, Törnberg S, Holland R, von Karsa L (eds.). European Commission, Office for Official Publications of the European Union, Luxembourg, pp. XIV–XX.

Corresponding author
L. von Karsa
Quality Assurance Group
Section of Early Detection and Prevention
International Agency for Research on Cancer
150 cours Albert Thomas
F-69372 Lyon cedex 08
France
Karsal@iarc.fr
QAS@iarc.fr

For women in Europe, breast cancer is currently the most common cancer and the most common cause of cancer-induced death. According to the most recently available estimates from cancer registry data, more than 330 000 new cases and nearly 90 000 deaths due to the disease were reported in the 27 Member States of the European Union (EU) in 2008 (Ferlay et al., 2010). Nine out of 10 of those deaths were in women aged 50 years or older, making breast cancer a leading cause of age-related mortality in women. As pointed out in the fourth edition of the European guidelines for quality assurance in breast cancer screening and diagnosis, published in 2006 (Perry et al., 2006; see also Perry et al., 2008), systematic early detection through screening, effective diagnostic pathways and optimal treatment can significantly lower current breast cancer mortality rates and reduce the burden of disease in the population. For these benefits to be obtained, high-quality services are essential. These may be achieved through the underlying basic principles of training, specialization, volume levels, multidisciplinary teamwork, the use of set targets and performance indicators and audit. Ethically, these principles should be regarded as applying equally to symptomatic diagnostic services and screening.

Since the fourth edition was published, the importance of continued efforts at the European level to implement and improve breast, cervical and colorectal cancer screening programmes in the EU Member States has been confirmed by resolutions of the European Parliament (European Parliament, 2006; European Parliament, 2008a; European Parliament, 2008b) and conclusions of the Council of the EU (Council of the European Union, 2008; Council of the European Union, 2010). A new European Partnership for Action Against Cancer (www.epaac.eu) has been established by the European Commission, and a third set of comprehensive European quality assurance guidelines, which deals with colorectal cancer screening, has been developed by experts and published by the European Commission (Segnan et al., 2010). These efforts underline the current consensus on the importance of continuously updating the comprehensive evidence-based standards and recommendations for best practice provided in the European guidelines.

A full revision of the European guidelines for quality assurance in breast cancer screening and diagnosis is now under way, and the new, fifth edition is planned for publication in the coming years. In the meantime, these supplements to the current, fourth edition have been developed to address those fields that the editors felt warranted an update of the standards, protocols and best practice recommendations. At a meeting of the editorial board in November 2009 in London, it was agreed that the supplements would be prepared by the same groups of experts engaged in the development of the fourth edition and according to the same methodology. The authors and editors welcome comments and feedback on these documents.

Digital mammography update

The first supplement (Digital mammography update) responds to the rapid technological development that has accompanied a wide increase in the use of digital imaging in mammography screening and diagnosis since the fourth edition was published. It has been prepared largely by the experts in the European Reference Organisation for Quality Assured Breast Screening and Diagnostic Services (EUREF) Physico-Technical Steering Group, who have been recruited from European medical physics quality assurance teams experienced in guideline development and in implementation of population-based mammography screening programmes. Part 1 of the first supplement (Acceptance and constancy testing; van Engen et al., 2013a) updates the digital mammography part of the European protocol for the quality control of the physical and technical aspects of mammography screening (van Engen et al., 2006) in the fourth edition (Perry et al., 2006). The authors indicate the changes compared with the version of the digital mammography protocol in the fourth edition, and in some sections further guidance and clarification is provided. It is therefore essential to consult the original text and the supplement together.

Only a few aspects of the updates to the digital mammography protocol can be highlighted here. Of prime importance are the changes to the system requirements that any mammography equipment must fulfil. They include, in addition to other elements, a fully automatic exposure control (AEC). A

system with solely manual exposure control (in which the user has to define anode material, filter, tube voltage and dose) is not acceptable; neither is a system with semi-automated exposure control (that adapts dose according to breast transparency, but still requires the user to define anode material, filter and tube voltage). For quality control purposes it should be possible to output unprocessed images in DICOM format from the acquisition workstation or computed radiography (CR) reader.

Part 2 of the first supplement (European type testing; van Engen et al., 2013b) adds a new dimension to the quality assurance philosophy anchored in the fourth edition, namely type testing of digital imaging systems. The need for a European type testing protocol based on the most recent version of the European guidelines became apparent during the preparation of the digital mammography update dealing with acceptance and constancy testing. Equipment users and manufacturers pointed out that an additional European protocol would facilitate avoidance of potential quality problems before equipment is delivered to users. The editorial board of the fourth edition of the guidelines therefore requested that the experts in the EUREF Physico-Technical Steering Group also develop a procedure and protocol suitable for standardizing type testing for mammography systems at the European level.

The standards and procedures presented in Part 2 of the first supplement take into account the updates to the digital mammography physico-technical protocol that are reported in Part 1. The European type testing protocol currently aims to verify whether imaging systems of a given type or brand are fundamentally capable of fulfilling the acceptance criteria of the European physico-technical protocol. Furthermore, guidelines are provided on best practice in controlling dose and (clinical) image quality. After a successful type test, individual mammography units of the same type or brand still need to undergo an acceptance test before clinical use. Passing the type test merely guarantees that the system is in principle capable of meeting the requirements of the European physico-technical protocol; the type test report may also provide suggestions for optimal use, or may specify conditions to be avoided in practice.

The chapter on the physico-technical aspects of quality control in the fourth edition was originally developed for acceptance and constancy testing and not for type testing. Therefore, some differences exist between Part 1 and Part 2 of the first supplement; these differences are reported in Part 2 of the supplement. The new European protocol for type testing (Part 2) has been developed for digital mammography (direct radiography and computed radiography) systems. In the future, the protocol may be expanded to include type testing of image-processing algorithms, workstations and film digitizers.

The preparation of the new and revised standards, protocols and recommendations in the first supplement was coordinated by the EUREF office based at the National Expert and Training Centre for Breast Cancer Screening in Nijmegen, The Netherlands. It involved extensive planning, empirical measurements and numerous scientific studies in the work package on digital mammography in the project 'European Cooperation on Development and Implementation of Cancer Screening and Prevention Guidelines (ECCG)', in which key quality assurance centres in Belgium, France, Germany, Italy, The Netherlands and the United Kingdom collaborated. Experience in the practical application of the new and further developed standards was gathered in site visits and training courses attended by medical physicists and technicians from several EU Member States, including those that acceded to the EU in 2004 and 2007. This experience fed into the final formulation of the digital mammography supplements. The highly effective professional and scientific cooperation in the project was facilitated by the collaborative platform provided by participation of most of the authors in the EUREF Physico-Technical Steering Group.

Pathology update

The second supplement (Pathology update; Wells et al., 2013) deals with several topics in the quality assurance of pathology in breast cancer screening and diagnosis in which problems and practical solutions as well as new techniques and other advances have emerged in recent years. The authors

felt that these topics warranted urgent attention pending a full revision of the pathology chapter. They include the classification of early forms of neoplastic changes in the breast and the differential diagnosis of columnar cell lesions (CCLs), atypical ductal hyperplasia (ADH) and ductal carcinoma in situ (DCIS). Other sections provide an update on the classification of invasive carcinoma and on the diagnosis and differential diagnosis of microinvasive carcinoma. Special attention is paid to the assessment of the axilla, and the updated guidance on axillary dissection and preoperative staging takes into account problems and pitfalls as well as the most recent revision of the TNM system (Sobin et al., 2009). The supplement also includes comprehensive practical recommendations on examination techniques and interpretation of sentinel node biopsy specimens. Best practice in the use of frozen sections, an update on vacuum-assisted needle core biopsy (VANCB) and recommendations on pathological reporting of post-chemotherapy specimens are also provided.

The supplement with updates on pathology has been prepared by the European Working Group on Breast Screening Pathology, which also prepared the chapter on pathology in the fourth edition (Wells et al., 2006) and in previous editions of the guidelines. The group consists of 38 breast pathologists in 23 EU Member States, Switzerland and Turkey and has been led by the editor of the chapter on pathology in the guidelines. All of the pathologists in the group are leaders in the field of breast pathology in their countries. The recommendations in the supplement have been drafted by selected members of the group and have been discussed and agreed upon at twice-yearly meetings held in Hungary, Switzerland, Italy, France, The Netherlands and Finland since the fourth edition was published.

Implementation of population-based cancer screening programmes

In Annex 1, the supplements volume also includes recommendations on implementation of population-based cancer screening programmes that are relevant not only to breast cancer screening but also to cervical and colorectal cancer screening. The Annex consists of a brief statement (von Karsa et al., 2013) and a longer article (Lynge et al., 2012) summarizing the results of a pan-European workshop organized by the European Science Advisory Network for Health (EuSANH, www.EuSANH.eu) in Stockholm in February 2011. At the meeting, a multidisciplinary group of scientists and professionals experienced in implementation and quality assurance of cancer screening programmes and in development of scientific advice on health policy reviewed the available evidence. The conclusions and comprehensive recommendations that emerged identify the major determinants of successful implementation of population-based cancer screening programmes.

The authors emphasize that any policy decision in Europe to implement a cancer screening programme should take into account EU recommendations and guidelines based on the available evidence and the experience in Europe in implementing population-based cancer screening programmes. Key references in this regard are the Recommendation on Cancer Screening of 2 December 2003 of the Council of the EU (Council of the European Union, 2003), the European guidelines for quality assurance in breast, cervical and colorectal cancer screening (Perry et al., 2006; Arbyn et al., 2008; Segnan et al., 2010; see also Perry et al., 2008 and Arbyn et al., 2010) and recent reports dealing with the implementation of cancer screening programmes in the EU (von Karsa et al., 2008; Commission of the European Communities, 2008; Anttila et al., 2009).

The experience in Europe shows that quality-assured implementation of population-based cancer screening programmes generally involves a very long process, extending over 10 or more years and beginning with planning followed by feasibility testing, piloting and gradual rollout of the programme across a country. Successful completion of the quality-assured process requires engagement of civil society, including societal debate; effective cancer registration; good governance providing long-term political commitment; adequate and sustainable resources and competent oversight; autonomous programme management with coordination of numerous stakeholders and activities; organizational development and control of resources (dedicated budget and staff); as well as international collaboration in quality assurance. In a fully established programme, the proportion of the expenditure

devoted to quality assurance should be no less than 10–20%, depending on the scale of the programme. This investment is cost-effective and will save lives.

Computerized information systems and accessible registries are necessary for the management of effective and efficient screening services (e.g. for call and re-call systems and fail-safe procedures in follow-up of participants with abnormal test results). They are also needed to monitor and evaluate the performance and the outcome of the screening programme, e.g. through linkage of individual data on cancer occurrence and morbidity, screening history, diagnosis and treatment.

Once the political decision has been taken to establish a population-based cancer screening programme, a competent programme coordinator should receive the mandate to manage and thereby adequately assure quality throughout the process of programme implementation. This is necessary to fulfil the European principles and guidelines and relevant national standards and regulations. The coordinator should be provided with sufficient organizational and financial resources to effectively manage the screening programme and take further decisions as necessary.

The same conclusions apply to efforts at the European level to support the EU Member States in implementing the Council Recommendation on cancer screening. Their success in making evidence-based screening programmes recommended by the EU available to all citizens who may benefit will depend on the availability of comprehensive, evidence-based guidelines that are updated continuously. Successful implementation of the guidelines across the EU will likewise require engagement of civil society and the availability of an autonomous coordination under competent oversight, with adequate, sustainable resources and long-term political support.

References

Anttila A, Ronco G, Working Group on the Registration and Monitoring of Cervical Cancer Screening Programmes in the European Union; within the European Network for Information on Cancer (EUNICE) (2009). Description of the national situation of cervical cancer screening in the member states of the European Union. *Eur. J. Cancer*, 45:2685–2708.
http://www.ejcancer.info/article/S0959-8049(09)00576-0/abstract

Arbyn M, Anttila A, Jordan J, Ronco G, Schenck U, Segnan N, Wiener H, Herbert A, von Karsa L (2010). European guidelines for quality assurance in cervical cancer screening, second edition – summary document. *Ann. Oncol.*, 21:448–458.

Arbyn M, Anttila A, Jordan J, Schenck U, Ronco G, Segnan N, Wiener H, Herbert A, Daniel J, von Karsa L (eds.) (2008). *European guidelines for quality assurance in cervical cancer screening. Second edition*. European Commission, Office for Official Publications of the European Communities, Luxembourg.
http://bookshop.europa.eu/is-bin/INTERSHOP.enfinity/WFS/EU-Bookshop-Site/en_GB/-/EUR/ViewPublication-Start?PublicationKey=ND7007117

Commission of the European Communities (2008). Report from the Commission to the Council, the European Parliament, the European Economic and Social committee and the Committee of the Regions – Implementation of the Council Recommendation of 2 December 2003 on cancer screening (2003/878/EC) Brussels, Report no. COM(2008) 882 final.
http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=COM:2008:0882:FIN:EN:PDF

Council of the European Union (2003). Council Recommendation of 2 December 2003 on cancer screening (2003/878/EC). *Off. J. Eur. Union* L 327:34–38.
http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2003:327:0034:0038:EN:PDF

Council of the European Union (2008). Council Conclusions on reducing the burden of cancer. 2876th Employment, Social Policy, Health and Consumer Affairs Council meeting. Luxembourg, 10 June 2008.

Brussels, Belgium, Press Office of the Council of the European Union. 10414/08 (Presse 166), http://www.consilium.europa.eu/ueDocs/cms_Data/docs/pressData/en/lsa/101031.pdf

Council of the European Union (2010). Council Conclusions on Action Against Cancer, 3032nd Council meeting. General affairs. Press release 236, 13420/1/10 REV 1, Brussels, Belgium, 13 September 2010, Press Office of the Council of the European Union. http://www.consilium.europa.eu/press/press-releases/latest-press-releases/newsroomrelated?lang=en&bid=72&grp=17287&id=1851

European Parliament (2006). European Parliament Resolution on Breast Cancer in the Enlarged European Union. P6_TA(2006)0449.

European Parliament (2008a). European Parliament resolution of 10 April on combating cancer in the enlarged European Union. P6_TA-PROV(2008)0121.

European Parliament (2008b). European Parliament resolution of 9 October 2008 on 'Together for Health: A Strategic Approach for the EU 2008-2013'. P6_TA(2008)0477.

Ferlay J, Shin HR, Bray F, Forman D, Mathers C, Parkin DM (2010). GLOBOCAN 2008, Cancer Incidence and Mortality Worldwide: IARC CancerBase No. 10 [Internet] International Agency for Research on Cancer, Lyon, France. http://globocan.iarc.fr/

Lynge E, Törnberg S, von Karsa L, Segnan N, van Delden JJM (2012). Determinants of successful implementation of population-based cancer screening programmes. *Eur. J. Cancer*, 48:743–748.

Perry N, Broeders M, de Wolf C, Törnberg S, Holland R, von Karsa L (2008). European guidelines for quality assurance in breast cancer screening and diagnosis, Fourth edition – Summary document. *Ann Oncol.*, 19:614–622.

Perry N, Broeders M, de Wolf C, Törnberg S, Holland R, von Karsa L, Puthaar E (eds.) (2006). *European guidelines for quality assurance in breast cancer screening and diagnosis. Fourth edition*. European Commission, Office for Official Publications of the European Communities, Luxembourg.

Segnan N, Patnick J, von Karsa L (eds.) (2010). *European guidelines for quality assurance in colorectal cancer screening and diagnosis. First edition*. European Commission, Office for Official Publications of the European Union, Luxembourg.

Sobin LH, Gospodarowicz MK, Wittekind C (eds.) (2009). *TNM Classification of Malignant Tumours*, 7th edition. Wiley-Blackwell.

van Engen RE, Bosmans H, Dance DR, Heid P, Lazzari B, Marshall N, Schopphoven S, Thijssen M, Young KC (2013a). Digital mammography update. European protocol for the quality control of the physical and technical aspects of mammography screening. S1, Part 1: Acceptance and constancy testing. In: *European guidelines for quality assurance in breast cancer screening and diagnosis. Fourth edition, Supplements.* Perry N, Broeders M, de Wolf C, Törnberg S, Holland R, von Karsa L (eds.). European Commission, Office for Official Publications of the European Union, Luxembourg, pp. 1–54.

van Engen RE, Bosmans H, Heid P, Lazzari B, Schopphoven S, Thijssen M, Young KC (2013b). Digital mammography update. European protocol for the quality control of the physical and technical aspects of mammography screening. S1, Part 2: European type testing. In: *European guidelines for quality assurance in breast cancer screening and diagnosis. Fourth edition, Supplements*. Perry N, Broeders M, de Wolf C, Törnberg S, Holland R, von Karsa L (eds.). European Commission, Office for Official Publications of the European Union, Luxembourg, pp. 55–71.

van Engen RE, Young KC, Bosmans H, Thijssen M (2006). European protocol for the quality control of the physical and technical aspects of mammography screening. Chapter 2b: Digital mammography. In: *European guidelines for quality assurance in breast cancer screening and diagnosis. Fourth edition.* Perry N, Broeders M, de Wolf C, Törnberg S, Holland R, von Karsa L (eds.). European Commission, Office for Official Publications of the European Communities, Luxembourg, pp. 105–165.

von Karsa L, Anttila A, Primic Žakelj M, de Wolf C, Bielska-Lasota M, Törnberg S, Segnan N (2013). Stockholm statement on successful implementation of population-based cancer screening

programmes. Annex 1a. In: *European guidelines for quality assurance in breast cancer screening and diagnosis. Fourth edition, Supplements*. Perry N, Broeders M, de Wolf C, Törnberg S, Holland R, von Karsa L (eds.). European Commission, Office for Official Publications of the European Union, Luxembourg, pp. 123–128.

von Karsa L, Anttila A, Ronco G, Ponti A, Malila N, Arbyn M, Segnan N, Castillo-Beltran M, Boniol M, Ferlay J, Hery C, Sauvaget C, Voti L, Autier P (2008). *Cancer screening in the European Union, Report on the implementation of the Council Recommendation on cancer screening – First Report*. European Commission, Office for Official Publications of the European Communities, Luxembourg. http://ec.europa.eu/health/ph_determinants/genetics/documents/cancer_screening.pdf

Wells CA, Amendoeira I, Apostolikas N, Bellocq JP, Bianchi S, Boecker W, Borisch B, Bussolati G, Connolly CE, Cserni G, Decker T, Dervan P, Drijkoningen M, Ellis IO, Elston CW, Eusebi V, Faverly DR, Heikkilä P, Holland R, Kerner H, Kulka J, Jacquemier J, Lacerda M, Martinez-Penuela J, de Miguel C, Nordgren H, Peterse JL, Rank F, Regitnig P, Reiner A, Sapino A, Sigal-Zafrani B, Tanous AM, Thorstenson S, Zozaya E (2006). Chapter 6: Quality assurance guidelines for pathology. In: *European guidelines for quality assurance in breast cancer screening and diagnosis. Fourth edition*. Perry N, Broeders M, de Wolf C, Törnberg S, Holland R, von Karsa L (eds.). European Commission, Office for Official Publications of the European Union, Luxembourg, pp. 219–311.

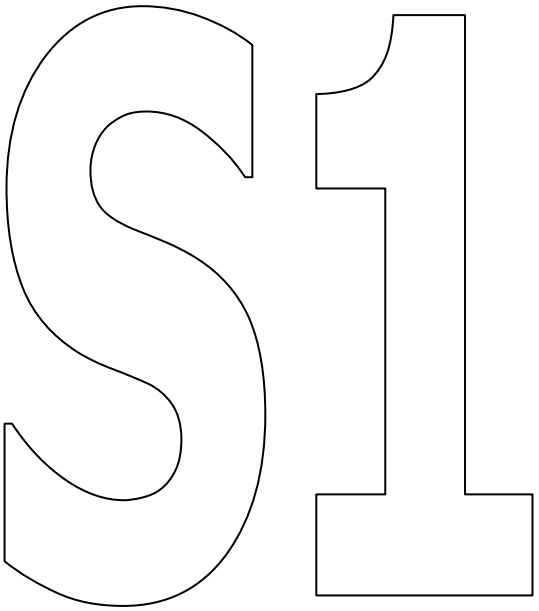
Wells CA, Amendoeira I, Bellocq JP, Bianchi S, Boecker W, Borisch B, Bruun Rasmussen B, Callagy GM, Chmielik E, Cordoba A, Cserni G, Decker T, DeGaetano J, Drijkoningen M, Ellis IO, Faverly DR, Foschini MP, Frković-Grazio S, Grabau D, Heikkilä P, Iacovou E, Jacquemier J, Kaya H, Kulka J, Lacerda M, Liepniece-Karele I, Martinez-Penuela J, Quinn CM, Rank F†, Regitnig P, Reiner A, Sapino A, Tot T, Van Diest PJ, Varga Z, Wesseling J, Zolota V, Zozaya-Alvarez E (2013). S2: Pathology update. Quality assurance guidelines for pathology. In: *European guidelines for quality assurance in breast cancer screening and diagnosis. Fourth edition, Supplements*. Perry N, Broeders M, de Wolf C, Törnberg S, Holland R, von Karsa L (eds.). European Commission, Office for Official Publications of the European Union, Luxembourg, pp. 73–120.

European guidelines for quality assurance in breast cancer screening and diagnosis

Fourth Edition

Supplement

Digital mammography update



European protocol for the quality control of the physical and technical aspects of mammography screening

Part 1

Acceptance and constancy testing

Part 2

European type testing

Introduction

The aim of the European Reference Organisation for Quality Assured Breast Screening and Diagnostic Services (EUREF) is to improve the quality of mammography in Europe and to disseminate knowledge on high-quality breast imaging. Within this context, EUREF has produced the European guidelines for quality assurance in breast cancer screening and diagnosis (Perry et al., 2001; de Wolf & Perry, 1996). The current version of these guidelines is the fourth edition, published by the European Commission in 2006 (Perry et al., 2006). It includes in Chapter 2 the European protocol for the quality control of the physical and technical aspects of mammography screening, an internationally recognized standard that is widely used in Europe and worldwide (van Engen et al., 2006a; van Engen et al., 2006b). In Part 2b of this protocol, the requirements for digital mammography imaging systems are defined and standards and protocols relevant to acceptance and constancy testing of equipment are described (van Engen et al., 2006b). Due to the rapid developments in imaging technology in recent years and further experience with digital mammography systems, some updating and extension of the scope of the protocol is warranted and is supplied here in the form of a supplement (S1: Digital mammography update).

Part 1 of the supplement (Acceptance and constancy testing; van Engen et al., 2013a) deals with relevant aspects of acceptance and constancy testing of digital mammography equipment that are addressed in Chapter 2b in the fourth edition. During the preparation of these aspects, equipment users and manufacturers expressed the need for type testing of digital imaging equipment based on the fourth edition. The availability of an additional protocol for type testing enables manufacturers to avoid potential quality problems before equipment is delivered to the user. The editorial board of the fourth edition therefore requested that the experts in the EUREF Physico-Technical Steering Group also develop a procedure and protocol suitable for standardizing type testing for mammography systems at the European level. The developed standards are presented in Part 2 of the supplement (European type testing; van Engen et al., 2013b); they take into account the updates presented in Part 1 of the supplement.

The type testing described in Part 2 of this supplement aims to verify whether imaging systems of a given type or brand are fundamentally capable of fulfilling the acceptance criteria of the European physico-technical protocol, and to provide guidelines for best practice in controlling dose and (clinical) image quality. After a successful type test, individual mammography units of the same type or brand still need to undergo an acceptance test before clinical use. Passing the type test merely guarantees that the system is in principle capable of meeting the requirements of the European physico-technical protocol; the type test report may also provide suggestions for optimal use or specify conditions to be avoided in practice.

The European protocol for type testing is currently developed for digital mammography (direct radiography [DR] and computed radiography [CR]) systems. In the future, the protocol may be expanded to include type testing of image processing algorithms, workstations and film digitizers.

This supplement has been prepared by the experts in the EUREF Physico-Technical Steering Group who have been recruited from European medical physics quality assurance teams experienced in guideline development and in implementation of population-based mammography screening programmes. The authors welcome all comments and feedback on this document to improve the standards and protocols. Future updates of the current version will be made available on the web sites of the European Commission (http://ec.europa.eu/health/index_en.htm) and EUREF (www.euref.org).

References

de Wolf C, Perry N (1996). *European guidelines for quality assurance in mammography screening. Second edition*. European Commission, Office for Official Publications of the European Communities, Luxembourg.

Perry N, Broeders M, de Wolf C, Törnberg S, Holland R, von Karsa L (eds.) (2006). *European guidelines for quality assurance in breast cancer screening and diagnosis. Fourth edition*. European Commission, Office for Official Publications of the European Communities, Luxembourg.

Perry N, Broeders M, de Wolf C, Törnberg S, Schouten J (eds.) (2001). *European guidelines for quality assurance in mammography screening. Third edition*. European Commission, Office for Official Publications of the European Communities, Luxembourg.

van Engen RE, Bosmans H, Dance DR, Heid P, Lazzari B, Marshall N, Schopphoven S, Thijssen M, Young KC (2013a). Digital mammography update. European protocol for the quality control of the physical and technical aspects of mammography screening. S1, Part 1: Acceptance and constancy testing. In: *European guidelines for quality assurance in breast cancer screening and diagnosis. Fourth edition, Supplements*. Perry N, Broeders M, de Wolf C, Törnberg S, Holland R, von Karsa L (eds.). European Commission, Office for Official Publications of the European Union, Luxembourg, pp. 1–54.

van Engen RE, Bosmans H, Heid P, Lazzari B, Schopphoven S, Thijssen M, Young KC (2013b). Digital mammography update. European protocol for the quality control of the physical and technical aspects of mammography screening. S1, Part 2: European type testing. In: *European guidelines for quality assurance in breast cancer screening and diagnosis. Fourth edition, Supplements*. Perry N, Broeders M, de Wolf C, Törnberg S, Holland R, von Karsa L (eds.). European Commission, Office for Official Publications of the European Union, Luxembourg, pp 55–71.

van Engen RE, van Woudenberg S, Bosmans H, Young KC, Thijssen M (2006a). European protocol for the quality control of the physical and technical aspects of mammography screening. Chapter 2a: Screen-film mammography. In: *European guidelines for quality assurance in breast cancer screening and diagnosis. Fourth edition.* Perry N, Broeders M, de Wolf C, Törnberg S, Holland R, von Karsa L (eds.). European Commission, Office for Official Publications of the European Communities, Luxembourg, pp. 61–104.

van Engen RE, Young KC, Bosmans H, Thijssen M (2006b). European protocol for the quality control of the physical and technical aspects of mammography screening. Chapter 2b: Digital mammography. In: *European guidelines for quality assurance in breast cancer screening and diagnosis. Fourth edition.* Perry N, Broeders M, de Wolf C, Törnberg S, Holland R, von Karsa L (eds.). European Commission, Office for Official Publications of the European Communities, Luxembourg, pp. 105–165.

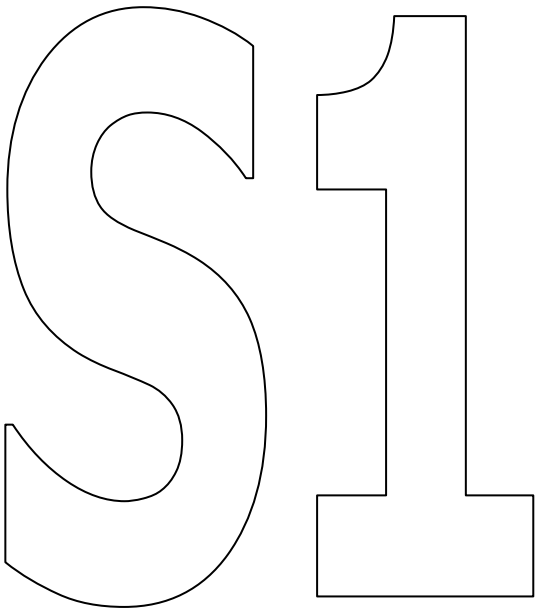
Corresponding author
R. van Engen
EUREF office
National Expert and Training Centre for Breast Cancer Screening
Radboud University Nijmegen Medical Centre
P.O. Box 6873
6503 GJ Nijmegen
The Netherlands
R.vanEngen@euref.org
info@euref.org

European guidelines for quality assurance in breast cancer screening and diagnosis

Fourth Edition

Supplement

Digital mammography update



European protocol for the quality control of the physical and technical aspects of mammography screening

Part 1

Acceptance and constancy testing

Authors
R. van Engen
H. Bosmans
D. Dance
P. Heid
B. Lazzari
N. Marshall
S. Schopphoven
M. Thijssen
K. Young

Authors
R. van Engen, The Netherlands*
H. Bosmans, Belgium*
D. Dance, United Kingdom
P. Heid, France*
B. Lazzari, Italy*
N. Marshall, Belgium
S. Schopphoven, Germany*
M. Thijssen, The Netherlands*
K. Young, United Kingdom*

* Member of the EUREF Physico-Technical Steering Group

Declarations of interest
An interest of H. Bosmans is reported on page IV.

Disclaimer
The views expressed in this document are those of the authors and do not necessarily reflect the official position of the European Commission. Neither the European Commission nor any other organization or any individual may be held responsible for any use that may be made of the information contained herein.

Acknowledgements
Financial support was provided by the European Union Public Health Programme (Project no. 2006322, European Cooperation on Development and Implementation of Cancer Screening and Prevention Guidelines [ECCG]).

The comments and suggestions received from numerous colleagues in the European Cancer Network for Screening and Prevention and the European Federation of Organisations for Medical Physics (EFOMP) are gratefully acknowledged.

The comments and suggestions received from the following companies are gratefully acknowledged: Adani, Agfa HealthCare, Barco, Carestream Health, Eizo, General Electric, Hologic, Konica Minolta, Philips Healthcare, Planmed, Sectra and Siemens.

Please cite this publication as follows
van Engen RE, Bosmans H, Dance DR, Heid P, Lazzari B, Marshall N, Schopphoven S, Thijssen M, Young KC (2013). Digital mammography update. European protocol for the quality control of the physical and technical aspects of mammography screening. S1, Part 1: Acceptance and constancy testing. In: *European guidelines for quality assurance in breast cancer screening and diagnosis. Fourth edition, Supplements*. Perry N, Broeders M, de Wolf C, Törnberg S, Holland R, von Karsa L (eds.). European Commission, Office for Official Publications of the European Union, Luxembourg, pp. 1–54.

Corresponding author
R. van Engen
EUREF office
National Expert and Training Centre for Breast Cancer Screening
Radboud University Nijmegen Medical Centre
P.O. Box 6873
6503 GJ Nijmegen
The Netherlands
R.vanEngen@euref.org
info@euref.org

Table of contents

This supplement contains **updates of the following sections and appendices** in Part 2b: Digital mammography in the European protocol for the quality control of the physical and technical aspects of mammography screening (van Engen et al., 2006b) originally published in the fourth edition of the European guidelines for quality assurance in breast cancer screening and diagnosis. It is therefore essential to consult the original text and the supplement together.

List of figures

List of tables

2b.1 Introduction

The European guidelines for quality assurance in breast cancer screening and diagnosis (Perry et al., 2006) include as Chapter 2 the European protocol for the quality control of the physical and technical aspects of mammography screening. In Part 2b of this protocol, the requirements for digital mammography imaging systems are defined (van Engen et al., 2006b). Due to the rapid developments in imaging technology in recent years and further experience with digital mammography systems, some updating of the protocol is required and is supplied here in the form of a supplement.

The authors welcome all comments and feedback on this document to improve the protocol. Updates of the current version will be made available on the EUREF web site (www.euref.org).

In this supplement, only the changes compared with the original version of the quality control protocol in the fourth edition are indicated. Items that are not mentioned in this supplement remain unchanged. In some sections, further guidance and clarification is provided.

2b.1.2 System requirements

A mammography unit should incorporate a fully automatic exposure control (AEC). A system with solely manual exposure control (in which the user has to define anode material, filter, tube voltage and dose) is not acceptable. Neither is a system with semi-automated exposure control (in which the user has to define anode material, filter and tube voltage but adapts dose according to breast transparency).

It should be possible to output unprocessed images in DICOM format from the acquisition workstation or CR reader for quality control purposes.

In addition to the parts of the DICOM standard mentioned in the fourth edition, mammography equipment should fulfil the recommendations of the IHE Mammography Image profile (MAMMO) and the IHE Mammography Acquisition Workflow profile (MAWF). The equipment must be CE marked and must be sold as 'usable for mammography purposes'. DICOM communication between X-ray unit and CR reader (complete transfer of exposure parameters) is recommended.

The acceptable level is the minimum acceptable level; however, it is recommended that systems operate as far as possible at a standard equal to or better than the achievable level. Slow screen-film systems are capable of operating close to the achievable image quality level. Faster screen-film systems are capable of operating near the acceptable level. For the next edition of the guidelines, the acceptable and achievable levels will be reviewed.

Section 2b.2.4.2, Modulation transfer function (MTF) and noise power spectrum (NPS), of the fourth edition has been moved to Appendix 7. This indicates that it is not recommended to perform these measurements on a routine basis for quality control (QC) purposes. In practice, however, linear system theory metrics are being used in system acceptance testing or in regular QC procedures. Appendix 7 represents an effort to harmonize the methods of determining linear system theory metrics in these cases.

Printer

This section applies when printed films are used for reading by either the first or the second reader (in current and subsequent screening rounds).

To obtain an optimal quality for reading hard-copy films on viewing boxes, requirements for printing mammographic images are:

- The hard-copy film type must be suitable for mammography.
- Printing two (or more) images on the same film is not recommended.
- To be able to print images with sufficient resolution, the pixel pitch of the printer should be similar to (or smaller than) the pixel pitch of the image and must always be ≤ 100 µm.
- All images must be printed directly without any manipulation by the user (contrast, luminosity, etc.) at the (diagnosis) workstation.
- Clinical information must be printed on each film: the name of the patient, laterality, date of exams, etc.

After each film refill in the hard-copy printer, the printer must be calibrated (if this is not done automatically). After a period of inactivity of more than one day, a calibration must be done before printing any diagnostic film.

Different film formats are available on the market. For mammography, only 2 sizes should be used:

- 200 mm × 250 mm (8 × 10 inches)
- 250 mm × 300 mm (10 × 12 inches).

Other formats are not recommended (i.e. 260 mm × 360 mm).

2b.1.5 Definition of terms

AEC sensor area (DR systems)
The area of a DR detector in which the exposure factors for an image are determined during or after a pre-exposure. For some AEC systems, the size and location of the AEC sensor area depends on the tissue (or test object) being imaged.

Dark noise image
An image acquired without exposure to the detector; cover the detector with a sufficiently thick metal plate, e.g. lead or stainless steel, and set the lowest possible tube loading.

Detector surface
The accessible area that is closest to the image receptor plane. Depending on removability, the detector may include e.g. breast support, covers, antiscatter grid.

Linearized pixel value
For images of systems with a non-linear response, the pixel values must be linearized before analysis using the response function of the system or the mathematical relationship provided by the manufacturer. Offset correction and linearization must be performed for noise evaluation and determination of linear system theory metrics. In principle, for some other tests (like the calculation of SDNR) the images need to be linearized. However, for those tests the images are acquired at almost the same dose so that the system can be supposed to be locally linear.

Modulation transfer function (MTF)
The MTF describes the response of a system to a sinusoidal input signal; see Appendix 7.

Noise power spectrum (NPS)
The NPS describes the variance of an image intensity divided among its frequency components; see Appendix 7.

Reference region of interest (ROI)
The size of the reference ROI is 5 mm × 5 mm, instead of 20 mm × 20 mm as originally stated in the fourth edition. The centre of the reference ROI is positioned 60 mm from the chest wall edge and centred laterally.

Standard image
Image made by an exposure of a 45 mm thick polymethyl methacrylate (PMMA) block plus 8 mm spacers covering the whole image receptor. A standard compression of 100 N should be applied, and the exposure should be made in the standard clinical AEC mode.

2b.2 Image acquisition

2b.2.1.1.1 Focal spot size
This measurement is omitted. Potential image quality problems are identified by image quality measurements.

2b.2.1.1.2 Source-to-image distance
This measurement is omitted.

2b.2.1.1.3 Alignment of X-ray field/image area
This measurement is optional.

2b.2.1.1.4 Radiation leakage
This measurement is omitted.

2b.2.1.1.5 Tube output
Tube output is measured only for calculating mean glandular doses and does not have to meet any limiting values. It must be known for all clinically used beam qualities. The number of different beam qualities can be substantial. The parametric approach as published by Robson (2001) allows the tube output to be estimated for any tube voltage in the range 25 kVp–32 kVp from a single measurement at 28 kVp that uses the same target/filter combination.

This is a 2-step procedure. First, the value of A is calculated from equation (1), in which *Output* is the measured air kerma (mGy) at 28 kVp, (kV) is the measured tube voltage and the parameter n is target/filter beam-specific and can be obtained from Table 1.1:

$$Output = A(\mathrm{kV})^n \quad (1)$$

Second, the tube output at another tube voltage (kV) is obtained from the following equation:

$$log_{10}(\text{air kerma}) = n\log_{10}(\mathrm{kV}) + \log_{10}(A) \quad (2)$$

Table 1.1: Calculated values of the parameters *n*, *a* and *b* for a range of target/filter combinations (Robson, 2001)

Target/filter combination	Measured filter thickness	*n*	*a*	*b*
Mo/30 µm Mo	36.1 µm	3.06	–0.000326	0.0273
Mo/25 µm Rh	29.9 µm	3.24	–0.000624	0.0445
Rh/25 µm Rh	29.9 µm	3.03	–0.000514	0.0425
W/50 µm Rh	58.9 µm	1.96	–0.000539	0.0403
Rh/1.0 mm Al	1.20 mm	4.39	–0.00113	0.0909
Mo/1.0 mm Al	1.20 mm	4.23	–0.000775	0.0593

2b.2.1.2.2 Half value layer

Half value layer (HVL) is measured to calculate glandular dose. HVL does not have to meet any requirements. Typical values are given in Table A5.3 in the fourth edition.

For patient dosimetry applications, HVLs of all clinically used beam qualities are required. The number of different beam qualities can be substantial. The parametric approach as published by Robson (2001) allows the HVL to be estimated for any tube voltage in the range 25 kVp–32 kVp from a single measurement at 28 kVp that uses the same target/filter combination.

As with the tube output, this is a 2-step procedure based on parameters from Table 1.1. First, the value of *c* is calculated from the equation

$$HVL = a(\mathrm{kVp})^2 + b(\mathrm{kVp}) + c \qquad (3)$$

in which *HVL* is the measured HVL at 28 kVp, (kVp) is the measured tube voltage and the parameters *a* and *b* are target/filter beam-specific and can be obtained from Table 1.1.

HVLs at other tube voltages can then be calculated from the same equation using *a*, *b* and *c*.

2b.2.1.3.1 Exposure control steps

If exposure control steps are available on a mammography system, then the control step should have known dose increments, e.g. 20% per step, and should be verified.

2b.2.1.3.3 Short-term reproducibility

The short-term reproducibility of the AEC system is calculated by the deviation of the linearized pixel value of 10 exposures in fully automatic mode of the standard test block.

If it is noticed that the system switches between two spectra, release the compression paddle and compress again or use another PMMA thickness (add, e.g., 5 mm PMMA) to force the choice of one single spectrum, and repeat the measurement.

Limiting value Deviations from the mean value of 10 exposures ≤ ±5%, achievable ≤ ±2%.
Frequency Every 6 months.
Equipment Standard test block.

2b.2.1.3.5 Breast thickness and composition compensation

Spacers are used to adjust the height of the compression paddle equal to the height of the compression paddle of the breast thickness with equivalent attenuation as given in Table 1.2. Dance and colleagues have verified that this table is accurate to within 1 mm or 2 mm across the wide range of spectra encountered in digital mammography (Dance et al., 2009).

The spacers should not cover the part of the detector in which exposure factors are determined (AEC sensor area).

Table 1.2: PMMA thickness and equivalent breast thickness in terms of attenuation and glandularity, accurate over a wide range of X-ray spectra (Dance et al., 2009)

PMMA thickness (mm)	Equivalent breast thickness (mm)	Glandularity of equivalent breast (%)
20	21	97
30	32	67
40	45	41
45	53	29
50	60	20
60	75	9
70	90	4

Images should be acquired in fully automatic mode; however, on some systems it may be convenient to mimic the exposure of the fully automatic mode in manual mode. In this case, it should be realized that the pre-exposure image is not used for the actual image on some types of system. This should be taken into account when determining signal-difference-to-noise ratio (SDNR) in manual mode. It should also be known whether the pre-exposure is included in the displayed mAs value in fully automatic mode.

The following method can be used to determine whether the tube load (mAs) of the exposure is included in the tube load indication (mAs) of the fully automatic exposure mode. Determine the relationship between tube load and pixel value (PV) in a reference ROI for a series of manual exposures over the clinical working range. Make an exposure in fully automatic mode. Record beam quality and mAs value. Verify whether the PV in the reference ROI of this image is compatible with the PV of a manual exposure with the same tube load (mAs).

The dimensions of the aluminium object are 10 mm × 10 mm and 0.2 mm thick. The object is positioned as shown in Figures 1.1 and 1.2. The aluminium object is positioned between the two lowest 10 mm thick plates of PMMA.

The ROI within the aluminium object is 5 mm × 5 mm in the middle of the object. The background ROIs are 4 ROIs of 5 mm × 5 mm, on all 4 sides of the aluminium object; see Figure 1.3.

Figure 1.1: Set-up for the breast thickness and composition measurements, top view and 3D view

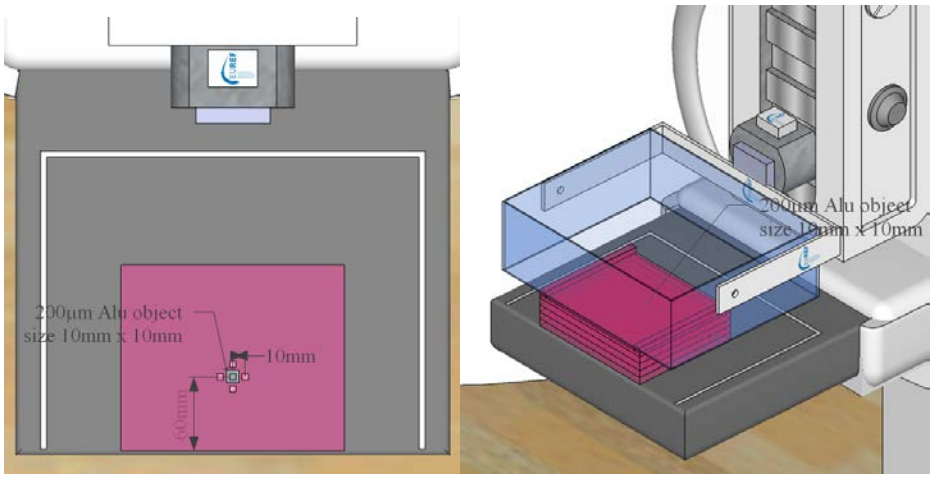


Figure 1.2: Set-up for the breast thickness and composition measurements, front view and side view

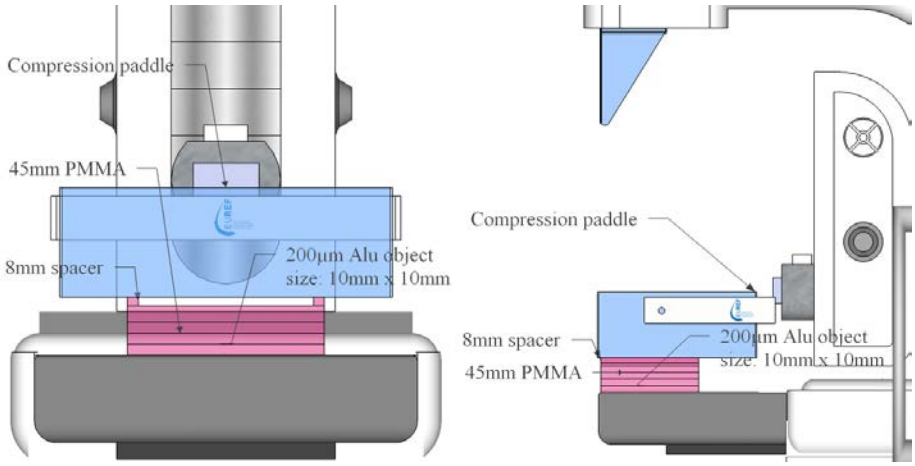
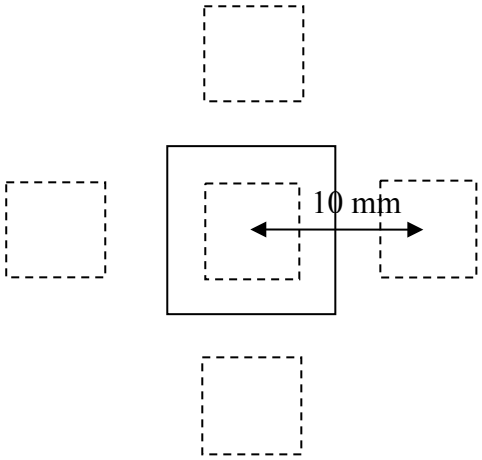


Figure 1.3: Position of the ROIs for calculating SDNR



Calculate PV(background) and the standard deviation SD(background) according to:

$$\mathrm{SD(background)} = \frac{\sum_1^4 \mathrm{SD(ROI_n)}}{4} \tag{4}$$

$$\mathrm{PV(background)} = \frac{\sum_1^4 \mathrm{PV(ROI_n)}}{4} \tag{5}$$

Calculate SDNR of the aluminium object:

$$SDNR = \frac{\mathrm{PV(signal)} - \mathrm{PV(background)}}{\sqrt{\frac{\mathrm{SD(signal)}^2 + \mathrm{SD(background)}^2}{2}}} \tag{6}$$

To apply the standards in the European protocol, the limiting values for SDNR have to be determined. Two methods are described here. The first is a simplified method using threshold gold thickness, and the second involves calculating the threshold contrast for the beam quality used. It is assumed in both methods that the exposure factors for the CDMAM phantom and the 50 mm of PMMA are the same. If for some reason they are not, then a small correction can be applied.

Simplified method using threshold gold thickness

The limiting values for SDNR are calculated using equations (7) and (8). These equations determine the SDNR values necessary to achieve the minimum and achievable threshold gold thickness in the image quality measurements for the 0.1 mm detail size at this thickness, i.e. 50 mm PMMA or equivalent.

$$SDNR_{\mathrm{minimum}} = SDNR_{\mathrm{measured}} \times \frac{Tg_{\mathrm{measured}}}{Tg_{\mathrm{minimum}}} \tag{7}$$

$$SDNR_{\mathrm{achievable}} = SDNR_{\mathrm{measured}} \times \frac{Tg_{\mathrm{measured}}}{Tg_{\mathrm{achievable}}} \tag{8}$$

where $Tg_{measured}$ is the threshold gold thickness for the 0.1 mm detail size predicted for a human observer, and $Tg_{minimum}$ and $Tg_{achievable}$ are the limiting values for threshold gold thickness for the 0.1 mm detail size in the European protocol.

It is estimated that the maximum error in estimating the target SDNR by this method as compared with calculations based on threshold contrasts (see next paragraph) is 4%. Thus, for routine QC this simplified method may be acceptable.

Method using threshold contrast

A more accurate method of defining the limiting values for SDNR involves estimating the contrast of the gold discs for the beam quality used. Such a transformation from gold thickness to contrast is necessary because of the non-linear relationship between these two properties. A simple tool for cal-

culating the limiting values is provided on the EUREF web site, and the method used is described here.

Limiting values for SDNR are calculated using equations (9) and (10). These equations determine the SDNR values necessary to achieve the minimum and achievable threshold gold thickness in the image quality measurements for the 0.1 mm detail size at this thickness, i.e. 50 mm PMMA or equivalent.

$$SDNR_{\text{minimum}} = SDNR_{\text{measured}} \times \frac{Tc_{\text{measured}}}{Tc_{\text{minimum}}} \quad (9)$$

$$SDNR_{\text{achievable}} = SDNR_{\text{measured}} \times \frac{Tc_{\text{measured}}}{Tc_{\text{achievable}}} \quad (10)$$

where

$$Tc_{\text{measured}} = 1 - e^{-\mu_{\text{eff}} \cdot Tg_{\text{measured}}} \quad (11)$$

$$Tc_{\text{minimum}} = 1 - e^{-\mu_{\text{eff}} \cdot Tg_{\text{minimum}}} \quad (12)$$

$$Tc_{\text{achievable}} = 1 - e^{-\mu_{\text{eff}} \cdot Tg_{\text{achievable}}} \quad (13)$$

where $Tc_{measured}$ is the threshold contrast for the 0.1 mm detail size calculated from the measured threshold gold thickness, $Tc_{minimum}$ and $Tc_{achievable}$ are the threshold contrasts calculated from the limiting values for threshold gold thickness for the 0.1 mm detail size in the European protocol and μ_{eff} is the effective attenuation coefficient used to estimate the contrast of gold discs in the CDMAM phantom depending on the beam quality used. A look-up table of values for μ_{eff} is included in the software tool on the EUREF web site.

The methods above determine the limiting values for SDNR for an attenuation equivalent to 50 mm PMMA. The European protocol adjusts $SDNR_{minimum}$ for other thicknesses of PMMA using equation (14) and Table 1.3. The $SDNR_{achievable}$ limit can be applied across other PMMA thicknesses.

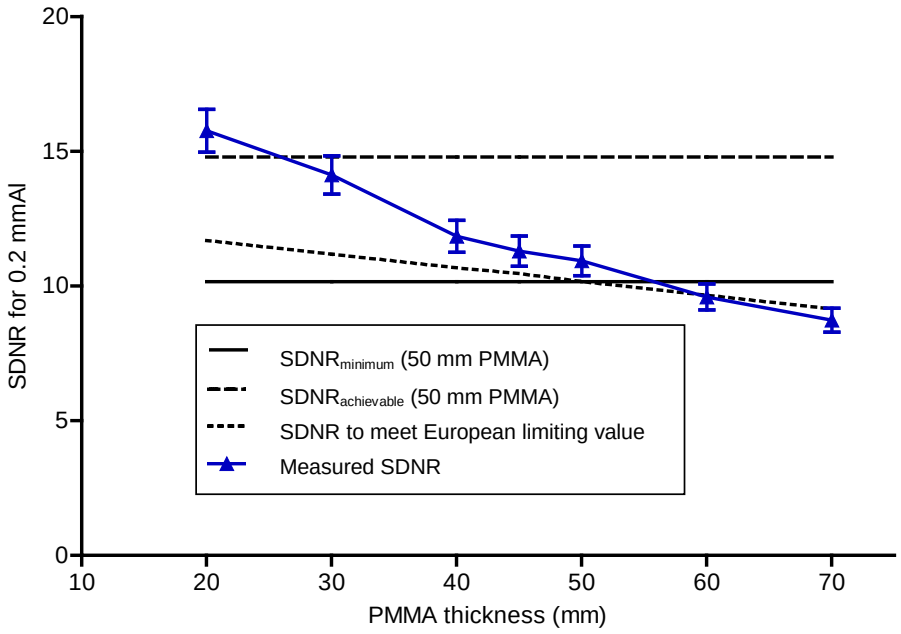
$$SDNR_{\text{minimum}} = SDNR_{\text{measured}} \times \frac{Tg_{\text{measured}}}{Tg_{\text{minimum}}} \times z \quad (14)$$

Table 1.3: z-Factors to calculate $SDNR_{minimum}$ at different thicknesses of PMMA

Thickness of PMMA (mm)	z-factor
20	1.15
30	1.10
40	1.05
45	1.03
50	1.00
60	0.95
70	0.90

To evaluate the AEC performance, the measured SDNR should be plotted against the PMMA thickness and compared with the limiting SDNR values, as shown in Figure 1.4. In this case, the system failed to exceed the minimum acceptable SDNR at the 60 mm and 70 mm thicknesses of PMMA.

Figure 1.4: Example of an evaluation of measured SDNR against target SDNR



Note that in general a change to a higher SDNR will be associated with better image quality if all other factors are unchanged. However, if such a change is associated with a change in image sharpness, the opposite may be true and an investigation of sharpness and image quality should be undertaken.

2b.2.1.3.6 Local dense area (only DR systems)

Most systems measure the attenuation of the imaged object during a pre-exposure. The areas with highest attenuation in the clinically relevant part of the image should determine the exposure factors for imaging. It is required that the signal-to-noise ratio (SNR) in the images be adjusted to the (relatively large) regions with highest density.

Put a 30 mm stack of PMMA plates on the bucky. Put spacers on top of the stack, such that the compression paddle is positioned at a height of 40 mm above the breast holder (compression force can be applied). The spacers should not cover the part of the detector in which exposure factors are determined (AEC sensor area). On the compression paddle, position a first, small PMMA plate, representing a relatively large area with higher density (20 mm × 40 mm, 2 mm thick), in the central part of the detector with its lower edge 50 mm from the chest wall side.[1] (It must be ensured that this plate is within the AEC sensor area. If this is not the case, another position should be chosen.) See Figures 1.5 and 1.6. Make an exposure, and record the exposure factors. Add another small PMMA plate on top of the previous one, and repeat the procedure until 10 small plates have been added.

Measure PV and standard deviation (SD) in the area of extra attenuation (20 mm × 40 mm PMMA plates) with a ROI of 5 mm × 5 mm. Calculate SNR for each image and the average SNR for all images. It should be checked whether the exposure of the images is increased with increasing thickness and whether the extra attenuation is detected. For this, the following value can be used as guidance: the SNR of each image should be within 20% of the average SNR (provisional).

[1] The small PMMA plates may also be positioned in a non-central region but must be positioned within the AEC sensor area.

Guidance	The SNR of each image should be within 20% of the average SNR (provisional).
Frequency	Every 6 months, or after AEC software upgrades.
Equipment	Three 150 mm × 180 mm PMMA plates (10 mm thick), two spacers (10 mm thick), ten 20 mm × 40 mm PMMA plates (2 mm thick).

Figure 1.5: Set-up for the local dense area measurement, top view and 3D view

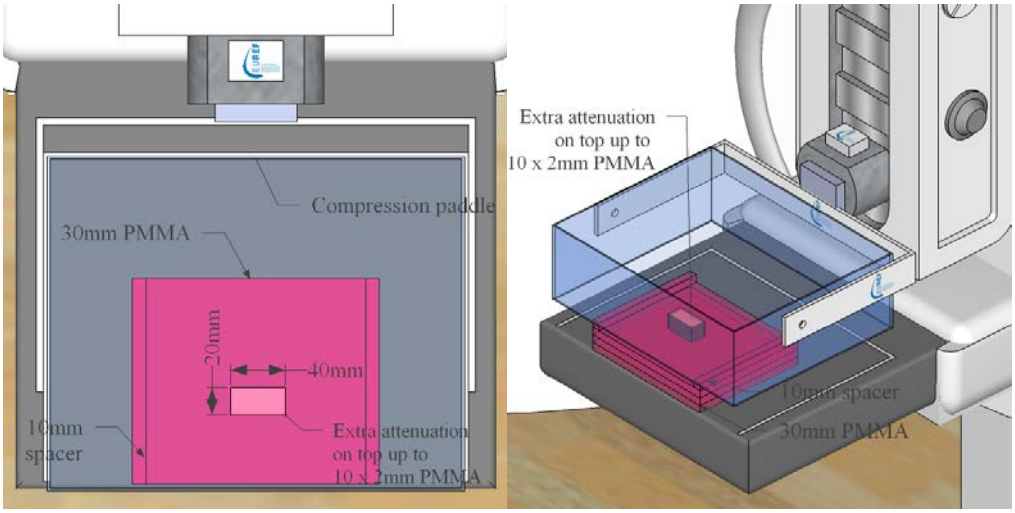
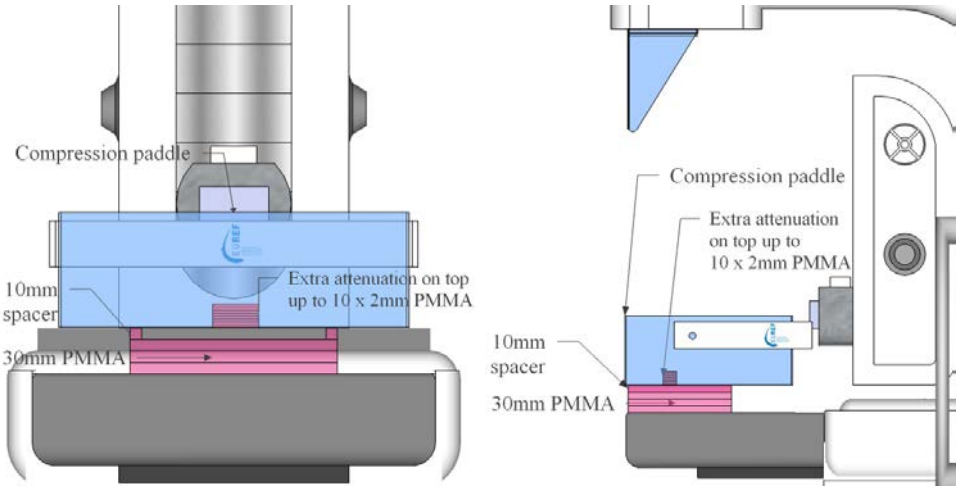


Figure 1.6: Set-up for the local dense area measurement, side views



2b.2.1.5 Antiscatter grid

2b.2.1.5.1 Grid system factor

This measurement is omitted.

2b.2.1.5.2 Grid imaging

This measurement is omitted.

2b.2.2.1.2 Noise evaluation

The scope of this test is the analysis of different noise components to provide additional information on the performance of the imaging system and to optimize troubleshooting in case of potentially decreased image quality. The method described here is reported in greater detail elsewhere (Evans et al., 2002; Borasi et al., 2003; Young et al., 2006b) and particularly in Bouwman et al. (2009).

General requirement: For systems with a non-linear response, the pixel data must be linearized before analysis.

Noise in images can be subdivided into electronic noise, quantum noise and structure noise:

$$SD^2 = k_e^2 + k_q^2 * p + k_s^2 * p^2 \quad (15)$$

SD = standard deviation in reference ROI
k_e = electronic noise coefficient
k_q = quantum noise coefficient
k_s = structure noise coefficient
p = average pixel value in reference ROI.

Electronic noise is assumed to be independent of the exposure level and arises from several sources: dark noise, readout noise and amplifier noise.

Quantum noise arises due to the variations in X-ray flux and (if present) secondary carrier flux.

Structure noise is present due to spatially fixed variations of the gain of an imaging system. The flat-fielding performed in DR systems will largely remove the effects of structure noise. Due to the limited number of images used for the flat-field mask and the associated noise in the mask, some structure noise will still be present.

Remove the compression paddle and all other removable parts (e.g. covers and antiscatter grid) from the X-ray beam. Position a 2 mm thick aluminium plate as close as possible to the X-ray tube.

Set the target/filter combination and the tube voltage that is chosen in fully automatic mode for a 45 mm thick PMMA object plus 8 mm spacers. In manual mode, set the minimum mAs value. Image the aluminium plate. Increase the mAs value and image the plate again. Take a large number of images at different mAs values (e.g. 15 values for acceptance tests, 8 values for subsequent tests) over the whole range of available values with a typical spacing of approximately 40%. Then, take a dark noise image.

It is optional to repeat the measurement for all target/filter combinations, with a clinically relevant tube voltage for each combination.

It is optional to measure or calculate the dose on the detector surface from tube output measurements for all spectra, to be able to use detector air kerma instead of PV in this evaluation.

Analysis steps:

1. Measure PV and SD in the reference ROI.
2. Plot PV against detector dose to determine the response function.
3. Plot SD^2 against PV (or detector dose).

4. Fit the points using equation (15), and determine the noise coefficients.
5. Determine the detector dose range for which quantum noise is the largest noise component.

Option: The calculated noise components can be used to plot PV (or detector dose) against the percentage of the total relative noise for all noise components. In this graph, the magnitude (in %) of all noise components is visualized for the range of PVs (or detector doses).

Limiting value Quantum noise should be the largest noise component for the pixel value range (detector dose range) that is used clinically.
Frequency Every 6 months.
Equipment Aluminium plate (2 mm thick) covering the whole X-ray field (near the tube), appropriate software tools.

2b.2.2.3.1 Image receptor homogeneity

In addition to the test described in the fourth edition, with an ROI size of 10 mm × 10 mm, another ROI, of dimensions 2 mm × 2 mm, is used to calculate variance in each ROI. Equation (16) is used to calculate variance. This parameter is sensitive to the occurrence of artefacts on the image and supports visual artefact analysis by indicating deviations automatically.

$$Var(X) = \frac{\sum_{i=1}^{N}(x_i - \mu)^2}{N-1} \quad (16)$$

More detailed information on the method described here is provided in van Engen et al. (2006a); see also Marshall (2006).

Limiting value The average variance of each ROI should be compared with the average variance of the neighbouring ROIs. If the variance in an ROI is ≥ 30% higher than the variance in neighbouring ROIs, the image should be investigated visually for an artefact at this position.
Frequency Weekly (it is recommended to perform this test before and after calibration); optional: daily.
Equipment Standard test block covering the complete detector, appropriate software tools.

2b.2.2.3.2 Detector element failure (DR systems)

The method and limiting values of the fourth edition are used. The bad pixel map should be easily accessible for all users and should be provided e.g. as a table including number, size and location of the defective elements, clusters and lines. If uncorrected bad pixels are visible on the images, this should be taken into account when evaluating detector element failure.

2b.2.2.4 Interplate sensitivity variations and plate uniformity (CR systems)

In addition to the test in the fourth edition: Take an image receptor homogeneity image (see Section 2b.2.2.3.1), and inspect the variance map for all plates. It may be necessary to clean the plates, following the standard cleaning procedure given by the manufacturers, to determine whether some artefact can be removed.

Limiting value No artefacts should be present. If the variance in an ROI is ≥ 10% higher than the variance in neighbouring ROIs, the image should be investigated visually for an artefact at this position.

Frequency Monthly.
Equipment Standard test object covering the complete detector.

2b.2.2.5 Influence of other sources of radiation

This test is omitted.

2b.2.2.6 Fading of latent image (CR systems)

This test is omitted.

2b.2.3 Dosimetry

This section makes it possible to apply the existing dosimetric procedures to the wider range of target/filter combinations found in modern X-ray sets. Typically, these involve using beam qualities with higher HVL than provided for previously. In addition, more specific guidance is given for the configuration to be used in determining the incident air kerma. Other aspects of the previous guidance are repeated here so that all the required information is available within this document.

2b.2.3.1 Dose to typical breasts simulated with PMMA

The doses to a range of typical breasts should be assessed with blocks of PMMA as breast substitutes using the usual clinically selected exposure factors, including any automatic selection of tube voltage and target/filter combination. This method relies on the equivalence in attenuation between different thicknesses of PMMA and typical breasts (Dance et al., 2000), as listed in Table A5.1 in Appendix 5. It should be noted that because PMMA is denser than breast tissue, any automatic selection of tube voltage, target or filter may be slightly different from that for real breasts. This can be corrected by adding spacers (e.g. expanded polystyrene blocks) to the PMMA to make up a total thickness equal to that of the equivalent breast. Small pieces of more attenuating materials can also be used as spacers provided they are outside the sensitive area of the AEC. On systems that determine the exposure factors using transmission, spacers should not be necessary.

Set the AEC to normally used clinical settings, and expose PMMA plates of 20 mm thickness. Record the exposure factors chosen by the AEC. Repeat this measurement for 30, 40, 45, 50, 60 and 70 mm PMMA thickness. (For routine testing it is sufficient to use only a 45 mm thickness of PMMA.) Calculate the average glandular dose (D) to a typical breast of thickness and composition equivalent to the thickness of PMMA by applying the following formula:

$$D = K g c s \quad (17)$$

where K is the incident air kerma (without backscatter) calculated at the upper surface of the PMMA using the method described later in this section. The g-factor corresponds to a glandularity of 50%, is derived from the values calculated by Dance (1990), Dance et al. (2000) and Dance et al. (2009) and is shown in Table A5.1 for a range of HVLs. The c-factor corrects for the difference in composition of typical breasts from 50% glandularity (Dance et al., 2000; Dance et al., 2011) and is given in Table A5.2 for typical breasts for women in the age group 50–64. Note that the g-factors and c-factors applied are those for the corresponding thickness of typical breast rather than the thickness of PMMA block used. Where necessary, interpolation may be made for different values of HVL. Typical values of HVL for various spectra are given in Table A5.3, but HVLs are normally measured at the same time as the measurements necessary to determine the incident air kerma. The s-factor shown in

Table A5.4a corrects for any difference due to the choice of X-ray spectrum (Dance et al., 2000; Dance et al., 2011). Note that for the W/Al target/filter combination, the s-factor varies with filter thickness and breast thickness (Tables A5.4b–e).

Maximum average glandular dose

Limiting value A maximum average glandular dose is set per PMMA thickness, as shown in Table 1.4.
Frequency Every 6 months.
Equipment Calibrated mammographic dosimeter, 20–70 mm thick blocks of PMMA.

Table 1.4: Dose levels for typical breasts simulated with PMMA

Thickness of PMMA (mm)	Equivalent breast thickness (mm)	Maximum average glandular dose to equivalent breasts (mGy)	
		Acceptable level	Achievable level
20	21	≤ 1.0	≤ 0.6
30	32	≤ 1.5	≤ 1.0
40	45	≤ 2.0	≤ 1.6
45	53	≤ 2.5	≤ 2.0
50	60	≤ 3.0	≤ 2.4
60	75	≤ 4.5	≤ 3.6
70	90	≤ 6.5	≤ 5.1

Incident air kerma

In the original publication by Dance (1990), the incident air kerma was calculated for a dosimeter in contact with and below a compression paddle. It follows that the determination of incident air kerma at the surface of PMMA test phantoms or breasts should be based on measurements made with this geometry to correctly include scatter from the paddle. The recommended geometry for the procedure is shown in Figures 1.7 and 1.8 and is described below.

Calculate the incident air kerma for each of the beam qualities used in exposing the blocks of PMMA by making an exposure of the dosimeter positioned as in Figures 1.7 and 1.8 using a manually selected tube loading (e.g. 50 mAs). Estimate the incident air kerma at the upper surface of the PMMA by using the inverse square law and scaling to the appropriate value of tube loading (mAs). The HVL should also be estimated at the same time using multiple layers of aluminium, as described in the fourth edition.

The chamber should be positioned on a line extending from the tube focus to a point on the mid-line of the breast support table 60 mm from the chest wall edge. If the chamber has backscatter correction, the recommended position is directly on the breast support with the paddle in contact. It would also be possible to make a measurement of air kerma with the dosimeter higher above the breast support and with the paddle in contact provided the appropriate inverse square law correction is made. This approach is recommended if the chamber does not have backscatter correction. The effect of scatter from the compression paddle on the measurement of incident air kerma is discussed in Dance et al. (2009), where it is shown that for the above geometry and a polycarbonate paddle 2.4 mm thick, scattered photons contribute 7% of the total measured air kerma.

Figure 1.7: Position of dosimeter to estimate incident air kerma for dose estimation, top view and 3D view

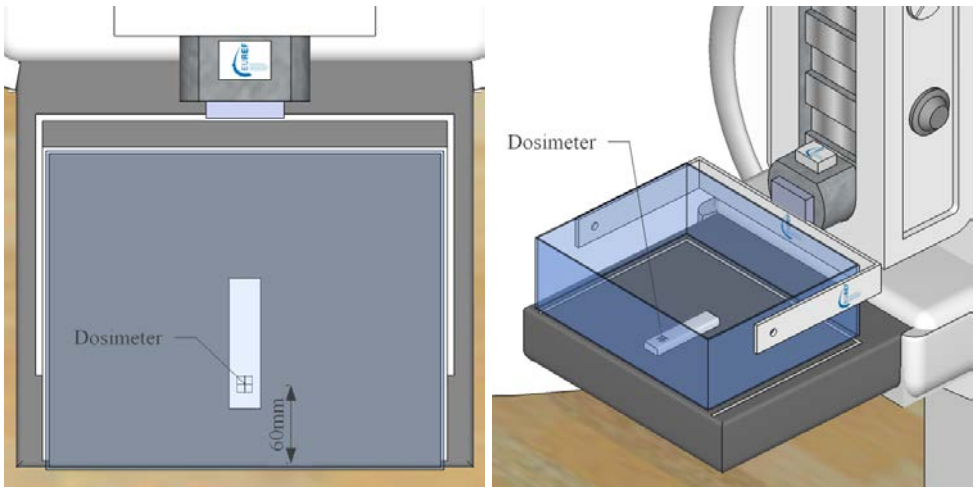
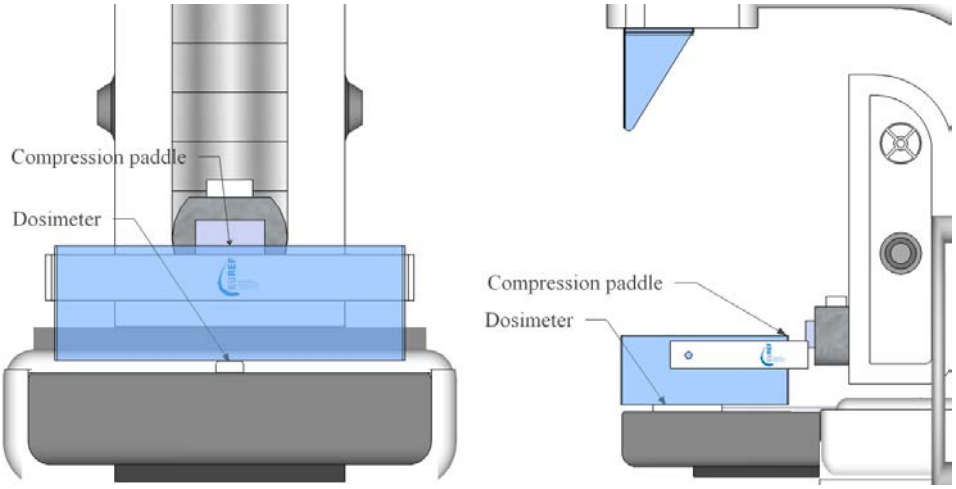


Figure 1.8: Position of dosimeter to estimate incident air kerma for dose estimation, front view and side view



2b.2.3.2 Clinical breast doses

It is also possible to measure the average glandular dose for a series of breast examinations on each mammography system. To do this, for each exposure the breast thickness under compression is measured and the exposure factors are recorded. From measurements of air kerma as described previously (Figures 1.7 and 1.8) at the tube voltage and target/filter combination used, the tube loading (mAs) may be used to estimate the incident air kerma and to determine the average glandular dose using equation (17). In this case, K is the incident air kerma calculated at the upper surface of the breast. For the appropriate breast thickness, g-factors should be interpolated from Table A5.5; c-factors for typical breast compositions for women in the age groups 50–64 and 40–49 are shown in Tables A5.6 and A5.7, respectively. Measurement of compressed breast thickness for this purpose is performed by the radiographer, by reading the displayed compressed thickness on the X-ray set. The

accuracy of the displayed thickness should be verified by applying a typical force (e.g. 100 N) to rigid material of known thickness. It may be necessary to apply correction factors if the displayed values are in error. An accuracy of ≤ ±2 mm is required. Software for making such dose calculations has been published by the United Kingdom Breast Screening Programme (Young, 2001).

2b.2.4.1 Threshold contrast visibility

In threshold contrast visibility analysis, the pixel value (PV) and SNR are assumed to be relatively constant over the imaging fields. If low-frequency trends in PV are present, it may be necessary to correct for such a trend before analysis.

Threshold contrast visibility is determined for cylindrical details with diameters in the range from 0.1 mm to 1 mm. The details are imaged on a background object with an attenuation equivalent to 50 mm of PMMA. The details must be positioned at a height of 20–25 mm above the breast support table. Use the exposure factors that are selected in fully automatic mode for 50 mm PMMA with 10 mm thick spacers, as found when measuring breast thickness and composition compensation.

Take at least 16 images of the details, and move the details slightly between the images to obtain images in which the relative positions of the details and the detector elements are different. The 'for processing' version of the images should be used for analysis.

Automated threshold contrast measurement using CDMAM

If the CDMAM phantom is used, it is recommended that automated computer reading be performed to determine the threshold gold thickness for each detail size. Software tools for doing this are downloadable from the EUREF web site (www.euref.org). This will involve downloading the latest version of CDCOM. Appendix 8 provides a description of how CDCOM works.

Using the output of CDCOM, a detection matrix is constructed and for each diameter of cylindrical details a psychometric curve is fitted (Veldkamp et al., 2003):

$$p(d) = \frac{0.75}{1 + e^{-f(C - Ct)}} + 0.25 \qquad (18)$$

C = logarithm of signal contrast; $C = \log(1 - e^{-\mu d})$
C_t = signal contrast at the threshold of 62.5%
f = fitting parameter
p(d) = the probability of detection of an object with size d.

A threshold at 62.5% correct response is used to determine the threshold contrast. Results for which the psychometric curve is fitted with only a few data points are disregarded.[2] To use the limiting values in the existing protocol (in the fourth edition), the resulting thresholds for each diameter have to be converted to human readout.

A tool called 'CDMAM analysis software tool' uses the output from CDCOM for a set of images to determine threshold gold thicknesses for the different detail sizes. A software description and software manual for 'CDMAM analysis software tool' is also available for download from the EUREF web site (www.euref.org). Because the automated analysis is more successful at locating the gold discs than human observers are, 'CDMAM analysis software tool' also provides the threshold gold thickness expected for a typical observer. Two basic methods of converting from automated thresholds to those predicted for human observers have been used. In the original approach described by Young et al.

[2] A typical range for which the psychometric curve can be fitted is 0.1 mm to 1.0 mm (CDMAM version 3.4).

(2006a), the thresholds are scaled up using a value for each detail diameter (referred to as the United Kingdom method). More recently, a formula such as that shown in equation (19) and described in Young et al. (2008) has been used, which scales the threshold gold thickness independent of diameter (referred to as the EU method). Both methods are currently implemented in 'CDMAM analysis software tool' to provide retrospective compatibility. However, it is recommended that the United Kingdom method be adopted.

$$T_{predicted} = a[T_{auto}]^n \quad (19)$$

$T_{predicted}$ = predicted human threshold gold thickness
T_{auto} = computer readout of threshold gold thickness
a and n = fitted parameters with a = 1.441, n = 0.895.

'CDMAM analysis software tool' also fits the resulting predicted threshold gold thicknesses with a third-order polynomial function (Equation (20)) to obtain the contrast-detail curve.

$$T = a + \frac{b}{x} + \frac{c}{x^2} \quad \frac{d}{x^3} \quad (20)$$

T = threshold gold thickness (µm)
x = detail diameter (mm)
a, b, c and d = coefficients adjusted to achieve a least-squares fit; all are ≥ 0

The values from the fitted curve should be checked against the limiting values for human readout of threshold gold thickness as published in the fourth edition.

Limiting value See fourth edition.
Frequency Every 6 months.
Equipment CDMAM structure plate compliant with version 3.4 (or higher) and four 10 ± 0.2 mm thick PMMA plates of the same size; appropriate software tools.

Comments and tips

The results of these measurements are related in a predictable manner to the exposure factors used for a given system in terms of radiation dose, all other things being equal. Therefore, it is important to record the exposure factors used and to calculate the corresponding radiation dose in terms of the mean glandular dose (MGD) to the standard breast simulated using a 50 mm thickness of PMMA, as described in the dosimetry section. Care should be taken when an exposure in fully automatic mode (from the thickness compensation measurement) is mimicked in manual mode. Account needs to be taken of whether the pre-exposure is included in the displayed mAs value in automatic mode. Once threshold gold thickness is known for one dose level, it is relatively straightforward to predict the results at other dose levels.

Whereas automated analysis works quite consistently for a given phantom, the results vary more between phantoms. Several methods to reduce this problem are being explored and are expected to result in improved versions of CDCOM and/or a calibration procedure for phantoms. The type and serial number of the phantom used should be recorded.

In threshold contrast visibility analysis, the PV and SNR are assumed to be relatively constant over the imaging fields. If large low-frequency trends in PV are present, it may be necessary to correct for such a trend before analysis e.g. by applying a flat-field correction to the images. However, this is rarely necessary because most manufacturers already apply such a correction to their unprocessed images.

If reasonable doubts exist about the automated readout of the phantom images, the images should be scored by human observers.

2b.2.4.2 Modulation transfer function (MTF) and noise power spectrum (NPS)

Moved to Appendix 7.

2b.2.4.3 Exposure time

The time for an exposure in all clinically used AEC modes is measured at 45 mm PMMA thickness.

2b.2.4.4 Geometrical distortion and artefact evaluation

Image the phantoms in fully automatic mode. Artefact evaluation should also be performed using the variance map mentioned in Section 2b.2.2.3.1.

2b.3 Image processing

It is not yet possible to perform objective quantitative measurements on image processing in the context of an acceptance test. Due to the importance of the subject, some guidance on the evaluation of image processing is given in this section. The method described here is reported in greater detail elsewhere (Van Ongeval et al., 2008; van Engen et al., 2013).

An unprocessed image (DICOM: 'for processing') with a linear relationship between detector dose and PV, which is the output of most current DR systems, is not the most suitable image for radiologists to read. In this type of image, the PV is related to the number of X-ray quanta interacting in the detector. A radiologist, however, is interested in structures in the breast and the amount of radiation attenuated by these structures. This 'attenuation image' is obtained by transforming PVs logarithmically. For most CR systems the images are already scaled logarithmically by the hardware of the reader, so the unprocessed images have a non-linear relationship between image receptor dose and PV.

In the next stage, the image is further processed to enhance the visibility of the clinically relevant information, yielding a processed image (DICOM: 'for presentation').

Processing techniques that are applied to mammographic images include:

- LUT and Bit operations
- thickness equalization at the edge of the breast
- sharpening of the image
- noise reduction
- contrast optimization.

Objective evaluation of image processing algorithms is very difficult. Image characteristics, like PV distribution (histogram), shape, etc., are used in image processing algorithms. This means that phantoms whose characteristics differ from those of a breast cannot be used to evaluate image processing. The difference in characteristics causes the processing of phantom images to be different

from that of images of breasts. Artefacts may also be introduced because image processing algorithms presume a breast edge (skin line) that may not be present in technical test objects.

Therefore, to date only subjective evaluation of image processing algorithms could be performed, by radiologists scoring mammograms.

After installation of a system and the acceptance test, it is recommended to carefully evaluate a series of clinical images. If possible, clinical images taken with the new system should be compared with previous images of the breasts of the same woman taken with a diagnostically established modality.

The following image characteristics may be taken into account when comparing images:

1. the visualization of the skin line
2. the visibility of vascular structures through dense parenchyma
3. the visualization of vascular and fibrous structures and pectoral muscle
4. the visualization of structures along the pectoral muscle
5. the visualization of Cooper's ligaments and vascular structures in the low- and high-PV areas of the image
6. the outlines of microcalcifications
7. the noise in the low- and high-PV areas of the image
8. the contrast in the low- and high-PV areas of the image
9. the appearance of glandular tissue
10. the appearance of the background area (the area directly exposed to the X-ray field without any attenuation by the imaged object)
11. the confidence of the radiologist with the representation of the image
12. the presence of artefacts.

It must be realized that large numbers of cases of different breast types, breast thicknesses and dose levels should be reviewed before conclusions about the quality of images can be drawn, because the visibility of structures may differ in individual cases, e.g. due to differences in positioning or variations in anatomical structures.

2b.4 Image presentation

2b.4.1 Monitors

2b.4.1.1 Ambient light

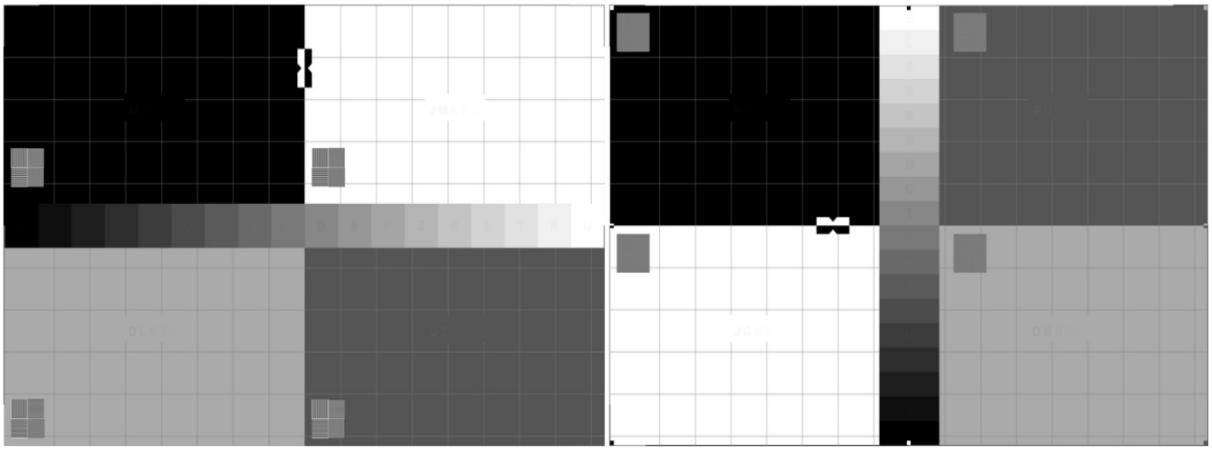
For LCD displays, the maximum ambient light value is increased to 20 lux. For CRT displays, the maximum ambient light level remains 10 lux.

2b.4.1.2–5 Constancy tests of monitor performance

In the fourth edition, it is stipulated that the following parameters of a monitor should be evaluated daily: geometrical distortion (on CRT displays) (2b.4.1.2), contrast visibility (2b.4.1.3) and display artefacts (2b.4.1.5). It was recommended that these tests should be performed by using the TG18-QC test pattern of the American Association of Physicists in Medicine (AAPM). In this section, alternative test patterns are described that can be used to test contrast visibility, geometrical distortion and display artefacts as efficiently as with the AAPM test patterns (Jacobs et al., 2007). The complete procedure includes the use of a newly generated test pattern for every evaluation and a fill-in sheet for comparing the readings with the true values. This procedure overcomes inattentive scoring and allows adherence to the QC procedures to be verified easily. The software for generating and scoring the test patterns is downloadable from the EUREF web site (www.euref.org).

The MoniQA pattern is divided into 4 equally sized rectangular segments with 4 uniform background values of different intensities. The values were chosen to be 0%, 33%, 66% and 100% of the maximum grey level. The position of these rectangles swaps randomly each time the pattern is generated, with one restriction: the rectangle with a grey level of 0% (L_{min}) will always have a mutual border with the rectangle with a grey level of 100% (L_{max}). This guarantees a black-to-white or white-to-black transition between the patches with the maximum and the minimum grey level. This transition can be either horizontal or vertical. Figure 1.9 shows two examples of the variable pattern.

Figure 1.9: Two examples of the MoniQA pattern. These patterns include checks for contrast visibility, geometrical distortion, spatial resolution, global image quality and artefacts (reprinted with permission from *Med. Phys.*)



Extra tests:

i. Low-contrast characters

In the centre of each rectangular segment there is a set of five characters that have a low contrast with the background PV (Figure 1.10a). Each time the pattern is generated, the characters are randomly chosen out of a subset of the Latin alphabet, i.e. *ABCDEHJKLMPSTUZ*. Each set of characters has PV differences of 7, 5, 3, 2 and 1 between background and character. The observer has to read as many characters as possible. It is suggested that the observer guess the identity of one character more than those seen with certainty.

Score criteria: If characters are not discriminated from the background, points are subtracted from the initial score of 100 according to the PV difference between character and background. If the least

visible character is not identified correctly, 1 point is deducted. The next character has a value of 2 points; the third character has a value of 3 points. For the fourth character, 5 points are deducted, and if the highest contrast character is not detectable, 7 points are deducted.

ii. Gradient bar of patches with increasing PVs and low-contrast characters

In the centre of the display, a gradient bar of 18 distinct greyscale steps is drawn, with PVs as used in the central rectangle of the AAPM TG18-LN patterns. This bar is horizontal or vertical but will never divide the rectangles with 0% and 100% of the maximum grey level. A randomly chosen character is placed on each step of the gradient. The bar is divided into 2 equally sized parts, a northern and a southern part or a western and an eastern part. In each part of the gradient bar, each character is unique. The characters are selected from the same subset of the Latin alphabet as used for the selection of the low-contrast characters. The greyscale value of each character is the same as the greyscale value of the preceding luminance patch (Figure 1.10b), with the whitest and darkest patches at the extremes.

To evaluate this pattern, characters are to be read starting in the middle and then reading, according to the orientation of the bar, towards the west and east or north and south. If a luminance character is visible, it is concluded that the underlying patch can be clearly distinguished from the adjacent patch. In the AAPM TG18-QC and the DIN test patterns, the purpose of this gradient bar is to verify whether the different steps are distinguishable. This is most critical for the lowest and highest PVs. When evaluating the MoniQA pattern, only the two last visible characters have to be registered.

Score criteria: 10 points are deducted for each incorrectly identified or invisible low-luminance patch, and this is done for both extremities of the gradient bar. If no character has been identified correctly, 9 × 10 points are deducted.

iii. The MoniQA pattern can be used to test geometrical distortion, whether all pixels of the test image are shown, high- and low-contrast spatial resolution and artefacts (black-to-white and white-to-black transition problems). Any imperfection lowers the score by 5 points, except a dead pixel, which lowers the score by 11 points.

The MoniQA pattern is highly variable. There are 16 combinations of background positions, 4 positions for the resolution pattern inside each background field and 2 resolution types (high- and low-contrast), which means the total number of possible configurations is 128. In addition, there is a very large number of combinations of characters for the low-contrast visibility checks.

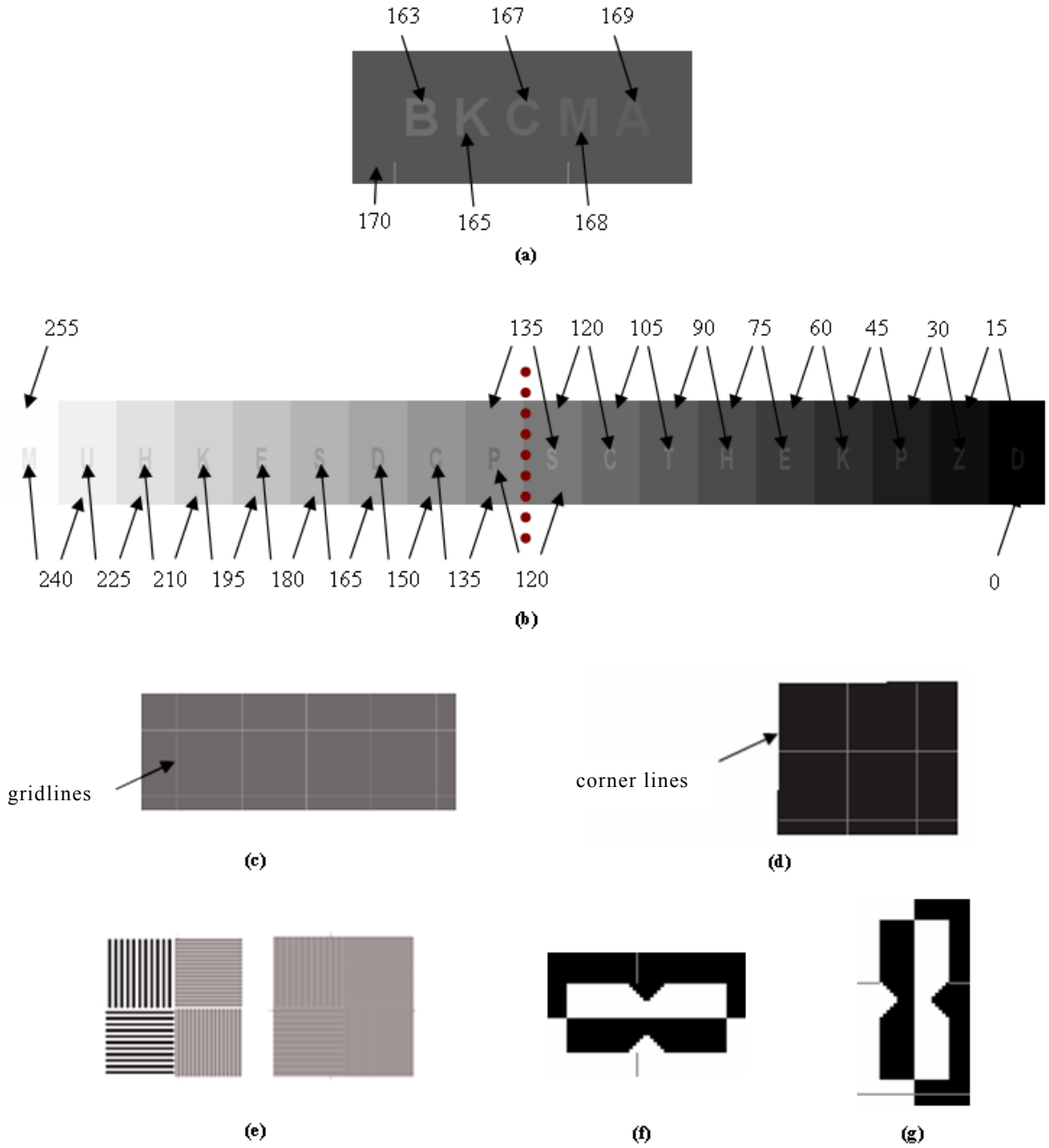
Limiting value Score obtained from MoniQA pattern should be ≥ 95.
Frequency Daily; optional: weekly.
Equipment MoniQA test pattern.

2b.4.1.8 Luminance uniformity

Limiting value of luminance deviation

	CRT	LCD
Test pattern TG18-UNL10	≤ 30%	≤ 30%
Test pattern TG18-UNL80	≤ 30%	≤ 15%

Figure 1.10: Examples of features of the MoniQA pattern: (a) example of a sequence of characters with a low-contrast luminance difference from the background; (b) gradient bar of patches with increasing pixel values and with randomly selected characters with a pixel value as in the adjacent patch; (c) grid pattern; (d) corner lines pattern; (e) resolution patterns: (left) high-contrast and (right) low-contrast; (f) horizontal and (g) vertical version of the hourglass object (all elements are shown with enhanced contrast for clarity)



2b.4.2 Printers

General remark: For all test items in this section, each test pattern should be printed 3 times. If optical densities (ODs) are to be measured, the average OD from the 3 prints should be used for further analysis.

The suggestions for QC made by AAPM Task Group 18 have been adapted slightly. The Task Group 18 measurements are based on measuring OD of a printed test pattern. From a QC point of view, a standard viewing box has been defined (luminance of the viewing box without film: 4000 cd/m²; luminance contribution due to ambient illuminance reflecting on the printout: 1 cd/m²). The ODs of the test pattern should be such that the printout in combination with this virtual viewing box conforms to the greyscale standard display function (GSDF). The luminance of the viewing boxes is controlled by the tests described in the screen-film section of the fourth edition.

- AAPM TG18 patterns must be in DICOM MG format.
- AAPM test patterns must be printed from the acquisition workstation (or printing server) and, if applicable, from the diagnosis workstation.
- AAPM test patterns must be printed for all film formats that are used.

2b.4.2.5 Optical density range

Print the TG18-PQC test pattern. Measure D_{min} and D_{max} on this image.

Limiting value $D_{min} \leq 0.25$ OD, $D_{max} \geq 3.60$ OD.
Frequency Every 6 months.
Equipment Suitable densitometer, TG18-PQC test pattern.

2b.4.3 Viewing boxes

If mammograms are read on printed images, check the viewing boxes using the method and limiting values described in the fourth edition (page 82).

2b.4.3.1 Ambient lighting level

The artefacts and loss of image quality associated with reflections from the display surface depend on the level of ambient lighting. It is important to verify that the ambient lighting in the room is below the maximum limit. The conditions for the tests should be similar to those during normal use of the equipment.

If mammograms are read on printed images, check ambient light using the method and limiting values described in the fourth edition (page 82).

Appendix 5: Tables for determination of average glandular dose

Most of the data in the following tables are from Dance (1990), Dance et al. (2000), Dance et al.. (2009) and Dance et al. (2011).

Table A5.1: g-Factors for breasts simulated with PMMA

PMMA thickness (mm)	Equivalent breast thickness (mm)	Glandularity of equivalent breast (%)	g-factor (mGy/mGy)										
			HVL (mm Al)										
			0.30	0.35	0.40	0.45	0.50	0.55	0.60	0.65	0.70	0.75	0.80
20	21	97	0.378	0.421	0.460	0.496	0.529	0.559	0.585	0.609	0.631	0.650	0.669
30	32	67	0.261	0.294	0.326	0.357	0.388	0.419	0.448	0.473	0.495	0.516	0.536
40	45	41	0.183	0.208	0.232	0.258	0.285	0.311	0.339	0.366	0.387	0.406	0.425
45	53	29	0.155	0.177	0.198	0.220	0.245	0.272	0.295	0.317	0.336	0.354	0.372
50	60	20	0.135	0.154	0.172	0.192	0.214	0.236	0.261	0.282	0.300	0.317	0.333
60	75	9	0.106	0.121	0.136	0.152	0.166	0.189	0.210	0.228	0.243	0.257	0.272
70	90	4	0.086	0.098	0.111	0.123	0.136	0.154	0.172	0.188	0.202	0.214	0.227
80	103	3	0.074	0.085	0.096	0.106	0.117	0.133	0.149	0.163	0.176	0.187	0.199

Table A5.2: c-Factors for breasts simulated with PMMA

PMMA thickness (mm)	Equivalent breast thickness (mm)	Glandularity of equivalent breast (%)	c-factor*										
			HVL (mm Al)										
			0.30	0.35	0.40	0.45	0.50	0.55	0.60	0.65	0.70	0.75	0.80
20	21	97	0.889	0.895	0.903	0.908	0.912	0.917	0.921	0.924	0.928	0.933	0.937
30	32	67	0.940	0.943	0.945	0.946	0.949	0.952	0.953	0.956	0.959	0.961	0.964
40	45	41	1.043	1.041	1.040	1.039	1.037	1.035	1.034	1.032	1.030	1.028	1.026
45	53	29	1.109	1.105	1.102	1.099	1.096	1.091	1.088	1.082	1.078	1.073	1.068
50	60	20	1.164	1.160	1.151	1.150	1.144	1.139	1.134	1.124	1.117	1.111	1.103
60	75	9	1.254	1.245	1.235	1.231	1.225	1.217	1.207	1.196	1.186	1.175	1.164
70	90	4	1.299	1.292	1.282	1.275	1.270	1.260	1.249	1.236	1.225	1.213	1.200
80	103	3	1.307	1.299	1.292	1.287	1.283	1.273	1.262	1.249	1.238	1.226	1.213

* for typical breasts for women in the age group 50–64

Table A5.3: Typical HVL measurements for different tube voltage (kV) and target/filter combinations (data include the effect on measured HVL of attenuation by a compression plate)

kV	HVL (mm Al) for target/filter combination						
	Mo Mo	Mo Rh	Rh Rh	W Rh	W Ag	W Al (0.5 mm)	W Al (0.7 mm)
25	0.32 ± 0.02	0.38 ± 0.02	0.37 ± 0.02	0.50 ± 0.03	0.51 ± 0.03	0.34 ± 0.03	0.42 ± 0.03
28	0.35 ± 0.02	0.42 ± 0.02	0.42 ± 0.02	0.53 ± 0.03	0.58 ± 0.03	0.39 ± 0.03	0.49 ± 0.03
31	0.38 ± 0.02	0.45 ± 0.02	0.45 ± 0.02	0.56 ± 0.03	0.61 ± 0.03	0.44 ± 0.03	0.55 ± 0.03
34	0.40 ± 0.02	0.47 ± 0.02	0.47 ± 0.02	0.59 ± 0.03	0.64 ± 0.03	0.49 ± 0.03	0.61 ± 0.03
37				0.62 ± 0.03	0.67 ± 0.03	0.53 ± 0.03	0.66 ± 0.03

Table A5.4a: s-Factors for clinically used spectra (Dance et al., 2000)

Target material	Filter material	Filter thickness (µm)	s-factor
Mo	Mo	30	1.000
Mo	Rh	25	1.017
Rh	Rh	25	1.061
W	Rh	50–60	1.042
W	Ag	50–75	1.042

Table A5.4b: s-Factors for a tungsten target filtered by 0.5 mm aluminium

PMMA thickness (mm)	Equiv. breast thickness (mm)	s-factor
20	21	1.075
30	32	1.104
40	45	1.134
45	53	1.149
50	60	1.160
60	75	1.181
70	90	1.198
80	103	1.208

Table A5.4c: s-Factors for a tungsten target filtered by 0.7 mm aluminium.

PMMA thickness (mm)	Equiv. breast thickness (mm)	s-factor
20	21	1.052
30	32	1.064
40	45	1.082
45	53	1.094
50	60	1.105
60	75	1.123
70	90	1.136
80	103	1.142

Table A5.4d: s-Factors for a tungsten target filtered by 0.5 mm aluminium

Breast thickness (mm)	Glandularity range (%)	Typical glandularity, age 50–64 (%)	Typical glandularity, age 40–49 (%)	kV range (kV)	s-factor
20	80–100	100	100	25–40	1.069
30	62–82	72	82	29–40	1.104
40	40–65	50	65	29–40	1.127
50	23–49	33	49	30–40	1.139
60	11–35	21	35	30–40	1.154
70	2–24	12	24	30–40	1.180
80	0.1–17	7	14	30–40	1.187
90	0.1–14	4	8	30–40	1.198
100	0.1–13	3	5	30–40	1.206
110	0.1–13	3	5	30–40	1.212

Table A5.4e: s-Factors for a tungsten target filtered by 0.7 mm aluminium

Breast thickness (mm)	Glandularity range (%)	Typical glandularity, age 50–64 (%)	Typical glandularity, age 40–49 (%)	kV range (kV)	s-factor
20	80–100	100	100	25–50	1.052
30	62–82	72	82	25–50	1.060
40	40–65	50	65	25–50	1.076
50	23–49	33	49	25–50	1.087
60	11–35	21	35	25–50	1.105
70	2–24	12	24	28–50	1.121
80	0.1–17	7	14	28–50	1.129
90	0.1–14	4	8	28–50	1.136
100	0.1–13	3	5	28–50	1.140
110	0.1–13	3	5	28–50	1.144

Table A5.5: Additional g-factors

HVL (mm Al)	Breast thickness (mm) 20	30	40	50	60	70	80	90	100	110
0.30	0.390	0.274	0.207	0.164	0.135	0.114	0.098	0.0859	0.0763	0.0687
0.35	0.433	0.309	0.235	0.187	0.154	0.13	0.112	0.0981	0.0873	0.0786
0.40	0.473	0.342	0.261	0.209	0.172	0.145	0.126	0.1106	0.0986	0.0887
0.45	0.509	0.374	0.289	0.232	0.192	0.163	0.14	0.1233	0.1096	0.0988
0.50	0.543	0.406	0.318	0.258	0.214	0.177	0.154	0.1357	0.1207	0.1088
0.55	0.573	0.437	0.346	0.287	0.236	0.202	0.175	0.1543	0.1375	0.1240
0.60	0.587	0.466	0.374	0.31	0.261	0.224	0.195	0.1723	0.1540	0.1385
0.65	0.622	0.491	0.399	0.332	0.282	0.244	0.212	0.1879	0.1682	0.1520
0.70	0.644	0.514	0.421	0.352	0.300	0.259	0.227	0.2017	0.1809	0.1638
0.75	0.663	0.535	0.441	0.371	0.317	0.274	0.241	0.2143	0.1926	0.1746
0.80	0.682	0.555	0.460	0.389	0.333	0.289	0.254	0.2270	0.2044	0.1856

The table is in units of mGy/mGy.
Additions to the table in Dance et al. (2000) are highlighted in grey.

Table A5.6: Additional c-factors for average breasts for women in the age group 50–64

Breast thickn. (mm)	Gland. (%)	HVL (mm Al) 0.30	0.35	0.40	0.45	0.50	0.55	0.60	0.65	0.70	0.75	0.80
20	100	0.885	0.891	0.9	0.905	0.91	0.914	0.919	0.923	0.928	0.932	0.936
30	72	0.925	0.929	0.931	0.933	0.937	0.94	0.941	0.947	0.950	0.953	0.956
40	50	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
50	33	1.086	1.082	1.081	1.078	1.075	1.071	1.069	1.064	1.060	1.057	1.053
60	21	1.164	1.160	1.151	1.15	1.144	1.139	1.134	1.124	1.117	1.111	1.103
70	12	1.232	1.225	1.214	1.208	1.204	1.196	1.188	1.176	1.167	1.157	1.147
80	7	1.275	1.265	1.257	1.254	1.247	1.237	1.227	1.213	1.202	1.191	1.179
90	4	1.299	1.292	1.282	1.275	1.27	1.26	1.249	1.236	1.225	1.213	1.200
100	3	1.307	1.298	1.29	1.286	1.283	1.272	1.261	1.248	1.236	1.224	1.211
110	3	1.306	1.301	1.294	1.291	1.283	1.274	1.266	1.251	1.240	1.228	1.215

Additions to the table in Dance et al. (2000) are highlighted in grey.

Table A5.7: Additional c-factors for average breasts for women in the age group 40–49

Breast thickn. (mm)	Gland. (%)	HVL (mm Al) 0.30	0.35	0.40	0.45	0.50	0.55	0.60	0.65	0.70	0.75	0.80
20	100	0.885	0.891	0.9	0.905	0.91	0.914	0.919	0.923	0.928	0.932	0.936
30	82	0.894	0.898	0.903	0.906	0.911	0.915	0.918	0.924	0.928	0.933	0.937
40	65	0.940	0.943	0.945	0.947	0.948	0.952	0.955	0.956	0.959	0.961	0.964
50	49	1.005	1.005	1.005	1.004	1.004	1.004	1.004	1.004	1.003	1.003	1.003
60	35	1.080	1.078	1.074	1.074	1.071	1.068	1.066	1.061	1.058	1.055	1.051
70	24	1.152	1.147	1.141	1.138	1.135	1.130	1.127	1.117	1.111	1.105	1.098
80	14	1.220	1.213	1.206	1.205	1.199	1.190	1.183	1.172	1.163	1.154	1.145
90	8	1.270	1.264	1.254	1.248	1.244	1.235	1.225	1.214	1.204	1.193	1.181
100	5	1.295	1.287	1.279	1.275	1.272	1.262	1.251	1.238	1.227	1.215	1.203
110	5	1.294	1.290	1.283	1.281	1.273	1.264	1.256	1.242	1.232	1.220	1.208

Additions to the table in Dance et al. (2000) are highlighted in grey.

Appendix 7: Linear system theory metrics and a practical guide to their measurement (optional)

A7.1 Modulation transfer function (MTF), noise power spectrum (NPS) and detective quantum efficiency (DQE)

Linear system theory metrics offer reproducible, objective estimates of X-ray detector noise and resolution properties and are sensitive to changes in detector performance over time. The International Electrotechnical Commission (IEC) standard 62220-1-2 describes the measurement of these parameters (IEC, 2004). However, this document is intended for use by the manufacturers; they can remove the X-ray detector from the system and perform a separate bench test. This is not possible for detectors in clinical use and hence a pragmatic approach, suitable for routine QC conditions, is presented here. Measurement geometry is likely to vary between systems, resulting in a loss of generality, and hence caution must be exercised when using these metrics to compare across systems. However, measured with care, these parameters offer significant insight into the performance of an individual detector; they can isolate performance changes over time and are useful when troubleshooting the entire imaging chain. Definitions of the equations used to calculate these parameters are given in the recommended literature; a firm grasp of the theory underlying these parameters is required before performing these measurements. Given that many QC physicists will not have the time to develop the required software, validated/verified software can be used for the calculations.

The IEC document prescribes standard measurements to be performed at 'the detector surface' (defined as 'the accessible area which is closest to the image receptor plane'); for routine QC measurements, there will be additional non-removable parts (e.g. breast support table, antiscatter grid and/or detector covers) in the X-ray beam during the measurement. With this in mind, a consistent geometry should be used for a given system/model. When comparing quantitative measurements from different physics centres, the data acquisition conditions must be stated explicitly for the sake of transparency. These include the geometry (position of antiscatter grid and breast support table, use of collimation/field area), beam energy and detector air kerma, along with the data conditioning parameters used in the calculation of the MTF and NPS (ROI dimensions, de-trending used, sectioning, etc.).

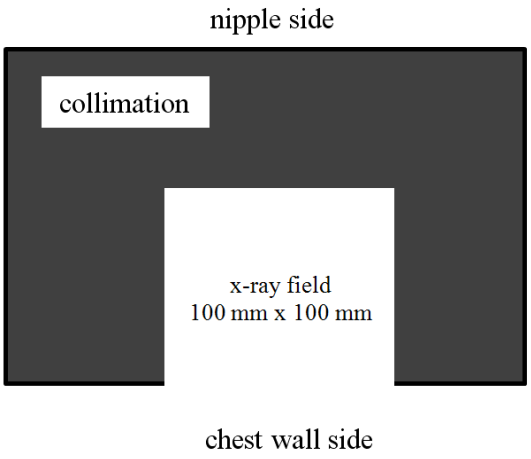
A7.2 Image type

The first step is to identify/select 'for processing' images on the system. These are images that have a fixed gain between detector output signal and air kerma at the detector (no 'autoranging' of the signal) and have minimal additional processing. For example, DR systems will generally apply detector offset and gain corrections together with pixel corrections; this is acceptable. Processing such as edge enhancement or proprietary image processing that prepares images 'for display' etc. must be disabled/not used. These image types are sometimes listed as a specific 'series' and 'study description' by the manufacturer.

A7.3 Collimation of X-ray field

Collimation can be used to reduce the influence of scattered radiation on the measurements. The use of collimation represents good practice but may be time-consuming to set accurately during routine QC (depending on the collimation type available). If collimation is used, a field area of 100 mm × 100 mm should be set at the beginning of the tests and kept in place for all the measurements (air kerma, detector response, NPS, MTF). Figure A7.1 shows suggested positioning of the collimation.

Figure A7.1: Position of the collimation – if used, this should be in place for all the measurements



A7.4 Detector response function

Quantitative analysis requires the measurement of the detector response function, relating air kerma at the detector input plane to pixel value (PV). This is used for linearization of the images from which the quantitative image quality metrics are calculated. Measurement of the detector response requires an estimate of the air kerma at the detector entrance plane.

For this measurement, set the same X-ray spectrum as used for a standard image, and place 2 mm Al at the X-ray tube. Remove the compression plate from the X-ray beam, protect the X-ray detector and measure air kerma as a function of mAs. Sample low mAs settings more finely because mAs linearity can be worse at low mAs values; furthermore, it is likely that low mAs values will be used in images for NPS estimation, given the filter efficiency/transmission of 2 mm Al. Calculate air kerma at the detector (K), applying an inverse square law correction to obtain the values at the level of the detector, and apply a grid transmission factor (use a factor appropriate for the geometry) if the grid is not removed. Fit a first-order polynomial function and confirm linearity of K with mAs; carefully examine mAs linearity at low mAs settings. Use this equation to calculate the mAs values needed for a range of detector air kerma values, e.g. 12.5, 25, 50, 100, 200, 400 and 800 µGy. This will enable testing of the detector at a fixed air kerma level for the lifetime of the detector and allow changes in performance to be tracked. Remove the grid, set the calculated mAs values (closest mAs station on the system) and acquire uniformly exposed (flood) images over the air kerma range. Measure PV and standard deviation at the standard position using the standard ROI size. Plot PV vs K, fit the appropriate curve for the system type (linear, logarithmic or power) and record the fit parameters. This function is used to linearize the PV data on a pixel-wise basis in the edge and flood images before calculating MTF and NPS. This must be done for all systems, even systems that produce linear 'for processing' images. After this step, the linearized images will have unity gain and zero offset (the mean PV in this image should be equal to the air kerma used to acquire the image).

A7.5 Noise power spectrum (NPS)

The NPS describes the variance of an image intensity (image PV) divided among its frequency components and is calculated from ROIs taken from a region of a uniformly exposed image. In practice, the NPS is calculated using:

$$NPS(u,v) = \frac{\Delta x \Delta y}{M \cdot 256 \cdot 256} \sum_{m=1}^{M} \left| \sum_{i=1}^{256} \sum_{j=1}^{256} (I(x_i, y_j) - S(x,y)) e^{-2\pi i (u_n x_i + v_k y_j)} \right|^2 \quad \text{(A7.1)}$$

where an ROI dimension of 256 × 256 pixels has been used. Here, M is the number of ROIs, Δx is the pixel spacing in the x direction, Δy is the pixel spacing in the y direction, $I(x_i,y_j)$ are the (linearized) pixel data, $S(x,y)$ is a 2-dimensional polynomial function fitted to the entire extracted region used for NPS analysis (not to the individual ROIs).

Select a flood image acquired at some reference detector air kerma, e.g. at 50 or 100 µGy, and calculate the NPS for this region. Use an image acquired at the same air kerma throughout the life of the detector. Linearize the image using the detector response curve. IEC 62220-1 defines an area of 50 mm × 50 mm for the NPS estimation, divided into ROIs of 256 × 256 pixels that overlap each other by 128 pixels. This strictly limits the physical region from which the NPS is calculated, reducing the effects of non-stationarity and large area non-uniformity on the NPS; however, several images are required to increase the number of spectra in the ensemble and hence reduce uncertainty of the spectral estimate. For QC purposes, a region of 100 mm × 100 mm can be used and ROIs of 256 × 256 pixels taken from this area. To reduce statistical uncertainty, the spectra from several identically acquired images can be averaged. It is recommended that a 2-dimensional polynomial function be fitted to and subtracted from the 100 mm × 100 mm area before extraction of the ROIs of 256 × 256 pixels. The final spectrum is sectioned from the ensemble. For systems with an isotropic NPS, this can be a radial average; for detectors with a non-isotropic NPS, the spectra sectioned from the 0° and 90° axes should be recorded separately. The axes (0° and 90° spatial frequency bins) contain information about the axial structured noise of the detector/image and should be included in the spectral estimate. The spectral ensemble (averaged over the number of individual spectra in the ensemble) is then normalized to give the normalized noise power spectrum (NNPS) by dividing by the mean PV of the linearized flood image used to calculate the spectral estimate, i.e. by dividing by the air kerma used to acquire the flood image.

The IEC standard specifies the use of collimation of 100 mm × 100 mm when acquiring the flood and edge images, to control the quantity of the scattered radiation in the image. A higher quantity of scattered radiation effectively leads to a higher detector air kerma per image and hence an increased NPS. Whereas collimation is essential for laboratory detector measurements, the value of collimation in a QC setting is limited and it should be considered optional (the collimation is heavy, and exactly the same collimator dimension must be used across QC visits).

The antiscatter grid can influence the measured NPS in several ways. First, the grid can introduce structured noise, predominantly of low spatial frequencies, which is often seen along the 0° and 90° NPS axes. Structured noise is multiplicative in nature and increases relative to other noise sources as detector air kerma is increased. The spatially periodic nature of the grid can also introduce spikes, indicating increased noise power at distinct spatial frequencies; this may indicate a grid motion problem. For example, a linear grid with 30 lines cm^{-1} will generate spikes at 3.0 mm^{-1} (and associated harmonics) in the NPS; these will be seen on the axis (0° or 90°) that is parallel to the direction of grid movement in the image.

The presence of the antiscatter grid in the X-ray beam during detector calibration presents a further complication. Some systems may have flat-field corrections explicitly for the case of grid in and grid

out of the X-ray beam, whereas others may have a single flat-field correction which presumes that the grid is present. For this latter system type, an 'imprint' of the flat-field correction will be applied to flood images that have been acquired with the grid removed, leading to a potential increase in the structured noise present in the image and hence in the NPS.

Record the NPS at 0.5 mm^{-1} and 2.0 mm^{-1} (either a radial average or the 0° and 90° axis values separately).

Limiting value Expect ≤ ±15% change in NPS at 0.5 mm^{-1} and 2.0 mm^{-1} from previous QC visit value and from baseline.
Frequency Every 6 months.
Equipment 2 mm Al filter (minimum 99% purity), calibrated dose meter, software for calculating objective image quality parameters.

A7.6 Pre-sampled modulation transfer function

The MTF describes the spatial frequency response of a linear, spatially invariant imaging system and is typically determined using a slanted radio-opaque edge placed at the detector input plane, a robust technique suitable for routine QC. This method generates the detector pre-sampled MTF, a parameter that describes the blurring due to the detector pixel aperture and the X-ray converter layer. The pre-sampled MTF is measured in two directions across the detector: the left-right and front-back (chest wall-nipple) directions. A stainless steel square of thickness 0.8 mm and dimensions 60 mm × 60 mm can be used; the provisos are that the edges must be sharp and straight and the edge must be radio-opaque. If the edge is not radio-opaque, then scattered radiation can influence the edge spread function (ESF) and hence the MTF. Alternative materials can be used, such as niobium, tantalum, tungsten, lead, etc. However, the edge must be of sufficient thickness (radio-opaque) and the material should not generate large quantities of characteristic radiation that may influence the ESF.

The pre-sampled MTF is calculated from a finely sampled ('over-sampled') ESF, generated by re-binning or re-projecting the (linearized) PV data in some region containing the edge. The over-sampled ESF is differentiated to generate the line spread function (LSF):

$$LSF\,(x) = \frac{d}{dx}\,ESF(x) \tag{A7.2}$$

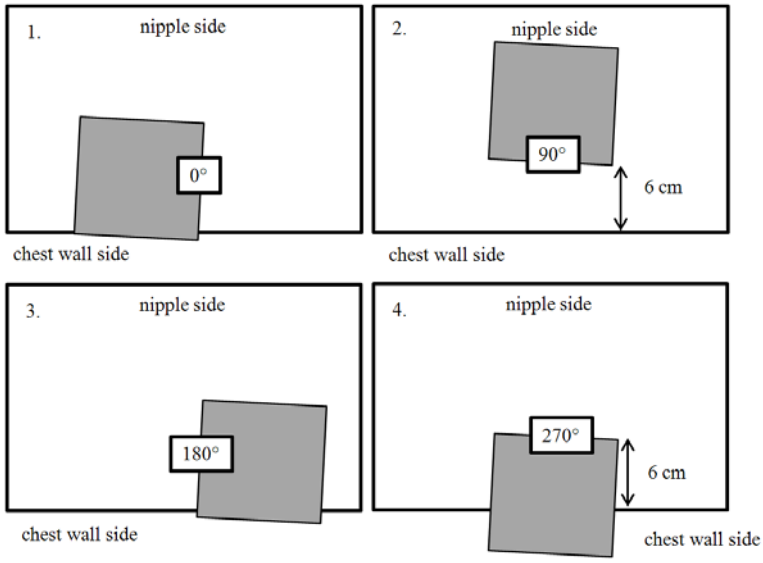
The pre-sampled MTF is obtained from the LSF by using the fast Fourier transform (FFT) and calculating the magnitude:

$$MTF = |F[LSF(x)]| \tag{A7.3}$$

With regard to practical measurements, set the same beam quality as used to acquire the flood images and an mAs setting to produce an air kerma at the detector that is approximately 3× the typical *K* value for AEC images. There must be no saturation in the high-signal region of the ESF, and there must be no clipping or truncation to zero linearized PV in the low-signal region of the ESF. This is generally achieved by ensuring that the PVs in the edge image (before linearization) all lie within the range covered by the detector response function. Place the edge on the breast support table ('detector surface') with an approximate angle of 1° to 3° with respect to the pixel matrix. In a full evaluation, 4 edge images are acquired, positioned to measure the MTF at the lateral centre of the detector and at 6 cm from the chest wall edge (Figure A7.2). An MTF is calculated for each acquisition: the left-right direction (low to high signal and high to low signal changes for the ESF) and similarly for the chest wall-nipple direction. The final MTF is an average of the two edge orientations;

e.g. the left-right MTF would be averaged from the 0° and 180° edge orientations in Figure A7.2. Extract a sufficiently large region containing the edge such that glare (low-frequency signal spread) within the detector is characterized. The actual ROI dimension will depend on the characteristic distance of the glare; however, an ROI of at least 40 mm × 40 mm should be used. Linearize the image PV data before calculating the MTF. Note the conditioning applied when obtaining the MTF result (smoothing, windowing, extrapolation of the LSF tails, etc.).

Figure A7.2: Positions of the edge for MTF measurement (any additional collimation that may be used is not shown)



If a metal square with at least two sharp, straight edges is available, then the MTF can be calculated from a single image for the two detector directions, although in this case the MTF will be evaluated at different positions on the detector. This is acceptable for routine QC measurements.

Corrections for non-uniformity in the MTF image can be applied to remove low-frequency trends using a uniformly exposed image (i.e. with the edge removed) acquired with exactly the same technique factors, although great care must be taken when making this kind of correction. As with the NPS, the X-ray beam can be collimated to 100 mm × 100 mm to reduce scattered radiation, although this is optional for QC measurements.

Record the spatial frequencies at which the MTF reaches 50% (left-right and chest wall-nipple directions).

Limiting value	Expect ≤ ±10% change in the spatial frequency for the 50% MTF point.
Frequency	Every 6 months.
Equipment	Radio-opaque edge of minimum dimensions 60 mm × 60 mm with sharp and straight edges, 2 mm Al filter, calibrated dose meter, software for calculating objective image quality parameters.

A7.7 Detective quantum efficiency (DQE)

The DQE describes the degradation of the input SNR by the X-ray detector due to detection, conversion and amplification of the X-ray signal. DQE is defined as the square of the ratio of the detector output SNR to the SNR at the detector input. It is common to compare the detector against a perfect photon counting device, and hence the reference SNR at the detector input is the number of X-ray photons per unit area (mm^{-2}).

$$DQE(u) = \frac{MTF^2(u)}{\left(K_a\, SNR_{in}^{\ 2}\right) NNPS(u)} \tag{A7.4}$$

This equation is given for the 1-dimensional case; *MTF(u)* is the pre-sampled MTF, *NNPS(u)* is the measured NNPS, K_a is the estimated air kerma at the detector surface and SNR_{in}^2 is the number of X-ray photons μGy^{-1} mm^{-2} for the beam quality used. Hence, $K_a \times SNR_{in}^2$ gives the total number of X-ray photons mm^{-2} at the detector input. Table A7.1a gives SNR_{in}^2 for some typical spectra (with added 2 mm Al at the X-ray tube) used by current X-ray systems, as provided also by IEC 62220-1. The data of Boone and colleagues (Boone et al., 1997) can be used to calculate SNR_{in}^2 for spectra that are not included in the IEC standard, using the formula:

$$SNR_{in}^{\ 2} = \int \frac{\Phi(E,V)}{K_a} dE \tag{A7.5}$$

where $\Phi(E,V)$ is the photon fluence at energy E when a tube voltage V is applied. Table A7.1b gives SNR_{in}^2 values for some other spectra, calculated according to Boone et al. (1997).

MTF and NPS are sensitive image quality parameters and are sufficient to track changes in detector performance for QC purposes. The DQE is an important image quality metric when comparing the absolute performance of detectors, either of a similar type or between manufacturers. However, calculation of the DQE requires an accurate estimate of the air kerma at the detector input plane and the true number of photons per unit air kerma for the photon spectrum used. Both of these can be difficult to achieve in practice. The DQE can be compared against manufacturer reference data.

Limiting value None.
Frequency Every 6 months.
Equipment 2 mm Al filter, calibrated dose meter, software for calculating objective image quality parameters, spectral modelling tool.

Table A7.1a: Number of photons μGy^{-1} mm^{-2} (SNR_{in}^2) for 2 mm Al according to IEC standard 62220-1-2

Tube potential (kV)	Anode	System filter	SNR_{in}^2 (μGy^{-1} mm^{-2})
28	Mo	Mo (32 µm)	4981
28	Mo	Rh (25 µm)	5439
28	Rh	Rh (25 µm)	5944
28	W	Rh (50 µm)	5975
28	W	Al (500 µm)	6575

Table A7.1b: Number of photons µGy–1 mm–2 (SNRin2) for 2 mm Al according to Boone et al. (1997) for some spectra not included in the IEC standard

Tube potential (kV)	Anode	System filter	SNR_{in}^2 (μGy^{-1} mm^{-2})
29	Rh	Rh (25 µm)	6248
35	W	Al (500 µm)	8823
32	W	Ag (50 µm)	7143

Further reading

The following publications are recommended for further reading: Cunningham (2000), Neitzel et al. (2004), Carton et al. (2005), Dobbins et al. (2006), Samei et al. (2006), Marshall (2007), Marshall (2009a), Marshall (2009b) and Mackenzie et al. (2010).

Appendix 8: Description of CDCOM version 1.6

A8.1 The CDMAM phantom

The CDMAM phantom was designed to derive contrast-detail visibility threshold information for mammography systems by conducting four alternative forced choice (4-AFC) experiments. It contains a grid marking of 205 cells, and each cell contains one gold disc at the centre and one disc in one of the corners (Figure A8.1). The corner in which a disc is located varies randomly between cells. The combination of disc diameter and thickness is unique to each cell. A total of 16 different diameters and 16 different thicknesses are present in the phantom (Bijkerk et al., 2001; Bijkerk et al., 2002). The CDMAM phantom is built up as an aluminium base on which the gold discs have been deposited by evaporation. The base is attached manually to a PMMA cover, which contains the grid lines and information about disc thickness and diameter (Figure A8.2).

After imaging the CDMAM phantom, the 4-AFC experiment is conducted by asking observers to mark the corner of each individual cell containing a gold disc. Afterwards, the observer output is compared with the true disc locations. This produces a grid with 'true' and 'false' scores. To reduce the impact of isolated true or false scores in the matrix, a nearest neighbour correction scheme should be applied on this matrix before the detection threshold is derived for each diameter.

Combining these detection thresholds for multiple readers and images for each disc diameter produces the contrast-detail threshold information (Figure A8.3). For more detailed information, see Young et al. (2006) and Young et al. (2008).

Figure A8.1: The CDMAM phantom

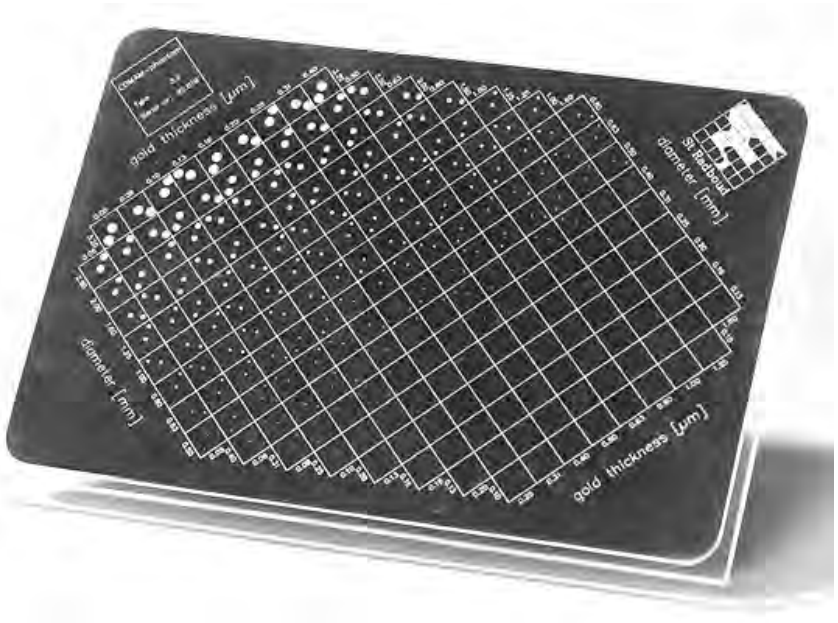


Figure A8.2: The aluminium base contains the gold discs of the CDMAM phantom; the PMMA cover contains the grid lines

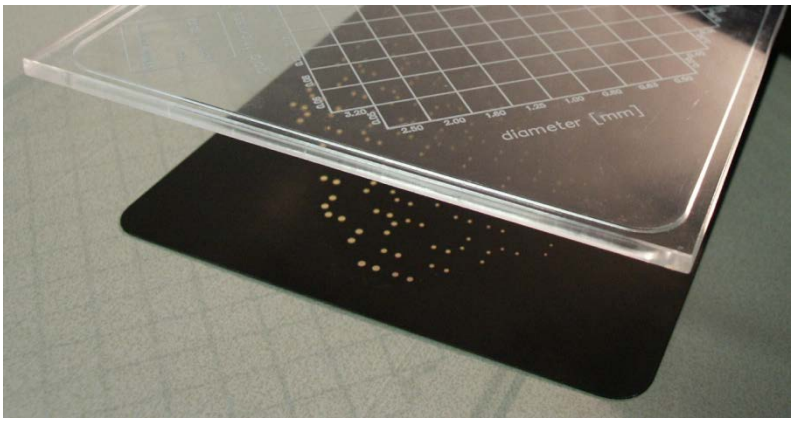
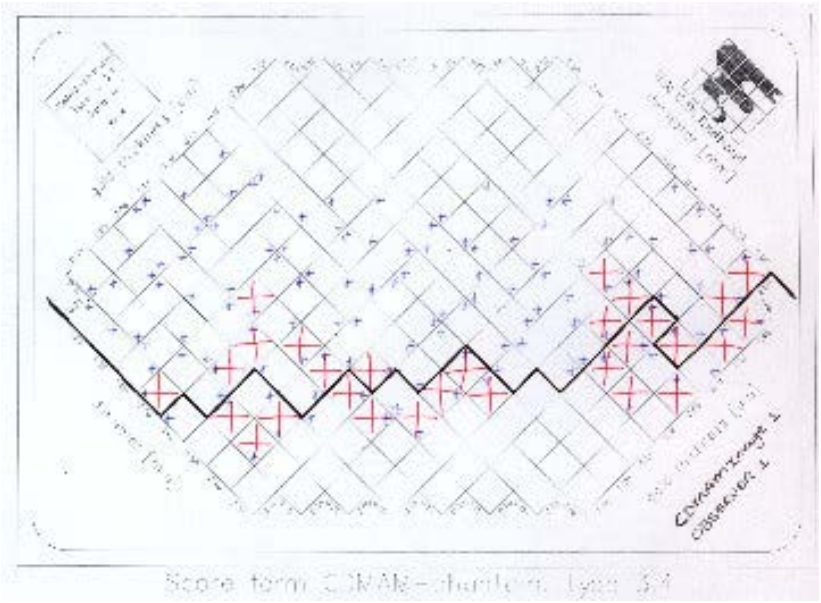


Figure A8.3: Example of human readout of the CDMAM phantom after nearest neighbour correction



A8.2 Disadvantages of human readout of the CDMAM phantom

Human readout of the CDMAM phantom does have several disadvantages:

- The reading of images is time-consuming.
- Due to the long reading time, in practice only small numbers of images and readers are used, which decreases reliability.
- Inter-reader variability, the variation in phantom image score from reader to reader.
- Intra-reader variability, the variation in phantom image score by one reader.
- There is a learning effect: readers may know the positions of the discs by heart. This influences their score.

- The nearest neighbour correction may not be the best scheme to obtain contrast-detail curves.

To eliminate and reduce these disadvantages, computer readout of CDMAM phantom images has been introduced.

A8.3 CDCOM

CDCOM is a software tool developed to automate the tasks of reading CDMAM version 3.4 phantom images and generating score sheets. **CDCOM does not read CDMAM version 3.2 images.** CDCOM does not aim to predict human readout, however. Instead, it tries to use the information present in the image more optimally to conduct a 4-AFC experiment. The results of CDCOM can therefore not be used directly in combination with contrast-detail visibility limits set for human observers. In Section 2b.2.4.1 of this protocol, the procedure is given to predict results for human observers from CDCOM results. To download software for automated readout of CDMAM images, see Visser & Karssemeijer (2012).

The automated image analysis by CDCOM can be separated into five phases:

- Analysis of the DICOM header
- Transformation to a standardized input
- Grid detection
- Phantom-specific and image-specific corrections
- Individual disc detection.

A8.4 Analysis of the DICOM header

CDCOM needs the following elements in the DICOM header to correctly assess the image:

- (0018,1164) Imager Pixel Spacing (used only as an initial estimator of image scale). If not present, use (0028,0030) Pixel Spacing.
- (0028,0010) Rows
- (0028,0011) Columns
- (0028,0100) Bits Allocated
- (0028,0101) Bits Stored
- (0028,1041) Pixel Intensity Relationship Sign (used to facilitate grid detection; can be overruled by manual selection). If not present, use (0028,0004) Photometric Interpretation.
- (7fe0,0010) Pixel Data.

A8.5 Transformation to a standardized input

The images of the CDMAM phantom are read by CDCOM and are transformed to a standardized form: (1) the image is cropped to the image of the phantom, (2) the image is rotated to the standardized orientation and (3) the image is scaled to a pixel size of 50 µm. This step is performed to eliminate differences in the detection of discs within the search area for images with different pixel size (see Section A8.8, Individual disc detection).

A8.6 Grid detection

As is the case for the human observer, the gridlines are used as a reference for global positioning. First, a linear Hough transform is performed at a reduced resolution of 400 µm per pixel, allowing the position of the gridlines to be detected as two sets of equidistant maxima in Euclidian space (Figure A8.4a). The cell corners can be calculated as the points where two gridlines intersect. Next,

the locations of the corners of each cell are determined more accurately by applying template matching (Figure A8.4b) to the area around the approximate locations of each cell corner. In this procedure, the template of the line crossing at the estimated position is correlated with the image data within a small area (21 × 21 pixels) around the predicted crossing. The location with the highest cross-correlation is the optimal location of the crossing. This step prevents image deformations (geometrical distortion), e.g. caused by a concave bucky shape, from influencing the results. No special template for the (partial) crossing at the boundaries is used; hence, this procedure may occasionally fail. Therefore, the consistency of the final result must be checked. If an outlying boundary crossing is found, this is corrected by interpolation from surrounding points (Veldkamp et al., 2003).

Figure A8.4a: Example of the result of a Hough transformation of a CDMAM image. The grid can be recognized as two columns with equidistant local maxima

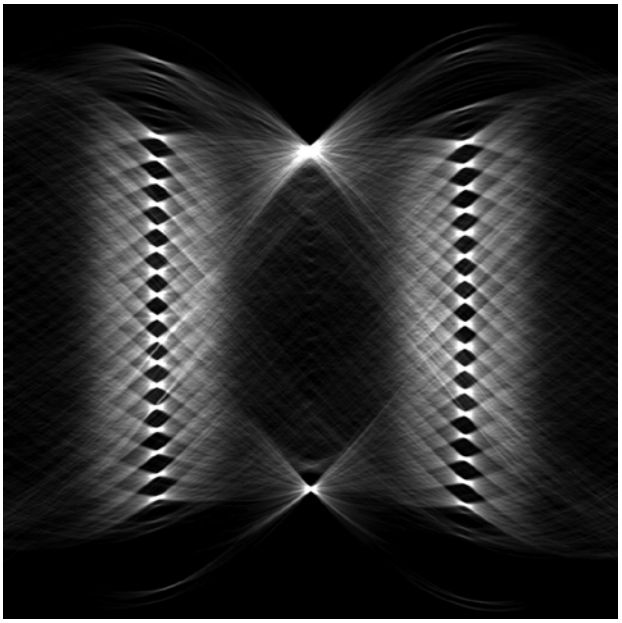
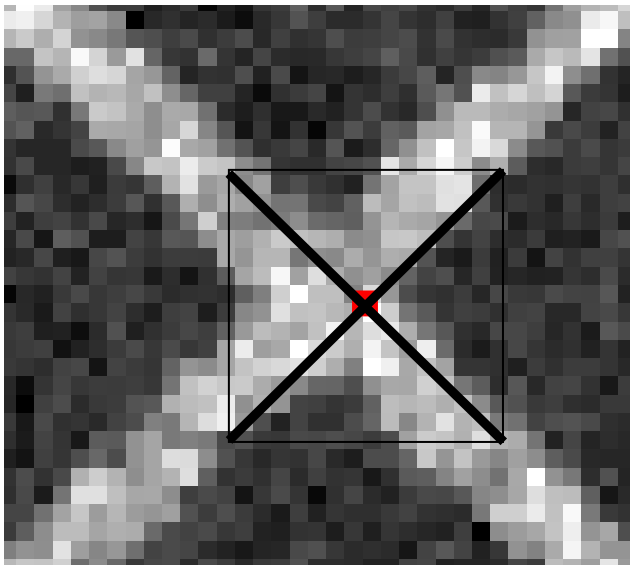


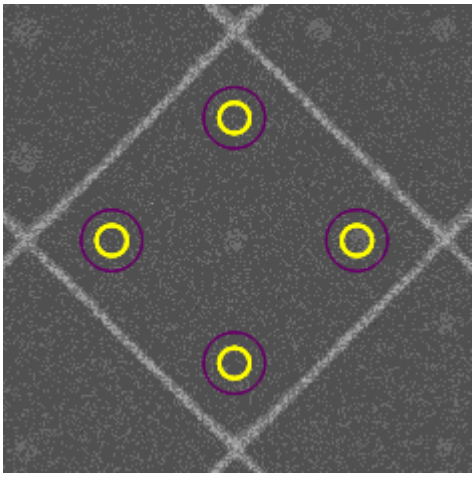
Figure A8.4b: Template matching: in an area of 21 × 21 pixels around the predicted crossing, the cross-correlation of the image with a template is determined



A8.7 Phantom-specific and image-specific corrections

CDCOM performs a 4-AFC experiment with an ideal observer model. It is based on the assumption that for each disc (with known diameter) there are four possible positions (yellow circles in Figure A8.5), of which the possible locations are exactly known. In practical situations, however, there will always be some inaccuracy in the determination of these positions (through human and computed observations) caused by the production process of the phantom, the geometrical distortion of the phantom image and the limited resolution of the imaging system. Therefore, CDCOM needs to vary slightly by applying the model observer in a search region around each calculated (theoretical) position (purple circles in Figures A8.5 and A8.6). It is crucial to understand that increasing the size of the search region will increase the influence of the image noise at the disc detection step. Therefore, increasing the search region will result in deterioration of the detection results, especially for smaller diameters. To limit the influence of the noise specific to every individual image, the disc search region has to be kept as small as possible, while avoiding the possibility of missing (part of) the disc under investigation.

Figure A8.5: Schematic illustration of the calculated disc positions (yellow inner circles) and the disc search areas defined around them (purple outer circles)



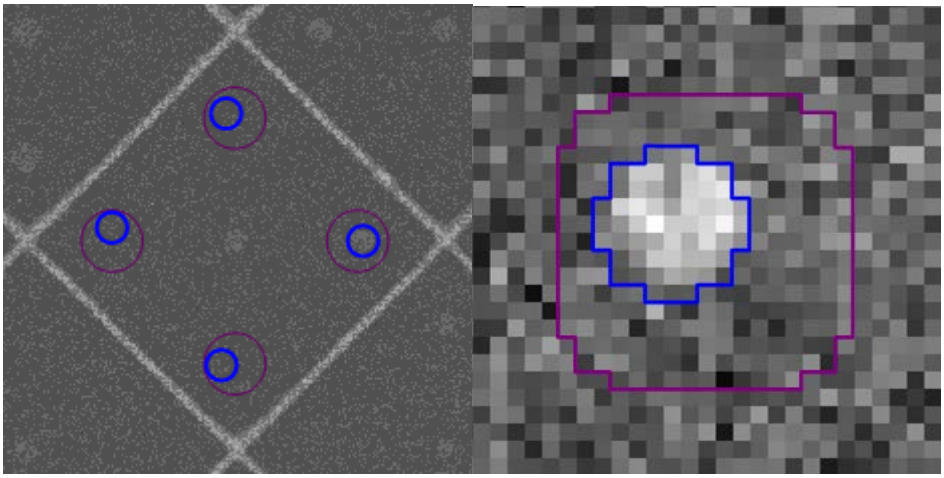
To avoid having to increase the search region, an estimate of the phantom-specific and image-specific translation, rotation and scaling is made by analysing the position of all easily detectable centre discs (large diameter and high contrast) using a relatively large search area of 500 μm^2. A template with the correct positions of all discs is rotated, translated and scaled to match the easily detectable discs, and calculated positions of all discs are obtained. Using a search area of 200 μm^2 around the calculated disc positions has proven to give reliable results (Visser et al., 2005).

A8.8 Individual disc detection

Due to the phantom-specific correction of the calculated disc positions, the disc search area for individual disc detection can be limited to 200 μm^2 around the calculated positions. Within this search area, the location in which a disc of the specified size is most likely to be located is determined.

This is done by finding the (by approximation) disc-shaped area that has the lowest or highest (depending on the image properties) total PV; see Figure A8.6a. In Figure A8.6b, the template is shown that is used to detect the circular discs on the phantom image. Due to the digital nature of the image, the shape of the circle is approximated by the template. This approximation would not be equal for different pixel sizes and would introduce differences in detection of the disc. Therefore, the CDMAM image is rescaled to a pixel size of 50 µm (see Section A8.5).

Figure A8.6: Individual disc detection: (left) the disc search areas in each of the cell corners (purple outer circles) and the positions within each search area (blue inner circles) where the disc is most likely to be located (based on total pixel value); (right) close-up of the detection (blue inner circle) of a 0.5 mm disc within a search area (purple outer circle)



The average values of the pixels in the four locations are then compared to decide which corner is most likely to contain the gold disc by determining the corner with the highest (or lowest, in the case of pixel values decreasing with higher object density) pixel value. If the corner selected is the corner actually containing the disc, CDCOM has correctly detected the disc; otherwise, it failed to detect it.

This process is repeated for the centre disc together with the 3 corners not containing a disc to obtain a second phantom image reading.

Appendix 9: Significance of test items

The test items described in the digital section of the European protocol for the quality control of the physical and technical aspects of mammography screening in the fourth edition and in Part 1 of the supplement to the digital section can be divided into four categories: (1) essential test items, which must be measured; (2) desirable test items, which should be measured; (3) optional test items, which could be measured and (4) test items that have been omitted in the supplement. This last category consists of (a) test items that have been proven to be of low importance, (b) test items that are (indirectly) covered by other tests and (c) test items that are difficult to measure by individual physicists and are covered by the supplier/manufacturer when installing a system.

Essential test items:

2b.2.2.1.1 Response function/2b.2.2.1.2 Noise evaluation
2b.2.1.3.3 AEC short-term reproducibility
2b.2.1.3.4 AEC long-term reproducibility (weekly/daily QC)
2b.2.1.3.5 AEC Breast thickness and composition compensation
2b.2.2.3.1 Image receptor homogeneity
2b.2.3 Dosimetry (requires 2b.2.1.2.2 Half value layer)
2b.2.4.1 Threshold contrast visibility
2b.2.4.5 Ghost image/erasure thoroughness

Desirable test items:

2b.2.1.2.1 Tube voltage
2b.2.1.3.1 Exposure control steps (if applicable)
2b.2.1.3.2 Back-up timer and security cut-off
2b.2.1.3.6 Local dense area
2b.2.1.4 Compression
2b.2.2.2 Missed tissue at chest wall side
2b.2.2.3.2 Detector element failure
2b.2.2.4 Interplate sensitivity variations
2b.2.4.3 Exposure time

Optional test items:

2b.2.1.1.3 Alignment of X-ray field/image area
2b.2.1.1.5 Tube output
2b.2.4.2 Modulation transfer function (MTF), noise power spectrum (NPS) and detective quantum efficiency (DQE)

Omitted test items:

2b.2.1.1.1 Focal spot size
2b.2.1.1.2 Source-to-image distance
2b.2.1.1.4 Radiation leakage
2b.2.1.5.1 Grid system factor
2b.2.1.5.2 Grid imaging
2b.2.2.5 Influence of other sources of radiation
2b.2.2.6 Fading of latent image

References

Bijkerk KR, Thijssen MAO, Arnoldussen ThJM (2001). Modification of the CDMAM contrast-detail phantom for image quality evaluation of full-field digital mammography systems. In: *IWDM 2000 5th International Workshop on Digital Mammography*. Yaffe MJ (ed.). Medical Physics Publishing, Madison, WI, pp. 633–640.

Bijkerk KR, Thijssen MAO, Arnoldussen ThJM (2002). *Manual CDMAM-phantom type 3.4*. University Medical Centre, Nijmegen, The Netherlands.

Boone JM, Fewell TR, Jennings RJ (1997). Molybdenum, rhodium, and tungsten anode spectral models using interpolating polynomials with application to mammography. *Med. Phys.*, 24:1863–1874.

Borasi G, Nitrosi A, Ferrari P, Tassoni D (2003). On site evaluation of three flat panel detectors for digital radiography. *Med. Phys.*, 30:1719–1731.

Bouwman R, Young K, Lazzari B, Ravaglia V, Broeders M, van Engen R (2009). An alternative method for noise analysis using pixel variance as part of quality control procedures on digital mammography systems. *Phys. Med. Biol.*, 54:6809–6822.

Carton AK, Vandenbroucke D, Struye L, Maidment AD, Kao YH, Albert M, Bosmans H, Marchal G (2005). Validation of MTF measurement for digital mammography quality control. *Med. Phys.*, 32:1684–1695.

Cunningham IA (2000). Applied linear system theory. In: *Physics and Psychophysics*, vol. 1. Beutel JKHL, Van Metter RL (eds.). SPIE, Bellingham, WA, pp. 79–159.

Dance DR (1990). Monte Carlo calculation of conversion factors for the estimation of mean glandular breast dose. *Phys. Med. Biol.*, 35:1211–1219.

Dance DR, Skinner CL, Young KC, Beckett JR, Kotre CJ (2000). Additional factors for the estimation of mean glandular breast dose using the UK mammography dosimetry protocol. *Phys. Med. Biol.*, 45:3225–3240.

Dance DR, Young KC, van Engen RE (2009). Further factors for the estimation of mean glandular dose using the United Kingdom, European and IAEA breast dosimetry protocols. *Phys. Med. Biol.*, 54:4361–4372.

Dance DR, Young KC, van Engen RE (2011). Estimation of mean glandular dose for breast tomosynthesis: factors for use with the UK, European and IAEA breast dosimetry protocols. *Phys. Med. Biol.*, 56:453–471.

Dobbins JT, Samei E, Ranger NT, Chen Y (2006). Intercomparison of methods for image quality characterization. II. Noise power spectrum. *Med. Phys.*, 33:1466–1475.

Evans DS, Workman A & Payne M (2002). A comparison of the imaging properties of CCD-based devices used for small field digital mammography. *Phys. Med. Biol.*, 47:117–135.

IEC (2004). Medical electrical equipment – characteristics of digital X-ray image devices – Part 1: Determination of the detective quantum efficiency. International Electrotechnical Commission, Geneva, Switzerland.

Jacobs J, Rogge F, Kotre J, Marchal G, Bosmans H (2007). Preliminary validation of a new variable pattern for daily quality assurance of medical image display devices. *Med. Phys.*, 34:2744–2758.

Mackenzie A, Doshi S, Doyle P, Hill A, Honey I, Marshall N, O'Neill J, Smail M (2010). *Measurement of the Performance Characteristics of Diagnostic X-Ray Systems: Digital Imaging Systems*. IPEM, York, United Kingdom, Report no. 32, part VII.

Marshall NW (2006). Retrospective analysis of a detector fault for a full field digital mammography system. *Phys. Med. Biol.*, 51:5655–5673.

Marshall NW (2007). Early experience in the use of quantitative image quality measurements for the quality assurance of full field digital mammography x-ray systems. *Phys. Med. Biol.*, 52:5545–5568.

Marshall NW (2009a). *Calculation of Quantitative Image Quality Parameters – Notes Describing the Use of OBJ_IQ_reduced*. NHSBSP, Sheffield, United Kingdom, Report no. NHSBSP Equipment Report 0902.

Marshall NW (2009b). Detective quantum efficiency measured as a function of energy for two full-field digital mammography systems. *Phys. Med. Biol.*, 54:2845–2861.

Neitzel U, Buhr E, Hilgers G, Granfors PR (2004). Determination of the modulation transfer function using the edge method: influence of scattered radiation. *Med. Phys.*, 31:3485–3491.

Perry N, Broeders M, de Wolf C, Törnberg S, Holland R, von Karsa L (eds.) (2006). *European guidelines for quality assurance in breast cancer screening and diagnosis. Fourth edition*. European Commission, Office for Official Publications of the European Communities, Luxembourg.

Robson KJ (2001). A parametric method for determining mammographic X-ray tube output and half value layer. *Br. J. Radiol.*, 74:335–340.

Samei E, Ranger NT, Dobbins JT, Chen Y (2006). Intercomparison of methods for image quality characterization. I. Modulation transfer function. *Med. Phys.*, 33:1454–1465.

van Engen R, Bosmans H, Heid P, Lazzari B, Schopphoven S, Thijssen M, Young KC (2013). Digital mammography update. European protocol for the quality control of the physical and technical aspects of mammography screening. S1, Part 2: European type testing. In: *European guidelines for quality assurance in breast cancer screening and diagnosis. Fourth edition, Supplements*. Perry N, Broeders M, de Wolf C, Törnberg S, Holland R, von Karsa L (eds.). European Commission, Office for Official Publications of the European Union, Luxembourg, pp. 55–71.

van Engen R, Swinkels MJ, Geertse TD, Oostveen LJ, Visser R (2006a). Using a homogeneity test as weekly quality control on digital mammography units. In: *Digital Mammography*. Astley S, Brady M, Rose M, Zwiggelaar R (eds.). Springer, Berlin, Heidelberg, pp. 259–265.

van Engen R, Young KC, Bosmans H, Thijssen M (2006b). European protocol for the quality control of the physical and technical aspects of mammography screening. Chapter 2b: Digital mammography. In: *European guidelines for quality assurance in breast cancer screening and diagnosis. Fourth edition.* Perry N, Broeders M, de Wolf C, Törnberg S, Holland R, von Karsa L (eds.). European Commission, Office for Official Publications of the European Communities, Luxembourg, pp. 105–165.

Van Ongeval C, Van Steen A, Geniets C, Dekeyzer F, Bosmans H, Marchal G (2008). Clinical image quality criteria for full field digital mammography: a first practical application. *Radiat. Prot. Dosimetry.*, 129:265–270.

Veldkamp WJ, Thijssen MA, Karssemeijer N (2003). The value of scatter removal by a grid in full field digital mammography. *Med. Phys.*, 30:1712–1718.

Visser R, Beckers B, Gennaro G, Baldelli P, van Engen R (2005). Comparison of individual CDMAM phantoms using automated image analysis. In: *Proceedings of the International Workshop in Digital Mammography (IWDM 2004)*. University of North Carolina, Chapel Hill, NC, pp. 123–130.

Visser R, Karssemeijer N (2012). CDCOM Manual: software for automated readout of CDMAM 3.4 images.

Young KC (2001). *Breast Dose Surveys in the NHSBSP: Software and Instruction Manual*. NHSBSP, Sheffield, United Kingdom, Report no. NHSBSP Report 01/10.

Young KC, Al Sager A, Oduko JJM, Bosmans H, Verbrugge B, Geertse T, van Engen R (2008). Evaluation of software in reading images of the CDMAM test object to assess digital mammography systems. In: *Proceedings of SPIE Medical Imaging 2008*. Hsieh J, Samei E (eds.), pp. 6913–6947.

Young KC, Cook JJH, Oduko JJM (2006). Automated and human determination of threshold contrast for digital mammography systems. In: *Digital Mammography*. Flynn MJ, Hsieh J (eds.). Springer, Heidelberg, Berlin, pp. 266–272.

Young KC, Cook JJH, Oduko JJM, Bosmans H (2006a). Comparison of software and human observers in reading images of the CDMAM test object to assess digital mammography systems. In: *Proceedings of SPIE Medical Imaging 2006*. Flynn MJ, Hsieh J (eds.), pp. 1–13.

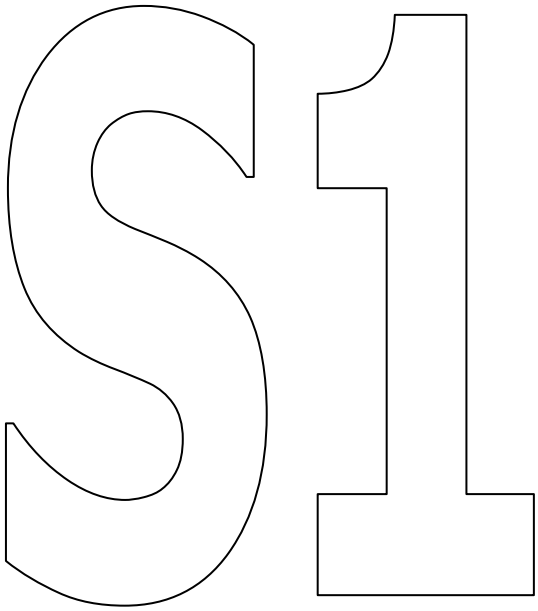
Young KC, Oduko JJM, Bosmans H, Nijs K, Martinez L (2006b). Optimal beam quality selection in digital mammography. *Br. J. Radiol.*, 79:981–990.

European guidelines for quality assurance in breast cancer screening and diagnosis

Fourth Edition

Supplement

Digital mammography update



European protocol for the quality control of the physical and technical aspects of mammography screening

Part 2

European type testing

Authors
R. van Engen
H. Bosmans
P. Heid
B. Lazzari
S. Schopphoven
M. Thijssen
K. Young

Authors
R. van Engen, The Netherlands*
H. Bosmans, Belgium*
P. Heid, France*
B. Lazzari, Italy*
S. Schopphoven, Germany*
M. Thijssen, The Netherlands*
K. Young, United Kingdom*

Contributors
J. A. Lelivelt, J. Jacobs, C. van Ongeval

* Member of the EUREF Physico-Technical Steering Group

Declarations of interest
An interest of H. Bosmans is reported on page IV.

Disclaimer
The views expressed in this document are those of the authors and do not necessarily reflect the official position of the European Commission. Neither the European Commission nor any other organization or any individual may be held responsible for any use that may be made of the information contained herein.

Acknowledgements
Financial support was provided by the European Union Public Health Programme (Project no. 2006322, European Cooperation on Development and Implementation of Cancer Screening and Prevention Guidelines [ECCG]).

Please cite this publication as follows
van Engen RE, Bosmans H, Heid P, Lazzari B, Schopphoven S, Thijssen M, Young KC (2013). Digital mammography update. European protocol for the quality control of the physical and technical aspects of mammography screening. S1, Part 2: European type testing. In: *European guidelines for quality assurance in breast cancer screening and diagnosis. Fourth edition, Supplements*. Perry N, Broeders M, de Wolf C, Törnberg S, Holland R, von Karsa L (eds.). European Commission, Office for Official Publications of the European Union, Luxembourg, pp. 55–71.

Corresponding author
R. van Engen
EUREF office
National Expert and Training Centre for Breast Cancer Screening
Radboud University Nijmegen Medical Centre
P.O. Box 6873
6503 GJ Nijmegen
The Netherlands
R.vanEngen@euref.org
info@euref.org

Table of contents

List of figures

1 Introduction

The protocol described in Part 2 of this supplement has been developed to test whether digital imaging systems of a given type or brand are fundamentally capable of fulfilling the acceptance criteria of the European protocol for the quality control of the physical and technical aspects of mammography screening (van Engen et al., 2006), which has been updated in Part 1 of this supplement (van Engen et al., 2013). The European type testing protocol takes into account the updates reported in Part 1 of this supplement and it provides guidelines for best practice in controlling dose and (clinical) image quality. Differences between the type testing protocol and the applicable standards and procedures for acceptance and constancy testing are also pointed out in this second part of the supplement. After a successful type test, individual mammography units of the same type or brand still need to undergo an acceptance test before clinical use.

The European protocol for type testing is currently developed for digital mammography (DR and CR) systems. In the future, the protocol may be expanded to include type testing of image-processing algorithms, workstations and film digitizers.

The authors welcome all comments and feedback on this document to improve the standards and protocols. Future updates of the current version will be made available on the web sites of the European Commission (http://ec.europa.eu/health/index_en.htm) and EUREF (www.euref.org).

2 Type testing procedure

2.1 Technical and clinical phases

A type test encompasses two evaluation phases: technical evaluation and clinical evaluation.

- **Phase 1: Technical evaluation**

In the first phase, two full physico-technical evaluations are performed on different systems in different locations. The medical physics experts testing the equipment are appointed by the medical physics team of the unit conducting the type test. They must be highly experienced in physico-technical quality control of mammography systems and in quality assurance of clinical aspects of breast cancer screening and diagnosis. The type test applicant (i.e. the manufacturer) must arrange in advance the locations and all logistic preparations with the unit conducting the type test. The systems must be available for the physico-technical evaluations for a period of at least 3 days. The physico-technical tests are primarily based on the measurements specified in Chapter 2b of the fourth edition (van Engen et al., 2006) and the updates to that chapter (van Engen et al., 2013). However, some tests have been modified for type testing, and some additional tests that are described in Section 3.2 must be performed.

After all results of the technical evaluation are available, the medical physics team of the unit conducting the type test will meet to discuss the results and decide whether phase 2, the clinical evaluation, can be initiated. To prepare the decision, the results of the physico-technical evaluation are compiled in a report and comments on the findings in the report are requested from the applicant and from the medical physics team of the unit conducting the type test. The medical physics team will consider whether problems have been detected that would need to be solved in order to pass the test, and if

so, whether the applicant is likely to be able to respond effectively. If the team determines that a solution is unlikely, the type test may be aborted by either the medical physics team or the applicant.

If the two physico-technical evaluations yield conflicting results, the medical physics team may decide that (some) tests must be repeated on a third system.

For CR systems, the physico-technical tests must be performed on two X-ray units of different brands. The applicant should provide appropriate information on the set-up of the X-ray units for the applicant's CR system and ensure that the mammography units are set up correctly for the applicant's CR plates at the sites of type testing.

- **Phase 2: Clinical evaluation**

After the technical evaluation phase is completed, the performance of the digital system is evaluated under clinical use for a period of at least 3 months at one of the sites at which the technical evaluation was performed. At this site, reading must be by soft copy and all workstations must pass the European protocol for viewing conditions. The site must also have sufficient workflow (i.e. similar to that used in European screening centres with an average of at least 50 clients or patients per system per day). The site is chosen by the applicant in cooperation with the unit conducting the type test.

Before the clinical test is allowed to start, images of 50 patients must be sent to the unit conducting the test for pre-review to determine whether the configuration of the image processing requires optimization. At the start of the clinical test, an application specialist of the manufacturer requesting the type test and a representative of the unit conducting the type test will be present at the clinical site for a period of at least 2 days to help start the clinical evaluation.

During the first week of the clinical evaluation, experts of the unit conducting the evaluation perform an initial assessment of the images. If image quality is not satisfactory or if dose exceeds permissible levels, adjustments to the equipment must be made by the applicant. If image quality and/or dose levels remain unsatisfactory, the clinical evaluation may be aborted by the unit conducting the evaluation. Further adjustments during the period of clinical testing are only allowed after consultation with the unit conducting the evaluation. Some adjustments may require an additional technical evaluation.

During the clinical evaluation period, a homogeneity image must be acquired every day in fully automatic mode to monitor the stability of the equipment. A record must be kept of all artefacts on clinical images and of all problems that occur with the equipment, by the radiographers/radiologists at the test site. At the end of the clinical evaluation period, a bad pixel map is obtained to determine the number of detector elements that became defective during the evaluation period.

When problems occur, the medical physics group of the unit conducting the type test has the right to extend the clinical evaluation period as necessary or to abort the clinical evaluation.

In addition to the stability test, a dose survey is conducted. For the dose survey, X-ray exposure data must be available, and the respective images must be available for verification. The mean glandular dose recorded in the DICOM header is compared with the values from the dose survey.

At the end of the clinical testing period, a set of clinical images is evaluated by a team of experts consisting of two radiologists with substantial experience in digital mammography, who are invited by the unit conducting the type test, and one of the members of the unit's medical physics team. The evaluated images consist of two groups: 50 patients or clients selected based on breast thickness, and 20 cases randomly selected out of referred patients. If a system with more than one mode is evaluated, at least 30 images of each mode are required.

In the event that the evaluation team is unable to reach a firm conclusion about the adequacy of overall image quality using the above-mentioned clinical images, additional evaluation may be required, e.g. through insertion of simulated lesions in unprocessed images to assess image processing.

2.2 Duration of a type test

After the two technical evaluations are performed, the results of both systems will be discussed by the unit conducting the type test, and permission to start the clinical evaluation will be given, depending on the adequacy of the technical results. The duration of the clinical evaluation is at least 3 months.

2.3 Final report

After completion of the clinical evaluation, the unit conducting the type test prepares a final report; it includes the overall result of the type test and is sent to the applicant. If a system passes, the results will be published on the web site of the unit conducting the test. If a system fails, the results will not be published on the web site (see Section 2.5).

2.4 Non-disclosure of confidential information

Type tests might be performed on systems using techniques that are different from currently existing systems. For such systems, current methods of measurement might be unsuitable and adaptations may be necessary and new test items may need to be added that are not specified in the current version of the type testing protocol. The applicant must therefore provide the medical physics team of the unit conducting the type test with all relevant information on the system being tested.

Relevant information includes: basic principles of the mammography system and particularly the image receptor, philosophy and operation of the AEC system, pre-exposure parameters, reconstruction technique of bad pixels, accepted number of bad pixels, etc. It is recommended that the unit conducting the type test consult the EUREF Physico-Technical Steering Group before initiating a type test, for advice on whether test items need to be adapted and/or new test items need to be added. If requested by the applicant, any relevant information provided can be regarded as confidential and a non-disclosure agreement can be signed with the parties involved.

2.5 Publication of results

The full report on the results of the type test will be made available to the applicant. If a system passes the type test, the results will be published on the web site of the unit conducting the type test. The applicant of the test will be able to comment on the report before publication on the web site. The unit conducting the type test will not use the information obtained from performing or reporting the results of the test in other publications without the consent of the test applicant.

In consensus with the applicant, it may be mentioned on the web site of the unit performing the test whether a specific system is under (technical or clinical) evaluation. If a system fails, its name will be

removed from the list. It will not be mentioned on the web site that a specific system has failed the type test.

3 Type testing protocol

By definition, type testing is performed on new types of equipment. Therefore, type tests may be performed on systems that use techniques that are different from those of existing systems. For such systems, the current methods of measurement could be unsuitable and adaptations may be necessary. It is also possible that new, additional tests will be required. Therefore the medical physics team of the unit conducting the test must be provided by the system manufacturer with all relevant information on the system being type tested. If measurement techniques require adaptation, this must be communicated in advance to the applicant whenever possible by the unit conducting the test. If differences are noticed during type testing, adaptations of methods of measurement will be made on the spot and discussed afterwards. This discussion must take place before the second physico-technical test is performed.

It is recommended that the unit conducting the type test consult the EUREF Physico-Technical Steering Group before initiating a type test, for advice on whether test items need to be adapted and/or new test items need to be added. Prior contact also facilitates further communication in the event that ad hoc advice is requested during performance of measurements.

Section 3.1 provides an overview of all technical tests performed in a European type test. Some additional clinical tests are described in Section 3.2.

For CR systems, at least 4 cassettes of standard size (18 × 24 cm) and 4 cassettes of large size (24 × 30 cm) should be available during the technical evaluation; at the site of the clinical evaluation, at least 8 cassettes should be available. If cassettes of large size are available, they may also be used at the clinical site.

3.1 Technical evaluation protocol

3.1.1 Relevant test items based on Chapter 2b of the fourth edition

The following test items described in Chapter 2b of the fourth edition (van Engen et al., 2006) and in Part 1 of this supplement (van Engen et al., 2013) are measured in a European type test technical evaluation:

2b.2.1.1.5 **Tube output:** Use the method described in Part 1 of this supplement. Do not use limiting values.

2b.2.1.2 **Tube voltage and beam quality**

2b.2.1.2.1 **Tube voltage:** Use the method described in Chapter 2b.

2b.2.1.2.2 **Half value layer:** Use the method described in Part 1 of this supplement.

2b.2.1.3 **AEC system**

2b.2.1.3.2 **Back-up timer and security cut-off:** Use the method and limiting values in Chapter 2b.

2b.2.1.3.3 **Short-term reproducibility:** Use the method and limiting values in Part 1 of this supplement.

2b.2.1.3.4 **Long-term reproducibility:** Use the method in Chapter 2b. Use the limiting values of Chapter 2b as action limits for further investigation. In the final report, present measurements on long-term stability from the clinical test period.

2b.2.1.3.5 **Breast thickness and composition compensation:** Use the method and limiting values in Chapter 2b (see 'Object thickness') and in Part 1 of this supplement.

2b.2.1.3.6 **Local dense area:** Use the method and limiting values in Part 1 of this supplement.

2b.2.2 **Image receptor**

2b.2.2.1 **Image receptor response**

2b.2.2.1.1 **Response function:** Use the method and limiting values in Chapter 2b.

2b.2.2.1.2 **Noise evaluation:** Use the method and limiting values in Part 1 of this supplement.

2b.2.2.2 **Missed tissue at chest wall side**: Use the method and limiting values in Chapter 2b.

2b.2.2.3 **Image receptor homogeneity and stability**

2b.2.2.3.1 **Image receptor homogeneity:** Use the method and limiting values in Chapter 2b and in Part 1 of this supplement.

2b.2.2.3.2 **Detector element failure (DR systems):** Use the method and limiting values in Chapter 2b. The bad pixel map should be easily accessible for all users. If uncorrected bad pixels are visible on the images, this should be taken into account when evaluating detector element failure.

2b.2.2.3.3 **Uncorrected defective detector elements (DR systems):** Use the method and limiting values in Chapter 2b.

2b.2.2.4 **Interplate sensitivity variations (CR systems):** Use the method and limiting values in Chapter 2b of the fourth edition and in Part 1 of this supplement. At least 4 cassettes of each size should be present.

2b.2.2.6 **Fading of latent image (CR systems):** Use the method in Chapter 2b.

2b.2.3 **Dosimetry:** Use the method and limiting values in Chapter 2b and in Part 1 of this supplement.

2b.2.4 **Image quality**

2b.2.4.1 **Threshold contrast visibility:** Use the method and limiting values in Chapter 2b and in Part 1 of this supplement.

Threshold contrast visibility is determined at several dose levels: the clinical glandular dose level and 2–4 other dose levels, which will be chosen such that a large range of dose levels is covered (e.g. between ½ and 2 times the clinical dose level).

For details of 0.1, 0.25, 0.5 and 1.0 mm diameter, threshold contrast is plotted against glandular dose. The dose is calculated at which the minimum acceptable image quality standard, and the achievable image quality standard, is obtained for the 0.1, 0.25, 0.5 and 1.0 mm diameter objects on the contrast threshold visibility phantom.

2b.2.4.2 **Modulation transfer function (MTF), noise power spectrum (NPS) and detective quantum efficiency (DQE):** Use the method described in Appendix 7 of Part 1 of this supplement.

2b.2.4.3 **Exposure time:** Use the method and limiting values in Chapter 2b.

2b.2.4.4 **Geometrical distortion and artefact evaluation:** Use the method and limiting values in Chapter 2b.

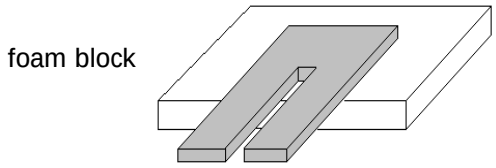
2b.2.4.5 **Ghost image/erasure thoroughness**: Use the method and limiting values in Chapter 2b.

3.1.2 Thickness indication

European type testing includes a clinical evaluation (see Section 3.2). In the clinical evaluation, a dose survey is conducted that requires prior checking of the indicator of the height of the compression paddle.

Two foam blocks with compressed thickness of about 20 mm and 40 mm are used for this measurement. A strip is cut out of the foam block to allow measurement of thickness during compression (see Figure 2.1). Thickness indication can be checked when the foam blocks (180 mm × 240 mm) are placed on the bucky. Position the blocks such that half of the block is positioned on the bucky and half of the block is positioned over the edge of the bucky at the chest wall side (see Figure 2.1). Apply compression (approximately 100 N), record the thickness indication and measure thickness at the reference point with an appropriate device (e.g. a calliper). Perform this measurement for the two foam blocks separately and together (so that measurements can be done at about 20, 40 and 60 mm compressed thickness).

Figure 2.1: Thickness indication check



3.2 Clinical evaluation protocol

3.2.1 Introduction

The imaging chain consists of several components, beginning with a patient or client who is being X-rayed and extending to the radiologist who is reading and reporting the images. During physico-technical evaluation, each component is checked separately. The clinical part of the European type testing protocol is essentially a global test. The quality of clinical images is affected interdependently by all components, including the image processing, which is not evaluated technically. The clinical part of the protocol therefore cannot be conducted separately; it must be performed on a system that has passed the technical evaluation.

The clinical evaluation includes an assessment of the image quality by a team of experts including radiologists, a stability test and a clinical dose evaluation. The required duration of the clinical evaluation is 3 months, but if problems occur, the unit conducting the clinical evaluation may decide to extend the period accordingly or to abort the evaluation.

The site at which the clinical images are obtained will be mutually decided by the unit conducting the clinical evaluation and the applicant. It must be one of the two sites at which the preceding technical evaluation was performed. It must have a sufficient workflow of 4 images per patient/client, and soft-copy reading is required. The clinical images will be sent to and scored by the unit conducting the clinical evaluation. The workstations used to evaluate the images must pass the European guidelines quality criteria. If an applicant prefers that the images be scored on the applicant's own workstation, the workstation must be provided to the unit conducting the evaluation. Before use, the applicant's workstation will be checked to verify that it passes the European guidelines quality criteria.

3.2.2 Procedure for acquiring the images

The clinical evaluation will be performed for a minimum of 3 months. Before the evaluation can start, anonymized images of 50 patients must be sent to the unit conducting the clinical evaluation for pre-review to determine whether the configuration of the image processing requires optimization. At the beginning of the clinical evaluation, a representative of the unit conducting the evaluation will be present at the clinical test site, together with a specialist of the applicant. The unit conducting the evaluation will install a hard disk onto which all images from the clinical test period will be copied. The DICOM headers of all images must be correctly filled in, but all patient or client data on the hard disk must be anonymized. Any image quality issues that may occur at the clinical test site, such as artefacts, must be recorded and reported to the unit conducting the evaluation.

3.2.3 Procedure for scoring the images

The images of 70 selected women will all be scored by a reading team consisting of radiologists and a physicist who have experience in digital mammography and in the assessment of image quality. These 70 cases will be selected from the images collected during the clinical test period as follows. First, the information on the DICOM headers is used to calculate the doses of all women who received a mammogram during the clinical test period. The data are also used to plot the mean glandular dose as a function of compressed breast thickness. This permits both determination of the dose distribution of clinical images and systematic selection of the cases to be scored by the reading team (all thicknesses, all types of tissues). The images are divided into 5 classes based on compressed breast thickness. In each thickness class, patients are ranked as a function of their dose; 10 patients are then selected in ranked series from each of the 5 classes using the image at the 10th percentile of the dose distribution, the image at the 20th percentile, the image at the 30th percentile, etc. A total of 50 cases are chosen in this manner. The other 20 cases are randomly selected from referred patients; the respective images must include opacities and microcalcifications.

Reading for image scoring is performed simultaneously by two radiologists and a physicist. A third radiologist with substantial experience in digital mammography and evaluating image quality is appointed in case of image quality problems. This could be either a radiologist from the unit conducting the clinical evaluation or an external radiologist.

3.2.4 The scoring form

The scoring form is based on experience with type testing in the Belgian and Dutch breast screening programmes (Van Ongeval et al., 2005; Van Ongeval et al., 2008). The first part (13 questions) requires 'yes' or 'no' answers from the panel of experts; it includes questions about anatomical

structures, noise, and contrast in low and high pixel value areas. In the second part (4 questions), the panel of experts scores the contrast and sharpness of the images on a scale between –2 and +2. In the third part (3 questions), the panel of experts rates the images on a scale from 1 to 10 according to certain characteristics and their potential impact on the interpretation of the image. The scoring is performed at the workstation where the reading team views the images.

3.2.5 Procedure for evaluating the data

The unit conducting the type test should have software for data entry during evaluation and analysis, including comparison with results of previous testing of other types. Ideally, the program should display the images and the score could be entered directly for the respective image. The software can be used to generate the median score for each question and also the number of 'yes' or 'no' answers for each question. The results are compared with other systems; any significantly lower value on a question for a system will be investigated in detail in collaboration with the third radiologist.

3.2.5.1 Dose evaluation

A dose survey is conducted during the clinical evaluation period; this requires access to the relevant X-ray exposure data and breast thickness data, which must be available in the image header. The mean glandular dose is compared with calculated values. When testing CR systems, the applicant should ensure that the mammography unit and reader are connected to fill in the DICOM header with the exposure values, breast thickness and dose indicator.

3.2.5.2 Stability test

During the clinical evaluation period, an image of a 45 mm homogeneous block of PMMA covering the whole image receptor must be acquired every day in fully automatic mode to monitor the stability of the equipment. A record must be kept of all artefacts on clinical images and of all problems that were noticed by the radiographers/radiologists at the test site. At the beginning and at the end of the clinical evaluation period, a bad pixel map may be obtained to determine the number of detector elements that became defective during this test period. These QC images and records will be sent to the unit conducting the type testing for evaluation.

3.2.5.3 Additional comments

Comments about potential special features, ergonomics and other relevant characteristics or conditions that may affect the results will be noted by the representative of the unit conducting the evaluation, who will be present at the start of the clinical evaluation period. In addition, comments may also be requested from the radiologists and radiographers at the clinical evaluation site.

3.2.6 Requirements for the set-up at the clinical test site

The software set-up at the clinical test sites is shown in Figure 2.2:

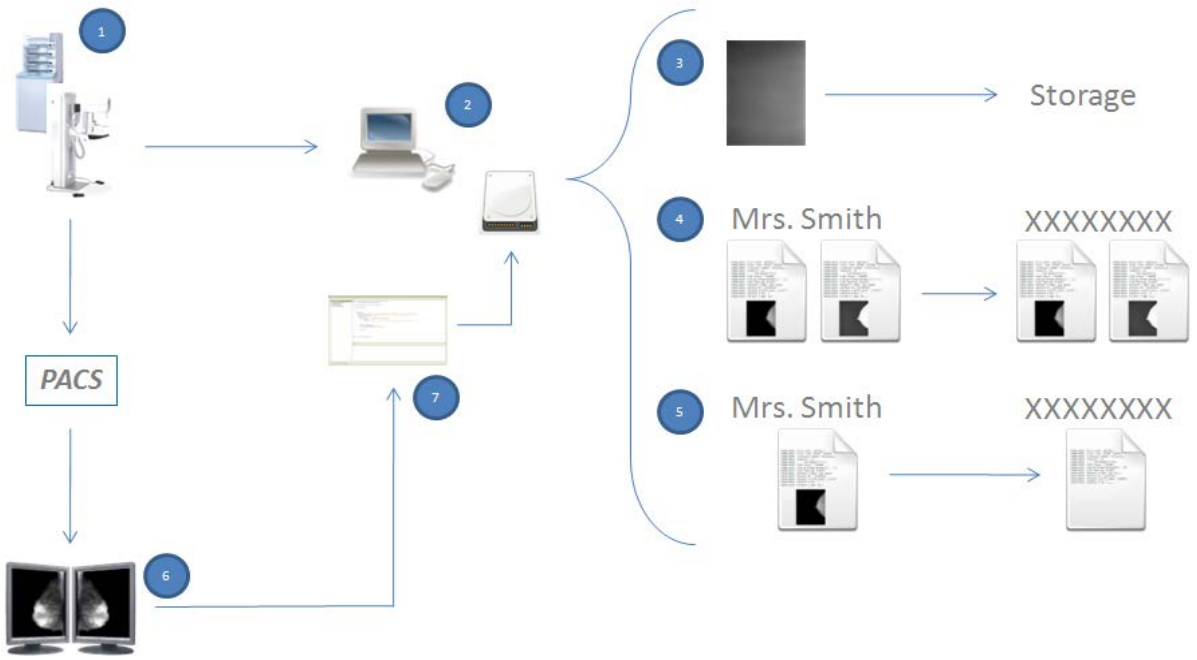
1. All image data (both clinical and technical) are sent in parallel to the Picture Archiving and Communication System (PACS) normally used at the site and also to a local, dedicated DICOM Service Class Provider (SCP). To enable this, the vendor must configure a DICOM output node at the mammography unit.

2. A software package (e.g. Gladys) (Jacobs et al., 2006) is installed on a local computer. This is a DICOM SCP that can be scripted to perform specific actions based on the values of the DICOM headers. All data are stored on an external hard drive, provided by the unit conducting the evaluation. In the paragraphs below, the requirements for Gladys are used as an example. Other software packages with similar functionality are also allowed. The requirements for these other packages may differ.

Requirements:

a. A computer that is connected to the mammography unit via the network
b. A free USB 2.0 port
c. A fixed IP address
d. An administrator access to this computer to install the software package as a Windows service.

Figure 2.2: Software set-up at clinical test site

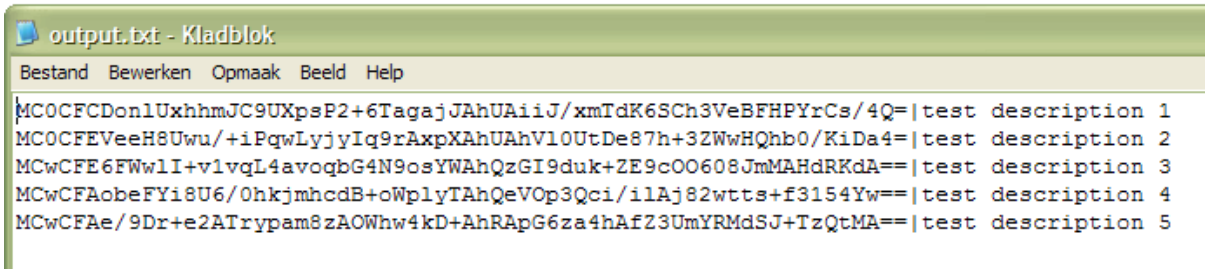


3. For the stability test, the homogeneous images (only 'for processing' images) are stored by Gladys on the external hard drive. The technical analysis is performed in a QC reference centre. To distinguish between patient data and the homogeneity data, the patient name ('QCMAMMO') is taken by convention and action is undertaken by Gladys based on this patient name.
4. All clinical images (both 'for presentation' and 'for processing' images) are also sent to Gladys. Using a one-way hash algorithm, these image data are anonymized and stored on the external hard drive in a specific folder structure. The DICOM headers that need to be anonymized can be different at each clinical test site, and Gladys can therefore be configured. This action will be performed on all data with a patient name that is different from 'QCMAMMO'.

5. For the collection of patient dosimetry data, a second copy of the anonymized patient data (only 'for presentation' images) is stored without any image information. These DICOM files are used by the unit conducting the evaluation to simplify the patient dosimetry study.
6. For the collection of the selected 20 referred cases, local radiologists are asked to note down the unique identifiers of at least 20 possible cases. They must include opacities and microcalcifications. To preserve confidentiality, the unique identifiers of the selected clinical cases must be anonymized at the clinical test site. For this purpose, the identifiers are saved as a text file in a format similar to that shown in Figure 2.3a. Gladys converts this text file using the same one-way hash algorithm as in (4). An example output file can be seen in Figure 2.3b. These steps assure that no client/patient-related information leaves the clinical test site. This procedure is performed by a representative of the unit conducting the evaluation.

Figure 2.3a: Example input format for selected patient cases

Figure 2.3b: Example output format for selected patient cases



3.2.7 Set-up at the reference site

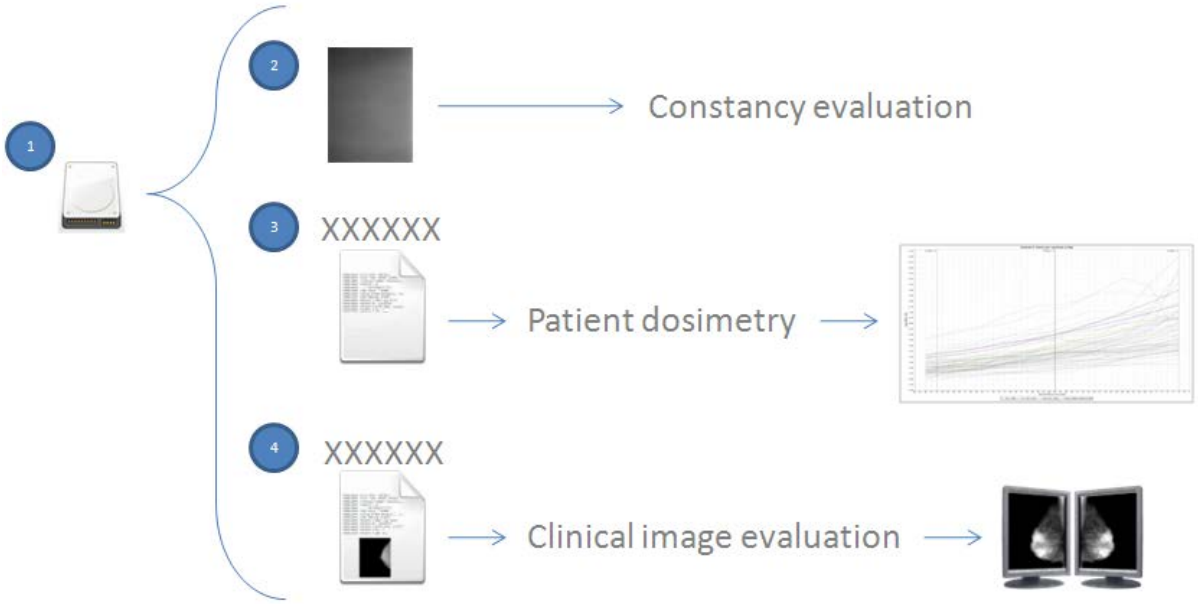
The set-up at the reference site of the unit conducting the evaluation is shown in Figure 2.4:

1. All data are transferred from the clinical test site to the reference site using the external hard disk provided by the unit conducting the evaluation.
2. The homogeneous acquisitions are evaluated for artefacts, and the stability over time of the mammography unit is monitored.
3. The DICOM headers without the image data are scanned for dosimetry-related DICOM data. These header values are used to calculate the patient dose using the method of Dance (Dance,

1990; Dance et al., 2000; Dance et al., 2009; Dance et al., 2011). Calculated values are compared with the current European guidelines levels.

4. The clinical image quality is evaluated using a dedicated software platform to perform observer studies by means of a visual grading analysis following the clinical evaluation form, which can be found at the end of this document (Section 3.2.8) (Jacobs et al., 2008). This task must be performed in a controlled environment using DICOM-calibrated viewing stations. If the applicant prefers to use a different workstation, that must be mutually agreed with a representative of the unit conducting the evaluation before the clinical evaluation period.

Figure 2.4: Software set-up at reference site



3.2.8 Clinical evaluation form for European type testing

Client number ☐

Please answer the following questions with 'yes' or 'no' (or N/A):

		Yes	No	N/A
1	Is the skin line visualized well?			
2	Are the vascular structures visible through the dense parenchyma?			
3	Is the pectoral muscle visualized sharply?			
4	Are the Cooper's ligaments and vascular structures in the subcutaneous and pre-pectoral area visualized well?			
5	Are the microcalcifications well visualized and well outlined?			
6	Is the contrast in the dark areas sufficient (e.g. no saturation of intensity of signals, no fully dark regions)?			
7	Is the contrast in the white areas sufficient (e.g. no fully white regions)?			
8	Is the glandular tissue sufficiently white?			
9	Is the background sufficiently dark?			
10	Do all images appear in the same way? (if not, please comment)			
11	Is there disturbing noise in the dark areas?			
12	Is there disturbing noise in the white areas?			
13	Are there any artefacts?			

For the following questions, please grade the images with a number from –2 (bad) to +2 (good). Please use the whole range.

		–2	–1	0	1	2
1	Contrast in the white regions					
2	Contrast in the dark regions					
3	Overall contrast					
4	Sharpness					

For the following questions, please rate the images with a number from 1 (bad) to 10 (good). Please use the whole range.

		1	2	3	4	5	6	7	8	9	10
5	How satisfied are you with the representation of micro-calcifications?										
6	How satisfied are you with the representation of opacities?										
7	How satisfied are you with the representation of the image?										

Remarks:

Thank you very much for providing this assessment.

References

Dance DR (1990). Monte Carlo calculation of conversion factors for the estimation of mean glandular breast dose. *Phys. Med. Biol.*, 35:1211–1219.

Dance DR, Skinner CL, Young KC, Beckett JR, Kotre CJ (2000). Additional factors for the estimation of mean glandular breast dose using the UK mammography dosimetry protocol. *Phys. Med. Biol.*, 45:3225–3240.

Dance DR, Young KC, van Engen RE (2009). Further factors for the estimation of mean glandular dose using the United Kingdom, European and IAEA breast dosimetry protocols. *Phys. Med. Biol.*, 54:4361–4372.

Dance DR, Young KC, van Engen RE (2011). Estimation of mean glandular dose for breast tomosynthesis: factors for use with the UK, European and IAEA breast dosimetry protocols. *Phys. Med. Biol.*, 56:453–471.

Jacobs J, Deprez T, Marchal G, Bosmans H (2006). The automatic analysis of digital images for quality control purposes made easy with a generic extendable and scriptable DICOM router. In: *Proceedings of the Scientific Assembly and Annual Meeting of the RSNA, Chicago, IL*.

Jacobs J, Zanca F, Marchal G, Bosmans H (2008). Implementation of a novel software framework for increased efficiency in observer performance studies in digital radiology. In: *Proceedings of the Scientific Assembly and Annual Meeting of the RSNA, Chicago, IL*.

van Engen RE, Bosmans H, Dance DR, Heid P, Lazzari B, Marshall N, Schopphoven S, Thijssen M, Young KC (2013). Digital mammography update. European protocol for the quality control of the physical and technical aspects of mammography screening. S1, Part 1: Acceptance and constancy testing. In: *European guidelines for quality assurance in breast cancer screening and diagnosis. Fourth edition, Supplements*. Perry N, Broeders M, de Wolf C, Törnberg S, Holland R, von Karsa L (eds.). European Commission, Office for Official Publications of the European Union, Luxembourg, pp. 1–54.

van Engen RE, Young KC, Bosmans H, Thijssen M (2006). European protocol for the quality control of the physical and technical aspects of mammography screening. Chapter 2b: Digital mammography. In: *European guidelines for quality assurance in breast cancer screening and diagnosis. Fourth edition.* Perry N, Broeders M, de Wolf C, Törnberg S, Holland R, von Karsa L (eds.). European Commission, Office for Official Publications of the European Communities, Luxembourg, pp. 105–165.

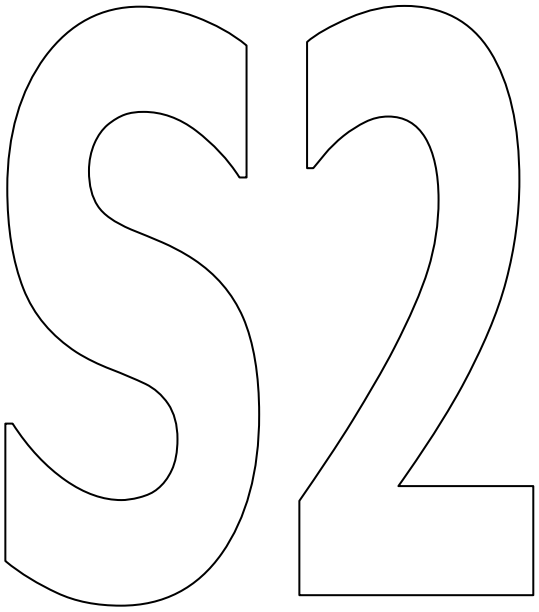
Van Ongeval C, Bosmans H, Van Steen A (2005). Current challenges of full field digital mammography. *Radiat. Prot. Dosimetry.*, 117:148–153.

Van Ongeval C, Van Steen A, Geniets C, Dekeyzer F, Bosmans H, Marchal G (2008). Clinical image quality criteria for full field digital mammography: a first practical application. *Radiat. Prot. Dosimetry.*, 129:265–270.

European guidelines for quality assurance in breast cancer screening and diagnosis

Fourth Edition

Supplement



Pathology update

Quality assurance guidelines for pathology

Editor
C.A. Wells

Produced by the European Working Group on Breast Screening Pathology

C.A. Wells, Dept. of Cytopathology, University College Hospital, London, United Kingdom
I. Amendoeira, Dept. of Pathology, Instituto de Patologia & Imunologia Molecular da Universidade do Porto and Hospital S. João, Porto, Portugal
J.P. Bellocq, Dept. of Pathology, University Hospital Strasbourg, Strasbourg, France
S. Bianchi, Division of Pathological Anatomy, Dept. of Medical and Surgical Critical Care, University of Florence, Italy
W. Boecker, Pathologie-Praxis, Hamburg, Germany
B. Borisch, Dept. of Pathology, University of Geneva, Switzerland
B. Bruun Rasmussen, Dept. Of Pathology, Herlev Hospital, Herlev, Denmark
G.M. Callagy, Dept. of Pathology, National University of Ireland, Galway, Ireland
E. Chmielik, Dept. of Pathology, Maria Sklodowska-Curie Memorial Cancer Center, Gliwice, Poland
A. Cordoba, Servicio de Anatomía Patológica, Complejo Hospitalario de Navarra, Pamplona, Spain
G. Cserni, Dept. of Pathology, University of Szeged, Hungary, and Dept. of Pathology, Bács-Kiskun County Teaching Hospital, Kecskemét, Hungary
T. Decker, Dept. of Pathology, Dietrich Bonhoeffer Medical Center, Neubrandenburg, Germany
J. DeGaetano, Dept. of Pathology Mater Dei Hospital, Malta
M Drijkoningen, Dept. of Pathology, University Hospitals Leuven, Leuven, Belgium
I.O. Ellis, Dept. of Histopathology, Nottingham City Hospital, Nottingham, United Kingdom
D.R. Faverly, CMP pathology laboratory and CCR Community Reference Center for cancer screening, Brussels, Belgium
M.P. Foschini, Dept. of Pathology, University of Bologna, at Bellaria Hospital, Bologna, Italy
S. Frković-Grazio, Dept. of Gynaecological Pathology and Cytology, University Clinical Hospital, Ljubljana, Slovenia
D. Grabau, Dept. of Pathology, Skåne University Hospital, Lund, Sweden
P. Heikkilä, Dept. of Pathology, University of Helsinki and HUSLAB, Helsinki, Finland
E. Iacovou, Dept. of Pathology, Nicosia General Hospital, Cyprus
J. Jacquemier, Dept. of Molecular Oncology, Institut Paoli Calmettes, Marseille, France
H. Kaya, Dept. of Pathology, Marmara University School of Medicine, Istanbul, Turkey
J. Kulka, 2nd Dept. of Pathology, Semmelweis University, Budapest, Hungary
M. Lacerda, Center for Neuroscience and Cell Biology, University of Coimbra, Portugal
I. Liepniece-Karele, Riga Eastern Clinical University Hospital, Riga, Latvia
J. Martinez-Penuela, Dept. of Pathology, Navarra Hospital Complex, Pamplona, Spain
C.M. Quinn, Dept. of Pathology, St Vincent's University Hospital, Dublin, Ireland
F. Rank†, Dept. of Pathology, Rigshospitalet, Copenhagen University Hospital, Denmark
P. Regitnig, Institute of Pathology, Medical University of Graz, Austria
A. Reiner, Institute of Pathology, Danube Hospital, Vienna, Austria
A. Sapino, Dept. of Biomedical Science and Human Oncology, University of Turin, Turin, Italy
T. Tot, Dept. of Pathology and Clinical Cytology, Central Hospital Falun, Falun, Sweden
P.J. Van Diest, Dept. of Pathology, University Medical Center Utrecht, The Netherlands
Z. Varga, Dept. of Pathology, Institute of Surgical Pathology, University Hospital Zurich, Switzerland
J. Wesseling, Dept. of Pathology, Netherlands Cancer Institute/Antoni van Leeuwenhoek Hospital, Amsterdam, The Netherlands
V. Zolota, Dept. of Pathology, Medical School, University of Patras, Rion, Patras, Greece
E. Zozaya-Alvarez, Navarra Hospital Complex, Pamplona, Spain

Declaration of interest

An interest of D. Grabau is reported on page VI.

Disclaimer

The views expressed in this document are those of the authors and do not necessarily reflect the official position of the European Commission. Neither the European Commission nor any other organization or any individual may be held responsible for any use that may be made of the information contained herein.

Acknowledgements
Financial support was provided by the European Union Public Health Programme (Project no. 2006322, European Cooperation on Development and Implementation of Cancer Screening and Prevention Guidelines [ECCG]).

Please cite this publication as follows
Wells CA, Amendoeira I, Bellocq JP, Bianchi S, Boecker W, Borisch B, Bruun Rasmussen B, Callagy GM, Chmielik E, Cordoba A, Cserni G, Decker T, DeGaetano J, Drijkoningen M, Ellis IO, Faverly DR, Foschini MP, Frković-Grazio S, Grabau D, Heikkilä P, Iacovou E, Jacquemier J, Kaya H, Kulka J, Lacerda M, Liepniece-Karele I, Martinez-Penuela J, Quinn CM, Rank F†, Regitnig P, Reiner A, Sapino A, Tot T, Van Diest PJ, Varga Z, Wesseling J, Zolota V, Zozaya-Alvarez E (2013). S2: Pathology update. Quality assurance guidelines for pathology. In: *European guidelines for quality assurance in breast cancer screening and diagnosis. Fourth edition, Supplements*. Perry N, Broeders M, de Wolf C, Törnberg S, Holland R, von Karsa L (eds.). European Commission, Office for Official Publications of the European Union, Luxembourg, pp. 73–120.

Corresponding author
C.A. Wells
Dept. of Cytopathology
University College Hospital
Level 2, Rockefeller Building
21 University Street
London WC1E 6JJ
United Kingdom
c.a.wells@me.com

Table of contents

List of tables

1 Introduction

The present supplement deals with several topics in the quality assurance of pathology in breast cancer screening and diagnosis in which problems and practical solutions as well as new techniques and other advances have emerged in recent years. The authors felt that these topics warranted urgent attention pending a full revision of the pathology chapter. They include the classification of early forms of neoplastic changes in the breast and the differential diagnosis of columnar cell lesions (CCLs), atypical ductal hyperplasia (ADH) and ductal carcinoma in situ (DCIS). Other sections provide an update on the classification of invasive carcinoma and on the diagnosis and differential diagnosis of microinvasive carcinoma. Special attention is paid to the assessment of the axilla, and the updated guidance on axillary dissection and preoperative staging takes into account problems and pitfalls as well as the most recent revision of the TNM system (Sobin et al., 2009). The supplement also includes comprehensive practical recommendations on examination techniques and interpretation of sentinel node biopsy specimens. Best practice in the use of frozen sections, an update on vacuum-assisted needle core biopsy (VANCB) and recommendations on pathological reporting of post-chemotherapy specimens are also provided.

The supplement has been prepared by the European Working Group on Breast Screening Pathology, which also prepared the chapter on pathology in the fourth edition (Wells et al., 2006) and in previous editions of the guidelines. The group consists of 38 breast pathologists in 23 EU Member States, Switzerland and Turkey and has been led by the editor of the chapter on pathology in the guidelines. The recommendations in the supplement have been drafted by selected members of the group and have been discussed and agreed upon at twice-yearly meetings held in Hungary, Switzerland, Italy, France, the Netherlands and Finland since the fourth edition was published.

2 Columnar cell lesions and atypical ductal hyperplasia

2.1 Background

In recent years there has been a growing interest and awareness of epithelium having a columnar morphology within breast biopsies. One of the common accompaniments of this change is the presence of secretions within lumina, which are often distinctly eosinophilic and associated with micro-calcifications (Fraser et al., 1998). These microcalcifications are detectable by mammography, and consequently more of these lesions are being detected with the introduction and expansion of breast screening programmes, especially since the implementation of digital mammography (Feeley & Quinn, 2008; Verschuur-Maes et al., 2011a).

Diverse terminology has been used for this epithelium, e.g. columnar cell lesions (CCLs) (Schnitt & Vincent-Salomon, 2003), flat epithelial atypia (FEA) (Tavassoli & Devilee, 2003), columnar cell alteration with apical snouts and secretions (CAPSS) (Fraser et al., 1998), enlarged lobular units with columnar alteration (McLaren et al., 2005), atypical cystic lobules (Oyama et al., 1999), ductal intraepithelial neoplasia flat type (Tavassoli, 1998), atypical cystic ducts (Kusama et al., 2000) and clinging carcinoma (Azzopardi et al., 1979). This diverse terminology, the increased detection of this lesion at screening and the uncertainties about its biology continue to cause concern among both diagnostic breast pathologists and their clinical colleagues in the breast care team. In this supplement,

we present a proposal for a simplified terminology for lesions with columnar epithelium and histological criteria for recognizing these (Walker et al., 2010). These proposals are not carved in stone and are likely to change as more knowledge is acquired about the pathobiology of these lesions and their natural course.

2.2 Terminology

In general, CCLs are characterised by enlarged terminal duct/lobular units (TDLUs) lined with "columnar" epithelium, meaning tall cells with eosinophilic cytoplasm and small cytoplasmic snouts protruding into the lumina. Intracytoplasmic vacuoles (like those in lobular neoplasia) are regularly seen. Usually, the nuclei are located basally within the cell, adjacent to the myoepithelium. Often, there is microcalcification in the lumina, and there may be extensive mucin production, which may even rupture into the stroma, producing a mucocoele-like lesion (Verschuur-Maes & Van Diest, 2011). Necrosis is extremely rare.

These lesions comprise a spectrum from the earliest columnar changes in otherwise almost-normal acinar epithelium to low-grade ductal carcinoma in situ (DCIS). Further, there may be a cross-over to lobular neoplasia, usual ductal hyperplasia and apocrine lesions.

Different classifications have been proposed for CCLs. The simplest classification and probably the most reproducible system for "pure" CCLs has been proposed by Schnitt et al. (Schnitt & Vincent-Salomon, 2003). This system discerns 4 groups of CCLs: columnar cell change (CCC), columnar cell hyperplasia, CCC with atypia and columnar cell hyperplasia with atypia. The morphology of those with atypia corresponds to FEA, as detailed in the latest WHO classification, and corresponds to ductal intraepithelial neoplasia 1a (DIN1a) (Tavassoli & Devilee, 2003). Sometimes CCLs may co-exist with lobular neoplasia.

2.3 Classification of columnar cell lesions (CCLs)

2.3.1 Columnar cell change (CCC)

Classic CCC consists of lobular acini, usually enlarged and lined by epithelial cells that are tall and snouted in a manner similar to that observed in tubular carcinoma. The cytoplasm is eosinophilic, and often there are luminal secretions and/or microcalcifications. There is morphological diversity within these groups. For example, the hyperchromasia of the nuclei can vary, as well as the nuclear shape and the "tallness" of the cells. In some cases, therefore, some cells are more cuboidal than columnar. The nuclei are regular and show no atypia and only small or no nucleoli. Mitoses are very rare, as is apoptosis.

2.3.2 Columnar cell hyperplasia

These lesions are like CCC, except that there is now a piling up of > 2 layers of epithelium, assuming that the stratification is real, as opposed to artefactual as a consequence of cross-cutting. This multi-layered epithelium may be flat or "hilly". There should be no complex cribriform architecture or micropapillary structures without stroma, as these favour a diagnosis of atypical ductal hyperplasia (ADH) or low-grade DCIS.

2.3.3 Columnar cell change with atypia

These lesions are like CCC, except that there is now “atypia”. Atypia may manifest in different ways. There may be relatively uniform rounded, evenly spaced nuclei with a similar cytomorphology to that displayed by the cells of low-grade DCIS and some intermediate-grade DCIS with larger, more irregular nuclei and/or nuclei with clear nucleoli. Here, polarisation of the cells is usually lost and the nuclei are often more centrally placed. Alternatively, there may be disturbance of the regular nuclear arrangement and/or mild to moderate nuclear atypia as reflected by larger, more irregular nuclei and/or nuclei with clear nucleoli. Mitoses are still very rare. Marked nuclear atypia favours a diagnosis of high-grade DCIS (e.g. cancerisation of lobules).

2.3.4 Columnar cell hyperplasia with atypia

These lesions are like columnar cell hyperplasia without complex architecture, except that there is now “atypia”. As above, there may be relatively uniform rounded, evenly spaced nuclei with a similar cytomorphology to that displayed by the cells of low-grade DCIS and some intermediate-grade DCIS with larger, more irregular nuclei and/or nuclei with clear nucleoli. Polarisation of the cells is, again usually lost and the nuclei are often more centrally placed. Again, there may be disturbance of the regular nuclear arrangement and/or mild to moderate nuclear atypia with larger, more irregular nuclei and/or nuclei with clear nucleoli. Mitoses are still very rare. Marked nuclear atypia favours a diagnosis of high-grade DCIS (e.g. cancerisation of lobules).

2.4 Differential diagnosis

2.4.1 Atypical ductal hyperplasia/low-grade ductal carcinoma in situ (DCIS)

By definition, CCLs do not have a complex architecture. If micropapillary or cribriform structures are seen, a diagnosis of ADH or low-grade DCIS should be considered, depending on the size of the lesion and how extensive the architectural complexity and regularity are. In practice, CCLs are often seen merging with the more elaborate architecture of ADH and DCIS, demonstrating the close relationship of atypical CCLs to these lesions. ADH and low-grade DCIS may well derive from CCLs. (See Table 1)

2.4.2 Ductal carcinoma in situ (DCIS)

Where flat proliferative epithelium within a TDLU has high cytonuclear grade nuclei (usually accompanied by more apoptosis and mitosis), the lesion should be categorised as high-grade DCIS.

2.4.3 Apocrine change

Apocrine epithelium shares with CCC the presence of apical snouts and, sometimes, associated secretions. However, the cytoplasm of apocrine cells is typically copious, granular and more pink than the eosinophilic appearance of CCLs. Further, unlike in most columnar cells, nucleoli are readily seen. Apocrine cells are usually androgen receptor (AR) positive but oestrogen receptor (ER) and progesterone receptor (PR) negative.

2.4.4 Lactational change

Another lesion that may be mistaken for atypical CCL is focal lactational change, which may appear both single-layered and mildly cytologically atypical with enlarged hyperchromatic nuclei. Typically, the epithelial cells are “hobnail”, or more cuboidal than tall, and the foamy appearance of the cytoplasm produced by finely divided lipid should alert the pathologist to this diagnosis.

2.5 Immunohistochemistry

In CCLs, there is extensive, and uniform, positivity of the nuclei for ER and PR in all lesional cells (Oyama et al., 1999; Allred & Mohsin, 2000; Vincent-Salomon, 2003; Tremblay et al., 2005). Most cells stain positively for keratin 19 (Oyama et al., 1999) and osteoprotegerin (Van Poznak et al., 2006), and a high proportion are positive with cyclin D1 (Oyama et al., 1999). CCLs do not express basal markers such as CK5/6 and CK14 (Otterbach et al., 2000) and are not uniformly positive for AR. p53 accumulation is lacking (Simpson et al., 2005). The utility of GCDFP15 (BRST2) to discriminate CCLs from apocrine change where it is extensively expressed has yet to be documented in the literature. High-grade DCIS with a flat morphology is positive for HER2, unlike CCC, which is uniformly negative for HER2 (Simpson et al., 2005).

2.6 Genetic aberrations

There is increasing evidence that there is progressive chromosomal damage in CCLs (with atypia) through to DCIS and invasive carcinoma, implying a molecular continuum from some lesions with a columnar morphology. This is not the case for ADH and carcinoma. Most importantly, in different studies a high concordance level of the genetic abnormalities was seen in CCLs with atypia as well as in the accompanying infiltrating and in situ components. With comparative genomic hybridisation, chromosomal abnormalities were identified with loss on 16q, 17p and X and gain on 15q, 16p and 19 (Simpson et al., 2005). Recently, with use of loss of heterozygosity (LOH), allelic losses were identified in CCL with atypia, low-grade DCIS and tubular carcinoma on the long arm of chromosome 16, as well as at chromosomes 8p21, 3p14, 1p36 and 11q14 (Aulmann et al., 2009). In another study, LOH at chromosomal loci 11q21-23.2, 16q23.1-24.2 and 3p14.2 was present in 50%, 45% and 41% of cases, respectively (Moinfar et al., 2000). Dabbs et al. have similarly shown progressive accumulation of allelic damage at 9q, 10q, 17p and 17q with increasing severity of the lesion (Dabbs et al., 2006).

This makes it very likely that CCLs are among the earliest forms of neoplastic change in the breast. Therefore, CCLs need to be seen as clonal and neoplastic proliferations, in which different chromosomes play a role, with loss on chromosome 16q seeming to be the most important.

2.7 Significance of columnar cell lesions

The overall biological significance of CCLs is not yet fully understood (Tremblay et al., 2005). There is supporting evidence for a precursor role within the spectrum of the low nuclear grade family of breast cancers (Abdel-Fatah et al., 2008). This includes an association with tubular carcinoma (Goldstein & O'Malley, 1997; Rosen, 1999; Aulmann et al., 2009) and the findings that these lesions are more commonly present in cancerous breasts than in non-cancerous breasts (Wellings et al., 1975), are often continuous with cancerous lesions (Kusama et al., 2000), commonly co-exist with lobular in situ neoplasia (Brogi et al., 2001) and share similar cytological characteristics and immunohistochemical profile and genetics with co-existent malignancies (Oyama et al., 1999; Moinfar et al., 2000; Tavassoli & Devilee, 2003; Simpson et al., 2005; Dabbs et al., 2006).

Despite the above, the certainty and speed of progression of these lesions to malignancy is not well known. A follow-up study performed by Boulos et al. showed a relative risk of 1.47 of developing invasive carcinoma after a needle core biopsy showing CCL (with or without atypia) (Boulos et al., 2008). The study had a follow-up interval of 17 years, and 1154 CCLs were included. Shaaban et al. described that 8 cases of CCL with atypia had a relative risk of 2.32 of developing subsequent invasive

carcinoma, compared with 11 controls with CCL without atypia, who did not develop cancer in the follow-up period (Shaaban et al., 2002). Verschuur-Maes et al. reported an 8-year progression risk for CCL with atypia and ADH emerging in CCL of about 20% (Verschuur-Maes et al., 2011b). However, other studies did not find evidence of the increased risk of breast cancer. Eusebi et al. discovered only 1 recurrence of the 25 CCLs with atypia (in their study, called low-grade clinging carcinoma of flat type) after 3 years, with a follow-up interval of 19 years (Eusebi et al., 1994). Bijker et al. and De Mascarel et al. also found no invasive carcinoma or DCIS in 59 and 84 CCLs with atypia, respectively, with a mean follow-up of 5.4 and 13.3 years, respectively (Bijker et al., 2001; de Mascarel et al., 2007).

It is worth noting that McLaren et al. found that patients with ER negative CCLs were at increased risk of malignancy compared with those whose lesions were ER positive (McLaren et al., 2005), but as yet there are no confirmatory studies.

Table 1: Characteristics of columnar cell lesions and ADH/DCIS

	Diagnosis			
Feature	**Columnar cell change**	**Columnar cell hyperplasia**	**Columnar cell lesion with atypia ****	**ADH/DCIS**
Topography	TDLU, acini may be mildly dilated or of normal size	TDLU, acini may be mildly dilated or of normal size	TDLU, often microcystically dilated acini	TDLU +/– adjacent ducts
Shape of acinar spaces	Irregularly shaped luminal margin	Irregularly shaped luminal margin	Often rounded acinar spaces, with smooth inner margin	Often rounded acini, but with complex structures extending into lumen (see Architecture, below)
Architecture	Flat	Tufts and mounds	Flat or tufted/mounds, not complex	Complex with micropapillary or cribriform structures
Stratification/ multi-layering	Not present	Present	May be present	May be present
Luminal secretions often with micro-calcifications	Present	Present	Present	May be present
Nuclear size	Small to medium	Small to medium	Small to medium	Small to medium
Nuclear shape	Oval, elongated	Oval, elongated	Often, but not always, rounded	Rounded
Nuclear texture	Bland	Bland	Speckled chromatin pattern may be present	Speckled chromatin pattern is common
Pleomorphism*	Uniform	Uniform	Uniform to moderately pleomorphic	Uniform
Position of nuclei within cell	Basally placed	Basally placed	Often central	Central
Nucleoli	Not conspicuous	Not conspicuous	Evident	May be evident
Mitoses	Generally absent	Generally absent	Generally scarce	Generally scarce
Extent	May be focal or extensive	May be focal or extensive	May be a focal area within background of non-atypical CCL	May be focal area within background of non-atypical CCL; by definition, ADH is small/microfocal

* If the cells display marked pleomorphism, then the lesion does not fall within the spectrum of CCLs but should be regarded as high-grade DCIS.

** Also referred to as flat epithelial atypia (FEA).

2.8 Reproducibility of current classification of columnar cell lesions

The reproducibility of the diagnosis of CCLs varies in the published literature. One study looking at the assessment of images by a mixed group of assessors showed poor agreement (Tan et al., 2005). Another demonstrated "excellent agreement" for a specified task within the CCL group after a PowerPoint-directed training session undertaken by a group of pathologists with an interest in breast pathology (O'Malley et al., 2006). Agreement was slightly better for determining the absence of CCLs with atypia (in their study, called FEA) (92.8%; 95% CI, 84.1–97.4%) than for determining its presence (90.4%; 95% CI, 79.9–96.7%). O'Malley et al. concluded that the diagnosis of CCLs with atypia and its distinction from those without atypia was highly reproducible if the available diagnostic criteria were adhered to. This simplified classification in the study of O'Malley et al. forms the basis of the recommendations presented here.

Issues of poor reproducibility can confound diagnosis and lead to inconsistent patient care. This is more likely when complex and diverse terminology with possible different interpretations exists. This is nowhere more apparent than in the case of lesions with columnar epithelial morphology and is the strongest argument for the simplified classification of these lesions presented here.

2.9 Clinical management of columnar cell lesions in needle core biopsies

Core biopsies bearing CCLs are typically sampled for the histological assessment of mammographic microcalcification. As with similar specimens, these should be examined at multiple (at least 3) levels. If CCC or columnar cell hyperplasia only is found, without atypia, the lesions should be regarded as benign, requiring only follow-up. Core biopsies with CCLs without atypia should be classified in the B2 category.

CCLs with atypia (whether CCC or columnar cell hyperplasia) should be regarded as of uncertain malignant potential, with a risk similar to ADH and atypical lobular hyperplasia. Atypical CCLs are associated with more advanced lesions in subsequent excisional biopsy more frequently than are those without atypia. (Guerra-Wallace et al., 2004; David et al., 2006; Kunju & Kleer, 2007; Martel et al., 2007; Chivukula et al., 2009; Senetta et al., 2009; Ingegnoli et al., 2010; Verschuur-Maes et al., 2011a). As for all such screen-detected lesions, multidisciplinary discussion should be undertaken to correlate radiological, clinical and histopathological findings to decide on surgical excision biopsy, vacuum-assisted biopsy or a wait-and-see policy. A recent systematic review concluded that, on the basis of the (in situ) carcinoma underestimation rates of patients with a core diagnosis of CCL with atypia and ADH associated with CCL, surgical excision should be considered (Verschuur-Maes et al., 2012). Core biopsies with CCLs with atypia should be classified in the B3 category.

2.10 Columnar cell lesions in surgical excision specimens

Thorough sampling and histological examination of surgical specimens bearing CCLs with atypia should be performed to search for more established neoplasia like ADH, DCIS, lobular neoplasia or even invasion. In association with some DCIS and/or invasive carcinoma, CCLs with atypia may be present and may be extensive and/or may extend to the margins of the specimen. This is particularly problematic for lesions that are well recognised to be associated with CCL with atypia, such as tubular carcinoma (Rosen, 1999). There is a very limited evidence base for guidance with regard to re-

excision or radiotherapy, but at present it is recommended that whole tumour size (DCIS plus invasive carcinoma) and margin status should include only those areas regarded as established DCIS or invasive tumour using conventional criteria. It should be recognised that these lesions have a very low risk of development of life-threatening disease.

2.11 Atypical ductal hyperplasia (ADH)

The distinction between low-grade DCIS and ADH is based on evidence derived from many series, including studies by David Page and co-workers (Page & Dupont, 1993). These have been supported by other studies, such as the Nurses' Health Study (Connolly & Schnitt, 1993; Marshall et al., 1997).

The significance of the diagnosis of ADH lies in the increased risk of invasive breast carcinoma, which is about 4–5 times that of the general population (Page et al., 1985; Ma & Boyd, 1992; London et al., 1992; Dupont et al., 1993; Page & Jensen, 1994) and may be even greater for pre-menopausal women (approaching 6 times increased risk) (London et al., 1992). This risk is further increased if the patient has a first-degree relative with breast cancer (10 times increased risk) (Page et al., 1985; Tavassoli & Norris, 1990; Page & Dupont, 1992).

The diagnostic criteria used to define ADH are imperfect. ADH was described initially based on exclusion rather than positive criteria, i.e. the recognition of some but not all of the features of DCIS (as well as the lack of the characteristics of usual type epithelial hyperplasia) (Page et al., 1985). This definition of ADH has been updated and, while the diagnosis still rests on an absence of all the features of DCIS, additional supporting features have been described (Page & Rogers, 1992; National Coordinating Group for Breast Screening Pathology, 1995). Page's view that the cellular changes of DCIS are present but occupy < 2 separate duct spaces is widely accepted. Others use a 2 mm cut-off; a lesion < 2 mm in maximum dimension is classified as ADH and a larger lesion as DCIS (Tavassoli, 1992). Others mention the involvement of a single TDLU. These criteria recognise essentially the same lesions. In essence, ADH is usually small and focal, measuring < 2–3 mm in maximum dimension. Larger foci are accepted as ADH if associated with a radial scar/complex sclerosing lesion or a papilloma.

There are three components to the diagnosis of ADH: the architectural pattern, cytology and disease extent. Because disease extent cannot be optimally assessed in core biopsy samples, the diagnosis of ADH cannot be made from core biopsy specimens, and the term "atypical intraepithelial proliferation of ductal type" is used in this setting (see the appropriate section in the fourth edition). ADH is formed from a uniform population of small or medium-sized round, cuboidal or polygonal hyperchromatic cells, which are regularly arranged. The nuclei are evenly distributed and may form a rosette-like pattern. Single small nucleoli only are present. Mitoses, particularly abnormal forms, are infrequently seen. Geometric spaces are noted and, in the cribriform type, the cells are arranged at right angles to the bridges formed. Micropapillary ADH is also recognised, and a solid pattern may very rarely be seen. Small foci of necrosis may rarely be identified in ADH and do not indicate that the process should be classified as DCIS.

At present, it is recommended that the diagnosis of ADH should be restricted to lesions that show the features described by Page et al. (Page et al., 1985; Page & Rogers, 1992), to which the quantified risk of developing breast carcinoma is linked. Even then, the diagnosis of ADH should be made with caution and only if low-grade DCIS has been seriously considered in the differential diagnosis. Lesser changes for which the possible classification lies between florid usual type hyperplasia (UEH) and ADH are less relevant with regard to a risk of developing breast carcinoma and should not be classified as ADH. However, it should also always be borne in mind that a proliferation at the edge of a biopsy may represent the periphery of a more established lesion of DCIS and further excision of the adjacent tissue may be warranted.

3 Acceptable use of frozen sections

3.1 Palpable breast tumours

Frozen section for intraoperative assessment of palpable breast tumours was invented several decades ago and has been widely used since then. It is a reliable method for palpable lesions > 1 cm in diameter (Fessia et al., 1984; Niemann et al., 1996). Discrepancies between frozen section diagnosis and subsequent paraffin section diagnosis consist mainly of false-negative results, with a low frequency of 1–2%. False-positive frozen section diagnoses are very rare events, occurring in < 0.1% of cases. In some cases the definitive diagnosis in frozen section has to be deferred to paraffin section. Overall sensitivity of frozen section diagnosis is reported to be > 90% and specificity about 97%. The main reasons for failure are due to macroscopic and microscopic sampling errors, histological interpretative errors, poor technical quality of frozen sections and lack of communication between pathologist and surgeon (Cserni, 1999; Laucirica, 2005). The introduction of preoperative diagnosis on core biopsies has led to a marked reduction in the use of frozen sections for assessment of breast lesions (Laucirica, 2005). Frozen section diagnosis may still be performed rarely, e.g. when core biopsy results are equivocal. Tumours < 1 cm in maximum diameter should not be subjected to frozen section diagnosis. In general, it would be expected that extensive preoperative interdisciplinary discussion between radiologists, surgeons and pathologists would decrease the need for frozen sections, which should be performed only where there is an immediate impact on surgical treatment.

3.2 Margin assessment

No consensus exists about the method of intraoperative margin assessment. Frozen sections are considered suitable for this application. The accuracy of frozen sections for margin assessment is described with a sensitivity and specificity of only 86% and 83%, respectively (Noguchi et al., 1995). This is mainly due to intraductal tumour components not included in frozen sections but detected on subsequent examination in paraffin sections. The fatty nature of the margins is often a technical hurdle for adequate frozen sections. However, frozen sections of margins for breast cancer may be unnecessary in cases where specimen radiology is performed and reported to an adequate standard.

3.3 Sentinel lymph node biopsy

Frozen sections for intraoperative assessment of sentinel lymph nodes (SLNs) should be performed only in cases with impact on immediate surgical treatment. The main purpose is to spare patients, with a positive SLN, 2-step surgery. Not all metastases will be identified in these lymph nodes in frozen sections due to the limited number of sections acceptable intraoperatively and freezing artefacts. The risk of false-negative results is reported to be between 9% and 52% (Cserni et al., 2003). Rarely, false positivity may occur. Overall, the accuracy is reported to be between 79% and 98%. In addition, during frozen sectioning, tissue loss may occur, which must be kept to a minimum (Cserni, 2006; Varga et al., 2008; Fritzsche et al., 2010). Imprint cytology is an acceptable alternative in centres with cytological expertise (Van Diest et al., 1999).

The findings of the ACOSOG Z0011 trial, suggesting that SLN micrometastases do not adversely impact on prognosis and challenging the view that all patients with a positive SLN require complete

axillary lymph node clearance, has led to a reduction in the need for intraoperative SLN evaluation (Giuliano et al., 2011).

3.4 Non-palpable breast tumours

Frozen section diagnosis is not recommended for non-palpable breast tumours. They are usually screen-detected and < 1 cm in diameter. Because of their small size, significant amounts of tissue may be lost during frozen section preparation, and definitive diagnosis in paraffin sections may become impossible.

3.5 Needle core biopsy and vacuum-assisted needle core biopsy

Core biopsies are small-volume specimens and are often used for screen-detected breast lesions, which may be complex lesions detected only by microcalcification in many cases and requiring special expertise in breast pathology. Overdiagnosis and underdiagnosis in this type of specimen are well-known phenomena, even in paraffin sections. For example, in cases of DCIS underdiagnosis is found in up to 20% of cases when compared with subsequent open biopsy (Verkooijen, 2002). Moreover, core biopsies are very small in volume and significant amounts of tissue may be lost during frozen sectioning, rendering paraffin section diagnosis impossible. Biomarkers, e.g. steroid hormone receptors, may need to be assessed on paraffin sections from carcinoma in core biopsies. If frozen sections are performed on core biopsies, reliable assessment cannot be guaranteed due to freeze and thaw artefacts. Therefore, frozen section examination of core biopsies is not appropriate under any circumstances.

4 Classifying invasive carcinoma

Typing invasive carcinomas has prognostic value and provides information on the pattern of metastatic spread and behaviour. Because (1) typing of breast carcinomas has been shown in external quality assurance EQA schemes to be relatively poorly reproducible, (2) the most frequent traditional type (ductal) can be correlated to prognosis only through additional grading and (3) important progress has been made in classifying breast cancers in a more molecular way that predicts response to current adjuvant therapeutics, the system has been revised to include current molecular insights.
Caution should be exercised in typing carcinomas in poorly fixed specimens or in specimens that have been removed from patients who have been treated by primary chemotherapy, hormonal therapy or radiotherapy before surgery.

4.1 Towards molecular classification of breast cancer

Microarray gene expression studies have suggested that breast cancers cluster in 3 main groups: (1) luminal cancers, which mainly have properties of glandular cells of the breast, driven by high expression of steroid receptors; (2) HER2-driven cancers, with amplification and overexpression of HER2/*neu* but lacking steroid receptor expression and (3) basal cancers, which have properties of the stem cells of the breast (Sorlie et al., 2001). Because these expression profiles reflect carcinogenetic

pathways and correlate with response to therapy, incorporating this new way of molecular thinking into breast cancer classification is becoming increasingly important. Because microarray gene expression is not easily performed in routine practice and, for the time being, requires frozen material, attempts have been made to translate this molecular classification into an immunohistochemical classification based on the expression of established markers such as ER, PR, HER2, basal cytokeratins, epidermal growth factor receptor (EGFR), and proliferation rate. This surrogate molecular classification is still evolving, and no real consensus has been reached. We nevertheless think it important to follow this trend and to attempt for the first time to come to some form of integration between traditional and molecular ("histomolecular") typing using our current state of knowledge (Weigelt & Reis-Filho, 2009).

Luminal cancers show expression of ER and/or PR, the luminal A subtype without and the luminal B subtype with overexpression of HER2 and/or high Ki67 score. Most tubular, invasive cribriform, mucinous, invasive micropapillary and lobular cancers fall into this group, as do most grade 1 and 2 ductal cancers. Luminal cancers especially metastasise to the bone, although luminal B cancers metastasise also to the brain with increased frequency.

The HER2 group shows overexpression of HER2 and lacks expression of ER and PR. This group comprises mainly ductal and some apocrine cancers. Apocrine cancers usually show expression of AR.

Basal cancers lack expression of ER, PR and HER2 (often denoted "triple negative") and express basal cytokeratins like CK5/6 and/or CK14 or EGFR. Please note here that "triple negative" does not equal basal per se! Although basal cancers may also express other markers associated with stem cells, like p63 and vimentin, these markers are not routinely used. The group of basal cancers usually comprises well-circumscribed cancers with high-grade features (atypical nuclei, many mitoses, geographic necrosis), lack of tubule formation with larger solid and often interconnecting islands, and abundant lymphocytoplasmic infiltrate around the lesion. There is often a central fibrotic/necrotic area. This group comprises high-grade cancers such as most medullary cancers and metaplastic cancers, *BRCA1* germline mutation-associated cancers and many grade 3 ductal cancers. Low-grade basal cancers comprise the salivary gland-like cancers (e.g. adenoid cystic carcinoma) (Foschini & Krausz, 2010). Some triple negative cancers may express AR and are possibly apocrine-derived.

In this supplement, we have attempted to relate the traditional histological types to this molecular classification to arrive at a synthesis of traditional morphological and new molecular typing (see Table 2). Inherently, this means that traditional histological typing has lost some of its importance because clinically it may no longer always be useful to discern some of the traditional WHO types that clinically behave similarly and, in fact, always overlapped considerably (Weigelt et al., 2008; Reis-Filho & Lakhani, 2008).

It should be realised, however, that invasive breast cancer forms a morphological and also a molecular spectrum, so not all established morphological types completely fall into one of the molecular groups, nor will all individual cancers be easily classifiable into one of the molecular types.

4.2 Special type versus mixed type and "no special type"

A special type cancer shows a classic morphology with all the hallmark histological features that will reproducibly be classified as such by different pathologists, requiring > 90% purity.

Mixed morphological differentiation with < 90% purity is relatively common, as is the mere absence of specific differentiation patterns, which mostly leads to a classification as "ductal – no special type" (ductal-NST). ER, PR and HER2 usually do not vary clearly between the histologically different parts of

these tumours, so that assessing these receptors in these mixed cancers does not pose specific problems. Prognosis, however, is dictated by the "worst" (least differentiated) part of the tumour.

Table 2: Molecular typing of breast cancer based on common immunohistochemical markers (Abd El-Rehim et al., 2005; Goldhirsch et al., 2011)

Molecular intrinsic subtype	Clinico-pathological definition	ER	PR	HER2	Ki67	Basal markers*
Luminal A	Luminal A	+	+ or –	–	Low	-
Luminal B	Luminal B (HER2 negative)	+	+ or –	–	High	-
Luminal B	Luminal B (HER2 positive)	+	+ or –	Overexpressed	Low or high	-
HER2	HER2 positive (non-luminal)	–	–	Overexpressed	Usually high	+/-
Basal	Triple negative (ductal)	–	–	–	Usually high	+

***** CK5/6 or CK14

4.3 Immunohistochemistry of breast cancer

Besides the markers used for molecular classification, we describe more specific markers for the different morphological types. In general, breast cancers are positive for AE1/3 (and other broad-spectrum keratins), CAM5.2, CK7 and EMA, and for BRST2 (GCDFP-15), which is usually focal (except in the case of apocrine differentiation).

4.4 Assessing predictive and prognostic markers

Proper assessment of prognostic and predictive factors requires adequate tissue processing in the first place. This includes freezing representative tissue when necessary (snap freezing in liquid nitrogen or isopentane, followed by storage in a –80° freezer) and adequate fixation of the rest in neutral buffered formaldehyde. Freezing or formaldehyde fixation must be done with as little delay as possible. To achieve adequate fixation, tissue must be sliced at 4–5 mm thickness, which should ensure full fixation in 24–48 hours. Longer fixation must be avoided to prevent false-negative or false-positive results for prognostic or predictive tests.

Prognostic features that are currently widely used include invasive tumour size (as measured in the fresh specimen, corrected when necessary on the microscopic sections), (sentinel) lymph node status, histological grade and proliferation rate (as measured by mitotic activity index or Ki67 labelling index). DNA microarrays are promising prognostic tools but are as yet insufficiently validated for routine clinical use as predictors of response to therapy. The cost of these tests is also worrisome.

Predictive tests for hormonal treatment are expression of ER and PR, whereas overexpression of HER2 predicts response to anti-HER2 therapy. The negative predictive value of these tests is generally higher than their positive predictive value, e.g. ER (and PR) negative tumours are unresponsive to hormonal therapies, but ER positive tumours are not necessarily responsive. For HER2, gene amplification tests (like fluorescence chromogenic, or silver in situ hybridisation, or PCR-based tests) are increasingly important, especially for equivocal cases or on core biopsies. Participation in quality assessment programmes for different molecular and immunohistochemical tests is essential.

Clearly, molecular assessments are the task of the pathologist within the regular framework of tumour workup. Pathologists should be and are ready to implement new molecular tests in their laboratory as soon as they are sufficiently clinically validated.

4.5 Infiltrating lobular carcinoma

Consideration of both the cytological appearance and the infiltration pattern leads to the diagnosis of lobular carcinoma and, in particular, to relevant subclassification. Typical cytology (small regular cells with clear boundaries, frequently having intracytoplasmic mucous vacuoles, and small nuclei with inconspicuous nucleoli) and a typical infiltration pattern (cells dissociated from each other or forming single files, often in targetoid patterns around uninvolved ducts) lead to the diagnosis of "classic type" lobular carcinoma. Mitotic rate is low. There are often satellite lesions around the dominant nodule, and infiltration is often multifocal. Several variants have been identified in addition to this classic form. At least 90% of the tumour should exhibit one or more of the classic or variant patterns to be classified as infiltrating lobular.

The alveolar variant exhibits small round aggregates of 20 or more cells with typical lobular cytology, and is better demarcated than the classic type.

The solid variant consists of more solid sheets of cells with little intervening stroma, often better demarcated than the classic type. Mitoses are usually also more frequent, and atypia may be more prominent.

The tubulolobular type exhibits microtubular formation as part of the classic pattern, thereby forming a spectrum with tubular cancer, and probably lies closer to tubular cancer in both morphological and immunohistochemical aspects. Tumours that show mixtures of typical tubular and classic lobular carcinoma should be classified as mixed (see below).

The pleomorphic variant is less common and exhibits the growth pattern of classic lobular carcinoma throughout, but the cells, although retaining lobular characteristics such as their discohesive character and intracytoplasmic vacuoles, have polygonal, eccentric pleomorphic nuclei. Mitoses are usually also more frequent, and expression of ER/PR is often less than in classic lobular carcinoma, reflected in a somewhat worse prognosis than classic lobular carcinoma (Eusebi et al., 1992; Weidner & Semple, 1992).

Lobular mixed type lesions consist of mixtures of the above subtypes of lobular carcinoma. Histiocytoid or myoid cellular differentiation may occur (Eusebi et al., 1995; Del Vecchio et al., 2005). It is currently insufficiently clear whether lymphoepithelioma-like cancer is a variant of lobular cancer. The pleomorphic variant can be associated with a prominent inflammatory infiltrate, a feature that may present a difficult differential diagnosis with lymphoepithelioma-like carcinoma (Dadmanesh et al., 2001).

These lesions are thought to derive especially from lobular neoplasia (atypical lobular hyperplasia/lobular carcinoma in situ [LCIS]) but may also be related to CCLs or low-grade DCIS (Abdel-Fatah et al., 2008).

Infiltrating lobular carcinoma is frequently non-palpable and/or mammographically occult. These lesions may also be difficult to visualise sonographically and are therefore over-represented in studies of interval cancer. Many of these are so-called "real" interval carcinomas without suspicious signs on the pre-diagnostic screening mammograms. For the same reasons, lobular carcinomas are also over-represented in the group of occult carcinomas initially presenting with metastasis.

Prognosis of lobular carcinoma as a group is intermediate, with a worse prognosis for pleomorphic and solid types, and there are often lymph node metastases that may require cytokeratin immunohistochemistry for recognition. Lobular cancers metastasise to the bone and gastrointestinal tract and other odd sites, like ovaries, bladder, orbit, serosal surfaces and meninges, in a higher frequency than other cancer types. Rare HER2-overexpressing lobular cancers may metastasise to the brain with increased frequency.

Immunohistochemically, typically E-cadherin and p120 are reduced, either aberrant (cytoplasmic) or negative; ER/PR are positive; HER2 is negative; low-molecular-weight keratins 7, 8, 18 and 19 are positive and high-molecular-weight keratins CK5/6 and CK14 are negative, with the exception of 34betaE12, which often shows positive staining. Most lobular cancers therefore fall into the luminal A molecular type, whereas a small minority will be luminal B when they overexpress HER2 or have high Ki67. Metastatic gastric cancer should be considered as a differential diagnosis in morphologically lobular cancers with no ER or PR expression.

4.6 Tubular carcinoma

Tubular carcinomas are composed of round, ovoid, teardrop-shaped or angulated single-layered tubules in an abundant cellular fibrous or fibro-elastotic stroma. The neoplastic cells are small and uniform and may show cytoplasmic apical snouting. Nuclei should not show a high degree of atypia. At least 90% of the tumour should exhibit the classic growth pattern to be classified as tubular. However, if the co-existent carcinoma is solely of invasive cribriform type then the tumour should be typed as tubular if the tubular pattern forms > 50% of the lesion. Some tubular differentiation may be seen in tubulolobular carcinomas (see above). Mitoses are rare.

These lesions are thought to derive from low-grade DCIS and CCLs, and are rarely associated with lobular neoplasia (Abdel-Fatah et al., 2008).

The clinical presentation is usually as a palpable, mammographically and sonographically visible lesion. The characteristic radiological appearance is that of a stellate abnormality, which can resemble a radial scar, or an asymmetric density.

Prognosis is excellent, but lymph node metastases may occur. Rare metastases are found especially in bone.

Immunohistochemically, typically E-cadherin is positive, ER is positive and PR usually so, HER2 is negative and CK5/6 and CK14 are negative. Almost all tubular cancers therefore fall into the luminal A molecular type.

4.7 Invasive cribriform carcinoma

This tumour is composed of cribriform islands of small, uniform neoplastic cells, with cytoplasmic apical snouting. Nuclei should not show a high degree of atypia. More than 90% of the lesion should exhibit a cribriform appearance, except in cases where the only co-existent pattern is tubular carcinoma (see above), when > 50% must be of a cribriform appearance to be classified as of invasive cribriform type. There are usually very few mitoses.

These lesions are thought to derive from low-grade DCIS and CCLs, and are rarely associated with lobular neoplasia (Abdel-Fatah et al., 2008).

The clinical presentation is usually as a well-circumscribed, palpable and mammographically and sonographically visible lesion, generally not stellate as in tubular carcinoma. Prognosis is quite good, but lymph node metastases may occur. Rare metastases are found especially in bone.

Immunohistochemically, typically E-cadherin is positive, ER/PR are positive, HER2 is negative and CK5/6 and CK14 are negative. Almost all invasive cribriform cancers therefore fall into the luminal A molecular type.

4.8 Mucinous carcinoma

This type has also been known as mucoid, gelatinous or colloid carcinoma. These lesions are relatively well circumscribed and show cribriform or papillary islands of uniform cells in lakes of extracellular mucin. The degree of nuclear atypia and mitotic activity may vary. At least 90% of the tumour must exhibit a mucinous appearance to be so classified.

Mucinous cancers can be divided into two groups: mucinous A (or paucicellular) and mucinous B (or hypercellular). Mucinous B cancers display histological and genetic features that significantly overlap with those of endocrine carcinomas (Sapino et al., 2000) (see below).

The precursor lesion is the mucinous form of DCIS (for both types A and B; for type B, there will usually be endocrine features) or the mucinous variant of CCLs (for type A).

The clinical presentation is usually as a palpable and well-circumscribed lesion on mammography and sonography, which can simulate a benign lesion.

Prognosis is good for pure (generally low-grade) lesions, but lymph node metastases may occur. Rare metastases are especially found in bone.

Immunohistochemically, typically ER/PR are positive, HER2 is negative and CK5/6 and CK14 are negative. Almost all mucinous type (types A and B) cancers therefore fall into the luminal A molecular type. Rarely, type B mucinous cancers may be luminal B.

4.9 Invasive micropapillary carcinoma

This type shows an infiltrative growth pattern of small regular micropapillary groups, not showing further gland formation (Zekioglu et al., 2004; Chen et al., 2008). Due to fixation and expression of mucins on the outside of the groups, they tend to detach from the surrounding stroma and often form a “halo” around the groups. Microcalcification is common. Atypia and mitotic activity may vary. These

tumours often show prominent lymphovascular invasion, and there are in fact cases where tumour is only found in lymph vessels.

The precursor lesion for this type of cancer is unclear. It is sometimes associated with micropapillary DCIS, but the common name does not imply a precursor relationship.

The clinical presentation is usually as a palpable lesion visible on mammography and sonography.

The degree of lymph node involvement is high. Prognosis is not essentially different from that of invasive ductal carcinomas of similar grade, but is therefore relatively poor because many are high-grade. There is a high rate of local recurrence.

Immunohistochemically, typically E-cadherin is positive, ER/PR are positive, HER2 is frequently positive and CK5/6 and CK14 are negative. These tumours can fall into the luminal A or luminal B molecular types, depending on the HER2 status and the Ki67 index. MUC1/EMA are present on the outer surface of the tumour islands (which can be stained by EMA), hence the original “inside-out” carcinoma designation.

4.10 Neuroendocrine breast cancer

Neuroendocrine breast cancers share the peculiar architecture of solid sheets of cells with a tendency to produce peripheral palisading or insular structures separated by a delicate fibro-vascular stroma common to most endocrine tumours in general. Even more specific are the cytological features: plasmacytoid, spindle or with signet ring appearance and with intracytoplasmic granules (in the apex of plasmacytoid cells or diffuse in the cytoplasm of mucous or spindle cells), which are particularly evident in Giemsa and Diff-Quik-stained smears. There is also a variant with true small-cell cytonuclear features like in other organs. The degree of nuclear atypia and mitotic activity varies, so that these cancers can be grade 1, 2 or 3. True small-cell neuroendocrine carcinomas should be regarded as grade 3 by definition. There may be some mucinous or apocrine differentiation in non-small-cell types.

The precursor lesion for this type of cancer is DCIS with similar morphological (sometimes mucinous) and neuroendocrine features.

The clinical presentation is usually as a palpable lesion visible on mammography and sonography in older patients. Serum levels of (neuro)endocrine markers may be increased.

Immunohistochemically, neuroendocrine breast carcinomas are exclusively those expressing chromogranin A, or chromogranin B, or synaptophysin in > 50% of their cells (Abd El-Rehim et al., 2005). Cancers that do not reach this threshold should be classified as “breast carcinomas with focal neuroendocrine differentiation”, which also do not show the typical morphology overall and have an older age range of patients. Expression of ER and PR is variable, and there is frequently expression of AR and GCDFP-15 (BRST2), pointing to apocrine differentiation. These tumours, therefore, can fall into any of the HER2 negative subtypes.

Axillary lymph nodes may be involved. Prognosis is not essentially different from that of invasive ductal carcinomas of similar grade. Favourable prognostic features in these cancers are mucinous differentiation and expression of ER and PR.

4.11 Apocrine carcinoma

These tumours are architecturally much like (high-grade) ductal cancers with variable morphology, but cytologically the cancer cells have abundant granular or homogeneously pink ("apocrine") cytoplasm and big round nuclei with conspicuous nucleoli. Mitotic activity may vary. Apocrine differentiation may also be seen in endocrine cancers (see above).

The precursor lesion of apocrine carcinoma is the apocrine variant of DCIS (which also may show variable architecture), and (atypical) apocrine hyperplasia may be the earliest precursor. To what extent (atypical) apocrine adenosis is a precursor is still unknown.

The clinical presentation is usually as a palpable lesion visible on mammography and sonography.

Prognosis and degree of lymph node involvement are not essentially different from those of invasive ductal carcinomas of similar grade. HER2 positive apocrine cancers metastasise to the brain with increased frequency.

Immunohistochemically, typically E-cadherin is positive, usually ER/PR are negative, HER2 is most often positive, CK5/6 and CK14 are negative and CEA, BRST2 and AR are positive. Most apocrine cancers therefore fall into the HER2 molecular class, whereas some expressing ER/PR as well as HER2 are classified as the luminal B molecular type. Occasional cases may be triple negative. It has been suggested that apocrine carcinoma may be a special molecular class driven by AR signalling (Weigelt et al., 2008).

4.12 Medullary-like carcinoma

Typical medullary carcinomas are well-circumscribed cancers composed of interconnecting solid epithelial islands, with a syncytial pattern and often with necrosis. The epithelium shows highly atypical nuclei and many mitoses, and there is abundant lymphoplasmocytic infiltrate in and around the lesion. Some squamous metaplasia may occur.

Cancers showing many but not all of these features are classified as medullary-like, and because there are few differences in molecular make-up and clinical behaviour between typical medullary and medullary-like cancers, they are placed together in this histomolecular classification. There is clearly overlap here with metaplastic cancers as well. Also, many of the *BRCA1* germline mutation-related cancers fall into this group, as well as cancers with sporadic inactivation of *BRCA1* (e.g. due to promoter hypermethylation).

These cancers are thought to derive from a similar type of DCIS expressing basal cytokeratins/EGFR, which are likely to originate from the stem cells of the breast. In typical medullary carcinoma, the presence of DCIS is, however, rare.

The clinical presentation is usually as a rapidly growing palpable and well-circumscribed lesion on mammography and sonography, which can simulate a benign lesion, often at relatively young age.

Prognosis is intermediate (not as bad as for usual basal cancers), and brain metastases occur commonly, but interestingly lymph node metastases are rare in typical medullary type cancers.

Immunohistochemically, typically E-cadherin is positive, ER/PR/HER2 are negative and often CK5/6 and/or CK14 are positive; most medullary-like cancers therefore qualify as (high-grade) basal cancers (Jacquemier et al., 2005). There is no consensus on the required level of CK5/6 and/or CK14

expression, but probably any expression of these basal keratins would suffice. Further, there is often expression of EGFR, c-kit, p63, vimentin, smooth muscle actin, p-cadherin and CD10, and accumulation of p53 (usually related to a p53 mutation).

4.13 Metaplastic carcinoma

Metaplastic carcinomas often have the features of medullary-like carcinomas, showing metaplasia of any kind. Squamous metaplasia is most often found, but undifferentiated mesenchymal areas, or areas showing (smooth or striated) muscle differentiation, cartilage or bone formation, or more rarely nerve type differentiation may also occur. The degree of metaplasia (especially that of mesenchymal metaplasia) may be so abundant that the primary epithelial nature of the cancer is morphologically no longer, or hardly, recognised and the diagnosis of carcinoma is primarily based on cytokeratin expression. Of these, especially the basal and high-molecular-weight cytokeratins may be useful.

These usually high-grade cancers clearly overlap with the medullary-like cancers, and thereby include cancers in *BRCA1* germline mutation carriers, as well as cancers with a sporadic inactivation of *BRCA1* (e.g. due to promoter hypermethylation).

These cancers are thought to derive from a similar type of DCIS expressing basal cytokeratins/EGFR, which are likely to originate from the stem cells of the breast, and are thereby basal (Weigelt et al., 2009).

The clinical presentation is usually as a rapidly growing palpable and well-circumscribed lesion on mammography and sonography, which can simulate a benign lesion, often at relatively young age.

Prognosis is among the worst, and lung and brain metastases occur commonly.

Immunohistochemically, typically E-cadherin is positive, ER/PR/HER2 are negative and CK5/6 and/or CK14 are positive. Most metaplastic cancers therefore qualify as (high-grade) basal cancers. There is no consensus on the required level of CK5/6 and/or CK14 expression, but probably any expression of these basal keratins would suffice. Further, there is often expression of EGFR, c-kit, p63, vimentin, smooth muscle actin, p-cadherin and CD10, and accumulation of p53.

4.14 Salivary gland-like cancers

This group of low-grade cancers comprises salivary gland-like cancers such as adenoid cystic carcinomas and rarer forms like acinic cell and oncocytic carcinomas. They show infiltrative growth, are often multinodular, and morphology may vary. The typical adenoid cystic types grow in rounded-off epithelial groups of varying size, with some showing a cribriform pattern. There are two types of lumina; the less abundant type is epithelial, whereas the lumina lined by myoepithelial cells contain greyish basement membrane substance. Apocrine snouts around the lumina are lacking, which helps to distinguish them from the lumina in tubular/cribriform carcinomas. Nuclear atypia is usually low, and mitoses are typically not frequent. Other cancers in this group may show morphologically prominent myoepithelial differentiation. Perineural invasion (which is prominent in this tumour type within salivary glands) is rare in the breast.

In situ precursors of similar (low-grade) morphology can be found in these cancers.

The clinical presentation is usually as a palpable lesion visible on mammography and sonography. Prognosis is very good, and lymph node metastases are rare.

Immunohistochemically, typically E-cadherin is positive, ER/PR/HER2 are negative, CK5/6 and/or CK14 are positive and there is expression of vimentin and myoepithelial markers such as smooth muscle actin, p63, calponin and caldesmon. Most salivary gland-like cancers are therefore classified as (low-grade) basal cancers (Foschini & Krausz, 2010). Most adenoid cystic carcinomas bear a t(6;9)(q22-23;p23-24) translocation, which leads to the formation of the MYB-NFIB fusion gene (Wetterskog et al., 2012).

4.15 Secretory carcinoma

Secretory type cancers may be considered to be almost a genetic entity due to the high frequency of a ETV6-NTRK3 fusion gene, which encodes a chimeric tyrosine kinase that has an oncogenic effect (Reis-Filho & Lakhani, 2008). Pathologically, these lesions are less well defined. Grossly, they usually form a circumscribed mass, architecturally resembling ductal carcinomas but with microcysts and/or secretory cytology with bubbly cytoplasm, signet cells and low-grade nuclei. DCIS precursors with similar cytology have been described.

The clinical presentation is usually as a palpable lesion visible on mammography and sonography at exceptionally young age and fairly often in males. Prognosis is very good, but lymph node metastases may occur.

Immunohistochemically, typically ER/PR are negative and HER2 is negative, but CK5/6 or CK14 may be positive, so that secretory cancers partly fall into the basal cancer group but should not be regarded as high-grade (Lae et al., 2009).

4.16 Ductal – no special type (ductal-NST)

This group contains infiltrating carcinomas that cannot be assigned to any other category or classified as any of the less common variants of infiltrating breast carcinoma. The tumour shows < 90% special type characteristics. Consequently, invasive ductal carcinomas exhibit great variation in appearance and are the most common carcinomas, accounting for up to 75% in published series.

These lesions are thought to derive from DCIS or CCLs, rarely from lobular neoplasia.

The clinical presentation is usually as a palpable, mammographically and sonographically visible lesion. Prognosis of ductal-NST carcinomas as a group is worse than that of most specific types, and is largely dependent on proliferation rate and grade. Lymph node metastases often occur. Systemic spread may occur to any site, with a preference for bone (especially ER positive/PR positive cases), liver, lung and, to a lesser extent, brain (especially HER2 positive cases).

Immunohistochemically, E-cadherin is usually positive and there is variable expression of ER/PR, HER2 and the basal cytokeratins CK5/6 and CK14. Molecularly, ductal cancers therefore spread over all molecular groups. Cases with ER and/or PR positivity and HER2 negativity are luminal A, cases with ER and/or PR positivity and HER2 positivity or high Ki67 are luminal B, cases with ER/PR negativity and HER2 positivity are HER2, and triple negative cases with CK5/6 and/or CK14 and/or EGFR expression are classified as basal.

Many cancers in patients with *BRCA2* mutations have “no special type” morphology, and most are molecularly classified as luminal A because they express ER and/or PR but usually lack HER2 expression.

Also, lymphoepithelioma-like cancers are included here as they are as yet difficult to group into any of the above categories, and are too rare to form a useful separate entity. Lymphoepithelioma-like carcinoma is considered by some to be a variant of invasive lobular cancer. It is usually well demarcated, with scattered pale individual cells, sometimes also forming solid groups, with vesicular nuclei showing only slight/moderate atypia and a small single nucleolus, mixed within a very marked infiltrate of lymphocytes and plasma cells. This tumour type, therefore, resembles the undifferentiated form of nasopharyngeal carcinoma, and has to be differentiated from medullary carcinoma. Mitotic activity may vary. Immunohistochemically, there is also variable expression of E-cadherin and ER/PR, HER2 is mostly negative and CK5/6 and CK14 are negative. Epstein-Barr virus (EBV) has not been convincingly demonstrated. Lymph node metastases are rare.

4.17 Other malignant tumours

Non-epithelial tumours and secondary carcinomas are included in this category. For purposes of convenience, malignant phyllodes tumours and malignant adeno-myoepitheliomas should also be recorded here.

4.18 Not assessable

This category should be used only if an invasive carcinoma cannot be assigned to any of the previous groups for technical reasons, e.g. the specimen is too small or poorly preserved.

5 Assessing the axilla

Axillary nodal status remains the single most important prognostic factor in patients with breast cancer. For many years, axillary dissection has been the preferred method for obtaining information about nodal status, but it has significant morbidity and node negative patients are hence disadvantaged. Axillary status can be assessed by various methods: clinical, radiological, pathological and surgical techniques, the most accurate of which is histological examination of sentinel lymph nodes (SLNs) or axillary dissection. Although imaging techniques (ultrasonography [US], colour Doppler, scintimammography, high-resolution computed tomography, etc.) have improved, no individual patient can be managed based exclusively on these results. Physical examination is inaccurate because node metastases are present in 15–60% of patients with non-palpable nodes and absent in up to one third of patients with palpable nodes.

5.1 Preoperative staging

Preoperative staging of axillary lymph nodes (ALNs) should be performed to support specific therapeutic decisions. For example, in patients with axillary metastases, the surgeon can save the expense and time needed to perform SLN biopsy (Boughey et al., 2010); more extensive staging can be planned, and “tailored” neoadjuvant chemotherapy can be instituted. US-guided biopsy of suspicious ALNs can be performed by fine-needle aspiration (FNA) using a 25 gauge needle (Bonnema et al., 1997; Sapino et al., 2003; Bedrosian et al., 2003; van Rijk et al., 2006) and/or by core needle

biopsy (CNB) using a 14–18 gauge needle (Damera et al., 2003; Nori et al., 2005; Abe et al., 2007; Rao et al., 2009); the choice is mainly operator- and institution-dependent. Specificity for both methods is high (95.7–100%), and sensitivity varies for both, depending on the selection criteria and number of passes. According to some authors (Rao et al., 2009), as far as diagnosis is concerned there is no clear advantage for either technique. Given the necessary expertise, FNA may allow equivalent sensitivity at a lower cost. FNA is also preferred by many units due to the proximity of large vessels and nerves. US-guided ALN biopsy can detect up to 64% of node positive patients preoperatively (Damera et al., 2003; Rao et al., 2009).

Image-guided and fine-needle aspiration cytology (FNAC) or CNB of suspicious nodes can improve the preoperative axillary staging of patients with breast cancer, avoiding a 2-stage axillary procedure and sparing as many as 1–26% of patients open lymph node biopsies (Krishnamurthy et al., 2002; Ciatto et al., 2007).

Handling of FNA and CNB material should follow the procedures in Chapters 6a and 6b of the fourth edition (Wells et al., 2006a).

Diagnosis is usually straightforward when representative material is available with good technical quality; most false-negative cases correspond to “small” metastases.

All cases of primary invasive breast cancer with negative results for metastatic disease on FNA or CNB are candidates for a SLN biopsy or other axillary procedure for definitive staging.

To validate the diagnosis of metastasis from breast cancer, the morphology of the neoplastic cells present in lymph node FNA or CNB material should be compared with that of the primary tumour by review of the preoperative FNA or CNB material of the primary breast carcinoma because metastases are usually similar to the primary tumour.

5.2 Problems and pitfalls

Although the aim of preoperative study is the search for metastases, caution should be taken to avoid overlooking other causes of axillary lumps.

5.2.1 Epidermal cysts and skin appendage tumours

Inflammatory processes can occur from skin appendages. In these cases, an inflammatory background and the presence of squamous epithelial cells should alert the pathologist to a benign lesion. Apocrine cells may also be included from neighbouring apocrine sweat glands and may cause some problems. Skin appendage tumours may also be a challenge because the morphology and immunohistochemical profile can be quite similar to those of breast tumours. The clinical setting combined with US should allow a correct diagnosis. Axillary breast tissue may also give rise to benign epithelial clusters in smears.

5.2.2 Lymph node inclusions

Benign epithelial or, rarely, mesothelial inclusions in lymph nodes can also occur in approximately 0.3% of cases and may be difficult to recognise due to their cytokeratin positivity. Naevus cell inclusions also occur.

5.2.3 Granulomatous inflammatory processes

Epithelioid macrophages can mimic carcinoma cells. They are associated with other inflammatory cells in the smear, and the presence of multinucleated macrophages should prevent a false-positive diagnosis. In case of doubt, cytokeratin immunocytochemistry may solve the differential diagnostic problem. Whenever necrosis is present, microbiological or serological studies must be done to rule out the possibility of a specific infection (tuberculosis, cat scratch disease, etc.).

5.2.4 Dermatopathic lymphadenopathy

Melanophages can mimic carcinoma cells if the load of melanin is low. Usually the smears are not very rich in these cells and they are isolated or aggregated to vascular stroma. In case of doubt, immunostaining (CD68, etc.) easily confirms the histiocytic origin of these cells.

5.2.5 Lipoma

Mature adipose tissue can originate from a lipoma or from an adipose lymph node; only the presence of lymphoid cells can indicate the origin of the lump. The absence of carcinoma cells does not rule out the possibility of metastases, although their occurrence in adipose lymph nodes is not frequent.

5.2.6 Metastatic tumours

Melanoma

ALN metastasis of melanoma is a difficult differential diagnosis. An amelanotic melanoma can mimic breast cancer metastasis and this differential diagnosis can be particularly difficult in FNA. However, metastases from melanoma frequently contain mono-, bi- and multinucleated malignant cells. The cells also are polyhedric and/or fusiform with intranuclear pseudo-inclusions and prominent nucleoli. Before looking at an ALN sample, a review of the primary breast carcinoma is important because metastases are usually similar to the primary tumour. Although melanoma can be a mimicker, it can be recognised and even in the absence of a clinical history, immunostaining (HMB45, pS100, etc.) is diagnostic.

Carcinoma (other than breast)

When a peculiar pattern, unusual for breast cancer, is seen, another metastatic tumour should be considered. A careful clinical history complemented by image study will resolve most problems.

Lymphoma

Hodgkin cells may mimic carcinoma cells; however, the comparison with the primary tumour, the relative paucity of malignant cells and the presence of eosinophils may alert the pathologist to the diagnosis of Hodgkin lymphoma.

Whenever a non-Hodgkin lymphoma, especially small-cell lymphocytic lymphoma or lymphomas with plasmacytoid differentiation, is considered, haematological assessment, immunophenotyping and/or open tissue biopsy is indicated.

5.3 Examination and interpretation of sentinel lymph node biopsy specimens

Since the publication of the fourth edition (Wells et al., 2006a), new problems and techniques have emerged in relation to nodal staging, reflected in the TNM classification, which is now in its seventh edition (Sobin et al., 2009). Although the basic principles have not changed, it was considered

necessary to update the pathology chapter (Chapter 6) with a supplement addressing some of the difficulties experienced since the fourth edition was published.

5.3.1 Isolated tumour cells (ITCs)

The aim of pathological nodal staging is to make a statement about the status of the lymph nodes, either node negative or node positive. The nodes are usually axillary, but very occasionally from other locations. On the basis of the SLN theory, the SLNs are the first nodes draining lymph from the tumour site and are therefore the most likely site of nodal metastasis. With selective removal of these lymph nodes, there is an opportunity for more detailed pathological work-up of these, and indeed the SLNs are subjected to more intensive investigations. However, the techniques used vary from institution to institution, and this heterogeneity is partly responsible for discrepancies in the rates of node positive breast cancer reported by different centres.

The pTNM is the staging system generally used. In its sixth edition, this classification introduced the concept of a type of nodal involvement (termed isolated tumour cells; ITCs) that should not be considered a metastasis from the point of view of staging and treatment. ITCs may be identified either by morphological methods [pN0(i+)] or by molecular studies [pN0(mol+)]. Whenever the lymph nodes are investigated for such lesions but the results are negative, the pN0(i-) and pN0(mol-) symbols are recommended by the TNM definitions. Although it is clear that pN0(i+) may reflect both ITCs detected by the standard haematoxylin and eosin (H&E) staining and those detected by immunohistochemistry (IHC), the seventh edition of the TNM classification (Edge et al., 2009; Sobin et al., 2009) clarifies that pN0(i-) should be used only for negative IHC findings.

It was found to be common practice to evaluate SLNs at multiple levels. *Gross slicing of the SLNs* (and, in some recommendations, any lymph node) *larger than 5–6 mm* and *H&E staining of one level per such slice* is recommended in nearly all protocols. When this investigation or step sections by H&E staining disclose no nodal involvement of any type, the pN0 category without any qualifiers should be used, and pN0(i-) should be reserved for negative investigations beyond this limit, i.e. negative findings by IHC. Because this still allows several levels of precision in staging, it is considered optimal to report the method of SLN investigation.

The definition of ITCs and their distinction from micrometastases was suboptimal and did not allow a reproducible segregation of low-volume nodal involvement into these two pTNM categories (Cserni et al., 2005; Cserni et al., 2006; de Mascarel et al., 2008). Although the classification of metastatic cells into the ITC category on the basis of the largest cluster not exceeding 0.2 mm in maximum diameter may have resulted in better reproducibility (Turner et al., 2008), this sometimes included rather large volume lesions, which were found to be associated with metastases beyond the SLNs more commonly than ITCs interpreted with a more restrictive approach (Cserni et al., 2008; Turner et al., 2008; Cserni, 2009; van Deurzen et al., 2010). The seventh edition of the TNM classification has introduced an alternative upper limit of 200 cells for ITCs (Edge et al., 2009; Sobin et al., 2009), and this should reduce the diagnostic discrepancies between different centres and also help to eliminate the unacceptable practice of labelling relatively extensive metastatic involvement of a lymph node as pN0(i+). Any lymph node involvement > 0.2 mm but < 2 mm in any of the 3 dimensions is categorised as a micrometastasis. According to the current TNM classification (Edge et al., 2009), the 0.2 mm size cut-off relates to the maximum diameter of the largest tumour cell cluster. There may be instances of nodal involvement with the largest cluster measuring < 0.2 mm in diameter but containing > 200 cells, and vice versa, clusters > 0.2 mm in diameter with < 200 cells. Size should be considered first, and the cell count only if the largest cluster is < 0.2 mm (Cserni et al., 2011).

5.3.2 Methods of nodal assessment

One of the aims of the European guidelines is to decrease heterogeneity in staging. These should, therefore, concentrate on the reliability and accuracy of determining a node negative status.

The guidelines state that the minimum aim of SLN investigation is to find all metastases greater than 2 mm in diameter (hereafter referred to as macrometastasis). It therefore follows that any SLNs reported to be pN0 or pN0(i-) should be interpreted that the SLN contains no metastasis greater than 2 mm in diameter (but may still contain micrometastasis or ITCs). Therefore, by devising a protocol to detect all macrometastases, the aim is to promote optimal accuracy for identifying node negative status related to macrometastases as near to 100% as possible (100% negative predictive value).

Similarly, when the guidelines state that the optimum aim in staging is to find (nearly) all micrometastases in the SLNs, they are also designed to identify a pN0 or pN0(i-) node negative status as near to a 100% predictive value for the lack of micrometastases as possible.

The identification of ITCs in the initial levels mandates a search for a larger metastasis in the same SLN, and levels should be cut to exclude the possibility of the ITC cluster being a smaller tangentially cut part of a larger metastasis. This would confirm the initial staging correctly as pN0(i+).

The pN0 or pN0(i-) category reflects that no ITCs were identified by the histological protocol used, although they may be present, just as occult metastases of other size may be present, depending on the sampling of the lymph node. This is why a systematic search is recommended and the method used should be reported, as stated previously. Optimally, the accuracy of staging should also be mentioned, at least in unit protocols (e.g. pN0(i-) with the examination of levels separated by 500 µm devised to identify metastases not smaller than 500 µm in greatest dimension; or pN0 with the examination of 4 levels separated by 250 µm from each slice approximately 2 mm in thickness, devised to identify metastases not smaller than 1 mm in greatest dimension).

No reasonable histological method can aim to identify ITCs with 100% accuracy in the SLNs, but some reports have suggested that ITCs identified incidentally with protocols devised to identify larger metastases also have some impact on prognosis (Colleoni et al., 2005; de Boer et al., 2009).

5.3.3 Intraoperative assessment

The sensitivity and positive predictive value of the intraoperative assessment of SLNs can be improved by specific techniques such as rapid immunohistochemistry and molecular testing of the lymph nodes.

Immunohistochemistry is most useful in cases of lobular carcinomas with micrometastatic involvement of the SLN, but can also help to identify larger metastases, and facilitate later paraffin section evaluation (Nahrig et al., 2003; Leikola et al., 2005; Choi et al., 2006).

Molecular (RT-PCR-based) assays may detect ITCs in greater proportion than histological methods, but cannot be optimally validated against the gold standard method of histology because of sampling biases. Validation can partially be made against other validated molecular assays.

The first group of RT-PCR-based nodal staging tests looked purely at the presence or absence of mammary epithelial (e.g. CK19, MUC-1, mammaglobin) or tumour cell markers (e.g. carcinoembryonic antigen, MAGE-3) and were confounded by many technical biases. Such tests should not be used for routine purposes (Cserni et al., 2003).

The second generation of tests are based on quantitative real-time or other mRNA amplification assays, and these, with a well-chosen cut-off value, may exclude false-positive cases with reasonable accuracy (Visser et al., 2008; Viale et al., 2008; Julian et al., 2008; Mansel et al., 2009; Castellano et al., 2012). Such assays, by means of being quantitative, can also be calibrated in a way to exclude the histological category of ITC (i.e. the pN0(i+) category) and have been successfully used in the intraoperative setting. They also allow more tissue to be sampled (and homogenised). The main problem remains that of sampling (Daniele et al., 2009). If the tissue is used intraoperatively for a molecular assay, it can no longer be used for histology, and vice versa. While the staging categories remain defined by size, this should be measured and therefore histology will remain the gold standard

unless equivalence between a metastasis measuring a given size and a given amount of mRNA can be established in the future.

It must be noted that a validated quantitative molecular assay is designed to detect nodal involvement greater than ITCs but the pN0(mol+) category recommended by both the sixth and seventh editions of the TNM classification does not take this into account. We therefore recommend that such cases be reported as "node positive (pN1) with … (named) molecular assay". Validation studies for some of the quantitative molecular assays have already been published (Cserni, 2012).

When testing SLNs in the intraoperative setting, a compromise should be made between the use of practically the whole lymph node for the molecular assay (aiming at the highest accuracy in staging) and the use of a part of the SLN for histology and using the residuum for the molecular assay (aiming at increasing the accuracy of staging, but also facilitating a more complex histological evaluation of the lymph node). Many pathologists would agree that no molecular assay should be carried out without a morphological examination of the nodal tissue because the molecular staging assay is simply a test for the presence or absence of metastases, whereas histology is a more complex diagnostic method capable of identifying other nodal disorders too. Thus, should the first approach of using the whole SLN for molecular assay be favoured, it is recommended that at least a frozen section or a (touch or scrape) cytology specimen be taken, examined and archived by pathologists for microscopic evaluation of the SLN tissue.

In the non-intraoperative setting, histology is the method of choice for SLN assessment (Khaddage et al., 2011; Cserni, 2012; Castellano et al., 2012).

The findings of the ACOSOG Z0011 trial, suggesting that SLN micrometastases do not adversely impact on prognosis and challenging the view that all patients with a positive SLN require complete ALN clearance, has led to a reduction in the need for intraoperative SLN evaluation (Giuliano et al., 2011).

6 Microinvasive carcinoma

A frequently encountered problem in the histological examination of in situ carcinoma is identifying the smallest focus or foci of invasive carcinoma (Boecker et al., 2006), so-called microinvasive carcinoma (MIC).

Although microinvasion is virtually almost exclusively associated with high nuclear grade comedo DCIS, it may also be associated with other types of DCIS and with LCIS. Microinvasion is reported to be related to the size/extension of associated in situ carcinoma. The incidence rate of MIC ranges from 0.68% to 2.4%.

The current prevailing view is that MIC appears to have an excellent prognosis with a low risk of associated ALN metastasis, but the reported incidence of ALN metastasis ranges from 0% to 20% (Bianchi & Vezzosi, 2008).

By general agreement, SLN biopsy is the standard procedure in the treatment of patients with this type of lesion due to the possibility of ALN metastasis in MIC. Because most MICs with a positive SLN have low-volume metastases and consequently a low risk of additional metastases in ALNs, the role of complete ALN dissection is still debated.

From a practical point of view, as mammographically detected microcalcifications considered to be associated with in situ breast carcinoma are preoperatively assessed by percutaneous core biopsy (NCB) or vacuum-assisted needle core biopsy (VANCB), an accurate histological diagnosis identifying microinvasion on core biopsy enables a SLN biopsy and excision of the primary tumour to be performed in a single surgical session (Bianchi & Vezzosi, 2008).

6.1 Definition

MIC is defined as a tumour in which the dominant lesion is in situ carcinoma (usually extensive high nuclear grade DCIS, but rarely other types of DCIS or LCIS) in which there are one or more clearly separate foci of infiltration, usually into non-specialised interlobular (Tavassoli & Eusebi, 2009) or interductal fibrous or adipose tissue, none measuring > 1 mm (about 2 high-power fields) in maximum diameter. When there are multiple foci of MIC, only the size of the largest focus is used to classify the microinvasion; the presence of multiple foci of microinvasion should, however, be noted and/or quantified. This definition is very restrictive, and tumours fulfilling the criteria are consequently rare (Sloane, 1991).

A focus of invasive carcinoma 1 mm or smaller without associated in situ carcinoma is not MIC but should be classified as invasive carcinoma and the maximum diameter measured, in contrast to the statement in the seventh edition of the *AJCC Cancer Staging Manual/Handbook* (Edge et al., 2009). The AJCC include small invasive tumours not exceeding 1 mm without in situ carcinoma in the microinvasive category.

The fifth edition of the *TNM Classification of Malignant Tumours*, published in 1997 (Sobin & Wittekind, 1997), was the first one that recognised a specific T substage for MIC, defined as "the extension of cancer cells beyond the basement membrane into the adjacent tissues with no focus more than 0.1 cm in greatest dimension" and formally reported it as pT1mic.

In the 2003 edition of *WHO Classification of Tumours, Pathology and Genetics of Tumours of the Breast and Female Genital Organs* (Sobin & Wittekind, 1997; Tavassoli & Devilee, 2003), it is reported, in spite of the pT1mic category officially being recognised by the fifth edition of the *TNM Classification of Malignant Tumours* (Sobin & Wittekind, 1997), that there is no generally accepted agreement on the definition of MIC, particularly about the maximum diameter compatible with a diagnosis of microinvasion. On this basis, MIC is still considered an evolving concept that has not reached the status of a WHO-endorsed disease entity. In the seventh edition of the TNM classification, this is referred to as pT1mi.

6.2 Pathological diagnosis

The tumour focus/foci must usually invade into non-specialised interlobular or interductal stroma. The cells deemed to be invasive should be distributed in a non-organoid pattern that does not represent tangential sectioning of a duct or lobular structure involved by in situ carcinoma. Tangentially sectioned foci of in situ carcinoma that simulate microinvasion are within the specialised intralobular and periductal stroma and usually occur as compact groups of tumour cells that have a smooth border surrounded by a circumferential layer of myoepithelial cells and stroma or a thickened basement membrane (Rosen, 2009).

At the sites of microinvasive foci, tumour cells are distributed singly or as small groups that have irregular shapes reminiscent of conventional invasive carcinoma with no particular orientation (Rosen, 2009).

The absence of basement membrane material around nests of tumour cells defines the process as being invasive. IHC for basement membrane components (laminin and type IV collagen) is helpful in demonstrating the presence or absence of basement membrane (Schnitt, 1998) even though IHC for laminin and type IV collagen is reported to be technically problematic in formalin-fixed, paraffin-embedded tissue (Schnitt, 1998). Unfortunately, cells of invasive cancer can still synthesise some components of basement membrane around invasive nests; therefore, the use of basement membrane markers for the detection of stromal invasion is not formally recommended and, when performed, should be interpreted with caution (Yaziji et al., 2000).

The presence of myoepithelial cells around nests of carcinoma cells defines the process as being in situ. IHC for myoepithelial cells has been used to help to determine whether a process represents in situ carcinoma or stromal invasion (Schnitt, 1998). A variety of markers have been used to detect myoepithelial cells. At present, the most commonly used antibodies are smooth-muscle myosin heavy chain (SMM-HC) and calponin; these are more specific for myoepithelial cells than actin antibodies (such as 1A4 and HHF-35 clones) and react less commonly with myofibroblasts. SMM-HC is not a perfect marker of myoepithelial cells because it manifests slightly lower sensitivity than calponin. Therefore, the optimal sensitivity and specificity of myoepithelial cell markers can be achieved when the SMM-HC marker is used in conjunction with the more sensitive but less specific marker calponin (Yaziji et al., 2000).

In a more recent study (Werling et al., 2003), antibodies to p63, a member of the p53 gene family, have been reported to offer excellent sensitivity and increased specificity for myoepithelial cells relative to antibodies to calponin and SMM-HC. p63 antibodies have the following diagnostic limitations: (1) they occasionally demonstrate an apparently discontinuous myoepithelial layer around nests of in situ lesions and (2) they react with a small but significant subset of breast carcinoma tumour cells, especially in tumours with a basal phenotype; however, this aberrant reactivity rarely causes diagnostic difficulty. p63 can complement or replace SMM-HC and/or calponin in the analysis of microinvasion because of its near-perfect sensitivity and near-absolute specificity in distinguishing myoepithelial cells from myofibroblasts. Detecting microinvasion can be difficult when there is marked periductal fibrosis or inflammation; in these cases, IHC for cytokeratin may be useful to confirm the presence of separate foci of neoplastic cells embedded in periductal fibrosis or inflammation.

The diagnosis of microinvasion still remains problematic, even with the use of ancillary techniques. If there is sufficient doubt about the presence of microinvasion (i.e. in cases with marked fibrosis or inflammation), the case should be classified as in situ carcinoma with possible microinvasion.

The diagnosis of MIC depends principally on the tissue sampling. MIC cannot be reliably excluded unless all tumour tissue is completely embedded and submitted for histological examination. This method is now recommended in clinical guidelines (Olivotto & Levine, 2001) as well as in breast screening programmes (National Coordinating Group for Breast Screening Pathology, 2005). However, it is well known that even with a large number of paraffin blocks, only a part of the tissue is examined microscopically and pathologists can never be absolutely certain that microinvasion is really absent.

Serial sections supported by IHC usually provide the best evidence of microinvasion. Care should be taken to perform IHC early in the evaluation of suspected microinvasion, before the blocks have been sectioned excessively (Rosen, 2009), first to confirm microinvasion and second to exclude the possibility of larger invasive foci.

6.3 Differential diagnosis

Microinvasion is one of the most commonly overdiagnosed events (if not the most commonly overdiagnosed event) in the pathology of breast carcinoma.

A variety of patterns in DCIS and, more rarely, in LCIS may be misinterpreted as stromal invasion. Schnitt has summarised lesions and artefacts commonly mistaken for microinvasion (Schnitt, 1998):

- DCIS involving lobules ("lobular cancerisation");
- chronic inflammatory reaction present in association with, and obscuring, involved ducts and acini;
- branching of ducts;
- distortion or entrapment of involved ducts or acini by fibrosis (due to a prior needling procedure);
- crush artefacts;
- cautery effects;
- artefactual displacement of DCIS or LCIS cells into the surrounding stroma or adipose tissue due to tissue manipulation or a prior needling procedure. In cases with a history of a prior needling procedure (FNA, NCB or VANCB), the diagnosis of MIC should be made with caution because artefactual disruption of the epithelial-stromal junction of glandular structures involved by in situ carcinoma is not infrequently encountered in the subsequent excision biopsy or therapeutic specimen. Granulation tissue, old or recent haemorrhage (iron-laden macrophages, cholesterol crystals), tissue tears and a degenerative appearance of the dislodged tumour cells can help in distinguishing pseudo-invasion from true invasion (Tavassoli, 1999);
- DCIS or LCIS involving benign complex sclerosing lesions such as radial scars, sclerosing adenosis, sclerosing papilloma, ductal adenoma.

7 Pathological reporting of post-chemotherapy specimens

Pathological assessment of resections after primary or neoadjuvant chemotherapy can be difficult if there has been a complete response, if there is insufficient information about the site of the original tumour and whether the lesion is uni- or multi-focal. When neoadjuvant chemotherapy is contemplated, the clinician who performs the core biopsy may insert a permanent marking device such as a metallic marker or tattoo (Dash et al., 1999), which can be identified in the specimen (see below). The pathologist should also have access to radiological information such that, where a marker may not have been placed, the original quadrant and site of the tumour can be assessed to accurately identify any residual tumour. In this respect, MRI findings are generally preferred because these are often more detailed than US or mammographic examination (Londero et al., 2004). It is, however, recommended that all cases undergoing preoperative chemotherapy should have some form of marker inserted to avoid problems with the subsequent localisation of the tumour. This is particularly important in cases with complete response having wide local excision because it can be impossible to be certain that the site of the tumour nidus has been removed with a conservation procedure. To ensure that the tumour site is completely removed, some units tattoo the skin with 4 tattoos to delineate the maximum tumour size. It is recommended that some form of marking of the extent of the tumour pre-chemotherapy be performed either by this method or an equivalent system. The pre-chemotherapy core biopsy should always be reviewed in parallel to assess response. Some units give an assessment of the number of cores involved.

Specimens from cases after primary or neoadjuvant chemotherapy should be orientated and sent to the laboratory in a similar way to the specimen delivery of other specimens and as described in the fourth edition (Wells et al., 2006b). Adequate prompt fixation is as important here as in any other breast specimen.

7.1 Macroscopic examination

Macroscopic examination of post-chemotherapy specimens should be performed in a similar manner to wide local excision or mastectomy specimens as recommended in the fourth edition (Wells et al., 2006a) and the margins painted according to the standard local protocol in place. Post-chemotherapy tumours may be obvious if there has been little response, and these specimens are really no different to handle from any primary resection specimen. In this case, reference to the protocol in the fourth edition (Wells et al., 2006a) or a local protocol based on these will suffice. If there has been a good response, however, the tumour may be soft, pale and/or ill-defined and identification and sizing of the residuum may be impossible. Often, an oedematous area of fibrosis may be all that can be seen or if there is associated calcification, radiographic examination of the specimen and/or radiography of the subsequent slices may be helpful (Pinder et al., 2007). If a clip or wire coil has been used, this can also be identified by specimen radiography and attention directed to this area. The clip or coil should be macroscopically visible and should be removed before taking blocks of the area. If a tattoo method has been used to identify the site of the tumour then this area should be easily identified macroscopically. If there has been no pre-chemotherapy marking and there has been a good response, it may be very difficult to identify the tumour. In this case, the original site of the tumour should be identified based on the MRI or radiology as above, and this area should be extensively sampled.

Post-chemotherapy specimens should be sliced at regular intervals (not more than 1 cm for mastectomy and 0.5 cm for wide local excision specimens) such that reconstruction of the residual tumour size can be performed accurately in the third dimension by multiplying the number of slices in which tumour is seen by the thickness of the slices (Apple & Suthar, 2006). Any multicentricity should be sought by reference to the radiology and by careful macroscopic examination. Large blocks may also be useful in this regard.

Lymph nodes should be blocked in the routine manner recommended in the fourth edition, depending on the procedure (see Section 6b.3 in the fourth edition). Often preoperative nodal staging will have been attempted, sometimes by preoperative US-guided FNA or biopsy or by a pre-treatment SLN procedure. The latter may make the axillary assessment more difficult due to fibrosis. There is a suggestion that a decreased yield of nodes may be found (Baslaim et al., 2002).

7.2 Microscopic examination

If residual tumour is identified then there is no problem, but if there has been a complete response, there may be difficulty in confirming the original tumour site. As noted above, this can be critical to ensure that the nidus has been removed in those tumours being treated by wide local excision. Microscopy of the original tumour site often shows a scarred area with distortion of the normal architecture. Normal breast structures are effaced in this area, and there is often a macrophage and chronic inflammatory response, with the macrophages often containing iron pigment, haemosiderin and debris. Oedema or necrosis may be present, and the stroma is also frequently altered from normal breast stroma (Pinder et al., 2007). It is also important to assess both the cellularity and extent of the residual tumour. An indication of the extent can be assessed by documenting the number of involved blocks or slices per total number examined from the abnormal area.

Preoperative chemotherapy can extensively alter tumours such that identifying isolated residual tumour cells can be difficult. Immunocytochemical staining with a broad-spectrum cytokeratin such as AE1/AE3 can be helpful in this respect.

Chemotherapy may effectively sterilise the invasive tumour but may leave in situ changes or lymphatic vessel emboli relatively unchanged (Honkoop et al., 1997). It should be also noted that normal lobules may show apocrine changes and severe nuclear atypia post-chemotherapy.

7.3 Reporting of prognostic and predictive factors

Tumour size, grade, typing and predictive factors (ER, PR and HER2) are extensively altered by chemotherapy, and in general the pre-treatment core biopsy is recommended for assessment of these prognostic factors. One should, however, realise that IHC for HER2 may yield both false-negative and false-positive results on core biopsy and that grading is a little less reliable on core biopsy alone (Harris et al., 2003; Rakha & Ellis, 2007). Chemotherapy should not be given without a representative core biopsy of the tumour before treatment, and trial or research protocols also recognise this (e.g. NEOTANGO). Although grading may retain prognostic value, in general these prognostic factors are best assessed on the preoperative core biopsy until large amounts of data become available from neoadjuvant trials to contradict this advice. As regards ER, PR and HER2 status, there are conflicting results in the literature and because of this and the possibility of complete response, it is recommended to assess these parameters on the pre-treatment core biopsy.

7.4 Microscopic reporting of lymph nodes

Assessment of lymph nodes should take into account the number of nodes containing residual tumour, the number showing fibrosis with other associated features of tumour regression, such as haemosiderin-containing macrophages, and alteration of nodal architecture. All nodes identified macroscopically should be submitted for histology, and a careful search for occasional residual cells in fibrotic nodes should be undertaken, including the use of IHC for broad-spectrum cytokeratins, e.g AE1/AE3 or MNF116, where necessary. There is some evidence that patients with evidence of response of nodal metastases and regression of metastases have better disease-free survival than those with persistent metastases (Newman et al., 2003). Because of this, the number of nodes assessed to have evidence of tumour regression should also be recorded.

7.5 Reporting of tumour response

There are several histopathological classifications of response to chemotherapy, but there is no current consensus favouring any particular system. This hampers the comparison of response between different studies. Systems evaluating both the response of the primary tumour and the response in nodes are conceptually most likely to be more accurate in predicting the outcome, but hard evidence for this is lacking. Most systems acknowledge that complete histopathological response is an excellent prognostic sign, and this was the basis for Feldman's classification in 1986 (Feldman et al., 1986). This has been largely superseded by classifications such as those of Chevallier et al. (Chevallier et al., 1993) and Sataloff et al. (Sataloff et al., 1995). In some countries, the Miller-Payne system is used (Ogston et al., 2001).

Pinder et al. (Pinder et al., 2007) proposed a classification of response along the lines of that proposed by Sataloff et al. (Sataloff et al., 1995) (see Table 3). An approach such as this is logical and includes both nodal and tumour response. This would seem to be sensible especially in cases where there is total or almost total tumour response in the breast but little effect on nodal disease. Where

there is a mixture of categories, e.g. one node with a metastasis showing no response and one node showing fibrosis, the worst category should be used (e.g. category 4).

Table 3: Recommended classification of response to chemotherapy

Tumour response

1. Complete pathological response, either (i) no residual carcinoma or (ii) no residual invasive tumour but DCIS present.
2. Partial response to therapy, either (i) minimal residual disease/near total effect (e.g. < 10% of tumour remaining) or (ii) evidence of response to therapy but with 10–50% of tumour remaining or (iii) > 50% of tumour cellularity remains evident, when compared with the previous core biopsy sample, although some features of response to therapy present. Points (ii) and (iii) are somewhat subjective, especially when the core biopsy cannot be reviewed.
3. No evidence of response to therapy.

Nodal response

1. No evidence of metastatic disease and no evidence of changes in the lymph nodes.
2. Metastatic tumour not detected but evidence of response/down-staging, e.g. fibrosis.
3. Metastatic disease present but also evidence of response, such as nodal fibrosis.
4. Metastatic disease present with no evidence of response to therapy.

In summary, it is recommended that pathologists report the following factors in post-chemotherapy specimens:

- Presence or absence of invasive and/or in situ carcinoma
- Histological type of residual invasive tumour (as in the fourth edition)
- Residual tumour size and size of fibrotic nidus where possible
- Margin status
- Extent of nodal involvement (number of nodes, number of involved nodes and number of nodes with evidence of tumour regression)
- Percentage of fibrosis, necrosis and inflammatory response
- Presence or absence of multi-focality
- Presence or absence of chemotherapy effect on both invasive and in situ carcinoma
- Reporting of response to chemotherapy by an accepted system. The Working Group recommends the use of the system in Table 3
- ypTNM.

Features that should be reported on the original core biopsy:

- Type
- Grade

- Presence of vascular invasion
- Presence of in situ disease
- ER, PR and HER2 status.

Ki-67 may be requested and used to inform therapy by some oncologists. IHC may be repeated after neoadjuvant treatment, and also in all relapses or metastases, because some changes in IHC expression may be found.

A formula for the estimation of residual cancer burden has recently been published (Symmans et al., 2007). This uses an algorithm that can be found online on the MD Anderson Cancer Center website (www.mdanderson.org/breastcancer_rcb).

8 Vacuum-assisted needle core biopsy (VANCB)

For certain types of mammographic abnormality, particularly microcalcification with a moderate or low-level suspicion of malignancy, a larger volume of tissue is required for accurate diagnosis (Meyer et al., 1997). This is possible with VANCB. The biopsy probe incorporates a vacuum channel, which applies negative pressure to the biopsy port and thereby sucks the adjacent breast tissue into the port for sampling. The biopsy probe is introduced into the breast and positioned using image guidance. The vacuum is activated and sucks breast tissue into the biopsy port; a cutting cylinder then passes down within the probe and separates the biopsy material from the surrounding tissue. The biopsy specimen is then delivered by applying negative pressure and while the main probe remains within the breast. Multiple specimens are obtained by rotating the biopsy probe within the breast so that the biopsy port is applied to different areas of breast tissue.

The potential advantages of this system are the ability to obtain a larger volume of tissue for histological examination and the rapid evacuation of any haematoma that collects at the site of biopsy. This ensures that the specimens obtained are of good quality and are not compromised by the presence of haematoma (Liberman et al., 1998a).

Under local anaesthetic, after a cutaneous incision of 5 mm, VANCB allows the radiologist to obtain 5–25 cores with needles usually 8–11 gauge. In cases where the whole lesion or a high proportion of the lesion has been removed, a small metal marker should be introduced through the biopsy probe and deployed at the biopsy site (Liberman et al., 1998b; Kettritz et al., 2004).

In cases of microcalcification, the cores obtained must undergo specimen radiography to identify those that contain representative microcalcifications. Afterwards, they should be put into a separate container with a fixative and labelled “with microcalcification”. The other cores should be put into another container with fixative and labelled “without microcalcification”. The cores should be embedded in separate cassettes (those with microcalcification separate from those without microcalcification), the number of cassettes depending on the number of samples. In each cassette the cores should be put in parallel and embedded in the same way. If significant calcification is not identified in the initial levels, the paraffin blocks can be X-rayed and the microcalcification identified. More levels of the block or blocks containing calcifications should be made until the calcified lesion is identified. The pathologist should identify the microcalcification biopsied by the radiologist in the respective slides and compare the microscopic morphology with the X-ray pattern. The multidisciplinary meeting is an ideal forum to correlate this. Specimen radiography using a dedicated

microfocus device with magnification at low kilovoltage (14–16 kV) improves the accuracy of specimen radiography.

In the histological report, B categories should be used and it should be clearly stated whether or not microcalcification is present (Heywang-Kobrunner et al., 2003; Kettritz et al., 2005).

As in US-guided core biopsies, cases where significant microcalcification is not identified in histological slides after several levels should be discussed by a multidisciplinary team (radiologist, pathologist and surgeon) (Heywang-Kobrunner et al., 2003; Kettritz et al., 2005; Tot, 2005; Tot & Gere, 2008).

Consumables are expensive, and this technique is often reserved for non-palpable lesions, in particular for microcalcification. It is usually performed under stereotaxis, on digital dedicated tables, but it can also be US-guided (Liberman et al., 1998a).

Published results on VANCB have demonstrated a lower equivocal sample rate and increased accuracy in the detection of small invasive tumours associated with an area of DCIS. The accuracy of histological diagnosis is 3% higher with 10 gauge needles than with 14 gauge needles (Salem et al., 2009). Consideration of the likely underlying histological nature of the lesion from the imaging features should therefore be taken into account when deciding on the sampling method to be used (Tot, 2005).

8.1 Multidisciplinary correlation

A simple imaging classification should be used to indicate the radiologist's degree of suspicion (see Section 4.4.2 in the fourth edition). This is useful for multidisciplinary management and audit purposes. The indication and the preferred method for non-operative biopsy can be decided according to the degree of suspicion and the nature of the lesion.

To optimise treatment planning, non-operative histological diagnosis should be the goal, replacing diagnostic open surgical biopsy. Management decisions for excision biopsy after needle biopsy (CNB or VANCB) should made based on radiological-pathological correlation, the presence of a detectable residual lesion in imaging after NB and the character of the lesion.

When dealing with radiological-pathological correlation in the multidisciplinary conference, B3 results other than atypical epithelial proliferations of ductal type are not always an indication for an excision biopsy. Reliable decisions on regular follow-up without further diagnostic intervention include: (a) findings of lobular neoplasia (except the pleomorphic type or those with necrosis, categorised as B5a) or (b) flat epithelial atypia, but only if they are associated with a benign histological lesion that correlates with the biopsy target (e.g. fibroadenoma, etc.), (c) papillary lesions without atypical findings that were completely removed by the diagnostic intervention and (d) radial scars detected as an additional microscopic finding to a benign lesion considered to be the biopsy target, provided no additional architectural distortion is present. The implementation of vacuum-assisted biopsy devices generally used for the assessment of calcification increases the chance that small lesions are removed completely. Consequently, the number of B3 lesions for which no further surgical procedure is necessary increases and therefore the positive predictive value for malignancy on surgical excision improves.

References

Abd El-Rehim DM, Ball G, Pinder SE, Rakha E, Paish C, Robertson JF, Macmillan D, Blamey RW, Ellis IO (2005). High-throughput protein expression analysis using tissue microarray technology of a large

well-characterised series identifies biologically distinct classes of breast cancer confirming recent cDNA expression analyses. *Int. J. Cancer*, 116:340–350.

Abdel-Fatah TM, Powe DG, Hodi Z, Reis-Filho JS, Lee AH, Ellis IO (2008). Morphologic and molecular evolutionary pathways of low nuclear grade invasive breast cancers and their putative precursor lesions: further evidence to support the concept of low nuclear grade breast neoplasia family. *Am. J. Surg. Pathol.*, 32:513–523.

Abe H, Schmidt RA, Sennett CA, Shimauchi A, Newstead GM (2007). US-guided core needle biopsy of axillary lymph nodes in patients with breast cancer: why and how to do it. *Radiographics*, 27 Suppl 1:S91-S99.

Allred DC, Mohsin SK (2000). Biological features of premalignant disease in the human breast. *J. Mammary Gland. Biol. Neoplasia*, 5:351–364.

Apple SK, Suthar F (2006). How do we measure a residual tumor size in histopathology (the gold standard) after neoadjuvant chemotherapy? *Breast*, 15:370–376.

Aulmann S, Elsawaf Z, Penzel R, Schirmacher P, Sinn HP (2009). Invasive tubular carcinoma of the breast frequently is clonally related to flat epithelial atypia and low-grade ductal carcinoma in situ. *Am. J. Surg. Pathol.*, 33:1646–1653.

Azzopardi JG, Ahmed A, Millis RR (1979). Problems in breast pathology. *Major. Probl. Pathol.*, 11:i–466.

Baslaim MM, Al Malik OA, Al-Sobhi SS, Ibrahim E, Ezzat A, Ajarim D, Tulbah A, Chaudhary MA, Sorbris RA (2002). Decreased axillary lymph node retrieval in patients after neoadjuvant chemotherapy. *Am. J. Surg.*, 184:299–301.

Bedrosian I, Bedi D, Kuerer HM, Fornage BD, Harker L, Ross MI, Ames FC, Krishnamurthy S, Edeiken-Monroe BS, Meric F, Feig BW, Akins J, Singletary SE, Mirza NQ, Hunt KK (2003). Impact of clinicopathological factors on sensitivity of axillary ultrasonography in the detection of axillary nodal metastases in patients with breast cancer. *Ann. Surg. Oncol.*, 10:1025–1030.

Bianchi S, Vezzosi V (2008). Microinvasive carcinoma of the breast. *Pathol. Oncol. Res.*, 14:105–111.

Bijker N, Peterse JL, Duchateau L, Julien JP, Fentiman IS, Duval C, Di Palma S, Simony-Lafontaine J, de Mascarel I, van de Vijver MJ (2001). Risk factors for recurrence and metastasis after breast-conserving therapy for ductal carcinoma-in-situ: analysis of European Organisation for Research and Treatment of Cancer Trial 10853. *J. Clin. Oncol.*, 19:2263–2271.

Boecker W, Parker S, Schulz-Wendtland R, Schnitt S, Buerger H (2006). Ductal carcinoma in-situ. In: *Preneoplasia of the breast. A new conceptual approach to proliferative breast disease*. Boecker W (ed.). Saunders Elsevier, Munich, pp. 407–466.

Bonnema J, van Geel AN, van Ooijen B, Mali SP, Tjiam SL, Henzen-Logmans SC, Schmitz PI, Wiggers T (1997). Ultrasound-guided aspiration biopsy for detection of nonpalpable axillary node metastases in breast cancer patients: new diagnostic method. *World J. Surg.*, 21:270–274.

Boughey JC, Moriarty JP, Degnim AC, Gregg MS, Egginton JS, Long KH (2010). Cost modeling of preoperative axillary ultrasound and fine-needle aspiration to guide surgery for invasive breast cancer. *Ann. Surg. Oncol.*, 17:953–958.

Boulos FI, Dupont WD, Simpson JF, Schuyler PA, Sanders ME, Freudenthal ME, Page DL (2008). Histologic associations and long-term cancer risk in columnar cell lesions of the breast: a retrospective cohort and a nested case-control study. *Cancer*, 113:2415–2421.

Brogi E, Oyama T, Koerner FC (2001). Atypical cystic lobules in patients with lobular neoplasia. *Int. J. Surg. Pathol.*, 9:201–206.

Castellano I, Macri L, Deambrogio C, Balmativola D, Bussone R, Ala A, Coluccia C, Sapino A (2012). Reliability of whole sentinel lymph node analysis by one-step nucleic acid amplification for intraoperative diagnosis of breast cancer metastases. *Ann. Surg.*, 255:334–342.

Chen L, Fan Y, Lang RG, Guo XJ, Sun YL, Cui LF, Liu FF, Wei J, Zhang XM, Fu L (2008). Breast carcinoma with micropapillary features: clinicopathologic study and long-term follow-up of 100 cases. *Int. J. Surg. Pathol.*, 16:155–163.

Chevallier B, Roche H, Olivier JP, Chollet P, Hurteloup P (1993). Inflammatory breast cancer. Pilot study of intensive induction chemotherapy (FEC-HD) results in a high histologic response rate. *Am. J. Clin. Oncol.*, 16:223–228.

Chivukula M, Bhargava R, Tseng G, Dabbs DJ (2009). Clinicopathologic implications of "flat epithelial atypia" in core needle biopsy specimens of the breast. *Am. J. Clin. Pathol.*, 131:802–808.

Choi YJ, Yun HR, Yoo KE, Kim JH, Nam SJ, Choi YL, Ko YH, Kim BT, Yang JH (2006). Intraoperative examination of sentinel lymph nodes by ultrarapid immunohistochemistry in breast cancer. *Jpn. J. Clin. Oncol.*, 36:489–493.

Ciatto S, Brancato B, Risso G, Ambrogetti D, Bulgaresi P, Maddau C, Turco P, Houssami N (2007). Accuracy of fine needle aspiration cytology (FNAC) of axillary lymph nodes as a triage test in breast cancer staging. *Breast Cancer Res. Treat.*, 103:85–91.

Colleoni M, Rotmensz N, Peruzzotti G, Maisonneuve P, Mazzarol G, Pruneri G, Luini A, Intra M, Veronesi P, Galimberti V, Torrisi R, Cardillo A, Goldhirsch A, Viale G (2005). Size of breast cancer metastases in axillary lymph nodes: clinical relevance of minimal lymph node involvement. *J. Clin. Oncol.*, 23:1379–1389.

Connolly JL, Schnitt SJ (1993). Clinical and histologic aspects of proliferative and non-proliferative benign breast disease. *J. Cell Biochem. Suppl.*, 17G:45–48.

Cserni G (1999). Pitfalls in frozen section interpretation: a retrospective study of palpable breast tumors. *Tumori*, 85:15–18.

Cserni G (2006). Histopathologic examination of the sentinel lymph nodes. *Breast J.*, 12 Suppl 2:S152–S156.

Cserni G (2009). Isolated tumour cells versus micrometastases and non-sentinel node involvement in breast cancer. *Eur. J. Surg. Oncol.*, 35:897–898.

Cserni G (2012). Intraoperative analysis of sentinel lymph nodes in breast cancer by one-step nucleic acid amplification. *J. Clin. Pathol.*, 65:193–199.

Cserni G, Amendoeira I, Apostolikas N, Bellocq JP, Bianchi S, Bussolati G, Boecker W, Borisch B, Connolly CE, Decker T, Dervan P, Drijkoningen M, Ellis IO, Elston CW, Eusebi V, Faverly D, Heikkila P, Holland R, Kerner H, Kulka J, Jacquemier J, Lacerda M, Martinez-Penuela J, De Miguel C, Peterse JL, Rank F, Regitnig P, Reiner A, Sapino A, Sigal-Zafrani B, Tanous AM, Thorstenson S, Zozaya E, Wells CA (2003). Pathological work-up of sentinel lymph nodes in breast cancer. Review of current data to be considered for the formulation of guidelines. *Eur. J. Cancer*, 39:1654–1667.

Cserni G, Amendoeira I, Bianchi S, Chmielik E, Degaetano J, Faverly D, Figueiredo P, Foschini MP, Grabau D, Jacquemier J, Kaya H, Kulka J, Lacerda M, Liepniece-Karele I, Penuela JM, Quinn C, Regitnig P, Reiner-Concin A, Sapino A, Van Diest PJ, Varga Z, Vezzosi V, Wesseling J, Zolota V, Zozaya E, Wells CA (2011). Distinction of isolated tumour cells and micrometastasis in lymph nodes of breast cancer patients according to the new Tumour Node Metastasis (TNM) definitions. *Eur. J. Cancer*, 47:887–894.

Cserni G, Bianchi S, Boecker W, Decker T, Lacerda M, Rank F, Wells CA (2005). Improving the reproducibility of diagnosing micrometastases and isolated tumor cells. *Cancer*, 103:358–367.

Cserni G, Bianchi S, Vezzosi V, van Diest P, van Deurzen C, Sejben I, Regitnig P, Asslaber M, Foschini MP, Sapino A, Castellano I, Callagy G, Arkoumani E, Kulka J, Wells CA (2008). Variations in sentinel node isolated tumour cells/micrometastasis and non-sentinel node involvement rates according to different interpretations of the TNM definitions. *Eur. J. Cancer*, 44:2185–2191.

Cserni G, Sapino A, Decker T (2006). Discriminating between micrometastases and isolated tumor cells in a regional and institutional setting. *Breast*, 15:347–354.

Dabbs DJ, Carter G, Fudge M, Peng Y, Swalsky P, Finkelstein S (2006). Molecular alterations in columnar cell lesions of the breast. *Mod. Pathol.*, 19:344–349.

Dadmanesh F, Peterse JL, Sapino A, Fonelli A, Eusebi V (2001). Lymphoepithelioma-like carcinoma of the breast: lack of evidence of Epstein-Barr virus infection. *Histopathology*, 38:54–61.

Damera A, Evans AJ, Cornford EJ, Wilson AR, Burrell HC, James JJ, Pinder SE, Ellis IO, Lee AH, Macmillan RD (2003). Diagnosis of axillary nodal metastases by ultrasound-guided core biopsy in primary operable breast cancer. *Br. J. Cancer*, 89:1310–1313.

Daniele L, Annaratone L, Allia E, Mariani S, Armando E, Bosco M, Macri L, Cassoni P, D'Armento G, Bussolati G, Cserni G, Sapino A (2009). Technical limits of comparison of step-sectioning,immunohistochemistry and RT-PCR on breast cancer sentinel nodes: a study on methacarn-fixed tissue. *J. Cell Mol. Med.*, 13:4042–4050.

Dash N, Chafin SH, Johnson RR, Contractor FM (1999). Usefulness of tissue marker clips in patients undergoing neoadjuvant chemotherapy for breast cancer. *AJR Am. J. Roentgenol.*, 173:911–917.

David N, Labbe-Devilliers C, Moreau D, Loussouarn D, Campion L (2006). [Diagnosis of flat epithelial atypia (FEA) after stereotactic vacuum-assisted biopsy (VAB) of the breast: What is the best management: systematic surgery for all or follow-up?] *J. Radiol.*, 87:1671-1677.

de Boer M, van Deurzen CH, van Dijck JA, Borm GF, Van Diest PJ, Adang EM, Nortier JW, Rutgers EJ, Seynaeve C, Menke-Pluymers MB, Bult P, Tjan-Heijnen VC (2009). Micrometastases or isolated tumor cells and the outcome of breast cancer. *N. Engl. J. Med.*, 361:653–663.

de Mascarel I, MacGrogan G, Debled M, Brouste V, Mauriac L (2008). Distinction between isolated tumor cells and micrometastases in breast cancer: is it reliable and useful? *Cancer*, 112:1672–1678.

de Mascarel I, MacGrogan G, Mathoulin-Pelissier S, Vincent-Salomon A, Soubeyran I, Picot V, Coindre JM, Mauriac L (2007). Epithelial atypia in biopsies performed for microcalcifications. Practical considerations about 2,833 serially sectioned surgical biopsies with a long follow-up. *Virchows Arch.*, 451:1–10.

Del Vecchio M, Foschini MP, Peterse JL, Eusebi V (2005). Lobular carcinoma of the breast with hybrid myoepithelial and secretory ("myosecretory") cell differentiation. *Am. J. Surg. Pathol.*, 29:1530–1536.

Dupont WD, Parl FF, Hartmann WH, Brinton LA, Winfield AC, Worrell JA, Schuyler PA, Plummer WD (1993). Breast cancer risk associated with proliferative breast disease and atypical hyperplasia. *Cancer*, 71:1258–1265.

Edge SB, Byrd DR, Compton CC, Fritz AG, Greene FL, Trotti A (eds.) (2009). *AJCC Cancer Staging Handbook, from the AJCC Cancer Staging Manual. Seventh edition*. Springer, New York, Heidelberg.

Eusebi V, Feudale E, Foschini MP, Micheli A, Conti A, Riva C, Di Palma S, Rilke F (1994). Long-term follow-up of in situ carcinoma of the breast. *Semin. Diagn. Pathol.*, 11:223–235.

Eusebi V, Foschini MP, Bussolati G, Rosen PP (1995). Myoblastomatoid (histiocytoid) carcinoma of the breast. A type of apocrine carcinoma. *Am. J. Surg. Pathol.*, 19:553–562.

Eusebi V, Magalhaes F, Azzopardi JG (1992). Pleomorphic lobular carcinoma of the breast: an aggressive tumor showing apocrine differentiation. *Hum. Pathol.*, 23:655–662.

Feeley L, Quinn CM (2008). Columnar cell lesions of the breast. *Histopathology*, 52:11–19.

Feldman LD, Hortobagyi GN, Buzdar AU, Ames FC, Blumenschein GR (1986). Pathological assessment of response to induction chemotherapy in breast cancer. *Cancer Res.*, 46:2578–2581.

Fessia L, Ghiringhello B, Arisio R, Botta G, Aimone V (1984). Accuracy of frozen section diagnosis in breast cancer detection. A review of 4436 biopsies and comparison with cytodiagnosis. *Pathol. Res.Pract.*, 179:61–66.

Foschini MP, Krausz T (2010). Salivary gland-type tumors of the breast: a spectrum of benign and malignant tumors including "triple negative carcinomas" of low malignant potential. *Semin. Diagn. Pathol.*, 27:77–90.

Fraser JL, Raza S, Chorny K, Connolly JL, Schnitt SJ (1998). Columnar alteration with prominent apical snouts and secretions: a spectrum of changes frequently present in breast biopsies performed for microcalcifications. *Am. J. Surg. Pathol.*, 22:1521–1527.

Fritzsche FR, Reineke T, Morawietz L, Kristiansen G, Dietel M, Fink D, Rageth C, Honegger C, Caduff R, Moch H, Varga Z (2010). Pathological processing techniques and final diagnosis of breast cancer sentinel lymph nodes. *Ann. Surg. Oncol.*, 17:2892–2898.

Giuliano AE, Hawes D, Ballman KV, Whitworth PW, Blumencranz PW, Reintgen DS, Morrow M, Leitch AM, Hunt KK, McCall LM, Abati A, Cote R (2011). Association of occult metastases in sentinel lymph nodes and bone marrow with survival among women with early-stage invasive breast cancer. *JAMA*, 306:385–393.

Goldhirsch A, Wood WC, Coates AS, Gelber RD, Thurlimann B, Senn HJ (2011). Strategies for subtypes--dealing with the diversity of breast cancer: highlights of the St. Gallen International Expert Consensus on the Primary Therapy of Early Breast Cancer 2011. *Ann. Oncol.*, 22:1736–1747.

Goldstein NS, O'Malley BA (1997). Cancerization of small ectatic ducts of the breast by ductal carcinoma in situ cells with apocrine snouts: a lesion associated with tubular carcinoma. *Am. J. Clin. Pathol.*, 107:561–566.

Guerra-Wallace MM, Christensen WN, White RL, Jr. (2004). A retrospective study of columnar alteration with prominent apical snouts and secretions and the association with cancer. *Am. J. Surg.*, 188:395–398.

Harris GC, Denley HE, Pinder SE, Lee AH, Ellis IO, Elston CW, Evans A (2003). Correlation of histologic prognostic factors in core biopsies and therapeutic excisions of invasive breast carcinoma, *Am. J. Surg. Pathol.*, 27:11–15.

Heywang-Kobrunner SH, Schreer I, Decker T, Bocker W (2003). Interdisciplinary consensus on the use and technique of vacuum-assisted stereotactic breast biopsy. *Eur. J. Radiol.*, 47:232–236.

Honkoop AH, Pinedo HM, De Jong JS, Verheul HM, Linn SC, Hoekman K, Wagstaff J, Van Diest PJ (1997). Effects of chemotherapy on pathologic and biologic characteristics of locally advanced breast cancer. *Am. J. Clin. Pathol.*, 107:211–218.

Ingegnoli A, d'Aloia C, Frattaruolo A, Pallavera L, Martella E, Crisi G, Zompatori M (2010). Flat epithelial atypia and atypical ductal hyperplasia: carcinoma underestimation rate. *Breast J.*, 16:55–59.

Jacquemier J, Padovani L, Rabayrol L, Lakhani SR, Penault-Llorca F, Denoux Y, Fiche M, Figueiro P, Maisongrosse V, Ledoussal V, Martinez Penuela J, Udvarhely N, El Makdissi G, Ginestier C, Geneix J, Charafe-Jauffret E, Xerri L, Eisinger F, Birnbaum D, Sobol H (2005). Typical medullary breast carcinomas have a basal/myoepithelial phenotype. *J. Pathol.*, 207:260–268.

Julian TB, Blumencranz P, Deck K, Whitworth P, Berry DA, Berry SM, Rosenberg A, Chagpar AB, Reintgen D, Beitsch P, Simmons R, Saha S, Mamounas EP, Giuliano A (2008). Novel intraoperative molecular test for sentinel lymph node metastases in patients with early-stage breast cancer. *J. Clin. Oncol.*, 26:3338–3345.

Kettritz U, Morack G, Decker T (2005). Stereotactic vacuum-assisted breast biopsies in 500 women with microcalcifications: radiological and pathological correlations. *Eur. J. Radiol.*, 55:270–276.

Kettritz U, Rotter K, Schreer I, Murauer M, Schulz-Wendtland R, Peter D, Heywang-Kobrunner SH (2004). Stereotactic vacuum-assisted breast biopsy in 2874 patients: a multicenter study. *Cancer*, 100:245–251.

Khaddage A, Berremila SA, Forest F, Clemenson A, Bouteille C, Seffert P, Peoc'h M (2011). Implementation of molecular intra-operative assessment of sentinel lymph node in breast cancer. *Anticancer Res.*, 31:585–590.

Krishnamurthy S, Sneige N, Bedi DG, Edieken BS, Fornage BD, Kuerer HM, Singletary SE, Hunt KK (2002). Role of ultrasound-guided fine-needle aspiration of indeterminate and suspicious axillary lymph nodes in the initial staging of breast carcinoma. *Cancer*, 95:982–988.

Kunju LP, Kleer CG (2007). Significance of flat epithelial atypia on mammotome core needle biopsy: Should it be excised? *Hum. Pathol.*, 38:35–41.

Kusama R, Fujimori M, Matsuyama I, Fu L, Ishii K, Hama Y, Asanuma K, Shingu K, Kobayashi S, Tsuchiya S (2000). Clinicopathological characteristics of atypical cystic duct (ACD) of the breast: assessment of ACD as a precancerous lesion. *Pathol. Int.*, 50:793–800.

Lae M, Freneaux P, Sastre-Garau X, Chouchane O, Sigal-Zafrani B, Vincent-Salomon A (2009). Secretory breast carcinomas with ETV6-NTRK3 fusion gene belong to the basal-like carcinoma spectrum. *Mod. Pathol.*, 22:291–298.

Laucirica R (2005). Intraoperative assessment of the breast: guidelines and potential pitfalls. *Arch. Pathol. Lab Med.*, 129:1565–1574.

Leikola JP, Toivonen TS, Krogerus LA, von Smitten KA, Leidenius MH (2005). Rapid immunohistochemistry enhances the intraoperative diagnosis of sentinel lymph node metastases in invasive lobular breast carcinoma. *Cancer*, 104:14–19.

Liberman L, Dershaw DD, Rosen PP, Morris EA, Abramson AF, Borgen PI (1998a). Percutaneous removal of malignant mammographic lesions at stereotactic vacuum-assisted biopsy. *Radiology*, 206:711–715.

Liberman L, Smolkin JH, Dershaw DD, Morris EA, Abramson AF, Rosen PP (1998b). Calcification retrieval at stereotactic, 11-gauge, directional, vacuum-assisted breast biopsy. *Radiology*, 208:251–260.

Londero V, Bazzocchi M, Del Frate C, Puglisi F, Di Loreto C, Francescutti G, Zuiani C (2004). Locally advanced breast cancer: comparison of mammography, sonography and MR imaging in evaluation of residual disease in women receiving neoadjuvant chemotherapy. *Eur. Radiol.*, 14:1371–1379.

London SJ, Connolly JL, Schnitt SJ, Colditz GA (1992). A prospective study of benign breast disease and the risk of breast cancer. *JAMA*, 267:941–944.

Ma L, Boyd NF (1992). Atypical hyperplasia and breast cancer risk: a critique. *Cancer Causes Control*, 3:517–525.

Mansel RE, Goyal A, Douglas-Jones A, Woods V, Goyal S, Monypenny I, Sweetland H, Newcombe RG, Jasani B (2009). Detection of breast cancer metastasis in sentinel lymph nodes using intra-operative real time GeneSearch BLN Assay in the operating room: results of the Cardiff study. *Breast Cancer Res. Treat.*, 115:595–600.

Marshall LM, Hunter DJ, Connolly JL, Schnitt SJ, Byrne C, London SJ, Colditz GA (1997). Risk of breast cancer associated with atypical hyperplasia of lobular and ductal types. *Cancer Epidemiol. Biomarkers Prev.*, 6:297–301.

Martel M, Barron-Rodriguez P, Tolgay O, I, Dotto J, Tavassoli FA (2007). Flat DIN 1 (flat epithelial atypia) on core needle biopsy: 63 cases identified retrospectively among 1,751 core biopsies performed over an 8-year period (1992-1999). *Virchows Arch.*, 451:883–891.

McLaren BK, Gobbi H, Schuyler PA, Olson SJ, Parl FF, Dupont WD, Page DL (2005). Immunohistochemical expression of estrogen receptor in enlarged lobular units with columnar alteration in benign breast biopsies: a nested case-control study. *Am. J. Surg. Pathol.*, 29:105–108.

Meyer JE, Smith DN, Di Piro PJ, Denison CM, Frenna TH, Harvey SC, Ko WD (1997). Stereotactic breast biopsy of clustered microcalcifications with a directional, vacuum-assisted device. *Radiology*, 204:575–576.

Moinfar F, Man YG, Bratthauer GL, Ratschek M, Tavassoli FA (2000). Genetic abnormalities in mammary ductal intraepithelial neoplasia-flat type ("clinging ductal carcinoma in situ"): a simulator of normal mammary epithelium. *Cancer*, 88:2072–2081.

Nahrig JM, Richter T, Kuhn W, Avril N, Flatau B, Kowolik J, Hofler H, Werner M (2003). Intraoperative examination of sentinel lymph nodes by ultrarapid immunohistochemistry. *Breast J.*, 9:277–281.

National Coordinating Group for Breast Screening Pathology (1995). *Pathology Reporting in Breast Cancer Screening. Second edition*. NHSBSP, Sheffield.

National Coordinating Group for Breast Screening Pathology (2005). *Pathology Reporting of Breast Disease. Second edition*. NHSBSP, Sheffield.

Newman LA, Pernick NL, Adsay V, Carolin KA, Philip PA, Sipierski S, Bouwman DL, Kosir MA, White M, Visscher DW (2003). Histopathologic evidence of tumor regression in the axillary lymph nodes of patients treated with preoperative chemotherapy correlates with breast cancer outcome. *Ann. Surg. Oncol.*, 10:734–739.

Niemann TH, Lucas JG, Marsh WL, Jr. (1996). To freeze or not to freeze. A comparison of methods for the handling of breast biopsies with no palpable abnormality. *Am. J. Clin. Pathol.*, 106:225–228.

Noguchi M, Minami M, Earashi M, Taniya T, Miyazaki I, I, Mizukami Y, Nonomura A, Nishijima H, Takanaka T, Kawashima H, Saito Y, Takashima C, Nakamura S, Michigishi T, Yokoyama K (1995). Pathologic assessment of surgical margins on frozen and permanent sections in breast conserving surgery. *Breast Cancer*, 2:27–33.

Nori J, Bazzocchi M, Boeri C, Vanzi E, Nori BF, Mangialavori G, Distante V, Masi A, Simoncini R, Londero V (2005). Role of axillary lymph node ultrasound and large core biopsy in the preoperative assessment of patients selected for sentinel node biopsy. *Radiol. Med.*, 109:330–344.

O'Malley FP, Mohsin SK, Badve S, Bose S, Collins LC, Ennis M, Kleer CG, Pinder SE, Schnitt SJ (2006). Interobserver reproducibility in the diagnosis of flat epithelial atypia of the breast. *Mod. Pathol.*, 19:172–179.

Ogston KN, Miller I, Payne S, Hutcheon WA, Sarkar TK, Heys DS (2001). A novel grading system to assess pathological response and predict survival in patients receiving primary chemotherapy for breast cancer. *Eur. J. Cancer*, 37:27.

Olivotto I, Levine M (2001). Clinical practice guidelines for the care and treatment of breast cancer: the management of ductal carcinoma in situ (summary of the 2001 update). *CMAJ.*, 165:912–913.

Otterbach F, Bankfalvi A, Bergner S, Decker T, Krech R, Boecker W (2000). Cytokeratin 5/6 immunohistochemistry assists the differential diagnosis of atypical proliferations of the breast. *Histopathology*, 37:232–240.

Oyama T, Maluf H, Koerner F (1999). Atypical cystic lobules: an early stage in the formation of low-grade ductal carcinoma in situ. *Virchows Arch.*, 435:413–421.

Page DL, Dupont WD (1992). Indicators of increased breast cancer risk in humans. *J.Cell Biochem. Suppl.*, 16G:175–182.

Page DL, Dupont WD (1993). Anatomic indicators (histologic and cytologic) of increased breast cancer risk. *Breast Cancer Res. Treat.*, 28:157–166.

Page DL, Dupont WD, Rogers LW, Rados MS (1985). Atypical hyperplastic lesions of the female breast. A long-term follow-up study. *Cancer*, 55:2698–2708.

Page DL, Jensen RA (1994). Evaluation and management of high risk and premalignant lesions of the breast, *World J. Surg.*, 18:32–38.

Page DL, Rogers LW (1992). Combined histologic and cytologic criteria for the diagnosis of mammary atypical ductal hyperplasia. *Hum. Pathol.*, 23:1095–1097.

Pinder SE, Provenzano E, Earl H, Ellis IO (2007). Laboratory handling and histology reporting of breast specimens from patients who have received neoadjuvant chemotherapy. *Histopathology*, 50:409–417.

Rakha EA, Ellis IO (2007). An overview of assessment of prognostic and predictive factors in breast cancer needle core biopsy specimens. *J. Clin. Pathol.*, 60:1300–1306.

Rao R, Lilley L, Andrews V, Radford L, Ulissey M (2009). Axillary staging by percutaneous biopsy: sensitivity of fine-needle aspiration versus core needle biopsy. *Ann. Surg. Oncol.*, 16:1170–1175.

Reis-Filho JS, Lakhani SR (2008). Breast cancer special types: why bother? *J. Pathol.*, 216:394–398.

Rosen PP (1999). Columnar cell hyperplasia is associated with lobular carcinoma in situ and tubular carcinoma. *Am. J. Surg. Pathol.*, 23:1561.

Rosen PP (2009). *Rosen's Breast Pathology. Third edition*. Lippincott Williams & Wilkins, Philadelphia.

Salem C, Sakr R, Chopier J, Marsault C, Uzan S, Darai E (2009). Accuracy of stereotactic vacuum-assisted breast biopsy with a 10-gauge hand-held system. *Breast*, 18:178–182.

Sapino A, Cassoni P, Zanon E, Fraire F, Croce S, Coluccia C, Donadio M, Bussolati G (2003). Ultrasonographically-guided fine-needle aspiration of axillary lymph nodes: role in breast cancer management. *Br. J. Cancer*, 88:702–706.

Sapino A, Righi L, Cassoni P, Papotti M, Pietribiasi F, Bussolati G (2000). Expression of the neuro-endocrine phenotype in carcinomas of the breast. *Semin. Diagn. Pathol.*, 17:127–137.

Sataloff DM, Mason BA, Prestipino AJ, Seinige UL, Lieber CP, Baloch Z (1995). Pathologic response to induction chemotherapy in locally advanced carcinoma of the breast: a determinant of outcome. *J. Am. Coll. Surg.*, 180:297–306.

Schnitt SJ (1998). Microinvasive carcinoma of the breast: a diagnosis in search of a definition. *Adv. Anat. Pathol.*, 5:367–372.

Schnitt SJ, Vincent-Salomon A (2003). Columnar cell lesions of the breast. *Adv. Anat. Pathol.*, 10:113–124.

Senetta R, Campanino PP, Mariscotti G, Garberoglio S, Daniele L, Pennecchi F, Macri L, Bosco M, Gandini G, Sapino A (2009). Columnar cell lesions associated with breast calcifications on vacuum-assisted core biopsies: clinical, radiographic, and histological correlations. *Mod. Pathol.*, 22:762–769.

Shaaban AM, Sloane JP, West CR, Moore FR, Jarvis C, Williams EM, Foster CS (2002). Histopathologic types of benign breast lesions and the risk of breast cancer: case-control study. *Am. J. Surg. Pathol.*, 26:421–430.

Simpson PT, Gale T, Reis-Filho JS, Jones C, Parry S, Sloane JP, Hanby A, Pinder SE, Lee AH, Humphreys S, Ellis IO, Lakhani SR (2005). Columnar cell lesions of the breast: the missing link in breast cancer progression? A morphological and molecular analysis. *Am. J. Surg. Pathol.*, 29:734–746.

Sloane JP (1991). Pathology reporting in breast cancer screening. Royal College of Pathologists Working Group. *J. Clin. Pathol.*, 44:710–725.

Sobin LH, Gospodarowicz MK, Wittekind C (eds.) (2009). *TNM Classification of Malignant Tumours. Seventh edition*, Wiley-Blackwell.

Sobin LH, Wittekind C (eds.) (1997). Breast Tumours. In: *TNM Classification of Malignant Tumours. Fifth edition*. John Wiley & Sons, New York.

Sorlie T, Perou CM, Tibshirani R, Aas T, Geisler S, Johnsen H, Hastie T, Eisen MB, van de Rijn M, Jeffrey SS, Thorsen T, Quist H, Matese JC, Brown PO, Botstein D, Eystein Lonning P, Borresen-Dale AL (2001). Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. *Proc. Natl. Acad. Sci. U.S.A*, 98:10869–10874.

Symmans WF, Peintinger F, Hatzis C, Rajan R, Kuerer H, Valero V, Assad L, Poniecka A, Hennessy B, Green M, Buzdar AU, Singletary SE, Hortobagyi GN, Pusztai L (2007). Measurement of residual breast cancer burden to predict survival after neoadjuvant chemotherapy. *J. Clin. Oncol.*, 25:4414–4422.

Tan PH, Ho BC, Selvarajan S, Yap WM, Hanby A (2005). Pathological diagnosis of columnar cell lesions of the breast: are there issues of reproducibility? *J. Clin. Pathol.*, 58:705–709.

Tavassoli FA (1992). Intraduct hyperplasias, ordinary and atypical. In: *Pathology of the Breast*. Appleton and Lange, Connecticut, pp. 155–191.

Tavassoli FA (1998). Ductal carcinoma in situ: introduction of the concept of ductal intraepithelial neoplasia. *Mod. Pathol.*, 11:140–154.

Tavassoli FA (1999). *Pathology of the Breast. Second edition*. Appleton & Lange, Stanford.

Tavassoli FA, Devilee P (2003). *WHO Classification of Tumours. Pathology and Genetics of Tumours of the Breast and Female Genital Organs*. IARC Press, Lyon.

Tavassoli FA, Eusebi V (2009). Microinvasive carcinoma. In: *Tumors of the Mammary Gland*. AFIP Atlas of Tumor Pathology: Series 4. American Registry of Pathology, Silver Spring, Maryland.

Tavassoli FA, Norris HJ (1990). A comparison of the results of long-term follow-up for atypical intraductal hyperplasia and intraductal hyperplasia of the breast. *Cancer*, 65:518–529.

Tot T (2005). Correlating the ground truth of mammographic histology with the success or failure of imaging. *Technol. Cancer Res.Treat.*, 4:23–28.

Tot T, Gere M (2008). Radiological-pathological correlation in diagnosing breast carcinoma: the role of pathology in the multimodality era. *Pathol. Oncol. Res.*, 14:173–178.

Tremblay G, Deschenes J, Alpert L, Quenneville LA (2005). Overexpression of estrogen receptors in columnar cell change and in unfolding breast lobules. *Breast J.*, 11:326–332.

Turner RR, Weaver DL, Cserni G, Lester SC, Hirsch K, Elashoff DA, Fitzgibbons PL, Viale G, Mazzarol G, Ibarra JA, Schnitt SJ, Giuliano AE (2008). Nodal stage classification for breast carcinoma: improving interobserver reproducibility through standardized histologic criteria and image-based training. *J. Clin. Oncol.*, 26:258–263.

van Deurzen CH, Cserni G, Bianchi S, Vezzosi V, Arisio R, Wesseling J, Asslaber M, Foschini MP, Sapino A, Castellano I, Callagy G, Faverly D, Martin-Martinez MD, Quinn C, Amendoeira I, Kulka J, Reiner-Concin A, Cordoba A, Seldenrijk CA, Van Diest PJ (2010). Nodal-stage classification in invasive lobular breast carcinoma: influence of different interpretations of the pTNM classification. *J. Clin. Oncol.*, 28:999–1004.

Van Diest PJ, Torrenga H, Borgstein PJ, Pijpers R, Bleichrodt RP, Rahusen FD, Meijer S (1999). Reliability of intraoperative frozen section and imprint cytological investigation of sentinel lymph nodes in breast cancer. *Histopathology*, 35:14–18.

Van Poznak C, Cross SS, Saggese M, Hudis C, Panageas KS, Norton L, Coleman RE, Holen I (2006). Expression of osteoprotegerin (OPG), TNF related apoptosis inducing ligand (TRAIL), and receptor activator of nuclear factor kappaB ligand (RANKL) in human breast tumours. *J. Clin. Pathol.*, 59:56–63.

van Rijk MC, Deurloo EE, Nieweg OE, Gilhuijs KG, Peterse JL, Rutgers EJ, Kroger R, Kroon BB (2006). Ultrasonography and fine-needle aspiration cytology can spare breast cancer patients unnecessary sentinel lymph node biopsy. *Ann. Surg. Oncol.*, 13:31–35.

Varga Z, Rageth C, Saurenmann E, Honegger C, von Orelli S, Fehr M, Fink D, Seifert B, Moch H, Caduff R (2008). Use of intraoperative stereomicroscopy for preventing loss of metastases during frozen sectioning of sentinel lymph nodes in breast cancer. *Histopathology*, 52:597–604.

Verkooijen HM (2002). Diagnostic accuracy of stereotactic large-core needle biopsy for nonpalpable breast disease: results of a multicenter prospective study with 95% surgical confirmation. *Int. J. Cancer*, 99:853–859.

Verschuur-Maes AH, van Deurzen CH, Monninkhof EM, Van Diest PJ (2012). Columnar cell lesions on breast needle biopsies: is surgical excision necessary? A systematic review. *Ann. Surg.*, 255:259–265.

Verschuur-Maes AH, Van Diest PJ (2011). The mucinous variant of columnar cell lesions. *Histopathology*, 58:847–853.

Verschuur-Maes AH, van Gils CH, van den Bosch MA, De Bruin PC, Van Diest PJ (2011a). Digital mammography: more microcalcifications, more columnar cell lesions without atypia. *Mod. Pathol.*, 24:1191–1197.

Verschuur-Maes AH, Witkamp AJ, De Bruin PC, van der Wall E, Van Diest PJ (2011b). Progression risk of columnar cell lesions of the breast diagnosed in core needle biopsies. *Int. J. Cancer*, 129:2674–2680.

Viale G, Dell'Orto P, Biasi MO, Stufano V, De Brito Lima LN, Paganelli G, Maisonneuve P, Vargo JM, Green G, Cao W, Swijter A, Mazzarol G (2008). Comparative evaluation of an extensive histopathologic examination and a real-time reverse-transcription-polymerase chain reaction assay for mammaglobin and cytokeratin 19 on axillary sentinel lymph nodes of breast carcinoma patients. *Ann. Surg.*, 247:136–142.

Vincent-Salomon A (2003). [Columnar lesions: a frequent diagnosis in breast pathology!] *Ann. Pathol.*, 23:593–596.

Visser M, Jiwa M, Horstman A, Brink AA, Pol RP, van Diest P, Snijders PJ, Meijer CJ (2008). Intra-operative rapid diagnostic method based on CK19 mRNA expression for the detection of lymph node metastases in breast cancer. *Int. J. Cancer*, 122:2562–2567.

Walker R, Research Subgroup of the NHSBSP Pathology Co-ordinating Group (2010). *Columnar Cell Lesions: NHSBSP Co-ordinating Group Recommendations for Terminology and Diagnostic Criteria*. https://telepathology.ucl.ac.uk/nccbp/docs/guidelines/ccl_final.pdf

Weidner N, Semple JP (1992). Pleomorphic variant of invasive lobular carcinoma of the breast. *Hum. Pathol.*, 23:1167–1171.

Weigelt B, Horlings HM, Kreike B, Hayes MM, Hauptmann M, Wessels LF, de Jong D, van de Vijver MJ, Van't Veer LJ, Peterse JL (2008). Refinement of breast cancer classification by molecular characterization of histological special types. *J. Pathol.*, 216:141–150.

Weigelt B, Kreike B, Reis-Filho JS (2009). Metaplastic breast carcinomas are basal-like breast cancers: a genomic profiling analysis. *Breast Cancer Res.Treat.*, 117:273–280.

Weigelt B, Reis-Filho JS (2009). Histological and molecular types of breast cancer: is there a unifying taxonomy? *Nat. Rev. Clin. Oncol.*, 6:718–730.

Wellings SR, Jensen HM, Marcum RG (1975). An atlas of subgross pathology of the human breast with special reference to possible precancerous lesions. *J. Natl. Cancer Inst.*, 55:231–273.

Wells CA, Amendoeira I, Apostolikas N, Bellocq JP, Bianchi S, Boecker W, Borisch B, Bussolati G, Connolly CE, Cserni G, Decker T, Dervan P, Drijkoningen M, Ellis IO, Elston CW, Eusebi V, Faverly DR, Heikkilä P, Holland R, Kerner H, Kulka J, Jacquemier J, Lacerda M, Martinez-Penuela J, de Miguel C, Nordgren H, Peterse JL, Rank F, Regitnig P, Reiner A, Sapino A, Sigal-Zafrani B, Tanous AM, Thorstenson S, Zozaya E (2006a). Chapter 6: Quality assurance guidelines for pathology. In: *European guidelines for quality assurance in breast cancer screening and diagnosis. Fourth edition*. Perry N, Broeders M, de Wolf C, Törnberg S, Holland R, von Karsa L (eds.). European Commission, Office for Official Publications of the European Union, Luxembourg, pp. 219–311.

Wells CA, Amendoeira I, Apostolikas N, Bellocq JP, Bianchi S, Boecker W, Borisch B, Bussolati G, Connolly CE, Cserni G, Decker T, Dervan P, Drijkoningen M, Ellis IO, Elston CW, Eusebi V, Faverly DR,

Heikkilä P, Holland R, Kerner H, Kulka J, Jacquemier J, Lacerda M, Martinez-Penuela J, de Miguel C, Nordgren H, Peterse JL, Rank F, Regitnig P, Reiner A, Sapino A, Sigal-Zafrani B, Tanous AM, Thorstenson S, Zozaya E (2006b). Chapter 6b: Open biopsy and resection specimens. In: *European guidelines for quality assurance in breast cancer screening and diagnosis. Fourth edition*. Perry N, Broeders M, de Wolf C, Törnberg S, Holland R, von Karsa L (eds.). European Commission, Office for Official Publications of the European Union, Luxembourg, pp. 257–311.

Werling RW, Hwang H, Yaziji H, Gown AM (2003). Immunohistochemical distinction of invasive from noninvasive breast lesions: a comparative study of p63 versus calponin and smooth muscle myosin heavy chain. *Am. J. Surg. Pathol.*, 27:82–90.

Wetterskog D, Lopez-Garcia MA, Lambros MB, A'Hern R, Geyer FC, Milanezi F, Cabral MC, Natrajan R, Gauthier A, Shiu KK, Orr N, Shousha S, Gatalica Z, Mackay A, Palacios J, Reis-Filho JS, Weigelt B (2012). Adenoid cystic carcinomas constitute a genomically distinct subgroup of triple-negative and basal-like breast cancers. *J. Pathol.*, 226:84–96.

Yaziji H, Gown AM, Sneige N (2000). Detection of stromal invasion in breast cancer: the myoepithelial markers. *Adv. Anat. Pathol.*, 7:100–109.

Zekioglu O, Erhan Y, Ciris M, Bayramoglu H, Ozdemir N (2004). Invasive micropapillary carcinoma of the breast: high incidence of lymph node metastasis with extranodal extension and its immunohistochemical profile compared with invasive ductal carcinoma. *Histopathology*, 44:18–23.

Annex 1

Successful implementation of population-based cancer screening programmes

Annex 1a

Stockholm statement on successful implementation of population-based cancer screening programmes

Authors
L. von Karsa
A. Anttila
M. Primic Žakelj
C. de Wolf
M. Bielska-Lasota
S. Törnberg
N. Segnan

Authors
L. von Karsa, IARC
A. Anttila, Finland
M. Primic Žakelj, Slovenia
C. de Wolf, Switzerland
M. Bielska-Lasota, Poland
S. Törnberg, Sweden
N. Segnan, Italy

Declarations of interest
Interests of C. de Wolf and N. Segnan are reported on pages V and VII respectively.

Disclaimer
The views expressed in this document are those of the authors and do not necessarily reflect the official position of the European Commission. Neither the European Commission nor any other organization or any individual may be held responsible for any use that may be made of the information contained herein.

Acknowledgements
The present statement is based on the results of a workshop held by the European Science Advisory Network for Health (EuSANH) on implementation and quality assurance of cancer screening programmes in Stockholm on 7–9 February 2011. A more detailed report has been published elsewhere and is reproduced in the second part of this annex with permission of Elsevier.

Members of the workshop expert committee were: Ahti Anttila, Finland; Magdalena Bielska-Lasota, Poland; Thomas Davidson, Sweden; Johannes JM van Delden, The Netherlands; Lawrence von Karsa, International Agency for Research on Cancer; Elsebeth Lynge, Denmark; Sue Moss, United Kingdom; Maja Primic Žakelj, Slovenia; Leo van Rossum, The Netherlands; Nereo Segnan, Italy; Sven Törnberg, Sweden; Chris de Wolf, Switzerland.

The workshop was also attended by the following observers: Euzebiusz Dziwinski, European Cancer Patient Coalition; Karl Freese, European Commission; Gunta Lazdane, World Health Organization; and by the following EuSANH representatives: Susanne V Allander, Sweden; Dorine Coenen, The Netherlands; Louise Gunning, The Netherlands; Monica Hultcrantz, Sweden; Måns Rosén, Sweden.

Comments were received from Jose Manuel Baena, Luc Bleyen, Mireille Broeders, Luis Bujanda, José Expósito, Roger Pla Farnós, Xandra Gravenstein, Rosella Hermans, Roland Holland, Harry de Koning, Chris Meijer, Nieves Ascunce Elizaga, Xavier Castells Oliveres, Julietta Patnick, Marina Pollán, Dolores Salas, Hélène Sancho-Garnier, and Jaroslav Volf.

The research leading to these results has received funding from the European Union's Seventh Framework Programme [FP7/2007-2013] under agreement number 229716 for the EuSANH project Improving Science Advice for Health in Europe (EuSANH-ISA) (www.eusanh.eu).

Preparation of this report has also been supported by the European Union Public Health Programme (Project no. 2006322, European Cooperation on Development and Implementation of Cancer Screening and Prevention Guidelines [ECCG]).

Please cite this publication as follows
von Karsa L, Anttila A, Primic Žakelj M, de Wolf C, Bielska-Lasota M, Törnberg S, Segnan N (2013). Stockholm statement on successful implementation of population-based cancer screening programmes. Annex 1a. In: *European guidelines for quality assurance in breast cancer screening and diagnosis. Fourth edition, Supplements*. Perry N, Broeders M, de Wolf C, Törnberg S, Holland R, von Karsa L (eds.). European Commission, Office for Official Publications of the European Union, Luxembourg, pp. 123–128.

Corresponding author
L. von Karsa
Quality Assurance Group
Early Detection and Prevention Section
International Agency for Research on Cancer
150, cours Albert Thomas
69372 Lyon Cedex 08
France
Karsal@iarc.fr

Stockholm statement on successful implementation of population-based cancer screening programmes

A multidisciplinary group of scientists and professionals experienced in implementation and quality assurance of cancer screening programmes and in development of scientific advice on health policy met at a pan-European workshop in Stockholm from 7 to 9 February 2011. The workshop was organized by the European Science Advisory Network for Health (EuSANH, www.EuSANH.eu). The list of experts attending the workshop is provided in this annex. The experts reviewed the available evidence on implementation and quality assurance of cancer screening programmes with a focus on organization and reduction of barriers to participation. After comprehensive discussion, the experts reached the following, mutually agreed conclusions:[1]

Any policy decision in Europe to implement a cancer screening programme should take into account European Union (EU) recommendations and guidelines based on the available evidence and the experience in Europe in implementing population-based cancer screening programmes. Key references in this regard are the Council Recommendation of 2 December 2003 on cancer screening of the Council of the European Union [1], the European guidelines for quality assurance in breast, cervical and colorectal cancer screening [2–4] and recent reports dealing with the implementation of cancer screening programmes in the EU [5–7]. These references recognize that societal values in addition to professional, technical and scientific standards are of prime importance in any decision to implement cancer screening programmes. Furthermore, there is no doubt that the population-based approach to programme implementation as recommended by the Council of the EU and the European guidelines is more equitable, more effective and more cost-effective than an opportunistic approach. The latter usually leads to overuse of health resources by a portion of the target population with lower cancer risk, and underuse by less advantaged groups with higher cancer risk.

The experience in Europe shows that successful implementation of population-based cancer screening programmes requires long-term political commitment, a comprehensive quality management programme and sustainable resources. In a fully established programme, the proportion of the expenditure devoted to quality assurance should be no less than 10–20%, depending on the scale of the programme. In the initial years, this proportion may be substantially higher due to the low volume of screening examinations compared with the situation after complete rollout of a nationwide programme. This investment is cost-effective and will save lives.

Once the political decision has been taken to establish a population-based cancer screening programme, a competent programme coordinator should receive the mandate to manage the entire process of programme implementation, beginning with a planning phase and followed by feasibility testing, piloting and, depending on the interim results, subsequent gradual rollout of a programme fulfilling the principles and standards recommended in the Council Recommendation [1] and the European guidelines [2–4] and relevant national standards and guidelines. The coordinator should be provided with sufficient organizational and financial resources to effectively manage the screening programme and take further decisions as necessary. These decisions should enable the coordinator

[1] The present statement summarizes key results of the workshop. A more detailed report has been published elsewhere [8] and is reproduced with permission of Elsevier in the second part of this annex.

and the coordination team to establish the screening programme in the respective health services context, taking into account the need for the professional and organizational management to control the quality of the entire screening process, including informing and inviting the target population, performance of the screening test, diagnosis, therapy and subsequent care. The existing expertise in Europe in implementation of population-based cancer screening programmes should be available for exchange of information and experience, such as through the European cancer screening networks and the European guidelines development activities coordinated by the International Agency for Research on Cancer (IARC, www.iarc.fr), and related initiatives such as the European Partnership for Action Against Cancer (EPAAC, www.epaac.eu).

Additional tools, including computerized information systems and accessible registries, are necessary for the management of effective and efficient screening services (e.g. for call and re-call systems and fail-safe procedures in follow-up of participants with abnormal test results). They are also needed to monitor and evaluate the performance and the outcome of the screening programme, e.g. through linkage of individual data on cancer occurrence and morbidity, screening history, diagnosis and treatment.

Furthermore, key performance and quality indicators of the screening process must be recorded and monitored and the results must be analysed and used for quality management processes. Monitoring and evaluation reports must be published regularly to inform the public and decision makers and to permit timely modification of programme policy, if necessary. The experience of EuSANH in developing advice for health policy making, taking into account not only scientific and professional but also societal aspects, could play an important role in this regard in the future.

References

1. Council of the European Union (2003). Council Recommendation of 2 December 2003 on cancer screening (2003/878/EC). *Off. J. Eur. Union.*, L 327:34–38. http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2003:327:0034:0038:EN:PDF
2. Segnan N, Patnick J, von Karsa L (eds.) (2010). *European guidelines for quality assurance in colorectal cancer screening and diagnosis. First edition*. European Commission, Office for Official Publications of the European Union, Luxembourg. http://bookshop.europa.eu/is-bin/INTERSHOP.enfinity/WFS/EU-Bookshop-Site/en_GB/-/EUR/ViewPublication-Start?PublicationKey=ND3210390
3. Arbyn M, Anttila A, Jordan J, Schenck U, Ronco G, Segnan N, Wiener H, Herbert A, Daniel J, von Karsa L (eds.) (2008). *European guidelines for quality assurance in cervical cancer screening. Second edition*. European Commission, Office for Official Publications of the European Communities, Luxembourg. http://bookshop.europa.eu/is-bin/INTERSHOP.enfinity/WFS/EU-Bookshop-Site/en_GB/-/EUR/ViewPublication-Start?PublicationKey=ND7007117
4. Perry N, Broeders M, de Wolf C, Törnberg S, Holland R, von Karsa L, Puthaar E (eds.) (2006). *European guidelines for quality assurance in breast cancer screening and diagnosis. Fourth edition*. European Commission, Office for Official Publications of the European Communities, Luxembourg. http://bookshop.europa.eu/is-bin/INTERSHOP.enfinity/WFS/EU-Bookshop-Site/en_GB/-/EUR/ViewPublication-Start?PublicationKey=ND7306954
5. Anttila A, Ronco G, Working Group on the Registration and Monitoring of Cervical Cancer Screening Programmes in the European Union; within the European Network for Information on Cancer (EUNICE) (2009). Description of the national situation of cervical cancer screening in the member states of the European Union. *Eur. J. Cancer*, 45:2685–2708. http://www.ejcancer.info/article/S0959-8049(09)00576-0/abstract

6. von Karsa L, Anttila A, Ronco G, Ponti A, Malila N, Arbyn M, Segnan N, Castillo-Beltran M, Boniol M, Ferlay J, Hery C, Sauvaget C, Voti L, Autier P (2008). *Cancer screening in the European Union, Report on the implementation of the Council Recommendation on cancer screening – First Report*. European Commission, Office for Official Publications of the European Communities, Luxembourg. http://ec.europa.eu/health/ph_determinants/genetics/documents/cancer_screening.pdf

7. Commission of the European Communities (2008). Report from the Commission to the Council, the European Parliament, the European Economic and Social Committee and the Committee of the Regions – Implementation of the Council Recommendation of 2 December 2003 on cancer screening (2003/878/EC) Brussels, Report no. COM(2008) 882 final. http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=COM:2008:0882:FIN:EN:PDF

8. Lynge E, Törnberg S, von Karsa L, Segnan N, van Delden JJM (2012). Determinants of successful implementation of population-based cancer screening programmes. *Eur. J. Cancer*, 48:743–748. http://www.sciencedirect.com/science/article/pii/S0959804911004898

Annex 1b

Determinants of successful implementation of population-based cancer screening programmes

Authors
E. Lynge
S. Törnberg
L. von Karsa
N. Segnan
J.J.M. van Delden

Authors
E. Lynge, Denmark
S. Törnberg, Sweden
L. von Karsa, IARC
N. Segnan, Italy
J.J.M. van Delden, The Netherlands

Declarations of interest
Interests of E. Lynge and N. Segnan are reported on page VII.

Disclaimer
The views expressed in this document are those of the authors and do not necessarily reflect the official position of the European Commission. Neither the European Commission nor any other organization or any individual may be held responsible for any use that may be made of the information contained herein.

Elsebeth Lynge is undertaking a comparative study of new-generation HPV tests, involving collaboration with Roche Diagnostics A/S, Genomica S.A.U., Qiagen Gaithersburg Ltd., and Gen-Probe Inc., and has served as unpaid adviser for Gen-Probe and Norchip. Nereo Segnan participated at an advisory board meeting for Colorectal Cancer Screening in January 2011, as a paid expert, on colorectal cancer blood screening assay, organized by Roche Diagnostics Ltd.

Acknowledgements
This part of Annex 1 (*Determinants of successful implementation of population-based cancer screening programmes*) has been previously published in the European Journal of Cancer (volume 48, pp 743-748) and is reproduced here with permission of Elsevier. The content of the article is based on the results of a workshop held by the European Science Advisory Network for Health (EuSANH) on implementation and quality assurance of cancer screening programmes in Stockholm on 7-9 February 2011. A brief statement summarizing the results of the workshop and a detailed list of the participants and other contributors is published in the first part of this annex.

The research leading to these results has received funding from the European Union's Seventh Framework Programme [FP7/2007-2013] under agreement number 229716 for the EuSANH project Improving Science Advice for Health in Europe (EuSANH-ISA) (www.eusanh.eu).

Preparation of this document has also been supported by the European Union Public Health Programme (Project no. 2006322, European Cooperation on Development and Implementation of Cancer Screening and Prevention Guidelines [ECCG]).

Please cite this publication as follows
Lynge E, Törnberg S, von Karsa L, Segnan N, van Delden JJM (2012). Determinants of successful implementation of population-based cancer screening programmes. *Eur. J. Cancer*, 48:743–748.

Corresponding author
E. Lynge
Department of Public Health
University of Copenhagen
Øster Farimagsgade 5
DK-1014 Copenhagen K
Denmark
Tel: + 45 35 32 76 35, Fax: + 45 35 32 73 83
elsebeth@pubhealth.ku.dk

EUROPEAN JOURNAL OF CANCER 48 (2012) 743–748

available at www.sciencedirect.com

ScienceDirect

journal homepage: www.ejconline.com

Determinants of successful implementation of population-based cancer screening programmes

Elsebeth Lynge [a,*], *Sven Törnberg* [b], *Lawrence von Karsa* [c], *Nereo Segnan* [d], *Johannes J.M. van Delden* [e]

[a] *Department of Public Health, University of Copenhagen, Denmark*
[b] *Department of Cancer Screening, Karolinska University Hospital, Stockholm, Sweden*
[c] *Lawrence von Karsa, International Agency for Research on Cancer, Lyon, France*
[d] *Department of Cancer Screening and Unit of Cancer Epidemiology, CPO Piemonte and S. Giovanni University Hospital, Torino, Italy*
[e] *Julius Center for Health Sciences and Primary Care, Medical Humanities, University Medical Center, Utrecht, The Netherlands*

ARTICLE INFO

Article history:
Available online 23 July 2011

Keywords:
Cancer
Screening
Population-based programme

ABSTRACT

To facilitate the future implementation of population-based cancer screening programmes in European countries, we summarised the experience gained from existing programmes across Europe. We listed points that citizens, advocacy groups, politicians, health planners, and health professionals should consider when planning, implementing and running population based cancer screening programmes. The list is general and is applicable to breast, cervical and colorectal cancer screening. It is based on evidence presented in the three European Union guidelines on quality assurance in cancer screening and diagnosis, supplemented with other literature and expert experience presented at a European Science Advisory Network for Health workshop. The implementation of a cancer screening programme should be divided into the following seven phases: (1) before planning, (2) planning, (3) feasibility testing, (4) piloting or trial implementation, (5) scaling up from pilot to service, (6) running of full-scale programme, and (7) sustainability. For each phase, a substantial number of specified conditions have to be met. Successful implementation of a cancer screening programme requires societal acceptance and local ownership along with the best evidence-based practise and verification of adequate performance in each phase of implementation.

1. Introduction

Screening and early detection of asymptomatic cases constitute important elements in the control of breast cancer, cervical cancer, and colorectal cancer. In accordance with the 2003 recommendation of the Council of the European Union (EU),[1] many European countries have implemented screening programmes for some or all of these three cancer sites. Additional programmes are currently being planned or established.

In principle, a good screening test should be simple and easy to use. However, the full preventive potential of screening tests will only be realised within a good screening programme, and such a programme is a complex organisation. To facilitate the future implementation of population-based screening programmes in European countries, it is therefore valuable to summarise the experiences gained from existing programmes across Europe. With this aim in mind, an expert group convened in Stockholm on 7–9 February 2011 under the

* *Corresponding author:* Address: Department of Public Health, University of Copenhagen, Øster Farimagsgade 5, DK-1014 Copenhagen K, Denmark. Tel.: +45 35 32 76 35; fax: +45 35 32 73 83.
E-mail address: elsebeth@sund.ku.dk (E. Lynge).
0959-8049/$ - see front matter
doi:10.1016/j.ejca.2011.06.051

auspices of the European Science Advisory Network for Health, and this paper reports on the outcome of this work.

The aim of cancer screening is to reduce the mortality from the disease screened for. When pre-cancer lesions are detected and treated, the incidence of the disease should also be reduced. For cancer screening to achieve its aim, a number of conditions need to be fulfilled. This report sets out the key points that all citizens, advocacy groups, politicians, health planners, and health professionals should consider when planning, implementing and running population-based cancer screening programmes. The following list is based on evidence from the scientific literature and expert experience. A major part of the evidence is reported in the EU guidelines on quality assurance of cancer screening and diagnosis.[2–4]

The list follows the steps in programme implementation; from the societal deliberations about new cancer-control measures to the sustainability of a well-implemented screening programme. Successful implementation of cancer screening encompasses many steps, here these are grouped into: before planning, planning, feasibility, piloting, roll-out (scaling-up from pilot to service and running a full-scale programme) and sustainability. The points are general and are applicable to breast, cervical, and colorectal cancer screening. In a few cases, points have been repeated if they are of prime importance in more that one phase.

2. Determinants

2.1. *Before planning*

The starting point must always be to promote professional and public understanding of the purpose, the benefits and the risks of screening. This implies that one has to organise a societal debate. The next step is to review existing evidence-based recommendations and guidelines taking into account the local setting. During the whole process international exchange of experience is encouraged. Key points at this stage are:

- Review of scientific literature.
- Collection of information on disease incidence, stage distribution, and survival.
- Collection of information on availability and quality of cure offered.
- Understanding the potential role of screening in cancer control.
- Assessment of evidence for adding screening to existing cancer control measures.
- Collection of experience from other countries.
- Building up professional and public understanding of the benefits and risks of screening.
- Political will, commitment, at all relevant levels (EU, Member States and regional).
- Decision on political responsibility for the process.
- Review of existing guidelines.
- Availability of treatments and facilities (both competence and resources).
- Assessment of facilitating factors/barriers for implementation of organised screening.
- Economic impact and cost-effectiveness of the programme.
- Formal decision and allocation of budget.
- Organisation of continuous societal debate and input.

2.2. *Comprehensive planning: feasibility of screening models, professional performance, organisation, financing, and quality assurance (QA)*

After the political decision has been taken to start the process of establishing a population-based cancer screening programme, the first step is comprehensive planning. This should cover the entire multidisciplinary screening process and the organisational aspects and will help to avoid unnecessary delays and costs later on.

The feasibility of screening models should be tested before detailed planning of pilot studies can begin. Planning should include: professional performance, organisational and financial aspects, as well as the scope and content of a comprehensive quality assurance programme. The initial plans should also consider the time frame within which the various issues need to be further developed. Key points at this stage are:

- Creation of professional dedication (understanding).
- Planning of infrastructure.
- Establishing of coordinating office with supervision mandate.
- Ensuring that screening is seen as a process.
- Designation of a process owner with mandate to run and manage the quality of the programme.
- Organisational development (self learning, quality driven).
- A separate coordination budget.
- Multidisciplinary case management.
- Collaboration between screening and treatment systems.
- Appropriate diagnostic assessment of patients.
- An appropriate screening monitoring IT-system with access and possibility to link registers e.g. population-, patient- and cancer registers.
- Comprehensive information system, serving all purposes.
- Development of a quality assurance plan, including technical QA.
- Adoption of approved QA plan.
- Definition of performance parameters and acceptable levels, including standards for health professionals.
- Contracts with health care providers.
- System for auditing, training and re-training.
- Assessment tools to exclude bad performers.
- Consideration of accreditation system or other comprehensive systems for ensuring competent service delivery.

2.3. *Preparation of all components of screening process, including feasibility testing*

Based on the comprehensive planning, the feasibility of the screening services and key components of programme management should be tested in small-scale studies that are designed to yield initial results with a limited amount of

EUROPEAN JOURNAL OF CANCER 48 (2012) 743–748

financial, technical, staff and time resources. The study results are taken into account in revising the initial plans, if necessary, prior to initiating pilot studies on a larger scale. Before the piloting phase can begin, the outcome of the feasibility phase should be thoroughly evaluated. Key points at this stage are:

- Scientific and ethical review of feasibility protocol.
- Correct and balanced information on 'benefit and risk'.
- Development of communication strategy.
- Societal input.
- Clearance of data protection and confidentiality issues.
- Creation of formal oversight for screening programmes.
- Scientific publication of feasibility results.

2.4. *Piloting and modification, if necessary, of all screening systems and components, including quality assurance in routine settings*

In England and many other European countries, implementation of breast, cervical and colorectal screening programmes started in pilot areas, and based on this experience the programmes were scaled up to national coverage.[5] In Finland, implementation of the programmes started in randomly selected cohorts, and was gradually extended to all targeted age groups.[6] This allows the outcome to be evaluated as a randomised controlled trial (randomised health policy).[7] The Finnish approach requires a national decision on screening implementation and the availability of a national population register. The approach permits evidence-based modification of the programme before it extends to the entire country.

The pilot implementation model starts with selection of one or a few pilot regions. Supervision and coaching is important in this phase in order to pick up problems in the screening process as soon as possible. The pilot phase also serves as a testing ground for the legal framework. The pilot outcome should be reported in the scientific literature and widely disseminated to health planners, politicians and health professionals. Based on the piloting, the financial implications of the roll out of the programme should be determined. Key points at this stage are:

- Budgeting.
- Ensuring financial commitment.
- Supervision and coaching of screening staff.
- Testing the legal framework.
- Ability to exclude bad performers.
- Scientific publication of outcome.

2.5. *Scaling up from pilot to service screening*

This is the actual implementation of the piloted intervention. All the points above need to be scaled-up to the size of the full programme. Effective communication of the experience gained to date in the implementation process should help to develop societal confidence in the programme. Key points at this stage are:

- Defining and contracting the local, regional and national programme teams, defining responsibilities.
- Setting-up infrastructure for coordination within health care settings.
- Identifying possible obstacles.
- Developing a plan for evaluation.
- Availability of staff (professional skills and numbers).
- Multidisciplinary case management.
- Special training, reference centre.
- Comprehensive information system, covering all steps in the screening process.
- Collaboration between screening, treatment and IT systems.
- Technical quality assurance.
- Reduction of barriers to participation.
- Tools to encourage compliance.
- Advocacy and collaboration with local civil society organisations.
- Population confidence.

2.6. *Running of a full-scale screening programme. Intensive monitoring of programme roll-out for early detection and correction of quality problems*

Maintaining high-quality of the screening service requires continuous supervision and rigorous scientific reporting. Attention must be paid to performance at each step in the screening process from information and invitation to performance of the screening test, assessment of abnormalities, and diagnosis and treatment of lesions detected in screening. Key points at this stage are:

- Supervision of all steps in the screening process.
- Ability to exclude bad performers.
- Testing grounds for new technologies.
- Monitoring the benefits and harms of screening.
- Scientific publication of outcome.

2.7. *Sustainability*

Sustainability is essential to achieve the potential impact of screening on the burden of disease in the population. This requires adequate, continuous financial support for preserving high programme quality. To maintain societal support, adequate communication of programme performance and impact is essential. This requires long-term evaluation in adequately planned studies with high-quality testing, reporting of performance and follow-up of screening outcomes. Key points at this stage are:

- Accurate and accessible communication of screening outcome.
- Population confidence.
- Organisational anchoring.
- Ensuring adequate financial resources and political commitment.

3. Discussion

The importance of screening as a tool for cancer control has been on the EU agenda for more than 20 years. In the European Code Against Cancer from 1989, women were advised to 'have a cervical smear regularly' and 'if possible, [to] undergo mammography at regular intervals above the age of 50'.[8] The need for organisation of screening into population-based programmes was stressed in the first quality assurance guidelines on breast[9] and cervical[10] cancer from 1993, and further developed in the preparatory work for the recommendation on cancer screening of the Council of the EU in 2003.[11] In the Council Recommendation, the EU Member States unanimously agreed on standards and principles for implementation of breast, cervical, and colorectal cancer screening programmes.[1]

However, the actual implementation of population-based screening programmes in the EU is still far from complete. By 2007, opportunistic cervical cancer screening was still the only available option for nearly half of the European target population, and 30% (8% without service and 22% outside groups served) of the women and men in the European target population were not offered colorectal cancer screening,[12] Table 1. The coverage by colorectal cancer screening programmes has, however, improved after 2007. For example, in Italy by the end of 2008, 36% of 50–69 year old men and women were invited to biennial screening with faecal occult blood test, and 1,171,000 persons were screened, attendence rate 47.5%.[13] The programmes in France and England became nationwide in 2009, and roll-out of programmes in the Netherlands and Denmark will start in 2013 and 2014, respectively. Furthermore, the screening activity is not yet standardised across the EU according to the European recommendations. For example by 2009, the lifetime number of recommended screening tests for cervical cancer varied from 6 to 50+ across the EU countries.[14]

Various obstacles in health care systems and in setting political priorities can inhibit the implementation of population-based screening. In 'new' Member States (that acceded to the EU after 2003), lack of resources is a serious problem. Adequate budgeting is a prerequisite for a successful programme as illustrated by recent experience from Poland.[15]

In 'old' Member States, organisation of screening may conflict with a traditional fee-for-service payment system. To decrease the number of opportunistic smears, the Netherlands stopped reimbursement of preventive smears taken outside the organised programme in 1996.[16] In England, target payments were introduced for general practitioners in 1990 to encourage them to include their female patients in the screening programme. There was no payment if the coverage was below 50%, a small payment if the coverage was between 50% and 79%, and a higher payment when the coverage reached 80%.[17]

Several countries have encountered problems with data confidentiality despite the fact that the EU directive on data protection[18] allows for linkage of health services data. The performance indicators listed in the EU guidelines may be used to monitor the programmes. However, it is clear that these indicators can be calculated only if access is provided to the necessary data. For example, calculation of the 'interval cancer rate as proportion of the underlying, expected, breast cancer incidence rate in the absence of screening'[2] requires access to data on all women with a normal screening mammogram and the individual follow-up of each of these women for incident breast cancer, death and emigration. Furthermore, a population-based breast cancer incidence rate for the period prior to initiation of screening must be available. It is encouraging that this indicator has been calculated for several European breast cancer screening programmes.[19]

Evaluation might also require the merging of datasets that have been separate in the past. In Sweden, the responsibility for cervical cancer screening rests with the counties that also keep the respective records. In order to perform a national audit of the screening programme data were retrieved and merged from 30 pathology and cytology laboratories throughout Sweden, and local cytology codes were converted to a common nomenclature.[20] Once national datasets have been established it is also possible to identify regional differences in screening uptake and outcome; a routine practise in the English[21] and Italian[22] programmes. Centralisation of activities also facilitates monitoring. The EU obligation to invite tenders for provision of large scale services has, however, in some cases impeded centralisation such as in the case of cytology services.

Local ownership and appropriate adaptation to the local health care system are important factors for success. In France, widespread opportunistic mammography screening was the norm until the organised programme became nation-

Table 1 – State of cancer screening in 27 Member States of the European Union by 2007 (adapted from von Karsa et al.[12]).

European recommendation* EU Target population	Breast cancer Women, aged 50–69 59 mio	Cervical cancer Women, aged 30–60 109 mio	Colorectal cancer Women and men, aged 50–74 136 mio
Proportion of target population covered by:			
Population-based, rollout complete	41%	22%	0%
Population-based, roll-out ongoing, piloting, planning	50%	29%	43%
Non-population-based	6%	47%	27%
Excluded from the regions and/or age groups offered screening	2%	2%	22%
no service	2%	<1%	8%

* Target ages recommended for breast and colorectal cancer screening recommended by European Union,[1] minimum target age recommended for cervical cancer screening by Arbyn et al.[3]

wide in 2004. The organised screening programme has therefore encompassed some of the features from the opportunistic practise. For example, a screening examination includes a clinical examination, a minimum of two views per breast, and a supplementary view if needed; all undertaken by an accredited radiologist.[23] In England, the organised programme started much earlier, in 1988. A radiographer with special training in mammography takes two views of each breast, and the whole visit lasts about half an hour.[24]

The comparison of performance indicators from different countries serves as a tool to monitor possible consequences of adapting screening programmes to local health care systems. For example, results from selected European breast cancer screening programmes showed that the cancer detection rate divided by the background incidence rate was somewhat higher in Copenhagen, 3.2, where the programme resembles the English set-up, than in Marseille, 2.5, and in Strasbourg, 2.9.[25]

To ensure the sustainability of a screening programme new evidence concerning screening methods should be reviewed regularly and the potential implications for programme policy should be considered, e.g. with regard to new technologies as flexible sigmoidoscopy,[26] or combination of screening with other preventive measures as human papillomavirus (HPV) vaccination.[27] Evidence-based guidelines should therefore be updated regularly.

Broad societal understanding of the benefits and risks of screening is also essential to the sustainability of an effective screening programme. Many of the performance parameters in the European guidelines are designed to provide an early indication of the benefits and risks of screening. However, long-term follow-up is needed to accurately assess both the benefit in reducing cancer specific mortality and the risk of over-diagnosis. The methodological challenges of such long-term studies using observational data[28] make it all the more important to provide sustainable support for accessibility and management of individual data.

Given the large number of individuals attending screening programmes, it is of utmost importance to avoid risks of low-quality screening. To ensure public confidence in the programme, it should be possible to identify and exclude poor performers. This is challenging because no screening service is infallible. A good example is the actions taken in England following identified failures in the cervical cancer screening programme.[29] Where resources are limited, it is better to implement one cancer screening programme at a time, than to start screening for all three cancer sites at once. Prioritisation may be based amongst other things on the number of cases, political will, and the availability of professional expertise and dedication. The detailed lists provided in this paper can serve as a guide to a gradual and successful implementation.

4. Conclusion

Screening programmes must be implemented effectively and operated in accordance with societal values and priorities. Prerequisites for a successful screening programme are societal acceptance, local ownership, and effective coordination along with the best evidence-based practice.

Given the complexity of the implementation process, it is not surprising that 10 or more years are commonly required to establish population-based cancer screening programmes. Effective, sustained coordination is required, beginning early in the process, with a clear vision of the multiple steps involved and adequate resources to provide leadership, develop consensus, and adapt to the evolving needs of the unfolding programme.

Conflict of interest statement

Elsebeth Lynge: Elsebeth Lynge is undertaking a comparative study of new-generation HPV tests, involving collaboration with Roche Diagnostics A/S, Genomica S. A. U., Qiagen Gaithersburg Ltd., and GenProbe Inc., and has served as unpaid advisor for GenProbe and Norchip.

Sven Törnberg: None declared.

Lawrence von Karsa: None declared.

Nereo Segnan: I participated to an advisory board meeting for Colorectal Cancer Screening in January 2011, as a paid expert, on colorectal cancer blood screening assay, organised by the Roche Diagnostics Ltd. I asked and received the formal permission by my employer, the S. Giovanni University Hospital of Turin.

Johannes J. M. van Delden: None declared.

Acknowledgements

We are indebted to Susanne V Allander, Ahti Anttila, Magdalena Bielska-Lasota, Dorine Coenen, Thomas Davidson, Euzebiusz Dziwinski, Karl Freese, Louise Gunning, Monica Hultcrantz, Gunta Lazdane, Sue Moss, Måns Rosén, Leo van Rossum, Chris de Wolf, and Maja Primic Žakelj for discussion of this paper at the Stockholm workshop. Valuable comments were received from Jose Manuel Baena, Luc Bleyen, Mireille Broeders, Luis Bujanda, José Expósito, Roger Pla Farnós, Xandra Gravenstein, Rosella Hermans, Roland Holland, Harry de Koning, Chris Meijer, Nieves Ascunce Elizaga, Xavier Castells Oliveres, Julietta Patnick, Marina Pollán, Dolores Salas, Hélène Sancho-Garnier, and Jaroslav Volf. The research leading to these results has received funding from the European Union's Seventh Framework Programme [FP7/2007-2013] under agreement number 229716.

REFERENCES

1. European Union. *Council recommendation of 2 December 2003 on cancer screening* (2003/878/EC).
2. Perry N, Broeders M. *De Wolf C, Törnberg S, Holland R, von Karsa L (Eds.)*. European guidelines for quality assurance in breast cancer screening and diagnosis. Fourth edition. Luxembourg: Office for official publications of the European Communities; 2006.
3. Arbyn M, Anttila A, Jordan J, et al. *Von Karsa L (Eds.)*. European guidelines for quality assurance in cervical cancer screening. Second edition. Luxembourg: Office for official publications of the European Communities; 2008.

4. Segnan N, Patnick J, von Karsa L. *European guidelines for quality assurance in colorectal cancer screening and diagnosis*. Luxembourg: Publication office of the European Union; 2010.
5. The UK CRC Screening Pilot Evaluation Team. *Evaluation of the UK colorectal cancer screening pilot. Final report*. February 2003, revised May 2003. http://www.cancerscreening.nhs.uk/bowel/pilot-evaluation.html [Accessed 3 March 2011].
6. Chiu SY, Malila N, Yen AM, et al. Analytical decision model for sample size and effectiveness projections for use in planning a population-based randomized controlled trial of colorectal cancer screening. *J Eval Clin Pract* 2011;**17**:123–9.
7. Hakama M, Pukkala E, Soderman B, Day N. Implementation of screening as a public health policy: issues in design and evaluation. *J Med Screen* 1999;**6**:209–16.
8. Sweet D, Vanvossel A. *(Eds.). Europe against cancer. Public health: initiatives and texts adopted in 1990*. Luxembourg: Office for Official Publications of the European Communities; 1990.
9. Kirkpatrick A, Törnberg S, Thijssen MAO. *European guidelines for quality assurance in mammography screening*. Luxembourg: Office for Official Publications of the European Communities; 1993.
10. Coleman D, Day N, Douglas G, et al. European guidelines for quality assurance in cervical cancer screening. *Eur J Cancer* 1993;**29A**(Suppl 4):138.
11. Advisory Committee on Cancer Prevention. Recommendations on cancer screening in the European Union. *Eur J Cancer* 2000;**36**:1473–8.
12. von Karsa L, Anttila A, Ronco G, et al. *Cancer screening in the European Union*. Luxembourg: European Communities, 2008.
13. Zorzi M, Baracco S, Fedato C, et al. Screening for colorectal cancer in Italy: 2008 survey. *Epidemiol Prev* 2010;**34**(5–6 Suppl 4):53–72.
14. Anttila A. Karsa L von, Aasmaa A, Fender M, patnick J, Rebolj M, Nicula F, Vass L, Valerianova Z, Voti L, Sauvager C, Ronco G. *Cervical cancer screening policies and coverage in Europe*. *Eur J Cancer* 2009;**45**:2649–58.
15. Spaczyński M, Karowicz-Bilinska A, Kedzia W, et al. Costs of population cervical cancer screening program in Poland between 2007–2009. *Ginekol Pol* 2010;**81**:750–6 [in Polish].
16. Rebolj M, van Ballegoijen M, Berkers L-M, Habbema D. Monitoring a national cancer prevention program: successful changes in cervical cancer screening in the Netherlands. *Int J Cancer* 2006;**120**:806–12.
17. Patnick J. Cervical cancer screening in England. *Eur J Cancer* 2000;**36**:2205–8.
18. Directive 95/46/EC of the European Parliament and of the Council of 24 October 1995 on the protection of individuals with regard to the processing of personal data and on the free movement of such. *Official Journal of the European Communities* 1995, 23 November, No L. 281 p. 31.
19. Törnberg S, Kemetli L, Ascunce N, et al. A pooled analysis of interval cancer rates in six European countries. *Eur J Cancer Prev* 2010;**19**:87–93.
20. Andrae B, Kemetli L, Sparén P, et al. Screening-preventable cervical cancer risks: evidence from a nationwide audit in Sweden. *J Natl Cancer Inst* 2008;**100**:622–9.
21. National Statistics. *Breast screening programme 2008-09*. http://www.ic.nhs.uk/statistics-and-data-collections/screening/breast-screening/breast-screening-programme-england-2008-09 [Accessed 8 June 2011].
22. Zappa M, Grazzini G, Naldoni C, et al. Introduction. The diffusion of screening programmes in Italy: 2008. *Epidemiol Prev* 2010;**34**(5–6 Suppl 4):5–8.
23. Séradour B, Ancelle-Park R. Breast cancer screening : are results of French and international programmes comparable ? *J Radiol 2006*;**87**:10098–14 [in French].
24. NHS Breast screening programme. *Breast screening a pocket guide*. http://www.cancerscreening.nhs.uk/breastscreen/publications/a-pocket-guide.html [Accessed 7 June 2011].
25. Broeders MJ, Scharpantgen A, Ascunce N, et al. European Breast Cancer Network. Comparison of early performance indicators for screening projects within the European Breast Cancer Network: 1989–2000. *Eur J Cancer Prev* 2005;**14**:107–16.
26. Atkin WS, Edwards R, Kralj-Hans I, et al. UK Flexible Sigmoidoscopy Trial Investigators. Once-only flexible sigmoidoscopy screening in prevention of colorectal cancer: a multicentre randomized controlled trial. *Lancet* 2010;**375**:1624–33.
27. Lynge E, Antilla A, Arbyn M, Segnan N, Ronco G. What's next? Perspectives and future needs of cervical screening in Europe in the era of molecular testing and vaccination. *Eur J Cancer* 2009;**45**:2714–21.
28. Duffy SW, Lynge E, Jonsson H, Ayyaz S, Olsen AH. Complexities in the estimation of overdiagnosis in breast cancer screening. *Br J Cancer* 2008;**99**:1176–8.
29. Cervical Screening Action Team: *The report*. NHS Executive, August 1998. http://www.dh.gov.uk/en/Publicationsandstatistics/Publications/PublicationsPolicyAndGuidance/DH_4005789 [Accessed 7 June 2011].

HOW TO OBTAIN EU PUBLICATIONS

Free publications:

- one copy:
  via EU Bookshop (http://bookshop.europa.eu);
- more than one copy or posters/maps:
  from the European Union's representations (http://ec.europa.eu/represent_en.htm);
  from the delegations in non-EU countries (http://eeas.europa.eu/delegations/index_en.htm);
  by contacting the Europe Direct service (http://europa.eu/europedirect/index_en.htm) or calling 00 800 6 7 8 9 10 11 (freephone number from anywhere in the EU) (*).

  (*) The information given is free, as are most calls (though some operators, phone boxes or hotels may charge you).

Priced publications:

- via EU Bookshop (http://bookshop.europa.eu).

Priced subscriptions:

- via one of the sales agents of the Publications Office of the European Union (http://publications.europa.eu/others/agents/index_en.htm).

ND-02-13-386-EN-C

Publications Office