University of Central Lancashire
UCLan

Central Lancashire Online Knowledge (CLoK)

Title	Improving the economic value of photographic screening for optical coherence tomography: detectable macular oedema - a prospective, multicentre, United Kingdom study
Type	Article
URL	https://clok.uclan.ac.uk/id/eprint/25099/
DOI	https://doi.org/10.3310/hta17510
Date	2013
Citation	Olson, J., Sharp, P., Goatman, K., Prescott, Gordon, Scotland, G., Fleming, A., Philip, S., Santiago, C., Borooah, S. et al (2013) Improving the economic value of photographic screening for optical coherence tomography: detectable macular oedema - a prospective, multicentre, United Kingdom study. Health Technology Assessment, 17 (51). pp. 1-141. ISSN 1366-5278
Creators	Olson, J., Sharp, P., Goatman, K., Prescott, Gordon, Scotland, G., Fleming, A., Philip, S., Santiago, C., Borooah, S., Broadbent, D., Chong, V., Dodson, P., Harding, S., Leese, G., Styles, C., Swa, K. and Wharton, H.

It is advisable to refer to the publisher's version if you intend to cite from the work.
https://doi.org/10.3310/hta17510

For information about Research at UCLan please go to http://www.uclan.ac.uk/research/

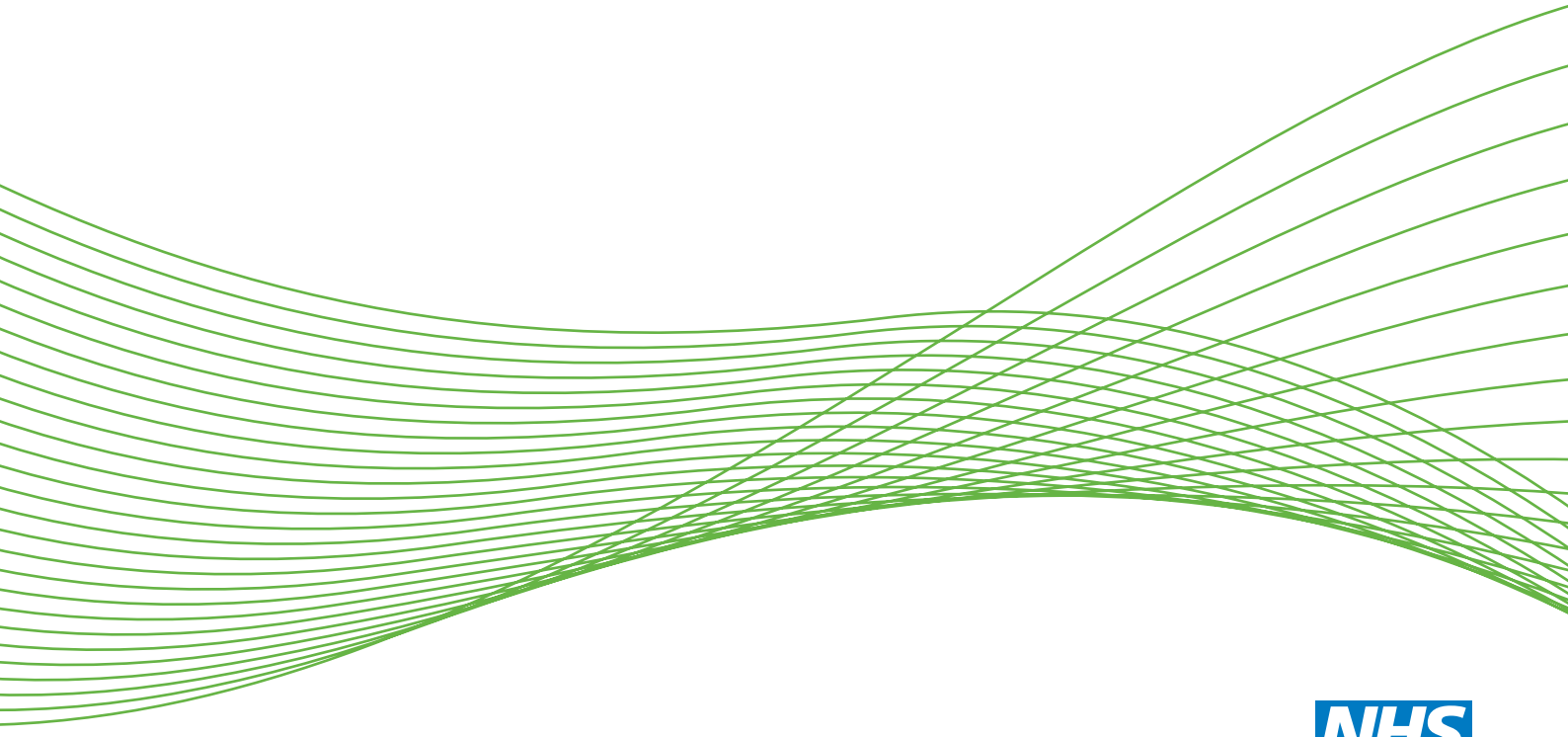
All outputs in CLoK are protected by Intellectual Property Rights law, including Copyright law. Copyright, IPR and Moral Rights for the works on this site are retained by the individual authors and/or other copyright owners. Terms and conditions for use of this material are defined in the http://clok.uclan.ac.uk/policies/

HEALTH TECHNOLOGY ASSESSMENT

VOLUME 17 ISSUE 51 NOVEMBER 2013
ISSN 1366-5278

Improving the economic value of photographic screening for optical coherence tomography-detectable macular oedema: a prospective, multicentre, UK study

J Olson, P Sharp, K Goatman, G Prescott, G Scotland, A Fleming, S Philip, C Santiago, S Borooah, D Broadbent, V Chong, P Dodson, S Harding, G Leese, C Styles, K Swa and H Wharton



National Institute for Health Research

DOI 10.3310/hta17510

Improving the economic value of photographic screening for optical coherence tomography-detectable macular oedema: a prospective, multicentre, UK study

J Olson,[1]* P Sharp,[2] K Goatman,[2] G Prescott,[3] G Scotland,[3] A Fleming,[2] S Philip,[4] C Santiago,[5] S Borooah,[6] D Broadbent,[7] V Chong,[8] P Dodson,[9] S Harding,[10] G Leese,[11] C Styles,[12] K Swa[13] and H Wharton[14]

[1]Grampian Retinal Screening Programme, Aberdeen, UK
[2]College of Life Sciences and Medicine, University of Aberdeen, Aberdeen, UK
[3]School of Medicine and Dentistry, Division of Applied Health Sciences, University of Aberdeen, Aberdeen, UK
[4]Grampian Diabetes Research Unit Diabetes Centre, Woolmanhill Hospital, Aberdeen, UK
[5]Eye Out-patient Department, Aberdeen Royal Infirmary, Aberdeen, UK
[6]Princess Alexandra Eye Pavilion, Edinburgh, UK
[7]Liverpool Diabetes Eye Centre and School of Clinical Sciences, Royal Liverpool University Hospital, Liverpool, UK
[8]Oxford Eye Hospital, Oxford, UK
[9]Heartlands Hospital and Aston University, Birmingham, UK
[10]Institute of Ageing and Chronic Disease, Department of Eye and Vision Science, University of Liverpool, Liverpool, UK
[11]Ninewells Hospital, Dundee, UK
[12]Department of Ophthalmology, Queen Margaret Hospital, Dunfermline, UK
[13]National Services Division, National Services Scotland, Edinburgh, UK
[14]Diabetes and Endocrinology Unit, Heartlands Hospital, Birmingham, UK

*Corresponding author

Declared competing interests of authors: John Olson, Peter Sharp and Alan Fleming have received funding for their institution from Medalytix Ltd. Graham Leese is a consultant for Novo Nordisk Ltd, Novartis Pharmaceuticals UK Ltd, Sanofi-aventis and Eli Lilly and Company. Simon Harding is a consultant for Novartis Pharmaceuticals UK Ltd. Victor Chong is a consultant for Novartis Pharmaceuticals UK Ltd, Bayer, Allergan Ltd and IRIDEX Corporation. Ken Swa sits on the Novartis Pharmaceuticals UK Ltd Advisory Board (Scotland).

Published November 2013
DOI: 10.3310/hta17510

This report should be referenced as follows:

Olson J, Sharp P, Goatman K, Prescott G, Scotland G, Fleming A, *et al.* Improving the economic value of photographic screening for optical coherence tomography-detectable macular oedema: a prospective, multicentre, UK study. *Health Technol Assess* 2013;**17**(51).

Health Technology Assessment is indexed and abstracted in *Index Medicus*/MEDLINE, *Excerpta Medica*/EMBASE, *Science Citation Index Expanded* (SciSearch®) and *Current Contents®*/ Clinical Medicine.

Health Technology Assessment

HTA/HTA TAR

ISSN 1366-5278 (Print)

ISSN 2046-4924 (Online)

Five-year impact factor: 5.804

Health Technology Assessment is indexed in MEDLINE, CINAHL, EMBASE, The Cochrane Library and the ISI Science Citation Index and is assessed for inclusion in the Database of Abstracts of Reviews of Effects.

This journal is a member of and subscribes to the principles of the Committee on Publication Ethics (COPE) (www.publicationethics.org/).

Editorial contact: nihredit@southampton.ac.uk

The full HTA archive is freely available to view online at www.journalslibrary.nihr.ac.uk/hta. Print-on-demand copies can be purchased from the report pages of the NIHR Journals Library website: www.journalslibrary.nihr.ac.uk

Criteria for inclusion in the *Health Technology Assessment* journal

Reports are published in *Health Technology Assessment* (HTA) if (1) they have resulted from work for the HTA programme, and (2) they are of a sufficiently high scientific quality as assessed by the reviewers and editors.

Reviews in *Health Technology Assessment* are termed 'systematic' when the account of the search appraisal and synthesis methods (to minimise biases and random errors) would, in theory, permit the replication of the review by others.

HTA programme

The HTA programme, part of the National Institute for Health Research (NIHR), was set up in 1993. It produces high-quality research information on the effectiveness, costs and broader impact of health technologies for those who use, manage and provide care in the NHS. 'Health technologies' are broadly defined as all interventions used to promote health, prevent and treat disease, and improve rehabilitation and long-term care.

The journal is indexed in NHS Evidence via its abstracts included in MEDLINE and its Technology Assessment Reports inform National Institute for Health and Care Excellence (NICE) guidance. HTA research is also an important source of evidence for National Screening Committee (NSC) policy decisions.

For more information about the HTA programme please visit the website: www.hta.ac.uk/

This report

The research reported in this issue of the journal was funded by the HTA programme as project number 06/402/49. The contractual start date was in May 2008. The draft report began editorial review in March 2012 and was accepted for publication in September 2012. The authors have been wholly responsible for all data collection, analysis and interpretation, and for writing up their work. The HTA editors and publisher have tried to ensure the accuracy of the authors' report and would like to thank the reviewers for their constructive comments on the draft document. However, they do not accept liability for damages or losses arising from material published in this report.

This report presents independent research funded by the National Institute for Health Research (NIHR). The views and opinions expressed by authors in this publication are those of the authors and do not necessarily reflect those of the NHS, the NIHR, NETSCC, the HTA programme or the Department of Health. If there are verbatim quotations included in this publication the views and opinions expressed by the interviewees are those of the interviewees and do not necessarily reflect those of the authors, those of the NHS, the NIHR, NETSCC, the HTA programme or the Department of Health.

© Queen's Printer and Controller of HMSO 2013. This work was produced by Olson *et al*. under the terms of a commissioning contract issued by the Secretary of State for Health. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

Published by the NIHR Journals Library (www.journalslibrary.nihr.ac.uk), produced by Prepress Projects Ltd, Perth, Scotland (www.prepress-projects.co.uk).

© Queen's Printer and Controller of HMSO 2013. This work was produced by Olson *et al*. under the terms of a commissioning contract issued by the Secretary of State for Health. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

Editor-in-Chief of *Health Technology Assessment* and NIHR Journals Library

Professor Tom Walley Director, NIHR Evaluation, Trials and Studies and Director of the HTA Programme, UK

NIHR Journals Library Editors

Professor Ken Stein Chair of HTA Editorial Board and Professor of Public Health, University of Exeter Medical School, UK

Professor Andree Le May Chair of NIHR Journals Library Editorial Group (EME, HS&DR, PGfAR, PHR journals)

Dr Martin Ashton-Key Consultant in Public Health Medicine/Consultant Advisor, NETSCC, UK

Professor Matthias Beck Chair in Public Sector Management and Subject Leader (Management Group), Queen's University Management School, Queen's University Belfast, UK

Professor Aileen Clarke Professor of Health Sciences, Warwick Medical School, University of Warwick, UK

Dr Tessa Crilly Director, Crystal Blue Consulting Ltd, UK

Dr Peter Davidson Director of NETSCC, HTA, UK

Ms Tara Lamont Scientific Advisor, NETSCC, UK

Professor Elaine McColl Director, Newcastle Clinical Trials Unit, Institute of Health and Society, Newcastle University, UK

Professor William McGuire Professor of Child Health, Hull York Medical School, University of York, UK

Professor Geoffrey Meads Honorary Professor, Business School, Winchester University and Medical School, University of Warwick, UK

Professor Jane Norman Professor of Maternal and Fetal Health, University of Edinburgh, UK

Professor John Powell Consultant Clinical Adviser, National Institute for Health and Care Excellence (NICE), UK

Professor James Raftery Professor of Health Technology Assessment, Wessex Institute, Faculty of Medicine, University of Southampton, UK

Dr Rob Riemsma Reviews Manager, Kleijnen Systematic Reviews Ltd, UK

Professor Helen Roberts Professorial Research Associate, University College London, UK

Professor Helen Snooks Professor of Health Services Research, Institute of Life Science, College of Medicine, Swansea University, UK

Please visit the website for a list of members of the NIHR Journals Library Board:
www.journalslibrary.nihr.ac.uk/about/editors

Editorial contact: nihredit@southampton.ac.uk

Abstract

Improving the economic value of photographic screening for optical coherence tomography-detectable macular oedema: a prospective, multicentre, UK study

J Olson,[1]* P Sharp,[2] K Goatman,[2] G Prescott,[3] G Scotland,[3] A Fleming,[2] S Philip,[4] C Santiago,[5] S Borooah,[6] D Broadbent,[7] V Chong,[8] P Dodson,[9] S Harding,[10] G Leese,[11] C Styles,[12] K Swa[13] and H Wharton[14]

[1]Grampian Retinal Screening Programme, Aberdeen, UK
[2]College of Life Sciences and Medicine, University of Aberdeen, Aberdeen, UK
[3]School of Medicine and Dentistry, Division of Applied Health Sciences, University of Aberdeen, Aberdeen, UK
[4]Grampian Diabetes Research Unit Diabetes Centre, Woolmanhill Hospital, Aberdeen, UK
[5]Eye Out-patient Department, Aberdeen Royal Infirmary, Aberdeen, UK
[6]Princess Alexandra Eye Pavilion, Edinburgh, UK
[7]Liverpool Diabetes Eye Centre and School of Clinical Sciences, Royal Liverpool University Hospital, Liverpool, UK
[8]Oxford Eye Hospital, Oxford, UK
[9]Heartlands Hospital and Aston University, Birmingham, UK
[10]Institute of Ageing and Chronic Disease, Department of Eye and Vision Science, University of Liverpool, Liverpool, UK
[11]Ninewells Hospital, Dundee, UK
[12]Department of Ophthalmology, Queen Margaret Hospital, Dunfermline, UK
[13]National Services Division, National Services Scotland, Edinburgh, UK
[14]Diabetes and Endocrinology Unit, Heartlands Hospital, Birmingham, UK

*Corresponding author john.olson@nhs.net

Objectives: To determine the best photographic surrogate markers for detecting sight-threatening macular oedema (MO) in people with diabetes attending UK national screening programmes.

Design: A multicentre, prospective, observational cohort study of 3170 patients with photographic signs of diabetic retinopathy visible within the macular region [exudates within two disc diameters, microaneurysms/dot haemorrhages (M/DHs) and blot haemorrhages (BHs)] who were recruited from seven study centres.

Setting: All patients were recruited and imaged at one of seven study centres in Aberdeen, Birmingham, Dundee, Dunfermline, Edinburgh, Liverpool and Oxford.

Participants: Subjects with features of diabetic retinopathy visible within the macular region attending one of seven diabetic retinal screening programmes.

Interventions: Alternative referral criteria for suspected MO based on photographic surrogate markers; an optical coherence tomographic examination in addition to the standard digital retinal photograph.

© Queen's Printer and Controller of HMSO 2013. This work was produced by Olson *et al*. under the terms of a commissioning contract issued by the Secretary of State for Health. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

Main outcome measures: (1) To determine the best method to detect sight-threatening MO in people with diabetes using photographic surrogate markers. (2) Sensitivity and specificity estimates to assess the costs and consequences of using alternative strategies. (3) Modelled long-term costs and quality-adjusted life-years (QALYs).

Results: Prevalence of MO was strongly related to the presence of lesions and was roughly five times higher in subjects with exudates or BHs or more than two M/DHs within one disc diameter. Having worse visual acuity was associated with about a fivefold higher prevalence of MO. Current manual screening grading schemes that ignore visual acuity or the presence of M/DHs could be improved by taking these into account. Health service costs increase substantially with more sensitive/less specific strategies. A fully automated strategy, using the automated detection of patterns of photographic surrogate markers, is superior to all current manual grading schemes for detecting MO in people with diabetes. The addition of optical coherence tomography (OCT) to each strategy, prior to referral, results in a reduction in costs to the health service with no decrement in the number of MO cases detected.

Conclusions: Compared with all current manual grading schemes, for the same sensitivity, a fully automated strategy, using the automated detection of patterns of photographic surrogate markers, achieves a higher specificity for detecting MO in people with diabetes, especially if visual acuity is included in the automated strategy. Overall, costs to the health service are likely to increase if more sensitive referral strategies are adopted over more specific screening strategies for MO, for only very small gains in QALYs. The addition of OCT to each screening strategy, prior to referral, results in a reduction in costs to the health service with no decrement in the number of MO cases detected.

Study registration: This study has been registered as REC/IRAS 07/S0801/107, UKCRN ID 9063 and NIHR HTA 06/402/49.

Source of funding: This project was funded by the NIHR Health Technology Assessment programme and will be published in full in *Health Technology Assessment*; Vol. 17, No. 51. See the HTA programme website for further project information.

Contents

List of abbreviations xi

Scientific summary xiii

Chapter 1 Introduction 1

Chapter 2 Study design and methods 3
Aims and objectives 3
Primary outcome measure 3
Other outcome measures 3
Study design 3
Recruitment and data collection 5
Image data analysis and grading 8

Chapter 3 Comparison of optical coherence tomography scanner thickness measurements 13
Subjects 13
Statistical analysis 13
Imaging protocols 14
Results 15
Conclusions 21

Chapter 4 Characteristics of study data 23
Introduction 23
Methods for demographics 23
Results 23
Conclusions 30

Chapter 5 Inferring the presence of macular oedema using retinal photographs 33
Introduction 33
Methods for statistical modelling 33
Results of statistical modelling 35
Strategies for detecting macular oedema 47
Classifiers 53
Results 54
Discussion and conclusions 58

Chapter 6 Health economic evaluation 61
Aims and objectives 61
Assessment of cost per case of macular oedema detected 61
Results 68
Assessing long-term cost-effectiveness 74
Results 82
Deterministic sensitivity analysis 82
Probabilistic sensitivity analysis 86
Discussion 86

© Queen's Printer and Controller of HMSO 2013. This work was produced by Olson *et al*. under the terms of a commissioning contract issued by the Secretary of State for Health. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

Chapter 7 Discussion **93**
Reason for study 93
Aim of the study 93
Study design 94
Cost-effectiveness 96
Strengths 97
Limitations 98
Implications for health care 99
Recommendations for research 99

Acknowledgements **101**

References **103**

Appendix 1 Techniques for measuring retinal thickness **109**

Appendix 2 Cost estimates for optical coherence tomography and slit-lamp examination within the screening programme **113**

Appendix 3 Demographics and statistical modelling – additional data **117**

Appendix 4 Protocol **125**

List of abbreviations

BH	blot haemorrhage
CAH	computer-assisted manual annotation strategy
CI	confidence interval
CWS	cotton wool spots
DD	disc diameter radius
EQ-5D	European Quality of Life-5 Dimensions
ETDRS	Early Treatment Diabetic Retinopathy Study
FA	fully automated annotation strategy
HRG	Healthcare Resource Group
ICC	intraclass correlation coefficient
log-MAR	logarithm of the minimum angle of resolution
M/DH	microaneurysm/dot haemorrhage
MO	macular oedema
MVL	moderate visual loss
QALY	quality-adjusted life-year
OCT	optical coherence tomography
OPCS	Office of Population Censuses and Surveys
OR	odds ratio
PbR	payment by results
PSSRU	Personal Social Services Research Unit
SD	standard deviation

© Queen's Printer and Controller of HMSO 2013. This work was produced by Olson *et al*. under the terms of a commissioning contract issued by the Secretary of State for Health. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

Scientific summary

Background

Compared with more established screening programmes, such as cervical and breast screening, diabetic retinopathy screening is still in its infancy: national screening programmes for diabetic retinopathy have been running throughout the UK for less than a decade.

The risk of retinopathy increases with disease duration. Type 2 diabetes is often diagnosed several years after onset (almost 40% of people with type 2 diabetes are found to have retinopathy at diagnosis). This is potentially sight-threatening in between 4% and 8% of cases. Sixty per cent of people with type 2 diabetes will have retinopathy 20 years from onset.

There are two main mechanisms responsible for visual loss from diabetic retinopathy. The first of these is proliferative retinopathy and the development of new blood vessels. The second is macular oedema (MO), the build-up of fluid involving the area of the retina associated with best vision, the macula. Although proliferative disease is most likely to lead to serious vision loss, MO is more common and is the leading cause of moderate visual loss in people with diabetes. However, laser treatment is only moderately effective, at best, for MO.

At the time of the introduction of national programmes for screening for diabetic retinopathy in the UK, screening for diabetic MO was not deemed to meet strict World Health Organization guidelines for screening. Diabetic retinopathy at the macula (maculopathy) is more common than proliferative diabetic retinopathy so retinal screening programmes have had to develop pathways to address this, although it is acknowledged that the evidence base for doing so is limited.

The current recommended method of retinal screening for diabetic retinopathy is digital fundus photography, a two-dimensional technology which cannot detect MO directly. Current photographic grading schemes rely on a combination of surrogate photographic markers, chosen by expert consensus, to infer the presence of MO. These manual grading schemes use combinations of features of retinopathy including microaneurysms/dot haemorrhages (M/DHs), blot haemorrhages (BHs), exudates and also visual acuity. At present, there is no consensus among the four nations as to which, and how many, features should be used to infer the presence of diabetic MO.

Clinical experience from the Grampian Retinal Screening Programme and Liverpool Screening Programme suggests that only 12–14% of these patients have evidence of MO, when examined by ophthalmologists using slit-lamp biomicroscopy.

Optical coherence tomography (OCT) is now recognised as the reference standard for measuring MO. It is an optical analogue of ultrasound imaging. The test is rapid, non-invasive and well tolerated by subjects. By acquiring a series of cross-sections it is possible to generate a thickness map of the macula. However, the equipment is still expensive.

Retinal ischaemia acts as a stimulus for the production of vascular endothelial growth factor. Vascular endothelial growth factor has very potent permeability-inducing properties, as well as stimulating angiogenesis. Currently there is a great deal of interest in the use of a range of treatments to counteract the action of vascular endothelial growth factor. Trials show, for the first time, that visual loss in people with diabetic MO can be reversed by intravitreal antivascular growth factor injection in approximately half of all treated patients, introducing a new paradigm in the treatment of diabetic MO and also now screening. It is probable that with these new treatments, screening for diabetic MO will meet the

© Queen's Printer and Controller of HMSO 2013. This work was produced by Olson *et al*. under the terms of a commissioning contract issued by the Secretary of State for Health. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

World Health Organization's criteria. Identifying patients, who will benefit most from these new therapies, is now an important issue for all diabetic retinopathy-screening programmes.

Objectives

The primary aim of the study was to determine the best photographic surrogate markers for detecting potentially sight-threatening diabetic MO, within English and Scottish national screening programmes. Specifically we wished to:

(a) investigate whether or not particular distributions and combinations of lesions (M/DHs, BHs and exudates), assessed manually or automatically, were more specific photographic surrogate markers of MO than current practice, using OCT as the reference standard
(b) assess the costs and consequences of using alternative distributions and combinations of these lesions to screen for MO, using either automated or manual detection of lesions
(c) model the long-term cost and quality-of-life implications of using alternative distributions and combinations of surrogate markers to screen for MO.

Once the study was under way, several screening programmes were found to be using OCT as part of the screening process to reduce false-positive referrals to the hospital eye service. Consequently, we added a further aim to assess the costs and consequences of using OCT within retinal screening programmes in addition to photographic surrogate markers, as this would affect how photographic markers would be used in future.

Methods

A total of 3540 patients with photographic signs of diabetic retinopathy visible within the macular region [exudates within two disc diameter radius (DD), M/DHs, BHs within one DD] were recruited from seven study centres at Aberdeen, Birmingham, Dundee, Dunfermline, Edinburgh, Liverpool and Oxford. Each subject had retinal photography and OCT on both eyes where possible.

Software was developed to assist the research nurse in the manual annotation of potential surrogate markers for MO (M/DHs, BHs and exudates), visible in the colour photographs.

Separate software was developed to analyse the distribution of retinal lesions, either annotated manually or automatically, to investigate whether or not particular distributions and combinations of lesions (M/DHs, BHs and exudates) within one or two DD from the centre of the fovea are more specific photographic surrogate markers of MO than current grading practice, using OCT as the reference standard.

The OCT images were analysed both quantitatively and qualitatively by the research nurse.

Statistical modelling was carried out to see if any of the manual grading schemes in England and Scotland included everything that might be considered important if starting from de novo. This information was used to inform the inclusion of various eye and subject characteristics within computer-assisted manual grading strategies and full automated grading strategies for detecting MO.

The grading strategies utilising photographic surrogate markers for detecting MO can be categorised as:

- *Manual grading strategies* These use photographic features in a binary fashion, similar to existing national criteria, which can be determined by visual inspection of retinal photographs by trained graders.

- *Computer-assisted manual annotation grading strategies* These use more detailed features obtained by manual annotation of retinal photographs which are then combined by a software classifier to determine a likelihood that MO is present.
- *Fully automated annotation grading strategies where no human intervention is required* These use features automatically annotated by image analysis software. As with computer-assisted manual annotation grading strategies, these are combined by a software classifier to determine a likelihood that MO is present.

A study was then carried out to assess the cost-effectiveness of alternative pathways for screening for MO using various combinations of photographic grading strategies, automation, visual acuity and OCT. An initial analysis focussed on the cost per true case of MO appropriately detected and referred. However, in most instances improvements in the sensitivity of the referral process come at the expense of decreased specificity. Thus, a secondary analysis assessed whether or not the increased costs associated with more sensitive screening strategies, resulting from more referrals to ophthalmology clinics (appropriate and inappropriate), are worth incurring given the potential health benefits. A Markov microsimulation model was developed for this purpose.

Results

The statistical analysis showed that the detection of MO in corresponding optical coherence tomograms was strongly related to the presence of retinal lesions in retinal photographs, being roughly five times more detectable among subjects with an exudate, a BH, or a M/DH present within one DD compared with the same lesion being absent. Having more than two M/DHs within one DD in an eye was of particular importance.

Subjects with worse visual acuity were about five times as likely to have MO in the relevant eye as those with better visual acuity.

The best-performing photographic grading strategy was a computer-assisted manual annotation grading strategy. This uses the results of manual annotation of the individual lesions in each image. Computer-assisted manual annotation is, however, a time-consuming procedure and so is unlikely to be considered for routine screening practice. Therefore, this strategy was not taken forward for economic analysis. Grading strategies relying on the manual annotation of retinal images without computer assistance of all retinal images were similarly not included.

The manual grading strategies chosen for comparison in the economic analysis included the current English grading scheme (manual grading strategy 1) and the current Scottish grading scheme (manual grading strategy 2). Manual grading strategy 2 is one of the higher-specificity, simple, manual grading strategies (59.5% sensitivity; 79% specificity) whereas manual grading strategy 1 represents a moderate-sensitivity, moderate-specificity, manual grading strategy (72.6% sensitivity; 66.8% specificity). A further manual grading strategy (strategy 16) was also included as a potential alternative to grading strategy 1 (having similar sensitivity, 73.3%, and higher specificity, 70.9%). Furthermore, consideration was given to the potential cost-effectiveness of utilising a fully automated annotation grading strategy which also utilised visual acuity (denoted FA2), with slightly higher sensitivity (75.9%) and better specificity (73.7%) than manual grading strategy 16. In addition, consideration was given to a grading strategy (grading strategy 8) that had 100% sensitivity, albeit with only 1.7% specificity, coupled with OCT prior to referral.

The impact of a screening pathway combining the selected grading strategies with OCT (within the screening programme prior to referral) to minimise false-positive referrals to eye clinics was also assessed.

The incremental costs per extra case of MO detected vary by grading strategy. Considering the more sensitive grading strategy 16 compared with grading strategy 2, the additional cost per extra case detected

© Queen's Printer and Controller of HMSO 2013. This work was produced by Olson *et al*. under the terms of a commissioning contract issued by the Secretary of State for Health. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

ranges from £1579 (English costs) to £985 (Scottish costs). When the grading strategies are coupled with OCT, in the screening pathway, these incremental costs fall to £636 and £528, respectively. Although the fully automated annotation grading strategy dominates manual grading strategies 1 and 16 under the base-case assumption of zero net increase in grading costs, it costs £897 per additional case detected and referred in comparison with manual grading strategy 2 (applying English screening and referral costs). This incremental screening cost drops to £405 when the grading strategies are coupled with OCT prior to referral. The 100% sensitive screening strategy 8 costs an additional £1510 per extra case detected in comparison with the most specific screening strategy (strategy 2 + OCT) when applying English referral costs, and £1360 per extra case detected when applying Scottish referral costs. In comparison with manual grading strategy 16 plus OCT, screening strategy 8 costs £1955 and £1784 per extra case detected and referred when applying English and Scottish referral costs, respectively.

The longer-term cost-effectiveness results mirror the pattern observed in the short term, in that the addition of OCT to the screening pathway results in cost-savings without reducing the health benefits. As with the short-term results, the fully automated annotation grading strategy also dominates manual grading strategies 1 and 16, under the assumption of zero net increase in grading costs associated with the introduction of automated grading. Furthermore, manual grading strategy 16 remains essentially dominant over, or very cost-effective in comparison with, strategy 1.

Conclusions

Compared with all current manual grading schemes, for the same sensitivity, a fully automated annotation grading strategy, using the automated detection of patterns of photographic surrogate markers, achieves a higher specificity for detecting OCT detectable MO in people with diabetes, especially if visual acuity is included in the automated strategy.

Overall costs to the health service are likely to increase if more sensitive referral strategies are adopted over more specific screening strategies for MO, for only very small gains in quality-adjusted life-years.

The addition of OCT to each grading strategy, as part of the screening pathway prior to referral, results in a reduction in costs to the health service with no decrement in the number of MO cases detected.

Study registration

This study has been registered as REC/IRAS 07/S0801/107, UKCRN ID 9063 and NIHR HTA 06/402/49.

Funding

Funding for this study was provided by the Health Technology Assessment programme of the National Institute for Health Research.

Chapter 1 Introduction

Compared with more established screening programmes, such as cervical and breast screening, diabetic retinopathy screening is still in its infancy: national screening programmes for diabetic retinopathy have been running throughout the UK for less than a decade.

The risk of retinopathy increases with disease duration.[1] Type 2 diabetes is often diagnosed several years after onset and consequently almost 40% of people with type 2 diabetes are found to have retinopathy at diagnosis. This is potentially sight-threatening in between 4% and 8% of cases. Sixty per cent of people with type 2 diabetes will have retinopathy 20 years from onset.

There are two main mechanisms responsible for visual loss from diabetic retinopathy. The first of these is proliferative retinopathy and the development of new blood vessels. The second is macular oedema (MO), the build-up of fluid involving the area of the retina associated with best vision, the macula. Although proliferative disease is most likely to lead to serious vision loss, MO is more common and is the leading cause of moderate visual loss (MVL) in people with diabetes. However, laser treatment is only moderately effective, at best, for MO.

At the time of the introduction of national programmes for screening for diabetic retinopathy in the UK, unlike screening for proliferative diabetic retinopathy, screening for diabetic MO was not deemed to meet strict World Health Organization guidelines for screening.[2] Diabetic retinopathy involving the macula (maculopathy) is more common than proliferative diabetic retinopathy so retinal screening programmes have had to develop pathways to address this, although it is acknowledged that the evidence base for doing so is limited.

Diabetes can affect the integrity and increase the permeability of the blood–retinal barrier. This is caused by a thickening of the basement membrane and fall in the number of supporting pericytes, making the blood–retinal barrier leaky.[3] This results in fluid accumulation within the outer plexiform and inner nuclear layers of the retina, and swelling of the cells within these layers. MO may be classified according to the fluid distribution: diffuse oedema is a general thickening of the central retina caused by either extensive capillary dilatation or capillary closure, whereas focal oedema is centred on specific vascular abnormalities, such as microaneurysms. Accumulated fluid defocuses the image on the retina, reducing visual acuity. If oedema persists, the increased pressure may lead to irreparable photoreceptor damage.

Retinal thickening is not visible directly on the retinal photographs used by screening programmes, so people are referred to ophthalmology clinics on the basis of a range of surrogate photographic markers, such as exudates within a certain distance of the foveal centre. A review of the evidence from the Early Treatment Diabetic Retinopathy Study (ETDRS)[4] suggests that exudates may be a sensitive marker of MO.[5] However, as this study excluded patients with mild retinopathy in the absence of exudate, the results are not directly applicable to screening programmes in the UK, where 60% of patients have no visible signs of retinopathy and > 30% have only mild retinopathy. Evidence from the Grampian Retinal Screening Programme suggests that only 12% of patients with surrogate markers referred to an ophthalmologist have indications of MO when examined by slit-lamp biomicroscopy. Similarly, a retrospective analysis from Liverpool, including 257 patients referred from the screening programme to the ophthalmology clinic between December 2001 and June 2002, found that only 14% had evidence of MO (unpublished data).

Macular oedema has traditionally been assessed clinically using a combination of slit-lamp biomicroscopy, stereoscopic photography and stereoscopic fluorescein angiography. However, these techniques have a number of limitations. Foremost is that they are only qualitative assessments, which are relatively insensitive to thickness changes. Furthermore, slit-lamp examination does not provide a pictorial record and, together with stereo photography, is known to be biased by the presence or absence of exudates.[6] Although the angiogram is a sensitive test for leakage, an objective measurement of thickening is not

© Queen's Printer and Controller of HMSO 2013. This work was produced by Olson *et al*. under the terms of a commissioning contract issued by the Secretary of State for Health. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

possible. Best corrected visual acuity has also been used as a surrogate indicator of thickening, but is neither sensitive nor specific, being affected by several factors besides macular thickness.

Optical coherence tomography (OCT) is recognised as the reference standard for measuring MO.[1,7] It is an optical analogue of ultrasound imaging, and uses low-coherence interferometry to acquire cross-sectional images of the retina. The resolution of the cross-sections is two orders of magnitude better than is achievable using ultrasound, with the axial resolution (in the plane of the retina) being between 8 and 16 µm and the lateral resolution (depth) between 10 and 15 µm. The test is rapid, non-invasive and well tolerated by subjects. By acquiring a series of cross-sections it is possible to generate a thickness map of the macula. OCT was first launched commercially by Carl Zeiss Meditec in 1996 and since then improvements have been made to the signal-to-noise ratio and axial and lateral resolution, and the acquisition time has been shortened.

Several investigators have compared OCT with slit-lamp biomicroscopy, concluding that OCT is the more sensitive test for macular thickening. It has been reported that the agreement between the slit-lamp and OCT is good for normal and extreme thicknesses, but equivocal between 200 and 325 µm.[8] Good agreement has also been found between thickening seen on stereo photographs and OCT.[9] Similarly, studies comparing OCT and fluorescein angiography indicate OCT to be at least as sensitive as angiography for detecting thickening.[10]

Since the ETDRS, focal and grid laser treatment have been the standard treatment for MO. A long-term follow-up study by Chew *et al*.[11] found that a median of 16.7 years after the original study, 42% of patients had a visual acuity of 20/20 [logarithm of the minimum angle of resolution (log-MAR) 0] or better in their best eye, and 84% had a visual acuity of 20/40 (log-MAR 0.3) or better in their best eye. Nevertheless, laser treatment is not without risk and is not effective in all cases.

Retinal ischaemia acts as a stimulus for the production of vascular endothelial growth factor.[12,13] Vascular endothelial growth factor has very potent permeability-inducing properties, as well as stimulating angiogenesis.[12] Over the last few years there has been a great deal of interest in the use of a range of adjunctive intravitreal treatments to counteract these pro-inflammatory and angiogenic stimuli, including corticosteroids like triamcinolone (Kenalog®, Bristol-Myers Squibb)[14,15] and antivascular endothelial growth factor therapies like pegaptanib (Macugen®, Pfizer Ltd),[16] bevacizumab (Avastin®, Roche Products Ltd)[17,18] and ranibizumab (Lucentis®, Novartis Pharmaceuticals UK Ltd).[19–22] These studies have shown significant visual and anatomic improvements in cases of diabetic MO using these novel treatments.

These trials show, for the first time, that visual loss in people with diabetic MO can be reversed in approximately half of all treated patients so introducing a new paradigm in the treatment of diabetic MO. It is probable that with these new treatments, screening for diabetic MO will meet the World Health Organization screening criteria.[23] Identifying patients, who will benefit most from these new therapies, is now an important issue for all diabetic retinopathy-screening programmes.

Chapter 2 Study design and methods

Aims and objectives

The primary aim the study was to determine the best method for detecting potentially sight-threatening MO in people with diabetes using photographic surrogate markers within the English and Scottish national screening programmes. Specifically we wished to:

(a) investigate whether or not particular distributions and combinations of lesions [microaneurysms/dot haemorrhages (M/DHs), blot haemorrhages (BHs) and exudates], assessed manually or automatically, are more specific photographic surrogate markers of MO than current practice, using OCT as the reference standard
(b) assess the costs and consequences of using alternative distributions and combinations of these lesions to screen for MO, using either automated or manual detection of lesions
(c) model the long-term cost and quality-of-life implications of using alternative distributions and combinations of surrogate markers to screen for MO.

Once the study was under way several screening programmes were found to be using OCT as part of the screening pathway to reduce false-positive referrals to the hospital eye service. Consequently, we added a further aim to assess the costs and consequences of using OCT within retinal screening programmes in addition to improving the photographic surrogate markers, as this would affect how photographic markers would be used in future.

Primary outcome measure

The primary outcome was the sensitivity and specificity of manual grading, computer-assisted manual annotation grading and fully automated annotation grading strategies, utilising photographic lesions to infer the presence of diabetic MO, compared with a reference standard based on the detection of diabetic MO using OCT.

Other outcome measures

The sensitivity and specificity estimates were used to assess the costs and consequences (i.e. the proportion of appropriate ophthalmology referrals) of using the alternative grading strategies for the detection of MO. The long-term costs and outcomes [visual loss and quality-adjusted life-years (QALYs)] of the alternative grading strategies were modelled using data from the epidemiological literature and available cost estimates. The effect of optionally including OCT within the screening pathway was also modelled.

Study design

The study was a multicentre prospective observational cohort study. The cohort consisted of subjects with features of diabetic retinopathy visible within the macular region attending one of seven diabetic retinal screening programmes. The specific diabetic retinopathy features of interest as surrogate markers for MO were M/DHs, BHs and exudates. All subjects were recruited and imaged at the participating centres in Aberdeen, Birmingham, Dundee, Dunfermline, Edinburgh, Liverpool and Oxford.

© Queen's Printer and Controller of HMSO 2013. This work was produced by Olson *et al*. under the terms of a commissioning contract issued by the Secretary of State for Health. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

Inclusion criteria

- Aged ≥ 18 years.
- One or more of the following features present in at least one eye identified using retinal photography:
  - M/DHs within one disc diameter radius (DD) of the centre of the macula.
  - BHs within one DD of the centre of the macula.
  - Exudates within two DD of the centre of the macula.
- Able and willing to provide signed informed consent.

Main exclusion criteria

- Any macular or pan-retinal laser treatment in the study eye(s) or any intraocular injection, since these interventions affect disease progression and characteristics.
- Intraocular surgery (e.g. cataract surgery) within 1 year of enrolment. Cystoid MO following cataract surgery, also known as Irvine–Gass syndrome, is the most common cause of decreased vision following cataract surgery.[24]
- Pregnancy. During pregnancy diabetic retinopathy can change much more rapidly.[25–27]
- An inadequate OCT image or two inadequate retinal photographs. Photographs are considered inadequate if either (1) the clarity is insufficient as the macular vessels are not clearly visible or (2) the field of view does not include a circular region with a radius of at least two disc diameters centred on the fovea.
- Contraindications to pupillary dilatation should pupil dilatation be necessary. Pharmacological pupillary dilatation is necessary where the pupil is too small to allow imaging, either retinal photography or OCT. Typically the pupil diameter must exceed 4 mm to allow imaging.

Reference standard for macular oedema

Optical coherence tomography was chosen as the reference standard for identifying the presence of diabetic MO.[1] Although other technologies are available for measuring retinal thickness (see *Appendix 1*), OCT has the highest resolution, good repeatability, and is the only method that provides detailed anatomical cross-sections.

Rationale for the macular oedema reference

The reference standard should identify all the cases with MO that would benefit from assessment (though not necessarily treatment) by an ophthalmologist in the eye clinic. The first large study to investigate MO (in the context of determining the efficacy of laser treatment for potentially sight-threatening retinopathy) was the ETDRS.[28] Stereoscopic fundus photography was used to determine whether or not retinal thickening was present and the following criteria were used to define clinically significant MO.

- Thickening within 500 µm of the centre of the macula.
- Exudates within 500 µm of centre of macula with adjacent thickening.
- Thickening of one disc area or larger where any part is within one disc diameter of the centre of the macula.

Optical coherence tomography now provides a more sensitive and specific test for MO than was available for the ETDRS. The ETDRS defined a circular grid consisting of nine regions[4] (see *Figure 29*) where the central region has a radius of 500 µm for the ETDRS central thickening region and the surrounding four regions have a diameter of a DD for the non-central thickening region. All current OCT scanners generate retinal thickness values for these regions of interest. Many studies have investigated the normal range of retinal thickness in these regions of interest. Subjects with abnormally thickened retinas may be selected based on a thickness threshold. For the Zeiss Stratus OCT™ (Carl Zeiss Meditec International, Jena, Germany) studies have shown that biomicroscopy is able to reliably identify central thicknesses > 300 µm,

and less reliably detect thicknesses between 250 and 300 µm. Setting a thickness threshold of 250 µm should therefore ensure that the majority of biomicroscopy-positive cases with central thickening will be included in the study. A Diabetic Retinopathy Clinical Research Network study that used Zeiss Stratus OCT selected eyes with previously untreated MO that were characterised by a central subfield mean thickness of at least 250 µm or an inner paracentral subfield mean thickness of at least 300 µm:[29] these thickness thresholds were used in this study. However, as different OCT scanners were used at the various centres it was necessary to correct for possible differences in thickness measurements, see *Chapter 3*.

Thickening seen on OCT, however, is not specific to diabetic retinopathy. When an abnormal thickness measure was found within the inner five ETDRS regions the OCT cross-sections were examined for intraretinal cysts or subretinal fluid. This is similar to the grading protocol in the Diabetic Retinopathy Clinical Research Network study[29] which states: 'Retinal morphology was assessed at baseline from OCT images for cystoid abnormalities and subretinal fluid'. Although the Diabetic Retinopathy Clinical Research Network study[29] used the selection criterion of a central field thickness threshold of at least 275 µm rather than the more common 250 µm. Other studies note that non-diabetic pathology is specifically excluded from the study (e.g. the RESTORE study[22] and the READ-2 study[21]).

In this study we used the Diabetic Retinopathy Clinical Research Network criteria for defining MO.[29] MO was deemed to be present where the following two criteria were met:

- The central ETDRS region thickness was > 250 µm, or any of the inner five regions were > 300 µm.
- A visible intraretinal cyst or area of subretinal fluid on the OCT cross-sections.

Sample size

A previous study found a 14% prevalence of macular lesions within a screening programme,[30] the subjects making up the target cohort for this study. Of those with macular lesions, 10% are expected to have MO. Thus, the majority of the study cohort will not have MO and hence the precision of the sensitivity measurement will be the determining factor in the calculation of study power and sample size.

To detect a 3% difference in sensitivity between two diagnostic tests with 80% power using the McNemar test requires a sample of 400 positive cases. Using the assumed prevalence of positive cases above of 10% this would mean recruiting 4000 subjects.

Regulatory approval

The study was approved by the North East Scotland Research Ethics Committee on 17 December 2007 (reference 07/S0801/107). Each participating centre obtained approval from their local ethics board. The study was cosponsored by the University of Aberdeen and NHS Grampian. All participants provided signed informed consent after reading the patient information sheet and following discussion with the local study representative. A Trial Steering Committee oversaw the conduct of the study. The study was registered with the United Kingdom Clinical Research Network (UKCRN reference 9063).

Recruitment and data collection

All patients were recruited at one of seven study centres in Aberdeen, Birmingham, Dundee, Dunfermline, Edinburgh, Liverpool and Oxford.

Visual acuity

Visual acuity was measured using either best corrected visual acuity or a pin hole (the method used was noted). Visual acuity was recorded using the log-MAR scale. Where subjects could not resolve characters on the vision chart they were asked if they could count fingers, see hand movement or perceive light, otherwise the eye was recorded as having no perception of light. Count fingers was assigned a log-MAR

© Queen's Printer and Controller of HMSO 2013. This work was produced by Olson *et al*. under the terms of a commissioning contract issued by the Secretary of State for Health. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

value of 2, and hand movement a log-MAR value of 3.[31] Where vision is worse than hand movement the subject is unable to resolve any object and therefore the visual acuity is undefined.

Pupillary dilatation

In the English centres all patients received mydriasis, while Scottish centres only used mydriasis if the pupil size was too small for imaging. Both retinal photograph and OCT scanning require a pupil size of at least 4 mm. Studies have shown that the use of pupillary dilatation does not affect OCT thickness measurements.[32,33]

Retinal photograph acquisition

All photographs were acquired by ophthalmic photographers or retinal screeners and the name of the photographer was noted for each image. A single retinal photograph was acquired for each eye meeting the following criteria:

- 45 degrees field of view
- macula centred
- colour digital photograph (between 3 and 8 megapixels)
- JPEG image compression (if used) set for high quality
- adequate field of view (the image should show a region having a radius of at least two disc diameters around the fovea)
- adequate clarity (i.e. adequate to see macular microaneurysms, if present)
- the fundus camera small pupil facility was acceptable providing a region of at least two DD was still visible around the fovea.

Optical coherence tomography reference image acquisition

All OCT scans were acquired by operators who had been accredited for the study. There was a maximum time limit of 4 weeks between the retinal photograph and OCT reference scan as the disease was unlikely to progress significantly during this period of time. Eighty-nine per cent of OCT scans were acquired on the same day as the retinal photograph.

Accreditation

As in other multicentre imaging studies, to avoid intercentre variation all OCT operators were required to be accredited before submitting data for the study. Operators submitted a portfolio that included the following images, collected using the OCT scanner they would be using for the study:

1. *Normal eye* Repeat macula maps of the same normal eye as per scanner model protocol.
2. *MO eye* Repeat macula maps of an eye showing obvious MO (i.e. central thickness of at least 300 µm).

The images were uploaded using the same website as the study data. Images were checked for:

1. *Foveal position (where visible)* The foveal minimum should be within 250 µm of the centre of the thickness map.
2. *Adequate image quality* Image quality was assessed by visual inspection as well as quantitative parameters such as signal strength and standard deviation (SD) in the central measurement, where present.
3. *Repeatability* Repeatability was assessed for each region as the absolute percentage difference between the repeat scans. These should be < 10% in all regions.

Optical coherence tomography data description

Although the precise OCT acquisition protocol depended on the model of scanner used, every scanner was set up to provide the following data.

- A nine-region (ETDRS style) map showing average regional thickness measured in microns.
- A horizontal cross-sectional view through the centre of the macula.
- (Optional.) If the cross-sectional view through the centre of the macula did not contain the region of greatest thickening then a second cross-section that included the region of greatest thickening was taken.

Other recorded patient information

No patient-identifiable information left the recruiting centre. Patient identifiers were removed from images and data such as subject age were recorded with insufficient granularity to be of help in identification. In addition to the retinal photograph, OCT scan and visual acuity assessment the following subject information was collected:

- age (rounded to the nearest year)
- gender
- amblyopia.

Amblyopia is an uncorrectable decrease in vision in one eye with no apparent structural abnormality seen to explain it. It is a diagnosis of exclusion, meaning that when a decrease in vision is detected, other causes must be ruled out. There is wide variation in reported estimates for its prevalence. A review of UK amblyopia studies in children aged ≤ 5 years assumed a prevalence of 4.8%.[34] A cohort study at an English retinal screening programme recorded a 10% amblyopia prevalence.[35] Since visual acuity is used as an indicator of MO in some screening programmes, amblyopia represents a confounding factor in the detection of MO. Other non-diabetic factors affecting visual acuity, such as lenticular opacities and macular degeneration, were not recorded separately as, unlike amblyopia, they directly affect the appearance or quality of the retinal photograph:

- type of diabetes (the type of diabetes was recorded as type 1, type 2, secondary or other)
- ethnicity (the ethnicity categories were Asian, Black, Caucasian, Chinese, mixed, other, unknown)
- first half of postcode
- glitazone use within the previous 6 months.

Glitazones, or thiazolidinediones, are a group of drugs that are prescribed to increase sensitivity to insulin in people with type 2 diabetes. The most common forms are rosiglitazone and pioglitazone. Rosiglitazone had its marketing authorisation suspended during the study in September 2010 by the European Medicines Agency because of concerns regarding an increased cardiovascular risk. Oedema is a known risk factor when using glitazones and studies have suggested they are associated with an increased risk of MO,[34–36] although another study found no association.[37] The recording of the use of glitazones was added to the study protocol in June 2009 because of the uncertainty surrounding their role in the development of MO.

Web submission and data validation

A website was developed to enable the transfer of both subject information and image data from each recruiting centre to the database in Aberdeen. The website also allocated the unique study identifier for each subject and printed out a reference sheet for the local study folder.

The website was designed to reduce data entry errors. Tick boxes, menus and calendar entry tools were used in preference to plain text entry (*Figure 1*). Checks were performed on the entered data, for instance that the subject was aged ≥ 18 years and therefore eligible for the study. Checks were also performed on the image data that were uploaded to ensure that the images were the expected size and format for the OCT scanner used at that centre.

© Queen's Printer and Controller of HMSO 2013. This work was produced by Olson *et al*. under the terms of a commissioning contract issued by the Secretary of State for Health. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

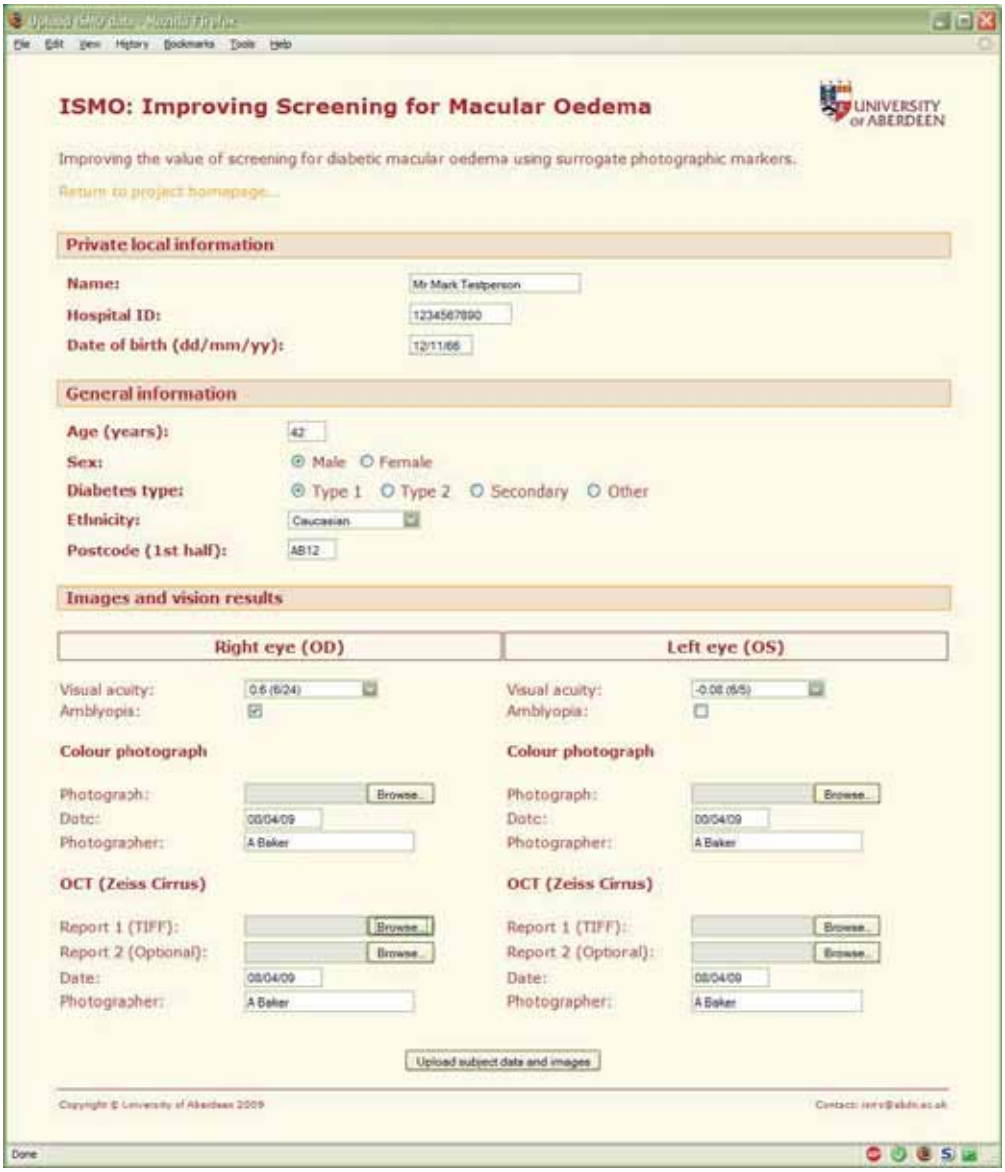


FIGURE 1 Web form used to upload study data.

Check sums were calculated for all images when they were uploaded. This served two purposes: first, it allowed the continuing integrity of the image data to be verified and, second, it enabled attempted duplicate image uploads to be rejected.

Image data analysis and grading

Retinal photographs

All the retinal photographs were graded and annotated by the same research nurse, who had 3 years' experience working as a retinal grader. A screenshot of the software used for grading is shown in *Figure 2*.

Image grading

All of the images were graded for image quality and severity of retinopathy and maculopathy following the Scottish Diabetic Retinopathy Grading Scheme.[38]

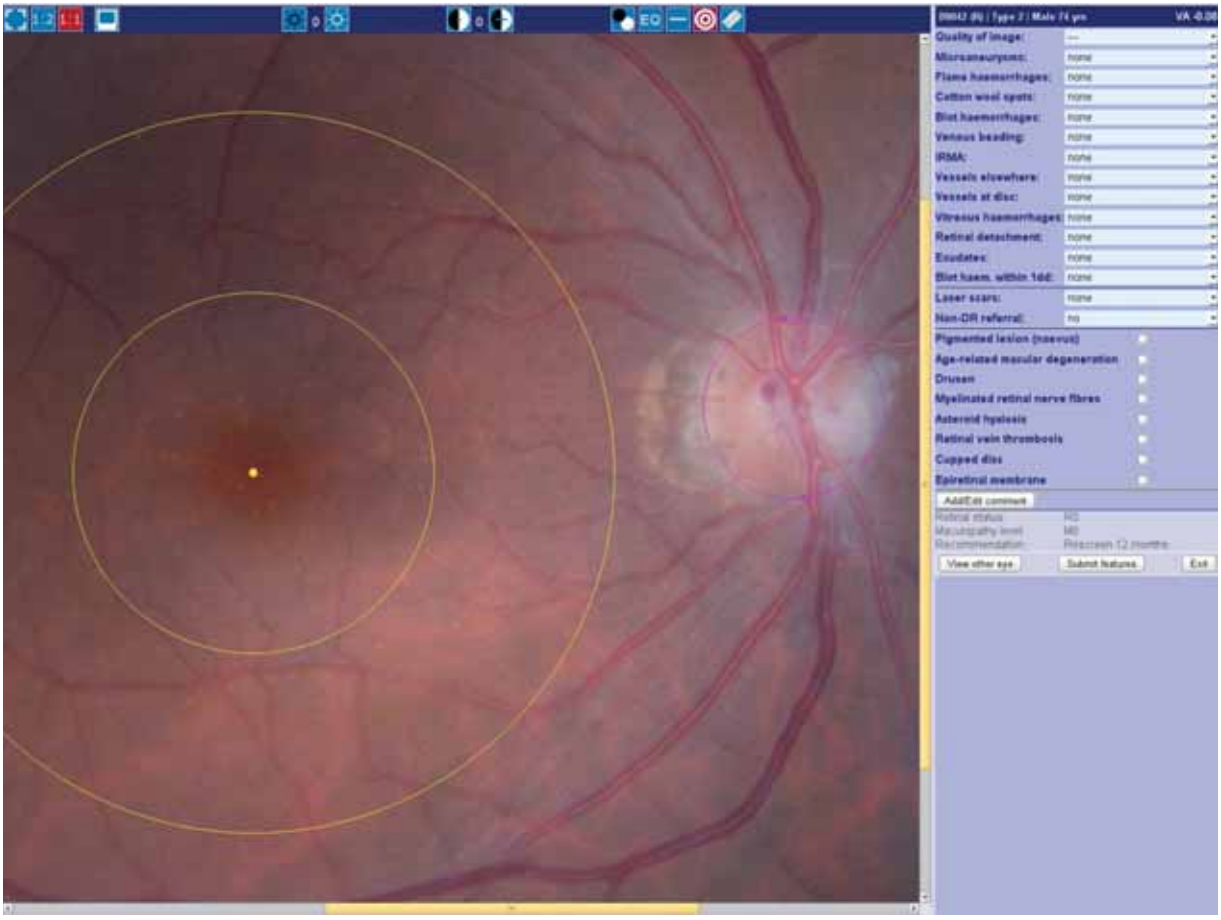


FIGURE 2 Screenshot of the software used to grade images.

Image annotation

Software was developed to enable the research nurse to annotate the retinal images, indicating the size and position of the optic disc, the position of the fovea, and the location, shape and size of individual lesions within two DD of the fovea (*Figure 3*). All the lesions associated with maculopathy (M/DHs, BHs and exudates) were annotated, as well as non-diabetic features with similar appearance which could confound the analysis [flame haemorrhages, drusen and cotton wool spots (CWS)].

Optical coherence tomography reference standard

Optical coherence tomography scans

The OCT scan was the study reference standard for determining whether or not oedema was present. The reference was based on the nine region thickness map and visual inspection of the cross-sectional images.

Thickness map

Each OCT scanner in the study displays the thickness map on a graphical report. Optical character recognition software was used to automatically store the thickness values in the database.

The retinal thickness was considered abnormal if the centre region thickness was > 250 µm or > 300 µm in any of the surrounding four regions (the outer four regions outside a 3 mm radius were not used). These thickness thresholds were taken from studies that used the Zeiss Stratus OCT, and were adjusted to account for the scanners used in the study (see *Chapter 3*). Images with abnormal thickening were then visually inspected for the presence of oedema.

© Queen's Printer and Controller of HMSO 2013. This work was produced by Olson *et al*. under the terms of a commissioning contract issued by the Secretary of State for Health. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

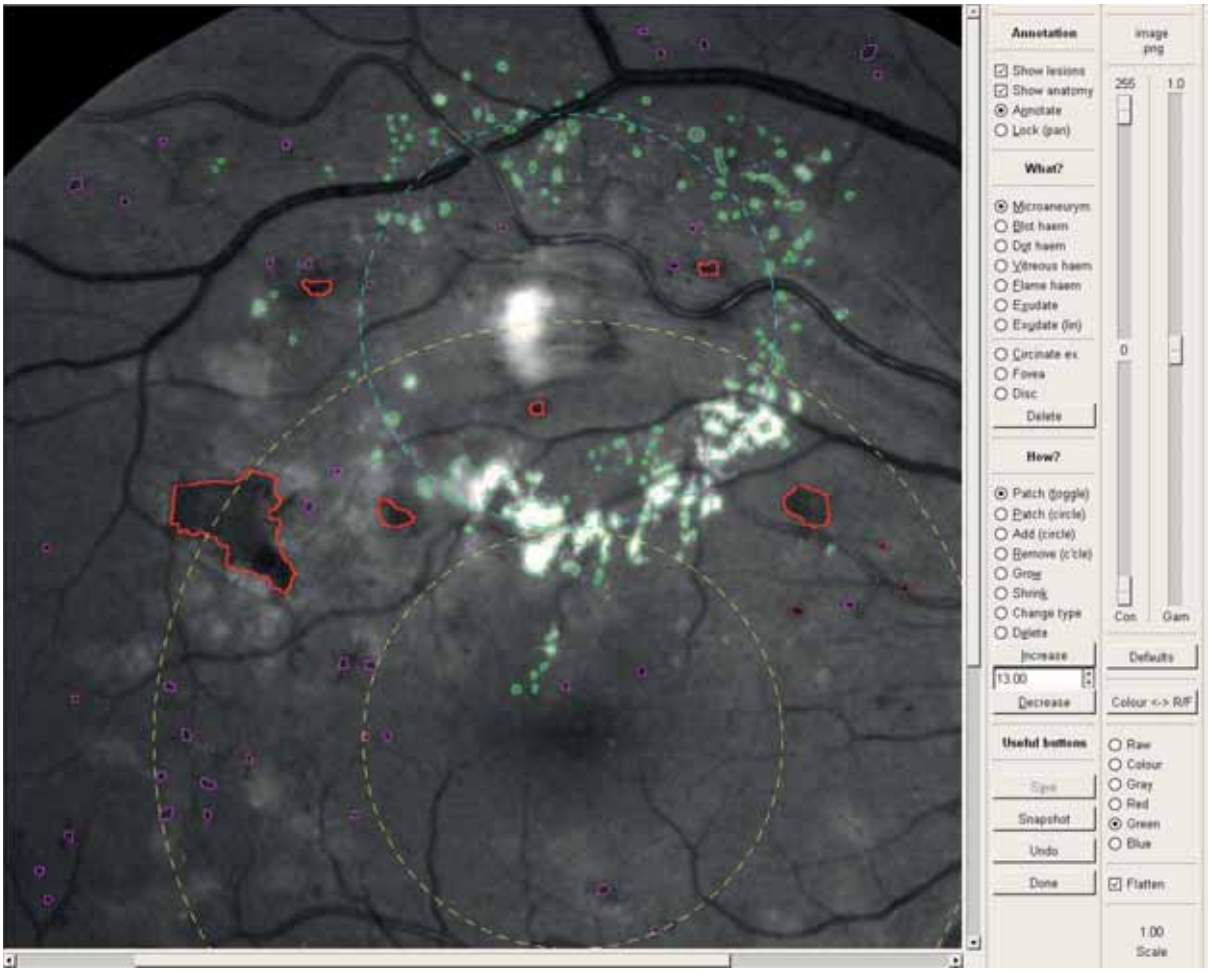


FIGURE 3 Screenshot of software used to annotate features on the retinal photographs.

Cross-sectional images

The cross-sectional images were visually assessed for the presence of oedema because several pathologies, besides diabetic MO, could affect retinal thickness. Confounding features include vitreo-macular traction, epiretinal membranes, macular holes, drusen, subretinal haemorrhage, pigment epithelial detachment and choroidal neovascular membrane. Example OCT thickness map and cross-sections are shown in *Figure 4*.

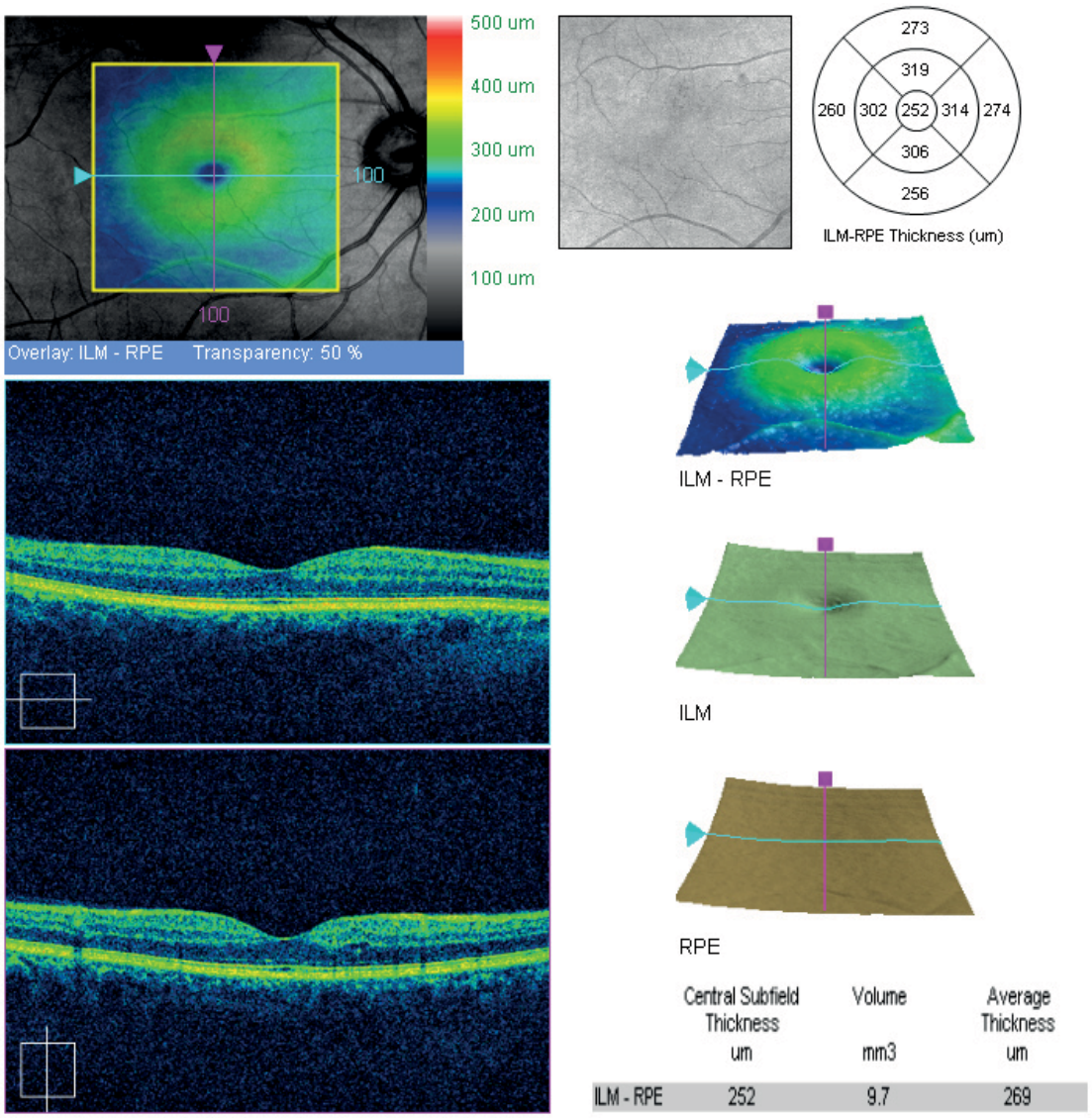


FIGURE 4 Example of OCT cross-sections and thickness map (report from Zeiss Cirrus OCT).

© Queen's Printer and Controller of HMSO 2013. This work was produced by Olson *et al*. under the terms of a commissioning contract issued by the Secretary of State for Health. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

Chapter 3 Comparison of optical coherence tomography scanner thickness measurements

When the study was originally planned the OCT market was dominated by a single device, the Zeiss Stratus OCT, which had been available since 2001. Subsequently, a number of new scanners came on the market offering faster acquisition, lower image noise and higher resolution.

As a result it was decided not to limit the study to one scanner. Furthermore, the advantages offered by the newer scanners meant that the Stratus OCT would rapidly become obsolete and so patient recruitment would become difficult.

In order to include results from different OCT scanners, experiments were undertaken before the study began to estimate the differences in thickness between the new scanners and the Stratus OCT, and so allow corrections to be made.

This study was considered a service evaluation by the North East of Scotland Research Ethics Committee on 20 August 2007. All subjects gave informed consent after the intent of the evaluation was explained to them.

Subjects

Forty-two volunteers aged $\geq$ 18 years took part. Sixteen of the volunteers were male and the median age was 37 years (interquartile range 28–43 years). None of the subjects had a history of eye disease or diabetes. Each subject had both eyes scanned twice by the same operator on at least one of the four OCT scanners used in the study. Fifteen subjects were scanned on a single scanner, 17 on two scanners and 10 on more than two scanners. Pharmacological pupillary dilatation was not used as other studies have shown it has no effect on thickness measurements.[32,33] There was no significant difference in the age (Wilcoxon signed-rank test) or male/female ratio (chi-squared test) between the subjects imaged on each scanner.

Measurements from left eyes were reflected left to right such that the nasal region was to the right and the temporal region to the left in both left and right eyes.

Statistical analysis

Both eyes were included in the analysis since no correlation was expected between repeat measurements in left and right eyes. This reduced the number of subjects required and also allowed investigation of temporal/nasal asymmetry. Repeatability was calculated as:

$$\text{Repeatability} = 1.96 \times \text{std}(D); \quad (1)$$

where std(D) is the SD of the repeat measurement differences.[39] Confidence intervals (CIs) were calculated for repeatability as $\pm (t_{n-1,0.05}/\sqrt{(2(n-1))}) \times 1.96 \times \text{std}(D)$, where n is the number of eyes and $t_{n-1,0.05}$ is the inverse cumulative t-distribution for a 5% probability.[40] The intraclass correlation coefficient (ICC) was also calculated as described by Shrout and Fleiss.[41] Interscanner agreement was calculated as for intrascanner repeatability using the equation above. Repeatability and ICC values were obtained from SPSS (version 17; SPSS Inc., Chicago, IL, USA). The difference between scanner measurements was plotted against the mean of the scanner measurements to detect any thickness dependence on the difference between scanners.

© Queen's Printer and Controller of HMSO 2013. This work was produced by Olson *et al*. under the terms of a commissioning contract issued by the Secretary of State for Health. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

As not all subjects were measured on all four scanners a mixed-effects model was used in preference to a standard analysis of variance.[42] It was implemented using the PROC MIXED procedure within SAS (version 9.1; SAS Institute Inc., Cary, NC, USA). The scanners were treated as fixed effects, using the Zeiss Stratus OCT as the reference scanner. Differences between volunteers' left and right eyes were also treated as a fixed effect, while the differences between volunteers were considered a random effect. The size of the differences between scanners could thus be estimated while allowing for differences between volunteers and between eyes within volunteers. Analyses were also carried out on the means of the annular inner four and outer four ETDRS regions.

Two schemes were compared for converting values between different scanners. The first applied a single additive constant, calculated as the mean difference between scanners, to all nine regions. The second used three separate additive constants for the central (region 1), inner regions (2, 3, 4 and 5) and outer regions (6, 7, 8 and 9). The mixed-model analysis was repeated in each case using the adjusted values to test whether or not the differences between the scanners remained significant.

Imaging protocols

Zeiss Stratus OCT

The 'fast macular protocol' was used on the Stratus OCT. It acquires six intersecting 6 mm radial cross-sections centred on the fovea. Each cross-section includes 128 A-scans of 1024 axial samples over a 2 mm depth range. The total acquisition takes 1.9 seconds. The scans were considered acceptable if (1) the SD of the central foveal measurements was < 10% of the mean central measurement, (2) the recorded signal strength was at least 4 out of a possible maximum of 10, (3) there were no warnings from the instrument regarding 'low analysis confidence', 'missing data' or 'high variance', (4) there were no visible boundary tracking errors and (5) the fixation point was within approximately 250 µm of the centre of the foveal pit. The scanner software version was 4.0.7.

Zeiss Cirrus OCT™ (Carl Zeiss Meditec International, Jena, Germany)

The Cirrus OCT (Carl Zeiss Meditec International, Jena, Germany) acquires A-scan data almost 70 times faster than the Stratus OCT. Scans were collected covering a 6 × 6 mm square area with a 512 × 128 raster pattern. Each A-scan consisted of 1024 samples over a 2 mm depth range. The total acquisition time was 2.4 seconds. A faster protocol is available, using a 200 × 200 raster pattern, which covers the same area in 1.5 seconds. The scans were considered acceptable if (1) the signal strength was at least 6 out of 10, (2) there were no obvious signs of movement, (3) there were no visible boundary tracing errors and (4) the fixation point was within 250 µm of the centre of the foveal pit. The scanner software version was 2.0.0.54.

Topcon 3D OCT-1000™ (Topcon Corporation, Tokyo, Japan)

Scans were acquired covering a 6 × 6 mm square area to a depth of 1.68 mm using a 256 × 128 raster pattern of A-scans. Each acquisition took approximately 2 seconds. The scans were considered acceptable if (1) there were no visible boundary tracing errors, (2) there were no obvious signs of movement, (3) < 10% of B-scan lines were missing in any ETDRS region and (4) the fixation point was within 250 µm of the centre of the foveal pit. The scanner software version was 2.11.

Heidelberg Spectralis™ (Heidelberg Engineering, Heidelberg, Germany)

The Spectralis A-scan rate of 40,000 A-scans/second is the fastest in this group. Its acquisition also differs from the other scanners in two significant respects. First, it includes an eye tracking facility which pauses the acquisition whenever the subject moves or blinks (consequently, the scan time can vary depending on subject movement). Second, each B-scan cross-section is acquired several times and combined to improve the signal-to-noise ratio and reduce speckle artefacts.

Scans were acquired for a nominal 30 × 30 degree region using a 768 × 256 raster pattern, with each B-scan cross-section acquired five times. Each A-scan consisted of 512 samples covering a 1.8 mm depth range. The scans were considered acceptable if (1) the 6 mm ETDRS region fitted within the acquired region and (2) there were no boundary tracking errors. Software allowed the regions to be translated to correct for minor fixation errors. Boundary tracings were not available for all images and so the exclusion criteria were less stringent than for the other scanners. The software version used was 3.0.

Results

Boundary measurements

Differences in how boundaries are identified by the various scanners were a significant problem. Boundary A (*Figure 5*) indicates reflections from the nerve fibre layer on the top surface of the retina; all the scanners begin their thickness measurements at this layer. However, there is disagreement about where to place the lower limit of the thickness measurement. Boundary B, probably reflections from the external limiting membrane, tends to be fainter than the three lower boundaries C, D and E. The Zeiss Stratus OCT takes the top of the bright reflective band (approximately layer C above) as the posterior limit (measurement M1). In contrast, the Heidelberg Spectralis takes the bottom of the bright reflective band (layer E, measurement M3). The Zeiss Cirrus OCT and Topcon 3D OCT-1000 use a limit between these two extremes, with the Topcon 3D OCT-1000 tending to find the lower edge of layer C and the Zeiss Cirrus OCT the top edge of layer D (measurement M2).

Scans meeting inclusion criteria

Eyes were excluded if either of the repeat scans failed to meet the inclusion criteria. The number of eyes included from each scanner were: the Zeiss Stratus OCT (45/60), the Zeiss Cirrus OCT (55/66), the Topcon 3D OCT-1000 (17/26) and the Heidelberg Spectralis (27/28).

Four Zeiss Stratus OCT scans had visible boundary errors (*Figure 6*), two had low analysis confidence warnings, one had a fixation error, one had missing data and two had high variance warnings. On the Zeiss Cirrus OCT, fixation errors affected 10 scans and one was excluded because of movement. On the Topcon 3D OCT-1000, 13 scans showed missing data, which in nine cases was serious enough for the scans to be excluded. One Heidelberg Spectralis scan was excluded owing to a missing thickness value in one region, probably because of a boundary tracing error.

Repeatability

Both eyes were included in the repeatability analysis since the repeat measurement differences are expected to be random and uncorrelated. In this study the correlation between thickness measurements in left and right eyes was strong (Pearson correlation coefficient 0.780), but there was no significant correlation in the repeat measurement differences in left and right eyes (Pearson correlation coefficient 0.014; $p = 0.9$).

Table 1 lists the repeatability and ICCs for the nine ETDRS regions. In general, repeatability was poorest in the smallest, central region (1) and best in the inner regions (2, 3, 4 and 5). The Zeiss Stratus OCT shows the poorest repeatability, with all regions being equal to, or worse than, the newer scanners.

Interscanner agreement

Box and whisker plots of the macular thickness in each of the nine regions are shown in *Figure 7*. By inspection it is clear that there were differences between all the scanners and that the sizes of these differences were similar for all regions.

The mean central thickness in volunteers was greatest on the Heidelberg Spectralis (277 µm, SD 15 µm) and least on the Zeiss Stratus OCT (201 µm, SD 19 µm). The Topcon 3D OCT-1000 (230 µm, SD 22 µm) and Zeiss Cirrus OCT (258 µm, SD 21 µm) were between these two extremes. In all scanners the average

© Queen's Printer and Controller of HMSO 2013. This work was produced by Olson *et al*. under the terms of a commissioning contract issued by the Secretary of State for Health. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

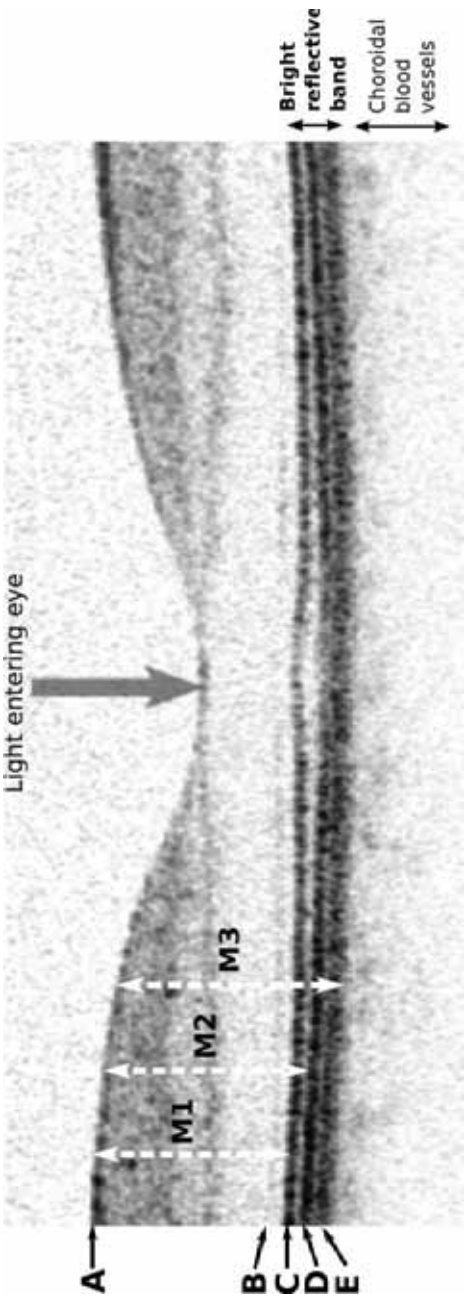


FIGURE 5 Optical coherence tomography cross-section through the centre of the macula showing the normal fovea (pit) at the region of highest acuity vision. As shown here, darker grey levels represent stronger signals. A number of highly reflective (i.e. dark) interface boundaries are visible on this image acquired with a Heidelberg Spectralis scanner. Boundary A indicates reflections from the nerve fibre layer on the top surface of the retina, whereas boundary B, which is probably reflections from the external limiting membrane, tends to be fainter than the three lower boundaries C, D and E. Earlier scanners were unable to resolve C, D and E, which were generally lumped together as the 'bright reflective band'.

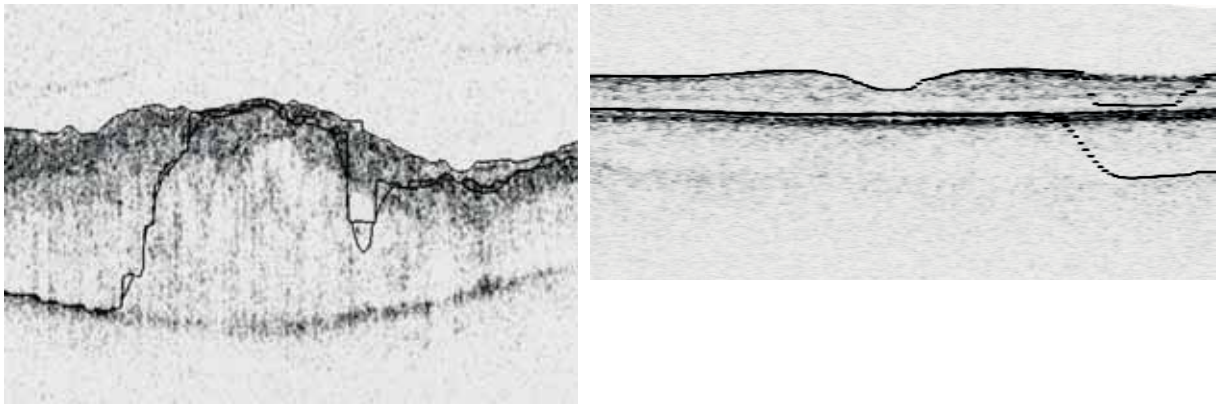


FIGURE 6 Example of an automatic boundary detection failure. Left, Topcon 3D OCT-1000 example from a patient with MO; right, a Zeiss Stratus OCT scan from a volunteer.

thickness in the superior region was greater than in the inferior region, and likewise greater in the nasal compared with the temporal region.

Table 2 lists the results of the mixed-model analysis. It shows the estimated differences between scanners calculated using scans from both eyes meeting the inclusion criteria. There were significant differences in thickness between all of the scanners in all nine regions ($p < 0.001$). In the central region the measurements were greater on the Zeiss Cirrus OCT, Topcon 3D OCT-1000 and Heidelberg Spectralis scanners than the Zeiss Stratus OCT by 58 µm, 28 µm and 78 µm, respectively. Compared with the central region differences, the other eight regional differences were slightly smaller on all scanners except the Topcon 3D OCT-1000, where the largest differences were in the inner regions.

As agreement in its technical sense is proportional to the SD of measurement differences,[39] any systematic difference is disregarded. The central region macular thickness agreement between the Zeiss Stratus OCT and Zeiss Cirrus OCT was 10.5 µm ($n = 30$; 95% CI 7.1 to 13.9 µm), between the Zeiss Stratus OCT and Topcon 3D OCT-1000 was 6.9 µm ($n = 13$; 95% CI 3.26 to 10.5 µm) and between the Zeiss Stratus OCT and Heidelberg Spectralis was 12.7 µm ($n = 14$; 95% CI 6.3 to 19.1 µm).

Interscanner thickness conversion

Using a single correction for all nine regions and rerunning the mixed-model analysis resulted in very similar measurements from all the scanners for the central region. However, statistically significant differences still remained; in the outer regions the Zeiss Cirrus OCT and Heidelberg Spectralis were up to 16 µm lower than the Zeiss Stratus OCT and the Topcon 3D OCT-1000 up to 7 µm higher. Refining the model and using three regional constants per scanner, where correction values were calculated for the central, inner and outer regions, resulted in very few statistically significant differences remaining and the average errors were now all ≤ 3 µm, with the largest errors in the inner nasal region.

The above strategy assumes additive correction alone is sufficient. The need for possible multiplicative correction was tested by plotting the difference between measurements from two scanners against the mean thickness measured on the two scanners. *Figure 8* shows the comparison of Zeiss Stratus OCT and Zeiss Cirrus OCT measurements for the central region. If an additive constant is sufficient to correct the difference then the line of regression should have a zero gradient. However, as *Figure 8* shows, the regression line has a gradient of –0.17, which is significantly different from a zero gradient ($p < 0.001$). Although only regions 1 and 9 are significantly different from zero gradient, probably owing to the small sample size, the bar chart in *Figure 9* of estimated gradients for all nine regions shows a clear trend, with negative gradients in the central and inner regions and positive gradients in the outer four regions.

© Queen's Printer and Controller of HMSO 2013. This work was produced by Olson *et al*. under the terms of a commissioning contract issued by the Secretary of State for Health. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

TABLE 1 Repeatability measures (*R*) and ICC for normal volunteers from the four scanners for each ETDRS region (95% CI shown in parentheses)

ETDRS region	Zeiss Stratus OCT (*n* = 45)		Zeiss Cirrus OCT (*n* = 55)		Topcon 3D OCT-1000 (*n* = 17)		Heidelberg Spectralis (*n* = 27)	
	R (μm)	ICC	*R* (μm)	ICC	*R* (μm)	ICC	*R* (μm)	ICC
1	12.16 (9.54 to 14.77)	0.948 (0.908 to 0.971)	6.74 (5.44 to 8.04)	0.987 (0.978 to 0.993)	6.10 (3.81 to 8.39)	0.990 (0.974 to 0.996)	6.03 (4.31 to 7.75)	0.981 (0.958 to 0.991)
2	7.23 (5.68 to 8.79)	0.955 (0.919 to 0.975)	7.68 (6.20 to 9.16)	0.954 (0.923 to 0.973)	3.29 (2.06 to 4.52)	0.982 (0.951 to 0.993)	4.44 (3.18 to 5.71)	0.980 (0.956 to 0.991)
3	8.15 (6.40 to 9.90)	0.949 (0.909 to 0.972)	5.74 (4.64 to 6.85)	0.979 (0.964 to 0.988)	5.15 (3.22 to 7.08)	0.965 (0.909 to 0.987)	3.15 (2.25 to 4.04)	0.991 (0.981 to 0.996)
4	8.85 (6.95 to 10.75)	0.924 (0.867 to 0.957)	6.76 (5.46 to 8.06)	0.968 (0.946 to 0.981)	4.11 (2.57 to 5.65)	0.972 (0.925 to 0.990)	4.78 (3.42 to 6.14)	0.977 (0.951 to 0.990)
5	9.57 (7.52 to 11.63)	0.908 (0.840 to 0.949)	6.26 (5.05 to 7.47)	0.972 (0.953 to 0.984)	3.33 (2.08 to 4.57)	0.980 (0.948 to 0.993)	3.54 (2.53 to 4.55)	0.988 (0.974 to 0.994)
6	9.56 (7.51 to 11.62)	0.923 (0.892 to 0.957)	8.44 (6.81 to 10.07)	0.936 (0.892 to 0.962)	5.55 (3.47 to 7.63)	0.970 (0.920 to 0.989)	4.79 (3.42 to 6.15)	0.979 (0.955 to 0.990)
7	7.99 (6.27 to 9.71)	0.963 (0.935 to 0.980)	6.33 (5.11 to 7.55)	0.970 (0.949 to 0.982)	4.44 (2.78 to 6.11)	0.980 (0.948 to 0.993)	4.59 (3.28 to 5.90)	0.983 (0.963 to 0.992)
8	9.78 (7.68 to 11.89)	0.942 (0.897 to 0.968)	6.91 (5.58 to 8.24)	0.956 (0.927 to 0.974)	5.48 (3.42 to 7.53)	0.978 (0.941 to 0.992)	5.45 (3.89 to 7.00)	0.971 (0.938 to 0.987)
9	10.87 (8.53 to 13.20)	0.903 (0.831 to 0.945)	6.13 (4.94 to 7.31)	0.974 (0.956 to 0.985)	4.54 (2.84 to 6.25)	0.960 (0.895 to 0.985)	6.70 (4.79 to 8.61)	0.953 (0.901 to 0.978)

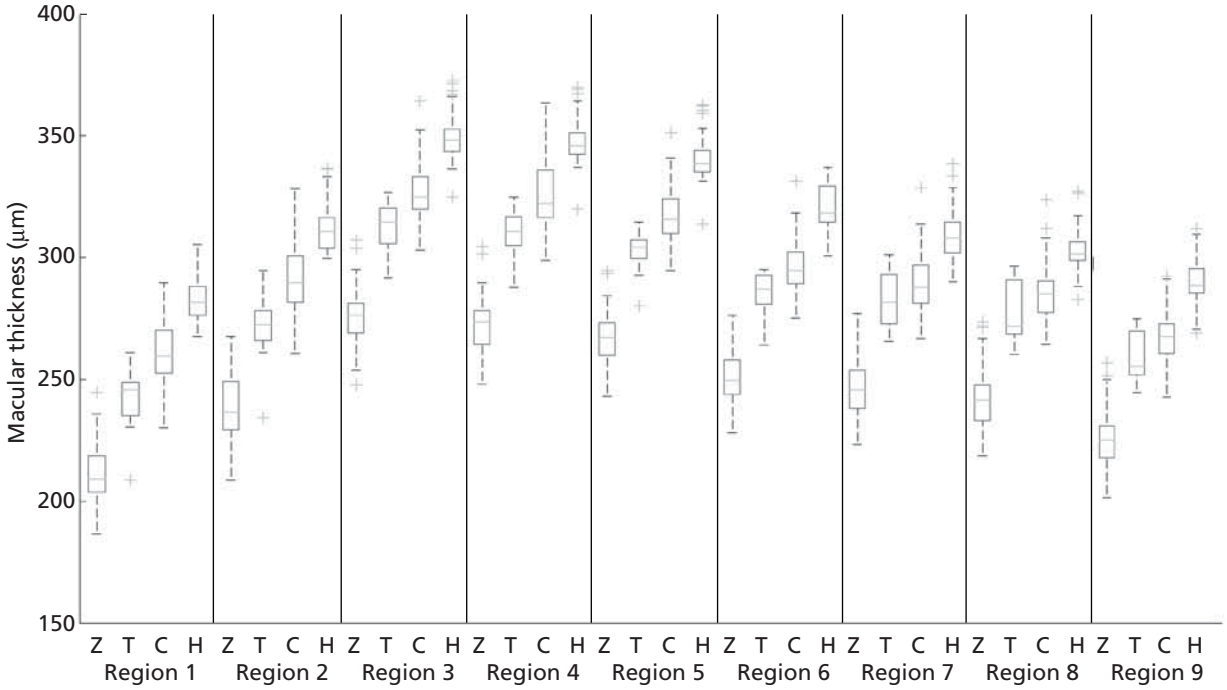


FIGURE 7 Box and whisker plots of macular thickness measurements for the nine ETDRS regions shown in *Figure 29* for the four scanners. The plots show the median (central stripe), interquartile range (box) and full range (whiskers), excluding outliers (+) which are defined as observations beyond 1.5 times the interquartile range from the box. C, Zeiss Cirrus OCT; H, Heidelberg Spectralis; T, Topcon 3D OCT-1000; Z, Zeiss Stratus OCT.

TABLE 2 Differences between other scanners and the Zeiss Stratus in each for each ETDRS region based on a mixed-model analysis

	Zeiss Cirrus OCT vs. Zeiss Stratus OCT (20 volunteers; 30 eyes)		Topcon 3D OCT-1000 vs. Zeiss Stratus OCT (7 volunteers; 13 eyes)		Heidelberg Spectralis vs. Zeiss Stratus OCT (10 volunteers; 14 eyes)	
ETDRS region	Diff. (µm)	95% CI	Diff. (µm)	95% CI	Diff. (µm)	95% CI
1	58.2	56.8 to 59.7	28.1	26.0 to 30.1	78.5	76.5 to 80.4
2	49.9	48.5 to 51.3	32.9	30.9 to 35.0	75.0	73.1 to 76.9
3	54.1	52.6 to 55.5	35.0	32.9 to 37.1	79.0	77.0 to 80.9
4	50.3	48.8 to 51.7	32.2	30.1 to 34.2	74.5	72.6 to 76.4
5	49.7	48.3 to 51.1	30.1	28.2 to 32.1	75.5	73.7 to 77.4
6	44.7	43.0 to 46.4	30.3	27.8 to 32.7	66.6	64.4 to 68.8
7	45.6	44.2 to 47.0	29.8	27.9 to 31.7	64.3	62.5 to 66.1
8	44.4	42.9 to 46.0	28.2	26.0 to 30.4	62.7	60.6 to 64.7
9	43.3	41.8 to 44.8	27.0	24.8 to 29.2	67.5	65.5 to 69.5
Mean inner	51.0	49.8 to 52.2	32.5	30.8 to 34.2	76.0	74.4 to 77.6
Mean outer	44.5	43.4 to 45.7	28.7	27.1 to 30.4	65.3	63.8 to 66.8

Diff., difference.

© Queen's Printer and Controller of HMSO 2013. This work was produced by Olson *et al*. under the terms of a commissioning contract issued by the Secretary of State for Health. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

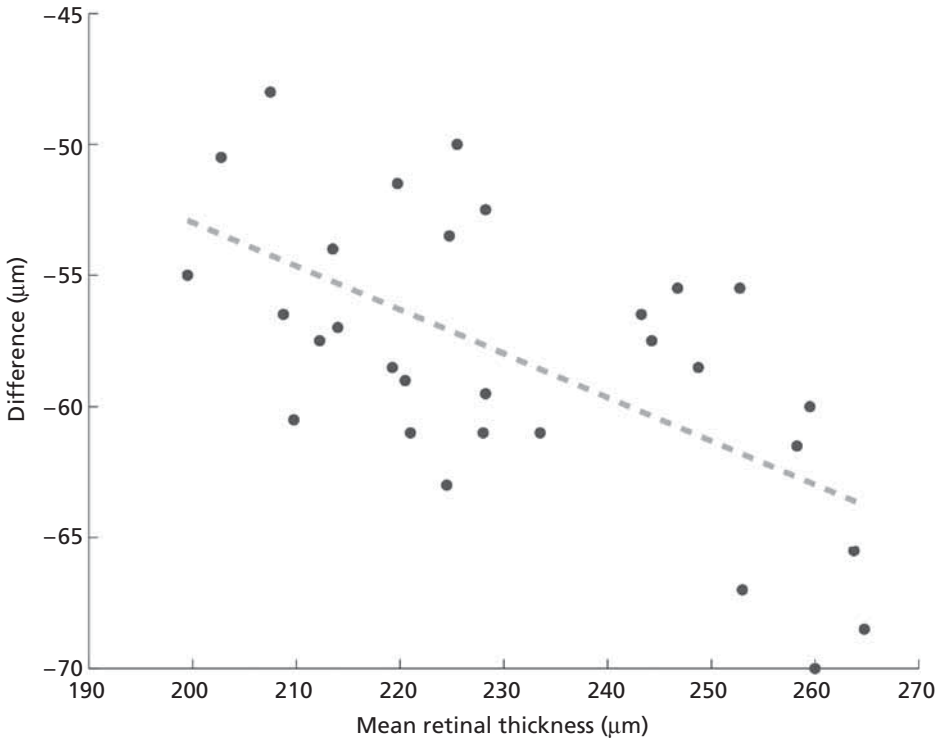


FIGURE 8 Graph of the difference between the Zeiss Stratus OCT and Zeiss Cirrus OCT measurements vs. average thickness. The dashed line shows the least squares regression.

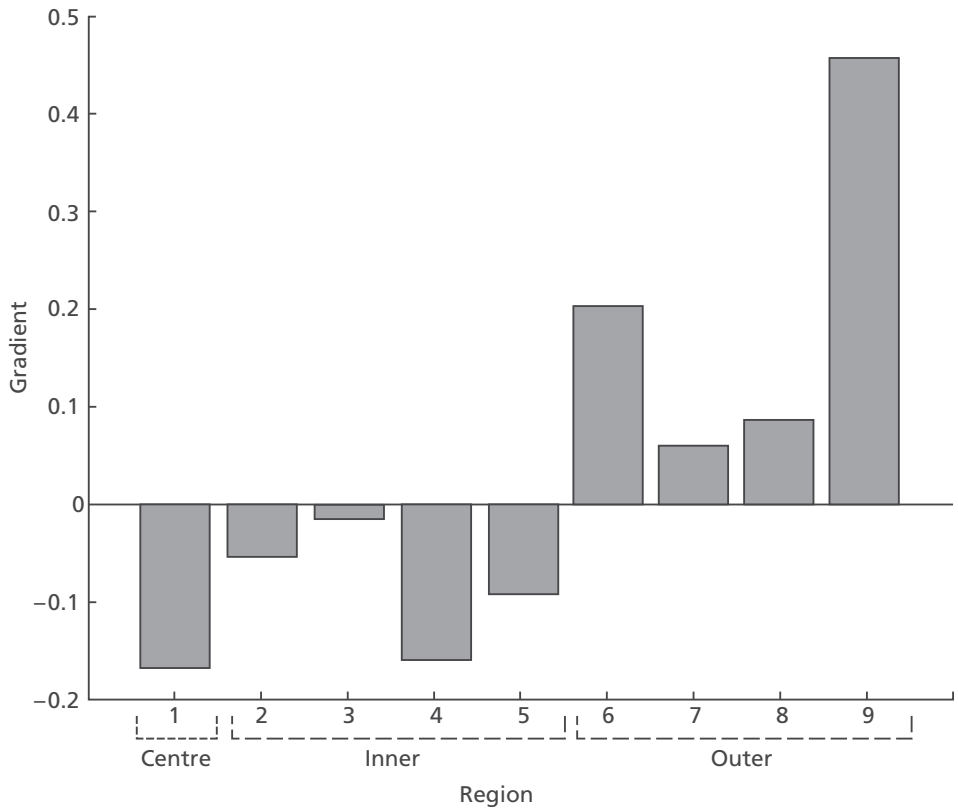


FIGURE 9 Bar chart showing the regression coefficients (slope) for each ETDRS region for plots of difference in Zeiss Stratus OCT and Zeiss Cirrus OCT vs. mean thickness.

Conclusions

Thickness measurements differ significantly between scanners. An additive correction using three separate correction values, for the central, inner and outer regions, resulted in average errors of ≤ 3 µm.

As explained in *Chapter 2*, there were two criteria for judging that MO was present. The first was that the central ETDRS region thickness should be > 250 µm, or any of the inner five regions should be > 300 µm. The final decision on the presence of MO was made by examination of the OCT images for the presence of intraretinal cyst or subretinal fluid. As the thickness measurement was primarily being used to identify those images that should be subjected to a visual examination, it was judged that the accuracy provided by using the additive correction was sufficient.

© Queen's Printer and Controller of HMSO 2013. This work was produced by Olson *et al*. under the terms of a commissioning contract issued by the Secretary of State for Health. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

Chapter 4 Characteristics of study data

Introduction

This chapter describes the data collected for this study. An initial analysis focuses on describing the subject characteristics and the lesions that were present. This is followed by comparisons of subjects with and without MO according to OCT.

Methods for demographics

The demographics of the data are described for all subjects and separately for those subjects whose MO status could and could not be determined based on the information from OCT of both eyes. Univariate methods are used to identify any substantial differences between them. Thereafter, subjects with undetermined MO status are excluded from all analyses. Demographic data for subjects with MO in either eye (or both) are compared with those without MO using simple univariate methods to identify potentially important subject characteristics.

The proportions of subjects identified as having retinal thickening in the central region and within ETDRS regions 1–5 in four different scanners were compared. The data from each centre were described in terms of the number of subjects included and excluded, and the presence of MO in one or both eyes. The presence or absences of different lesions per eye within one DD, and for exudates within one to two DD, are compared across centres.

Combinations of lesions (M/DHs, BHs and exudates) are then considered. Subjects are classified into mutually exclusive categories according to their most serious feature in the eye under consideration. The categories are: exudates present within one DD (regardless of other lesions), BHs present within one DD (but not exudates), M/DHs present within one DD (but not exudates or BHs), no lesions present within one DD, or lesions other than M/DHs, BHs or exudates present within one DD. These lesion combinations are also compared across centres by eye. Visual acuity is considered as a three category variable visual acuity better (log-MAR < 0.3; Snellen 6/9.5 or better), visual acuity worse (log-MAR ≥ 0.3; Snellen 6/12 or worse) and visual acuity missing. The presence of MO is considered in relation to these categories. More complex analyses are left for the statistical modelling described in *Chapter 5*.

Results

Of 3540 subjects recruited into the study, 370 could not be used for the study (*Figures 10* and *11*) according to the exclusion criteria given in *Chapter 2* and one subject was excluded from part of the study due to lost retinal photographs. Demographic data for all subjects, those included in further analyses and those excluded, are presented in *Table 3*. Where several mutually exclusive categories are presented, the percentages sum to 100% down the page. If there are only two possible categories, such as gender, then the percentage in the named group is presented (e.g. 60.1% of subjects are male).

Excluded subjects

The centres were not able to verify the quality of all aspects of a particular subject's information at the time of recruitment. However, checks of eligibility, and the quality and availability of data meant that 370 subjects were excluded for a variety of reasons, some for more than one reason. In 329 cases it was not possible to assess the thickness of the macula in one or both eyes and so the subject could not be classified as having MO present or absent, and in a further 41 cases retinal photographs from both eyes

© Queen's Printer and Controller of HMSO 2013. This work was produced by Olson *et al*. under the terms of a commissioning contract issued by the Secretary of State for Health. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

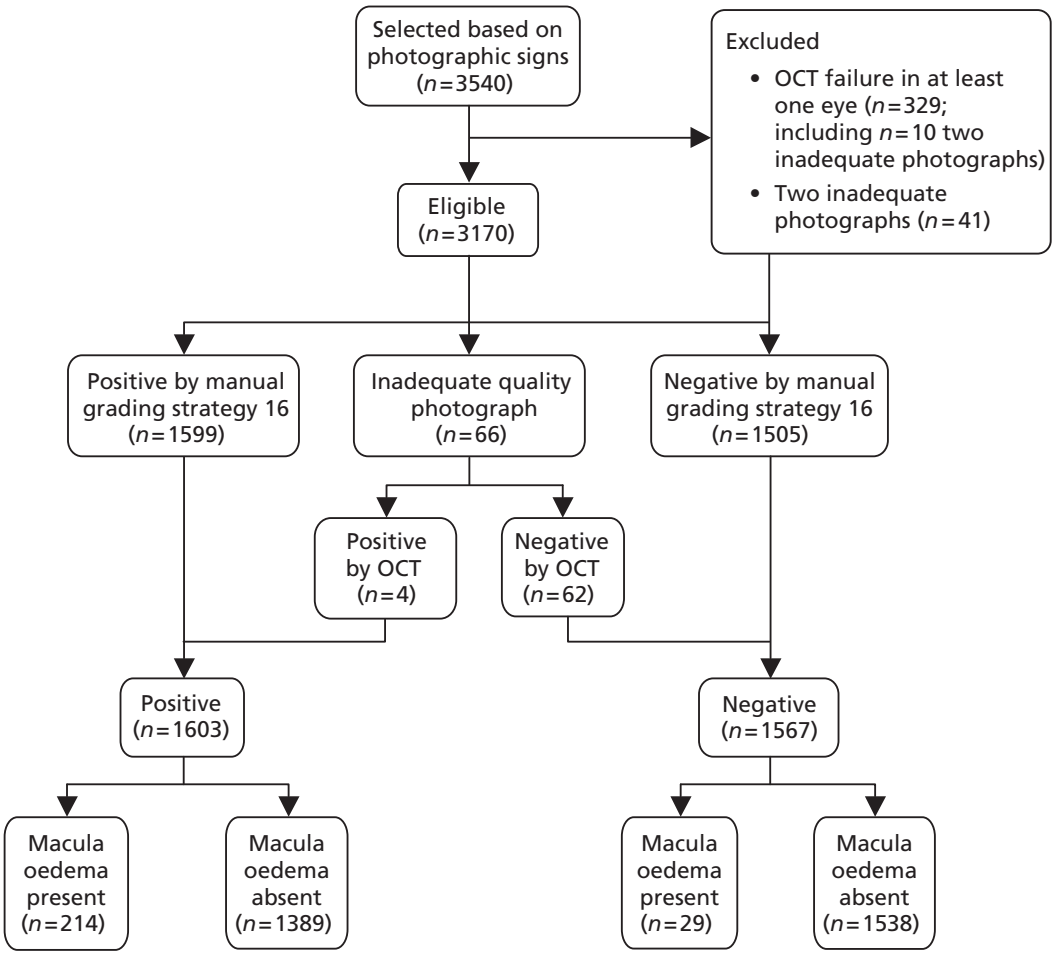


FIGURE 10 Study design for recruitment, with manual grading strategy 16 (see *Table 19*) as the diagnostic test and with the reference standard of MO presence. 'Positive' means that the image was judged to have MO.

were of inadequate quality. The reasons, sometimes overlapping, for OCT failure included an enucleated eye (one subject), errors in identifying the boundaries of the macula, low signal, evidence of movement of the eye, positional error of the image or loss of the image. Of the 370 subjects excluded, 12 subjects were found to have had previous laser treatment in one or both eyes. Sixty-eight subjects had poor clarity in one or both photographs and nine subjects had an incorrect field of view in one or both photographs. For some subjects, both OCT and photographs were inadequate.

There were some statistically significant differences between the 3170 subjects included in further analyses and the 370 excluded. The large size of the study meant that some differences of small magnitude were found to be statistically significant. The excluded subjects tended to be older by 6 years and were slightly more likely to be female. Asian or Black subjects were slightly more likely to be excluded (14.4% and 14.0%) rather than Caucasian (9.8%). A higher percentage of subjects with type 2 diabetes were excluded (11.2%) compared with subjects with type 1 diabetes (7.7%). Those with worse visual acuity (log-MAR ≥ 0.3) were more likely to be excluded. There were clear differences between percentages of subjects excluded on the four scanners: 15.9% of subjects on the Zeiss Stratus OCT, 13.9% on the Topcon 3D OCT-1000, 7.3% on the Heidelberg Spectralis and 1.4% on the Zeiss Cirrus OCT. These findings should not be interpreted in isolation. They could have been influenced by the fact that older people are more likely to have opacity and because different scanners (with different qualities) were used in different centres.

Prevalence of thickening and macular oedema

Among the 3170 subjects included in further analyses, there were significant differences in the percentages with thickening in region 1 [greatest for Zeiss Cirrus OCT (19.1%) and least for Zeiss Stratus

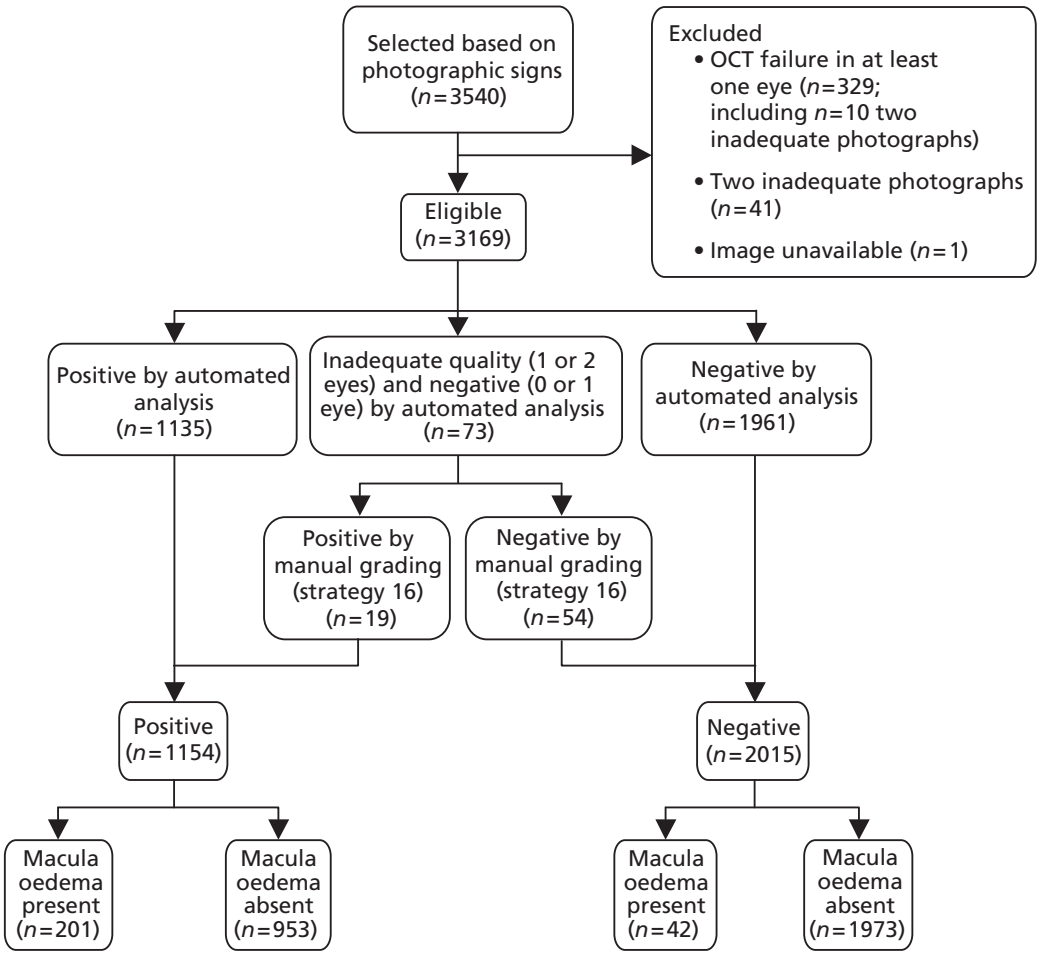


FIGURE 11 Study design for recruitment, with automated image analysis as the diagnostic test and with the reference standard of MO presence. 'Positive' means that the image was judged to have MO.

OCT (12.9%); comparison of all scanners $p = 0.002$] and also the percentages with thickness in at least one of the regions 1–5 [greatest for Zeiss Cirrus OCT (30.4%) and least for Heidelberg Spectralis (20.4%); comparison of all scanners $p < 0.001$]. The results are presented in *Table 4*. Note that the percentages sum to 100% across the page in every third column.

There were clear differences between percentages of subjects with MO in at least one eye on the four scanners: 11.8% of subjects scanned on Zeiss Cirrus OCT had MO, 8.7% on Heidelberg Spectralis, 6.5% on Topcon 3D OCT-1000 and 4.5% on Zeiss Stratus OCT (see *Table 4*). However, the type of scanner was almost entirely confounded with centre as most centres used a single type of scanner and some types of scanners were only used in a single centre. So the differences between scanners, as seen in the percentages with MO, could be due to differences between centres or other confounding factors.

All 243 subjects classified as having MO had thickening in at least one of the five inner regions. However, not all of those with thickness in at least one of these regions were classified as having MO after further investigation.

Counts of subjects and eye by centre

The numbers and percentages of subjects from each centre included and excluded from the study are presented in *Table 5*. The percentages of subjects sum to 100% across the table. Eyes may be excluded for more than one reason so percentages relating to eyes will not sum to 100% across the columns.

© Queen's Printer and Controller of HMSO 2013. This work was produced by Olson *et al*. under the terms of a commissioning contract issued by the Secretary of State for Health. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

TABLE 3 Demographic information for all 3540 subjects recruited

Characteristic	All subjects, *n*	%	Subject included in further analyses, *n*	%	Subject excluded from further analyses, *n*	%	*p*-value (included vs. excluded)
N	3540		3170		370		
Age, years (median, IQR)[a]	60	49, 70	60	49, 69	66	53, 75	< 0.001
Sex (male)	2128	60.1	1925	60.7	203	54.9	0.034
Ethnicity							
Caucasian	2968	83.8	2678	84.5	290	78.4	0.036[b]
Asian	431	12.2	369	11.6	62	16.8	
Black	86	2.4	74	2.3	12	3.2	
Other	29	0.8	25	0.8	4	1.1	
Unknown	26	0.7	24	0.8	2	0.5	
Diabetes:							
Type 1	768	21.7	709	22.4	59	15.9	0.013[c]
Type 2	2761	78.0	2452	77.4	309	83.5	
Unspecified	5	0.1	4	0.1	1	0.3	
Unknown	6	0.2	5	0.2	1	0.3	
Amblyopia (yes either eye)	110	3.1	86	2.7	24	6.5	
Glitazone use (yes)	198	5.6	177	5.6	21	5.7	1.000
Visual acuity better left[c]	3079	87.0	2807	88.5	302	81.6	0.001
Visual acuity worse left[c]	435	12.3	348	11.0	65	17.6	
Visual acuity left missing	18	0.7	15	0.5	3	0.8	
Visual acuity better right[c]	3079	87.0	2794	88.1	285	77.0	< 0.001
Visual acuity worse right[d]	435	12.3	361	11.4	74	20.0	
Visual acuity right missing	26	0.7	15	0.5	11	3.0	
Heidelberg Spectralis	508	14.4	471	14.9	37	10.0	< 0.001
Topcon 3D OCT-1000	870	24.6	749	23.6	121	32.7	
Zeiss Cirrus OCT	910	25.7	897	28.3	13	3.5	
Zeiss Stratus OCT	1252	35.4	1053	33.2	199	53.8	

IQR, interquartile range.
a For each group of subjects, the three numbers presented for age are the median age in years and 25th and 75th percentiles (the IQR) in years.
b Comparison of four categories as smallest combined.
c Comparison of three categories as smallest combined.
d Visual acuity better (log-MAR < 0.3; Snellen 6/9.5 or better), visual acuity worse (log-MAR ≥ 0.3; Snellen 6/12 or worse).

TABLE 4 Thickness and presence of MO by OCT scanner

	Thickness in ETDRS region 1 (either eye)			Thickness in any ETDRS region 1–5 (either eye)			MO in either eye		
Scanner	Yes	No	Total	Yes	No	Total	Yes	No	Total
Heidelberg Spectralis	72	399	471	96	375	471	41	430	471
	15.3%	84.7%	100.0%	20.4%	79.6%	100.0%	8.7%	91.3%	100.0%
Topcon 3D OCT-1000	110	639	749	165	584	749	49	700	749
	14.7%	85.3%	100.0%	22.0%	78.0%	100.0%	6.5%	93.5%	100.0%
Zeiss Cirrus OCT	171	726	897	273	624	897	106	791	897
	19.1%	80.9%	100.0%	30.4%	69.6%	100.0%	11.8%	88.2%	100.0%
Zeiss Stratus OCT	136	917	1053	238	815	1053	47	1006	1053
	12.9%	87.1%	100.0%	22.6%	77.4%	100.0%	4.5%	95.5%	100.0%
Total	489	2681	3170	772	2398	3170	243	2927	3170
	15.4%	84.6%	100.0%	24.4%	75.6%	100.0%	7.7%	92.3%	100.0%

Demographics by macular oedema status

There were 3170 subjects with known MO status according to OCT: 243 with MO and 2927 without MO. Demographics are presented in *Table 6*. Where several mutually exclusive categories are presented, the percentages will sum to 100% when read down the table. If there are only two possible categories, such as gender, then the percentage in the named group is presented (e.g. 60.5% of subjects with OCT MO present are male). The percentages in *Table 6* are calculated using the denominators of 243 and 2927, the total number with MO and the total number without MO. Percentages given in the text below are different and are calculated from the number with MO and a characteristic divided by the total number with that characteristic.

There were some statistically significant differences between the 243 subjects with MO and the 2927 subjects without MO. The subjects with MO tended have older median age by 8 years than those without MO. There were no differences in the percentage with MO by gender, glitazone use or amblyopia. A higher percentage of Caucasians (8.4%) had MO compared with other ethnic groups (3–4%). A higher percentage of subjects with type 2 diabetes (8.7%) had MO compared with subjects with type 1 diabetes (3.9%). Those with worse visual acuity (log-MAR ≥ 0.3) or missing visual acuity were more likely to have MO. The age, diabetes type and visual acuity results were likely to have been confounded with each other and this issue was addressed using multivariate analyses.

Counts of subjects and eyes with macular oedema by centre

There were differences in the percentages of subjects with MO by centre (*Table 7*). Aberdeen, Dundee and Oxford had a far higher prevalence of MO than the other centres. However, these could be due to differences in demographics, presence of lesions, different interpretations of the recruitment strategy or, as noted above, due to differences between scanners. The fact that differences between centres may be measured or unmeasured is why centre was included in the models as an adjustment variable. There were very few subjects with MO in both eyes (0–2.2%) across the seven different centres.

Presence of different types of lesion

Photographic feature information was missing for three left eyes and two right eyes so the denominator for left eyes was 3167 and for right eyes was 3168. Counts of subjects and eyes with MO and features in each eye are presented in *Table 8*. The same table split by centre is presented in in *Appendix 3* (see *Table 40*). Subjects can have several different lesions or none within each eye.

© Queen's Printer and Controller of HMSO 2013. This work was produced by Olson *et al*. under the terms of a commissioning contract issued by the Secretary of State for Health. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

TABLE 5 Counts of subjects and eye by centre (percentages of total recruited per centre)

	Subjects			Eyes					
Centre	Count of subjects recruited study	Count of subjects included in all analysis	Count of subjects excluded[a]	Count of left eyes with valid data	Count of left eyes excluded because of OCT failure	Count of left eyes excluded because of photo failure	Count of right eyes with valid data	Count of right eyes excluded because of OCT failure	Count of right eyes excluded because of photo failure
Aberdeen	927	909 (98.1)	18 (1.9)	898 (96.9)	9 (1.0)	22 (2.4)	904 (97.5)	7 (0.8)	17 (1.8)
Birmingham	958	842 (87.9)	116 (12.1)	888 (92.7)	59 (6.2)	13 (1.4)	886 (92.5)	66 (6.9)	9 (0.9)
Dundee	322	254 (78.9)	68 (21.1)	263 (81.7)	36 (11.2)	33 (10.2)	266 (82.6)	31 (9.6)	28 (8.7)
Edinburgh	298	218 (73.2)	80 (26.8)	245 (82.2)	51 (17.1)	2 (0.7)	250 (83.9)	44 (14.8)	4 (1.3)
Liverpool	455	416 (91.4)	39 (8.6)	409 (89.9)	18 (4.0)	33 (7.3)	412 (90.5)	15 (3.3)	29 (6.4)
Dunfermline	178	159 (89.3)	19 (10.7)	157 (88.2)	5 (2.8)	17 (9.6)	158 (88.8)	11 (6.2)	10 (5.6)
Oxford	402	372 (92.5)	30 (7.5)	374 (93.0)	15 (3.7)	15 (3.7)	375 (93.3)	14 (3.5)	15 (3.7)
Total	3540	3170[a] (89.5)	370 (10.5)	3234[b] (91.4)	193[b] (5.5)	135[b] (3.8)	3251[b] (91.8)	188[b] (5.3)	112[b] (3.2)

a Subjects may be excluded because the combination of results for both eyes does not allow classification of their MO status.
b Eyes may be excluded for more than one failure so percentages will not sum to 100% across the columns.

TABLE 6 Demographics by OCT MO status

Characteristic	MO+ve (either eye)		MO−ve (both eyes)		*p*-value (oedema Y/N)
	n	%	*n*	%	
N	243		2927		
Age, years (median, IQR)[a]	67	58, 75	59	48, 69	< 0.001
Sex (male)	147	60.5	1778	60.7	0.993
Ethnicity					0.015
Caucasian	223	91.8	2455	84.5	
Asian	16	6.6	353	12.1	
Black	2	0.8	72	2.5	
Other/unknown	2	0.8	47	1.6	
Diabetes					< 0.001[b]
Type 1	28	11.5	681	23.3	
Type 2	213	87.7	2239	76.5	
Unspecified	2	0.8	2	0.1	
Unknown	0	0.0	5	0.2	
Amblyopia (either yes)	83	2.8	3	1.2	0.204
Glitazone use (yes)	10	4.1	167	5.7	0.372
Visual acuity (L)					< 0.001
Better	163	67.1	2644	90.3	
Worse	77	31.7	271	9.3	
Missing	3	1.2	12	0.4	
Visual acuity (R)					< 0.001
Better	164	67.5	3630	89.9	
Worse	74	30.5	287	9.8	
Missing	5	2.1	10	0.3	
Scanner					< 0.001
Heidelberg Spectralis	41	16.9	430	14.7	
Topcon 3D OCT-1000	49	20.2	700	23.9	
Zeiss Cirrus OCT	106	43.6	791	27.0	
Zeiss Stratus OCT	47	19.3	1006	34.4	
Mutually exclusive categories					< 0.001
Exudates less than one DD	144	59.3	880	30.1	
Blots (no exudates) less than one DD	51	21.0	372	12.7	
Microaneurysms less than one DD	44	18.1	1327	45.3	
Exudates one to two DD	0	0.0	27	0.9	
No relevant diabetic retinopathy features more than two DD	4	1.6	321	11.0	

–ve, negative; +ve, positive; IQR, interquartile range; L, left; N, no; R, right; Y, yes.

a For those subjects with MO+ve and MO-ve, the three numbers presented for age are the median age in years and 25th and 75th percentiles (the IQR) in years.

b Comparison of three categories of diabetes (type 1, type 2 and other) as unknown and unspecified combined for statistical test.

© Queen's Printer and Controller of HMSO 2013. This work was produced by Olson *et al*. under the terms of a commissioning contract issued by the Secretary of State for Health. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

TABLE 7 Counts of subjects and eyes with MO by centre

				Subjects				
Centre	Count of subjects included in all analyses	Patients with MO in any eye (%)	Count of MO in both eyes	% of MO in both eyes	Count of MO in one eye	% of MO in one eye	Counts of MO in no eyes	% of MO in no eyes
Aberdeen	909	109 (12.0)	11	1.2	98	10.8	800	88.0
Birmingham	842	31 (3.7)	10	1.2	21	2.5	811	96.3
Dundee	254	31 (12.2)	4	1.6	27	10.6	223	87.8
Edinburgh	218	14 (6.4)	0	0.0	14	6.5	204	93.6
Liverpool	416	12 (2.9)	2	0.5	10	2.4	404	97.1
Dunfermline	159	7 (4.4)	1	0.6	7	3.8	152	95.6
Oxford	372	39 (10.5)	8	2.2	31	8.3	333	89.5
Total	3170	243 (7.7)	36	1.1	207	6.5	2927	92.3

Mutually exclusive groups of lesions

Mutually exclusive (non-overlapping) groups of features were identified so that every eye could be classified into one group and one group only. The groups for comparison were (1) M/DH only (not BH or exudates), (2) BH but not exudates (BH only + BH and M/DH only), and (3) exudates (regardless of any other feature being present). Prior to recruitment the three mutually exclusive groups of lesions were expected to be present in 11.3%, 1.4% and 3.5% of all scanned images with 83.8% of images having no M/DHs, BHs or exudates. Within the 16.2% of images expected to have lesions (the target population for this study), the three groups of lesions were expected to be found in 69.8%, 8.6% and 21.6% of images, respectively.

In this study, considering the left eyes only, the three groups of lesions were found in 40.3% (M/DHs only within one DD), 8.4% (BH only or BH and M/DHs within one DD) and 20.4% (exudates only or with M/DHs or BHs within one DD). In addition, 28.1% had no lesions within one DD and 2.8% had other lesions within one DD. Very similar percentages were found in the right eyes: 41.7%, 8.7% and 18.9% with 27.9% having no lesions and 2.8% having other lesions. This information is presented by centre in *Appendix 3* (see *Table 41*).

As these percentages did not match those expected, particularly those for M/DHs only, weighting of the subjects according to the features of their worst eye was considered necessary. This will be addressed in the next chapter.

The mutually exclusive features within one DD in the left and right eyes are compared in *Table 9*. The percentages shown in this table are out of the total of 3165 that had lesion information from both eyes. Two hundred and twenty-two subjects (7.0%) had exudates in both eyes, 39 subjects (1.2%) had BHs in both eyes and 566 subjects (17.9%) had M/DHs in both eyes. Of the 351 subjects (11.1%) that appear to have none or other features within one DD in both eyes, 27 had exudates between one and two DD in at least one eye, some had features such as flames, drusen or CWS and 268 did not appear to have any obvious features within one DD or exudates within two DD.

Conclusions

There were some significant differences between the 3170 subjects retained for further analyses and the 370 (10.5%) which were excluded. Excluded subjects were slightly older, more likely to be female, more likely to be Black or Asian and more likely to have type 2 diabetes or worse visual acuity. A higher

TABLE 8 Counts of subjects and eyes with MO and features in left and right eyes within one DD unless otherwise specified. Subjects can have several different lesions or none in an eye

Eye	Count (%) of presence of MO in that eye	Counts (%) of no MO in that eye	Count (%) of presence of BH	Count (%) of presence of CWS	Count (%) of presence of drusen	Count (%) of presence of exudate	Count (%) of presence of FH	Count (%) of presence of M/DH	Count (%) of presence of exudates in one to two DD
Left	144 (4.5)	3026 (95.5)	384 (12.1)	15 (0.5)	233 (7.4)	647 (20.4)	72 (2.3)	1965 (63.0)	569 (18.0)
Right	135 (4.3)	3035 (95.7)	388 (12.2)	14 (0.4)	252 (8.0)	598 (18.9)	69 (2.2)	2032 (64.1)	542 (17.1)

FH, flame haemorrhage.

© Queen's Printer and Controller of HMSO 2013. This work was produced by Olson *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

TABLE 9 Counts of subjects and percentages by mutually exclusive categories of features within one DD in left eye and right eye

	Right eye (mutually exclusive categories)				
Left eye (mutually exclusive categories)	Exudates in one DD	BHs (no exudates) in one DD	M/DHs (no BHs or exudates) in one DD	None/other in one DD	Total
Exudates in one DD	222 (7.0)	42 (1.3)	224 (7.1)	159 (5.0)	647 (20.4)
BHs (no exudates) in one DD	38 (1.2)	39 (1.2)	108 (3.4)	81 (2.6)	266 (8.4)
M/DHs (no BHs or exudates) in one DD	206 (6.5)	120 (3.8)	566 (17.9)	382 (12.1)	1274 (40.3)
None/other in one DD	132 (4.2)	76 (2.4)	419 (13.2)	351 (11.1)	978 (30.9
Total	598 (18.9)	277 (8.8)	1317 (41.6)	973 (30.7)	3165 (100.0)

percentage of subjects were excluded using the Zeiss Stratus OCT and Topcon 3D OCT-1000 than the Heidelberg Spectralis and Zeiss Cirrus OCT. There were also differences in exclusions by centre, but these were confounded by the scanner used.

The prevalence of MO differed greatly by centre ranging from 3.7% to 12.2%. It also differed between scanners, ranging from 4.5% to 11.8%.

In simple analyses, patient characteristics associated with MO were older age, Caucasian ethnicity, and having type 2 diabetes rather than type 1.There was no association between gender, amblyopia and glitazone use with MO.

The presence of MO was associated with worse visual acuity.

The types and distributions of lesions in subjects recruited for this study were not typical of what would be found in a cohort of people attending screening with lesions. Subjects with more severe lesions such as exudates and BHs were over-represented and those with more minor lesions such as M/DHs were under-represented. Consequently, weighting of the data was necessary for the more complex analyses to reflect what would occur in practice.

Chapter 5 Inferring the presence of macular oedema using retinal photographs

Introduction

This chapter addresses the first aim of the project, namely to investigate whether or not particular distributions and combinations of lesions (M/DHs, BHs and exudates), assessed manually or automatically, are more specific surrogate markers of MO than current grading practice, using OCT as the reference standard.

Initially statistical modelling was carried out to see if any of the current manual photographic screening strategies included everything that might be considered important. In the second part of this chapter this information is used to inform the inclusion of various eye and subject characteristics within computer-assisted manual photographic screening strategies and full automated strategies for detecting MO.

National screening programmes in the UK all agree that exudates within one DD of the centre of the fovea should be used to infer the presence of referable MO. However, there is disagreement as to how BHs or M/DHs within one DD of the centre of the fovea should be used. The value of BHs or M/DHs in relation to MO is therefore of particular interest.

The various lesions and lesion combinations within a single eye are investigated in relation to the presence of MO in that eye. Counts of the three main types of lesions are investigated in a similar way. The findings in *Chapter 4*, and from the analyses of single eyes, were used to guide the analyses predicting MO (in either eye) from information on the subject and the two separate eyes.

As noted in *Chapter 4*, the combinations of lesion types and better and worse visual acuity did not occur in this sample in the same proportions as expected from a cohort of all subjects attending retinal screening. More complex analyses were weighted so that the results might better reflect what is expected in such a cohort.

Methods for statistical modelling

Combinations of lesions (M/DHs, BHs and exudates) were considered in relation to MO. Subjects were classified into mutually exclusive categories according to their most serious feature in the relevant eye. The five categories were:

- exudates present within one DD (regardless of other lesions)
- BHs present within one DD (but not exudates)
- M/DHs present within one DD (but not exudates or BHs)
- other minor lesions (no M/DHs, BHs or exudates present) within one DD
- no lesions present within one DD.

These lesion combinations were also compared across centres by eye. Visual acuity was considered as a three category variable: better visual acuity (log-MAR < 0.3), worse visual acuity (log-MAR ≥ 0.3) and visual acuity missing.

© Queen's Printer and Controller of HMSO 2013. This work was produced by Olson *et al*. under the terms of a commissioning contract issued by the Secretary of State for Health. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

Single eye analyses

In order to investigate whether or not particular distributions and combinations of lesions identified by manual screening (M/DHs, BHs and exudates) were valuable markers of MO, single eyes were considered separately. Chi-squared tests were used to identify potential risk markers and confounding factors. Logistic regression analyses were used on data from one eye only to identify features identified by manual screening that might be important for predicting the presence of MO within the relevant eye. Different representations of features including presence, count and distribution were considered.

Initially the presence of features within one DD (and for exudates within one to two DD) was investigated in relation to MO within the same eye. In these initial analyses, only single features were considered, the other features being ignored. Following that, analysis combinations of features were considered. The presence of exudates between one and two DD was considered as a separate variable. These logistic regression analyses considered the lesion variable alone (unadjusted), after adjusting for centre, and after adjusting for centre, demographic variables and visual acuity in that eye.

Counts of the different types of lesions (exudates, BHs and M/DHs) within one DD (and for exudates within one to two DD) were first considered singly and then in combination in logistic regression analyses for MO in the relevant eye. These analyses were also adjusted first for centre, and then for demographic variables and visual acuity in that eye. The ranges of counts of lesions were not the same for different types of lesions and there was concern that results of analyses of counts could be dominated by high counts. As a sensitivity analysis, counts of lesions were collapsed into zero, one, two, three, four, five and more than five, and this count variable was included in the logistic regression analysis as if it was a discrete count. This was to ensure that any relationships observed were robust to the influence of very high counts.

There is particular interest in the contribution of M/DHs to the prediction of MO as M/DHs are included in current manual grading practice in England, but not in Scotland. In order to investigate whether or not larger numbers of M/DHs were useful in predicting MO, the counts of all three types of lesions were collapsed even further so they could be included in the logistic regression models as ordered categories: zero, one, two and more than two. The opportunity to observe non-linear increases in the odds of MO with increasing numbers of lesions was provided by this alternative representation of counts of lesions as four ordered categories.

As noted earlier, the combinations of lesion types and better and worse visual acuity did not occur in the proportions expected from a cohort of subjects attending retinal screening and so data weighting was necessary. Weighting of sample data is common in sample surveys either where the sampling scheme uses unequal probabilities by design or if the data are known to be unrepresentative of the population,[43] often due to disproportionate non-response.[44] If the sample is known to be substantially different to the population in the distribution of one or more key variables then an analysis of the raw data can produce biased estimates of prevalence. Reweighting can correct for bias in the estimates, but may have the effect of increasing the variances and complicating the analysis.[43] In the current study, if a statistic, such as the sensitivity of a diagnostic algorithm, differs between types of subjects, then the estimate of sensitivity based on the raw sample data could be biased towards the sensitivity within subjects who are over-represented and away from the sensitivity within those under-represented. The under-representation of subjects with only M/DHs was of particular concern. The simplest form of weight, sometimes called direct weights or post-stratification weights, was used.[44–46]

The weighting of subjects was performed as follows. The proportions of subjects falling into the five categories were calculated for the study sample, the expected proportions in the five categories being known from a previous study.[30] Weights were obtained by dividing the proportions of subjects in the study sample by the proportions in the population and then multiplying by a factor of 3170/2845 to account for the zero weighting of 325 subjects with either no lesions or only very mild lesions in either eye. Where appropriate, the analyses described above were repeated after weighting. Ideally weighting would have been done to correspond to the population proportions of the five groups within each centre, but these were not known at each centre. This weighting was also used in the health economics analyses in *Chapter 6*.

Both eyes analysis

Information on features from both eyes was combined in order to look at the prediction of presence of MO (in one or both eyes) per subject. The most serious feature in either eye was given priority in determining the classification of the subject by their lesions. The mutually exclusive categories chosen were:

- any exudates in either eye
- any BHs (but no exudates) in either eye
- more than two M/DHs in either eye (but no exudates or BHs)
- exactly two M/DHs in the eye with most M/DHs (but no exudates or BHs)
- exactly one M/DH in the eye with most M/DHs (but no exudates or BHs)
- those with no M/DHs in either eye and those subjects with lesions which were not M/DHs, BHs or exudates in either eye.

Poor vision was classified as worse (if visual acuity was log-MAR ≥ 0.3 in either eye or if visual acuity was missing in one eye and log-MAR ≥ 0.3 in the other), better (if visual acuity was log-MAR < 0.3 in both eyes) or missing (if visual acuity was missing in one eye and better in the other or missing in both eyes).

Logistic regression was used to investigate this representation of the lesions in relation to predicting the presence of MO in at least one eye. The analysis was repeated first for centre and visual acuity, and then for centre, visual acuity and demographic factors. The analyses of both eyes together were repeated after weighting.

Results of statistical modelling

Presence of individual lesion types in relation to macular oedema

In *Table 10* the presence or absence of a particular type of lesion in the relevant eye is considered. No adjustments were made for other lesions that might also have been present. A large difference between the percentages of subjects with MO and without MO and with a certain type of lesion present may be due to the presence of another type of lesion.

Macular oedema was roughly five times more prevalent among subjects with an exudate, a BH or a M/DH present within one DD compared with those with the same lesion absent (see *Table 10*). The prevalence of MO was also greater, albeit by a lesser amount, for subjects having an exudate within one to two DD, but no consideration was given whether or not there was an exudate within one DD as well. Too few eyes had CWS for a robust analysis. There were few eyes with flame haemorrhages or drusen so although there were sufficient eyes for these comparisons to be made, multivariate analyses were not considered appropriate. There did not appear to be any evidence of a relationship between the presence of drusen and the presence of MO in the relevant eye.

Subjects with worse visual acuity were about five times as likely to have MO in the relevant eye as those with better visual acuity.

Only a small number of subjects had amblyopia and, of these, very few had MO. There were no significant differences in the percentages with MO between those using and not using glitazone. In later analyses adjustment was made for glitazone use, but this was not possible for amblyopia because of the small numbers of subjects.

When mutually exclusive groups of lesions were considered, highly significant differences in the percentages with MO were found. More than 10% of subjects with exudates or with BHs (but not exudates) had MO in the same eye. The percentage with MO was just over 2% of those with just M/DHs present in the same eye, about 1% of those with lesions excluding M/DHs, BHs and exudates and < 1% with no lesions in that eye.

© Queen's Printer and Controller of HMSO 2013. This work was produced by Olson *et al*. under the terms of a commissioning contract issued by the Secretary of State for Health. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

TABLE 10 Relationship of individual lesions to MO. Results of single eye associations between each feature and MO

	Left eye					Right eye				
	Feature present		Feature absent			Feature present		Feature absent		
Feature	*n* with MO/total *N*	% with MO	*n* with MO/total *N*	% with MO	*p*-value	*n* with MO/total *N*	% with MO	*n* with MO/total *N*	% with MO	*p*-value
BH	63/384	16.4	81/2783	2.9	< 0.001	58/388	14.9	77/2780	2.8	< 0.001
CWS	0/15	0.0	144/3152	4.6	1.0	0/14	0.0	135/3154	4.3	1.0
Drusen	11/233	4.7	133/2934	4.5	1.0	6/252	2.4	129/2916	4.4	0.168
Exudate	81/647	12.5	63/2520	2.5	< 0.001	67/598	11.2	68/2570	2.6	< 0.001
FH	15/72	20.8	129/3095	4.2	< 0.001	11/69	15.9	124/3099	4.0	< 0.001
M/DH	123/1995	6.2	21/1172	1.8	< 0.001	118/2032	5.8	17/1136	1.5	< 0.001
Exudate within one to two DD	61/569	10.7	83/2598	3.2	< 0.001	56/542	10.3	79/2626	3.0	< 0.001
Visual acuity worse[a]	56/348	16.1	86/2807	3.1	< 0.001	56/361	15.5	75/2794	2.7	< 0.001
Amblyopia	0/51	0.0	144/3119	4.6	0.171	2/49	4.1	133/3121	4.3	1.0
Glitazone	5/177	2.8	139/2993	4.6	0.345	6/177	3.4	129/2993	4.3	0.691
Mutually exclusive groups	***n* with MO**	**% with MO**	***n* without MO**	**% without MO**	***p*-value**	***n* with MO**	**% with MO**	***n* without MO**	**% without MO**	***p*-value**
No lesions	7	0.8	882	99.2	< 0.001	5	0.6	880	99.4	< 0.001
M/DH only	28	2.2	1247	97.8		31	2.3	1289	97.7	
BH not exudate	27	10.2	239	89.8		31	11.2	246	88.8	
Exudate + any	81	12.5	566	87.5		67	11.2	531	88.8	
Other	1	1.1	89	98.9		1	1.1	87	98.9	
Total	144	4.5	3023	95.5		135	4.3	3033	95.7	

FH, flame haemorrhage.

a Visual acuity better (log-MAR < 0.3; Snellen 6/9.5 or better), visual acuity worse (log-MAR ≥ 0.3; Snellen 6/12 or worse).

Visual acuity was not available for 15 left eyes, two of which had MO, and for 15 right eyes, four of which had MO.

Logistic regression analyses of single eyes

Most of the following tables in this section are for the left eye only. Analyses of the right eyes were also completed but since they showed almost identical relationships these results have not been presented. *Table 11* shows the results of logistic regression analysis of a single eye with the outcome variable of MO in that eye. The odds ratio (OR) shows the increase in the odds of having MO in a subject with a characteristic compared with a subject without that characteristic. For a measured variable, such as age, the OR represents the change in the odds of having MO if that variable were to increase by one unit of measurement (1 year older) relative to the same subject without that increase. If multiple variables are included in the same logistic regression model then the OR for a single variable represents the increase in the odds of having MO if that variable were to increase by one unit of measurement, or to change to a different category, compared with a subject with all other characteristics and measurements held fixed.

In *Table 11*, the top category for each of the categorical variables is chosen as the reference group with OR shown as 1 and no CI displayed. The choice of the reference group is arbitrary and this does not affect statistical significance. The other ORs presented are relative to this reference group and a 95% CI is displayed in the next two columns. If this 95% CI includes 1 it shows that the OR is not significantly different to 1 (the null value) and so $p > 0.05$. If the 95% CI excludes 1 it indicates that this measured variable is significant for MO with $p < 0.05$ or that this categorical variable is significant with $p < 0.05$ and the relevant category is significantly different to the reference category in terms of the odds of having MO.

The first three columns of numbers show ORs and the lower and upper 95% confidence limits for each variable considered separately (unadjusted). The fourth to sixth columns show the ORs and 95% CIs for the variables adjusted for differences between centres. Ideally simultaneous adjustments would be made for centre and scanner type, but they are highly confounded with each other and so could not be included together in the same model. The feature results were similar regardless of whether they were adjusted for centre or scanner so adjustment has only been made for centre. The seventh to the ninth columns show the ORs and 95% CIs for the variables adjusted for centre, gender, age, glitazone and diabetes (type 1, type 2 and secondary/unknown). The last three columns show the ORs and 95% CIs for the variables adjusted for centre, gender, age, glitazone, diabetes (type 1, type 2 and secondary/unknown) and visual acuity.

When the mutually exclusive groups of lesions were included in the model, subjects with BHs (either on their own or BH with M/DHs) had 5.0 times the odds of having MO and subjects with exudates had 6.4 times the odds of having MO, both compared with those with only M/DHs. These ORs changed by only modest amounts as adjustments were made for centre and other characteristics to ORs of 3.6 and 6.7. Of these other variables, only centre, age and visual acuity were significant with older age being associated with higher odds of MO by 4% per additional year older (OR = 1.037; 95% CI 1.020 to 1.054). When visual acuity was included, having worse visual acuity significantly increased the odds of MO by a factor of 3.9 (95% CI 2.6 to 5.9).

An analysis of counts of lesions found that higher numbers of M/DHs, BHs and exudates within one DD significantly increased the odds of having MO when considered separately or together. The relationships were maintained after adjustment for centre, gender, age, glitazone, diabetes (type 1, type 2 and secondary/unknown) and visual acuity. Initially, counts of exudates between one to two DD appeared to be associated with greater odds of MO, but as this ceased to be significant after adjusting for exudates within one DD so it was dropped from subsequent models. It is important to note that the count of M/DHs within one DD was statistically significant after adjusting for centre, the count of BHs and the count of exudates. This suggested that having a lot of M/DHs was still useful for predicting MO.

The OR for counts of BHs was larger than for the other lesions. This does not mean BHs were more important, but that the increase in the odds of MO for each additional BH was larger than for each

© Queen's Printer and Controller of HMSO 2013. This work was produced by Olson *et al*. under the terms of a commissioning contract issued by the Secretary of State for Health. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

TABLE 11 Relationship of individual lesions to left eye MO (ORs; analysis unweighted)

Left eye												
Feature category (mutually exclusive)	**Unadjusted OR[a]**	**Lower 95% CI**	**Upper 95% CI**	**Adjusted OR[b]**	**Lower 95% CI**	**Upper 95% CI**	**Adjusted OR[c]**	**Lower 95% CI**	**Upper 95% CI**	**Adjusted OR[d]**	**Lower 95% CI**	**Upper 95% CI**
M/DH only	1											
BH not exudate	5.0	2.9	8.7	4.4	2.5	7.6	3.6	2.0	6.3	3.6	2.0	6.4
Exudate + any	6.4	4.1	9.9	5.8	3.7	9.2	7.2	4.5	11.5	6.7	4.2	10.8
Count of lesions of each type												
M/DH	1.19	1.15	1.22	1.12	1.08	1.17	1.13	1.09	1.18	1.13	1.09	1.18
BH	2.32	1.98	2.73	1.66	1.38	2.00	1.56	1.29	1.87	1.53	1.27	1.85
Exudate	1.17	1.13	1.20	1.14	1.11	1.17	1.16	1.12	1.20	1.16	1.12	1.19
Exudate within one to two DD	1.06	1.04	1.07									

a Unadjusted results.
b Adjusting for centre.
c Adjusting for centre, gender, age, glitazone and diabetes (type 1, type 2 and secondary/unknown).
d Adjusting for centre, gender, age, glitazone, diabetes (type 1, type 2 and secondary/unknown) and visual acuity.

additional M/DH or exudate. It was thought that this could be a consequence of there being a smaller range of counts of BHs across subjects than the range of counts of M/DHs or exudates.

The maximum numbers of exudates and M/DHs within a single eye of a single subject were 59 and 50, respectively, much higher than the maximum of 16 BHs within a single eye of a single subject. A sensitivity analysis was carried out with the counts of lesions fitted as a discrete measured variable taking a maximum value of six (so zero, one, two, three, four, five, and more than five). The ORs (not shown) were larger for each extra lesion of any type. The inferences of these results were consistent with those shown for raw counts meaning that differences in ranges of counts of different lesions were not responsible for the larger ORs for counts of BHs compared with ORs for counts of exudates or M/DHs.

See *Table 42* for the corresponding data for the right eye.

A simple descriptive analysis of count data was completed representing the counts of lesions of the three main types as ordered categories: zero, one, two and more than two (*Table 12*). The percentages of subjects with MO increased from 1% to 2% to > 3% when the number of M/DHs within one DD increased to more than two. These results suggested that this might be a useful representation when investigating the value of M/DHs in predicting MO.

TABLE 12 Macular oedema status per eye by presence of features in that eye

	MO status in relevant eye					
Lesions category within one DD in relevant eye	No OCT MO found in left eye	OCT MO found in left eye	Total in left eye	No OCT MO found in right eye	OCT MO found in right eye	Total in right eye
One M/DH, no exudate, no BH	488	5	493	492	8	500
	99.0%	1.0%	100.0%	98.4%	1.6%	100.0%
Two M/DHs, no exudate, no BH	251	5	256	298	6	304
	98.0%	2.0%	100.0%	98.0%	2.0%	100.0%
Three or more M/DHs, no exudate, no BH	508	18	526	499	17	516
	96.6%	3.4%	100.0%	96.7%	3.3%	100.0%
Only BH, no M/DH, no exudate	61	2	63	56	4	50
	96.8%	3.2%	100.0%	93.3%	6.7%	100.0%
BH + one M/DH, no exudate	45	2	47	43	1	44
	95.7%	4.3%	100.0%	97.7%	2.3%	100.0%
BH + two M/DHs, no exudate	25	0	25	41	1	42
	100.0%	0.0%	100.0%	97.6%	2.4%	100.0%
BH + three or more M/DHs, no exudate	108	23	131	106	25	131
	82.4%	17.6%	100.0%	80.9%	19.1%	100.0%
Exudate + any other lesion	566	81	647	531	67	598
	87.5%	12.5%	100.0%	88.8%	11.2%	100.0%
None/other	971	8	979	967	6	973
	99.2%	0.8%	100.0%	99.4%	0.6%	100.0%
Total	3023	144	3167	3033	135	3168
	95.5%	4.5%	100.0%	95.7%	4.3%	100.0%

© Queen's Printer and Controller of HMSO 2013. This work was produced by Olson *et al*. under the terms of a commissioning contract issued by the Secretary of State for Health. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

This pattern was also reflected in the results presented in *Table 13*, which shows that the odds of having MO increased by a non-significant factor with one or two M/DHs within one DD in comparison with having no M/DHs. However, the OR of having MO dramatically increased to 6.4 when there were more than two M/DHs. This relationship was maintained (OR = 5.2; 95% CI 3.0 to 8.9) after adjusting for the counts of other lesions represented in the same way, and also centre, gender, age, glitazone, diabetes and visual acuity. The fact that the count of several M/DHs within one DD was statistically significant after adjusting for centre, the count of BHs, the count of exudates and visual acuity, suggested that while having any M/DHs might be of little value in predicting MO, having a lot of M/DHs in an eye was still useful for predicting MO in that eye. See *Table 43* for the corresponding data for the right eye.

The unweighted analyses presented in *Tables 11* and *13* were repeated using weighted analyses, the results are presented in *Tables 14* and *15*. These analyses had the effect of up-weighting data from subjects with M/DHs and down-weighting data from subjects with more serious lesions. This small inflation had no influence on the ORs, but made CIs narrower by a very small amount. See *Tables 44* and *45*, respectively, for the corresponding data for the right eye.

TABLE 13 Relationship of individual lesions within one DD to left eye MO (ORs; analysis unweighted)

Left eye									
Feature category within DD (mutually exclusive)	Unadjusted OR[a]	Lower 95% CI	Upper 95% CI	Adjusted OR[b]	Lower 95% CI	Upper 95% CI	Adjusted OR[c]	Lower 95% CI	Upper 95% CI
Zero M/DHs	1.00			1.00			1.00		
One M/DH	0.87	0.41	1.85	0.78	0.35	1.72	0.90	0.41	2.01
Two M/DHs	1.85	0.90	3.79	1.69	0.79	3.60	1.90	0.89	4.10
More than two M/DHs	6.27	3.89	10.11	4.18	2.45	7.13	5.16	2.97	8.94
Zero BHs	1.00			1.00			1.00		
One BH	4.75	3.11	7.24	3.12	1.96	4.98	2.82	1.76	4.54
Two BHs	6.54	3.21	13.34	2.46	1.07	5.61	2.29	0.98	5.34
More than two BHs	20.45	11.08	37.72	6.84	3.39	13.78	5.30	2.61	10.79
Zero exudates	1.00			1.00			1.00		
One exudate	3.30	1.70	6.41	2.39	1.15	4.99	2.81	1.33	5.92
Two exudates	3.25	1.45	7.31	2.83	1.12	7.11	3.81	1.51	9.64
More than two exudates	6.98	4.84	10.07	4.09	2.69	6.19	5.04	3.26	7.78
Visual acuity better[d]	1.00			1.00			1.00		
Visual acuity worse[d]	6.07	4.24	8.68	5.32	3.51	8.07	3.95	2.55	6.12
Visual acuity unknown/missing	4.87	1.08	21.91	7.78	1.39	43.43	4.02	0.66	24.58

a Unadjusted results.
b Adjusting for centre, other features and visual acuity.
c Adjusting for centre, gender, age, glitazone, diabetes (type 1, type 2 and secondary/unknown) and visual acuity. There were too few with amblyopia to adjust for this.
d Visual acuity better (log-MAR < 0.3; Snellen 6/9.5 or better), visual acuity worse (log-MAR ≥ 0.3; Snellen 6/12 or worse).

TABLE 14 Relationship of individual lesions to left eye MO (analysis weighted)

Left eye												
Feature category (mutually exclusive)	Unadjusted OR[a]	Lower 95% CI	Upper 95% CI	Adjusted OR[b]	Lower 95% CI	Upper 95% CI	Adjusted OR[c]	Lower 95% CI	Upper 95% CI	Adjusted OR[d]	Lower 95% CI	Upper 95% CI
M/DH	1.0			1.0			1.0			1.0		
BH not exudate	4.4	2.1	9.1	3.6	1.7	7.6	3.1	1.5	6.7	3.1	1.4	6.8
Exudate + any	6.0	3.9	9.2	5.2	3.3	8.2	6.3	3.9	10.1	5.8	3.6	9.4
Count of lesions of each type												
M/DH	1.22	1.17	1.27	1.17	1.11	1.22	1.17	1.12	1.23	1.17	1.12	1.23
BH	2.52	2.01	3.16	1.51	1.18	1.94	1.43	1.11	1.84	1.41	1.09	1.83
Exudate	1.19	1.14	1.23	1.14	1.10	1.18	1.15	1.11	1.20	1.15	1.11	1.20
Exudate in a one to two DD[e]	1.05	1.03	1.07									

a Unadjusted results.
b Adjusting for centre.
c Adjusting for centre, and the other features.
d Adjusting for centre, gender, age, glitazone, diabetes (type 1, type 2 and secondary/unknown) and visual acuity. There were too few with amblyopia to adjust for this.
e Counts of exudates between one and two DD is non-significant after adjusting for the presence of exudates within one DD and so is dropped from subsequent models.

© Queen's Printer and Controller of HMSO 2013. This work was produced by Olson *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

TABLE 15 Relationship of individual lesions within one DD to left eye MO (OR; analysis weighted)

Left eye									
Feature category (mutually exclusive)	Unadjusted OR[a]	Lower 95% CI	Upper 95% CI	Adjusted OR[b]	Lower 95% CI	Upper 95% CI	Adjusted OR[c]	Lower 95% CI	Upper 95% CI
Zero M/DHs	1.00			1.00			1.00		
One M/DH	0.81	0.36	1.81	0.85	0.37	1.98	0.91	0.39	2.10
Two M/DHs	1.74	0.79	3.79	1.94	0.86	4.37	2.06	0.91	4.69
More than two M/DHs	5.18	2.94	9.14	4.32	2.33	8.00	4.96	2.64	9.32
Zero BHs	1.00			1.00			1.00		
One BH	6.06	3.41	10.77	2.78	1.43	5.40	2.49	1.27	4.88
Two BHs	8.82	3.19	24.43	2.21	0.66	7.39	1.97	0.58	6.76
More than two BHs	23.37	10.00	54.60	5.95	2.20	16.06	4.85	7.79	13.18
Zero exudates	1.00			1.00			1.00		
One exudate	3.87	1.65	9.08	2.39	0.92	6.21	2.78	1.06	7.31
Two exudates	3.81	1.33	10.90	2.74	0.83	9.04	3.41	1.02	11.36
More than two exudates	8.18	5.28	12.67	3.96	2.36	6.66	4.67	2.73	7.99
Visual acuity better[d]	1.00			1.00			1.00		
Visual acuity worse[d]	7.40	4.91	11.14	6.22	3.91	9.89	4.85	2.99	7.88
Visual acuity unknown/missing	4.58	0.52	40.08	9.09	0.69	119.12	5.38	0.37	78.04

a Unadjusted results.
b Adjusting for centres and visual acuity.
c Adjusting for centre, gender, age, glitazone, diabetes (type 1, type 2 and secondary/unknown) and visual acuity. There were too few with amblyopia to adjust for this.
d Visual acuity better (log-MAR < 0.3; Snellen 6/9.5 or better), visual acuity worse (log-MAR ≥ 0.3; Snellen 6/12 or worse).

When the mutually exclusive groups of lesions were included in the model, subjects with BHs (alone or BH with M/DHs) had 4.4 times the odds of having MO and subjects with exudates had 6.0 times the odds of having MO, both compared with those with only M/DHs. After adjusting for centre and other characteristics, these ORs changed by only modest amounts to 3.1 and 5.8, respectively. Of these other variables, only centre, age and visual acuity were significant with older age being associated with higher odds of MO by 3% per additional year older (OR = 1.031; 95% CI 1.011 to 1.051). When visual acuity was included, having worse visual acuity significantly increased the odds of MO by a factor of 3.8 (95% CI 2.3 to 6.3).

An analysis of counts of lesions found that higher counts of M/DHs, BHs and exudates within one DD, considered separately and together, significantly increased the odds of having MO. The relationships were maintained after adjustment for centre, gender, age, glitazone, diabetes (type 1, type 2 and secondary/unknown) and visual acuity. Initially, counts of exudates between one to two DD appeared to be associated with greater odds of MO, but this ceased to be significant after adjusting for exudates within one DD so this count was dropped from subsequent models.

It is important to note here, as with the unweighted analyses, that the count of M/DHs within one DD is statistically significant after adjusting for centre and the count of BHs and the count of exudates. This suggests that having a lot of M/DHs is still useful for predicting MO.

The OR for counts of BHs was again larger than for the other lesions. This does not mean that BHs were more important, but that the increase in the odds of MO for each additional BH was larger than for each additional M/DH or exudate. It was thought that this could be a consequence of there being a smaller range of counts of BHs across subjects than the range of counts of M/DHs or exudates, but as with the unweighted analyses, a sensitivity analysis showed that this was not the case.

This pattern of the increasing prevalence of MO with increasing numbers of lesions was also reflected in the results presented in *Table 13*. The odds of having MO increased by a non-significant factor with one or two M/DHs within one DD in comparison to having no M/DHs and the OR of having MO dramatically increased to 5.2 when there were more than two M/DHs. This relationship was maintained (OR = 5.0; 95% CI 2.6 to 9.3) after adjusting for the counts of other lesions represented in the same way: centre, gender, age, glitazone, diabetes and visual acuity.

Some of the ORs presented in *Table 15* were large with very wide CIs, particularly those for having several BHs. This was because weighting had the effect of reducing the apparent number of subjects with BHs and subjects with MO. This would tend to make the estimates of ORs and their CIs more volatile. After adjustment for centre, gender, age, glitazone use, diabetes type and visual acuity these estimates of the ORs for having BHs were more stable and had inferences more similar to the earlier unweighted analyses. The CIs for missing visual acuity were very wide and the width varied greatly because of the very small numbers with missing visual acuity.

Relationship of individual lesions in both eyes to macular oedema in either eye (analysis unweighted)

Single eye analyses are simple to interpret and are helpful in establishing factors which might contribute to a prediction model. However, they do not show how a strategy might work on a subject where information about the possibility of MO comes from both eyes. Mutually exclusive groups were defined taking information about lesions within one DD from both eyes (*Table 16*). The most serious lesions in either eye took priority in assigning subjects to six groups. Any exudates within one DD in either eye meant the subject was assigned to the exudates group. A BH within one DD in either eye, without any exudates, meant the subject was assigned to the BH group. For those with M/DHs in both eyes (and no BHs or exudates) or M/DHs (and no BHs or exudates) in one eye and no exudates or BH in the other, the group was defined by the number of M/DHs within one DD in the eye with the most M/DHs. A subject with no M/DHs, BHs or exudates within one DD was assigned to the 'None/other both eyes' group. Subjects were also classified into three groups according to visual acuity [better visual acuity (log-MAR < 0.3; Snellen 6/9.5 or better) in both eyes, worse visual acuity (log-MAR ≥ 0.3; Snellen 6/12 or worse) in at least one eye, missing visual acuity in both eyes or missing visual acuity in one and better visual acuity in the other]. The percentage of subjects with MO was calculated for each combination of visual acuity and lesions. There was a substantial two- to fourfold increase in the percentage with MO if visual acuity was worse rather than better in all of the lesion groups. There were very few subjects with MO among those subjects with better visual acuity and zero to two M/DHs within one DD in either eye.

Logistic regression of the relationship of individual lesions in both eyes to macular oedema in either eye (analysis unweighted)

Odds ratios for three models for predicting MO are presented in *Table 17*. From left to right these models include: only lesion type and the count of M/DHs (the unadjusted ORs for visual acuity are also presented here); lesion type and the count of M/DHs adjusted for centre and visual acuity; and lesion type and the count of M/DHs adjusted for centre, visual acuity and demographic variables. Exudates, BHs or more than two M/DHs within one DD in either eye were all predictive of having MO, as was worse visual acuity. Having one or two M/DHs within one DD in an eye did not appear to be any different to having none/other mild lesions. The odds of having MO were increased by factors of 11.2 for having at least one eye with an exudate, 5.9 for having at least one eye with a BH (but no exudates), 3.4 for having more than two M/DHs relative to having eyes with a maximum of one M/DH in either. Having a M/DH in one or both eyes was not significantly different to having a maximum of two M/DHs in either eye or having no lesions or mild lesions (not a M/DH, BH or exudate).

© Queen's Printer and Controller of HMSO 2013. This work was produced by Olson *et al*. under the terms of a commissioning contract issued by the Secretary of State for Health. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

TABLE 16 Presence of MO in either eye by most severe lesion present within one DD, count of M/DHs present within one DD and visual acuity

	Better visual acuity in both eyes			Worse visual acuity in at least one eye			Visual acuity missing in both eyes or missing in one and better in the other		
Lesion grouping	MO either eye	MO absent	Total	MO either eye	MO absent	Total	MO either eye	MO absent	Total
None/other both eyes	1	278	279	3	67	70	0	2	2
	0.4%	99.6%	100.0%	4.3%	95.7%	100.0%	0.0%	100.0%	100.0%
At most one M/DH, either eye	4	416	420	4	63	67	0	2	2
	1.0%	99.0%	100.0%	6.0%	94.0%	100.0%	0.0%	100.0%	100.0%
At most two M/DHs in either eye	6	258	264	2	39	41	0	0	0
	2.3%	97.7%	100.0%	4.9%	95.1%	100.0%	0.0%	0.0%	0.0%
More than two M/DHs, either eye	17	475	492	11	69	80	0	1	1
	3.5%	96.5%	100.0%	13.8%	86.2%	100.0%	0.0%	100.0%	100.0%
BH (not exudate), either eye	22	296	318	29	74	103	0	3	3
	6.9%	93.1%	100.0%	28.2%	71.8%	100.0%	0.0%	100.0%	100.0%
Exudate, either eye	82	752	834	61	126	187	1	1	2
	9.8%	90.2%	100.0%	32.6%	67.4%	100.0%	50.0%	50.0%	100.0%
Total	132	2475	2607	110	438	548	1	9	10
	5.1%	94.9%	100.0%	20.1%	79.9%	100.0%	10.0%	90.0%	100.0%

Visual acuity better (log-MAR < 0.3; Snellen 6/9.5 or better), visual acuity worse (log-MAR ≥ 0.3; Snellen 6/12 or worse).

Having worse visual acuity was predictive of MO and appeared to increase the odds of having MO by a factor of 3.3. Having worse visual acuity increased the size of the odds of MO for all types of lesion and the size of this increase was not significantly different between lesions types.

The ORs for the fully adjusted model including lesions, centre, visual acuity and demographic variables are given in the final three columns. Other variables which were significant in this fully adjusted model were age and centre. For every additional year older the odds of having MO increased by a factor of 1.035 (95% CI 1.022 to 1.049) in other words by 3.5% per year older.

Subjects with one M/DH within one DD in one or both eyes were taken as the reference group so ORs for having MO in either eye are calculated compared with this group. It was not appropriate to take the group with no lesions/other lesions as the reference group for the analyses of both eyes because many of this group would be zero weighted in the weighted analyses.

The 95% CI for the OR of having MO in those with, at most, two M/DHs within one DD in either eye relative to at most one M/DH in either eye included 1. This was also true of the CI for no lesions or lesions other than M/DHs, BHs or exudates in either eye. This meant that there was no significant difference in the odds of having MO in these subjects compared with subjects with at most one M/DH within one DD in either eye. The fact that the count of several M/DHs within one DD was statistically significant after adjusting for centre, the count of BHs and the count of exudates suggested that having a lot of M/DHs within one DD was still useful for predicting MO. This remained true after adjusting for other variables including visual acuity.

TABLE 17 Relationship of individual lesions within one DD to both eyes MO (OR; analysis unweighted)

Both eyes									
Feature category (mutually exclusive)	Unadjusted OR[a]	Lower 95% CI	Upper 95% CI	Adjusted OR[b]	Lower 95% CI	Upper 95% CI	Adjusted OR[c]	Lower 95% CI	Upper 95% CI
At most one M/DH, either eye[d]	1.00			1.00			1.00		
At most two M/DHs, either eye	1.62	0.60	4.36	1.62	0.59	4.41	1.59	0.58	4.38
More than two M/DHs, either eye	3.09	1.40	6.84	3.00	1.34	6.71	3.35	1.49	7.55
BH (not exudate), either eye	8.22	3.85	17.53	6.19	2.85	13.42	5.89	2.70	12.85
Exudate, either eye	9.85	4.79	20.25	8.77	4.20	18.33	11.21	5.31	23.66
None/other both eyes	0.69	0.21	2.32	0.55	0.16	1.85	0.49	0.15	1.67
Visual acuity better[e]	1.00			1.00			1.00		
Visual acuity worse[e]	4.70	3.58	6.17	4.25	3.18	5.70	3.35	2.46	4.56
Visual acuity unknown/missing	1.44	0.19	11.11	3.03	0.36	25.73	2.09	0.25	17.86

a Unadjusted results.
b Adjusting for centre and visual acuity.
c Adjusting for centre, gender, age, glitazone, diabetes (type 1, type 2 and secondary/unknown) and visual acuity. There were too few with amblyopia to adjust for this.
d Subjects with one M/DH within one DD in one or both eyes are taken as the reference group.
e Visual acuity better (log-MAR < 0.3; Snellen 6/9.5 or better), visual acuity worse (log-MAR ≥ 0.3; Snellen 6/12 or worse).

If the 11 subjects with missing visual acuity were excluded, the results were almost identical to those above. With this restriction it was possible to include an interaction between visual acuity and the feature category. There was no evidence of an interaction between these variables and this suggested that both lesion category and visual acuity were important in predicting presence of MO, but the strong effect of having worse visual acuity did not vary between different lesion groups.

The receiver operating curves shown in *Figure 12* correspond to the three models presented in *Table 17*. The model that included only the lesion categories within one DD in at least one eye [at least one M/DH, two M/DHs, more than two M/DHs, at least one BH (but no exudates), at least one exudate] had a receiver operating characteristic curve made up of straight lines between the points (sensitivity and 1-specificity) corresponding to the threshold being changed to include an extra category. For example, if only subjects with at least one exudate were referred, then the sensitivity and specificity would be 0.59 and 0.70, respectively. If subjects with at least one exudate and subjects with at least one BH were referred, then the sensitivity and specificity would be 0.80 and 0.57, respectively (1 – specificity = 0.43). The other two receiver operating characteristic curves correspond to the other two models presented in *Table 17* where one model included lesion type and the count of M/DHs, centre and visual acuity, and the other model included lesion type and count of M/DHs, centre, visual acuity and demographic variables. There were gains in sensitivity and specificity from choosing more complex models as would usually be expected.

Relationship of individual lesions in both eyes to macular oedema in either eye (analysis weighted)

Similar models to those in *Table 17* are presented in *Table 18* for weighted data. Subjects with no lesions in either eye or only lesions other than exudates, BHs and M/DHs within one DD were removed from this analysis. For the weighted data exudates, BHs or more than two M/DHs within one DD in either eye were

© Queen's Printer and Controller of HMSO 2013. This work was produced by Olson *et al*. under the terms of a commissioning contract issued by the Secretary of State for Health. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

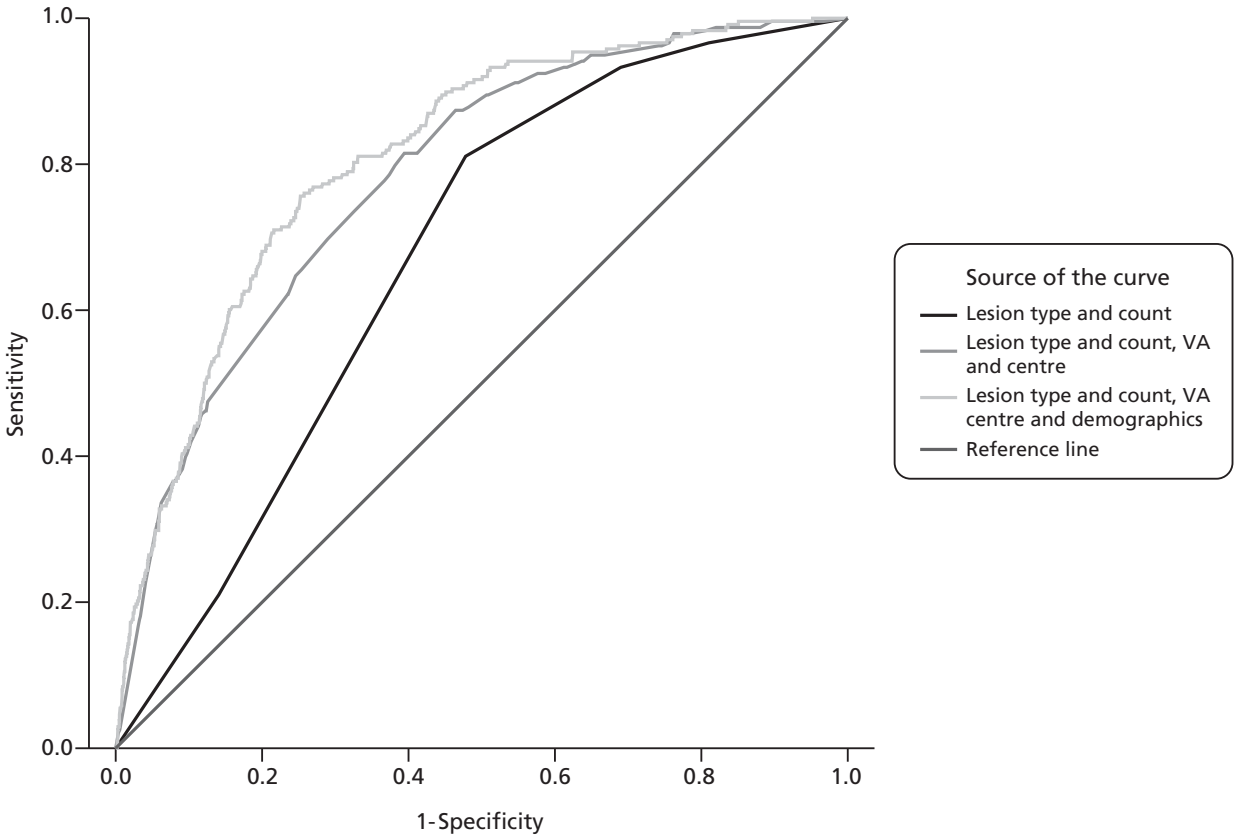


FIGURE 12 Receiver operating characteristic curves for models that are presented in *Table 17* for MO in one eye or both eyes. VA, visual acuity.

TABLE 18 Relationship of individual lesions within one DD to both eyes MO (analysis weighted)

Both eyes									
Feature category (mutually exclusive)	Unadjusted OR[a]	Lower 95% CI	Upper 95% CI	Adjusted OR[b]	Lower 95% CI	Upper 95% CI	Adjusted OR[c]	Lower 95% CI	Upper 95% CI
At most one M/DH, either eye	1.00			1.00			1.00		
At most two M/DHs, either eye	1.62	0.77	3.43	1.61	0.75	3.45	1.63	0.75	3.52
More than two M/DHs, either eye	3.09	1.69	5.64	3.09	1.67	5.71	3.45	1.86	6.42
BH (not exudate), either eye	8.21	3.96	17.01	6.23	2.92	13.27	5.91	2.75	12.70
Exudate either, eye	9.85	5.53	17.56	8.82	4.83	16.11	10.96	5.92	20.29
Visual acuity better[d]									
Visual acuity worse[d]	4.70	3.42	6.45	4.01	2.85	5.62	3.28	2.29	4.68
Visual acuity unknown/missing	2.09	0.14	31.31	3.28	0.20	53.19	2.34	0.14	38.17

a Unadjusted results.
b Adjusting for centre and visual acuity.
c Adjusting for centre, gender, age, glitazone, diabetes (type 1, type 2 and secondary/unknown) and visual acuity.
d Visual acuity better (log-MAR < 0.3; Snellen 6/9.5 or better), visual acuity worse (log-MAR ≥ 0.3; Snellen 6/12 or worse).

all predictive of having MO, as was worse visual acuity (see *Table 18*). Having one or two M/DHs within one DD in an eye did not appear to change the odds of MO.

There were some large ORs observed in the table, again due to the down-weighting of the subjects with more severe groups of lesions so the estimates were more volatile than for the unweighted analyses. This was less pronounced after adjustment for other variables. The odds of having MO were increased by factors of 11.0 for having at least one eye with an exudate, 5.9 for having at least one eye with a BH (but no exudates) and 3.5 for having more than two M/DHs within one DD relative to having eyes with a maximum of one M/DH within one DD in either. Having a maximum of two M/DHs within one DD in either eye was not significantly different to having a M/DH in one or both eyes.

Having worse visual acuity was predictive of MO and appeared to increase the odds of having MO by a factor of 3.3. Having worse visual acuity increased the size of the odds of MO for all types of lesion and the size of this increase was not significantly different between lesions types.

The ORs for the fully adjusted model including lesions, centre, visual acuity and demographic variables are given in the final three columns. Other variables which were significant in the fully adjusted model were age and centre. For every additional year older the odds of having MO increased by a factor of 1.029 (95% CI 1.014 to 1.044); in other words by 3.0% per year older.

As for the unweighted analysis, the 95% CI for the OR of having MO in those with at most two M/DHs within one DD in either eye relative to at most one M/DH within one DD in either eye included 1. This was also true of the CI for no lesions or lesions other than M/DHs, BHs or exudates in either eye.

This meant that there was no significant difference in the odds of having MO in for these subjects compared with subjects with at most one M/DH in either eye. The fact that the count of several M/DHs within one DD was statistically significant after adjusting for centre, the count of BHs and the count of exudates suggested that having a lot of M/DHs was still useful for predicting MO. This remained true after adjusting for other variables including visual acuity.

If the data were restricted to exclude those with missing visual acuity as well as those with no features or features other than M/DHs, BHs and exudates, the results were almost identical to those described above. With this restriction it was again possible to include an interaction between visual acuity and feature category.

There was no evidence of an interaction between these variables and this suggested that both lesion category and visual acuity were important in predicting presence of MO, but the effect of having worse visual acuity did not vary between different lesion groups.

Strategies for detecting macular oedema

Findings from the statistical analysis were used to inform the inclusion of various eye and subject characteristics that might be considered within computer-assisted manual strategies and fully automated strategies. The strategies for detecting MO modelled in this study can be categorised as follows.

1. *Manual grading strategies* These use features, similar to those in existing national criteria, which can be identified by visual inspection by trained graders.
2. *Computer-assisted manual annotation grading strategies* These use more detailed features obtained by manual annotation which are then combined by a software classifier to determine a likelihood that MO is present.

© Queen's Printer and Controller of HMSO 2013. This work was produced by Olson *et al*. under the terms of a commissioning contract issued by the Secretary of State for Health. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

3. *Fully automated annotation grading strategies where no human intervention is required* These use features determined by image analysis software. As with computer-assisted manual strategies, these features are combined by a software classifier to determine a likelihood that MO is present.

Manual grading strategies

The manual strategies used the following photographic features:

- M/DH within one DD
- M/DH between one DD and two DD
- BH within one DD
- BH between one and two DD
- exudates within one DD
- exudates between one DD and two DD.

In addition, the following feature was used:

- visual acuity worse than or equal to Snellen 6/12 (log-MAR 0.3).

Eighteen different combinations of these features were modelled as the manual grading strategies and are listed in *Table 19*. Manual grading strategies 1 and 2 model the current practice in the English and the Scottish national programmes, respectively. The statistical analysis in the earlier part of this chapter showed that exudates, BHs and M/DHs within one DD, as well as reduced visual acuity, all significantly increase the risk of having MO. Furthermore, the risk increases with the number of lesions present within one DD. Lesions outside of one DD, including exudates, were not found to be predictive of MO. Manual grading strategy 4 (M/DHs within one DD only), manual grading strategy 5 (M/DHs or BHs within one DD), manual grading strategy 9 (M/DHs within two DD) and manual grading strategy 10 (M/DHs or BHs within two DD) investigate red lesions alone. Manual grading strategy 6 (exudates within one DD) and manual grading strategy 11 (exudates within two DD) examine bright lesions alone. Manual grading strategy 15 considers visual acuity alone. The remainder of the manual grading strategy 16 strategies explore combinations of red and bright lesions with visual acuity.

Computer-assisted manual annotation grading strategies

A computer-assisted manual annotation grading strategy differs from manual grading in that an automated classifier (described later) is used to combine the image information. This means that improved performance can be achieved because of the association between features. Also, it is possible to include more complex features, such as the number and area of lesions, although this would require additional grading time. As shown in *Table 20*, the computer-assisted manual annotation grading strategy also used computer intensity measurements within the macula (as explained below), visual acuity and other patient information.

Fully automated annotation grading strategies

Like the computer-assisted annotation grading strategies, the fully automated annotation grading strategies take advantage of richer information within the image, but with all the information derived automatically from the image without any human intervention. Clearly, as the automated lesion detection system approaches the performance of the manual observer, the performance of the fully automated annotation systems should be similar to the best computer-assisted annotation grading strategies. Three fully automated annotation grading systems were tested, two of which incorporated visual acuity and other patient information (see *Table 20*). The automated analysis produces technical failures (i.e. images of insufficient quality for automated analysis) and these were handled by using manual grading strategy 16 as illustrated in the flowchart in *Figure 13*.

TABLE 19 Simple manual strategies

Manual strategy	M/DH within one DD	M/DH within two DD	BH within one DD	BH within two DD	Exudate within one DD	Exudate within two DD	VA ≥ 0.3
1	VA		VA		✗	Group[a]	
2			✗		✗		
3	VA (3 +)[b]		VA	VA	✗	VA	
4	✗						
5	✗		✗				
6					✗		
7	VA		✗		✗	✗	
8	✗		✗		✗		
9	✗	✗					
10	✗	✗	✗	✗			
11					✗	✗	
12			✗	✗	✗	✗	
13	✗	✗	✗	✗	✗	✗	
14	✗		✗		✗	✗	
15							✗
16	VA		✗		✗		
17	VA	VA	✗	VA	✗	✗	
18	VA (3+)[b]		✗		✗		

VA, visual acuity.
a Exudates between one and two DD are only included if they comprise a 'group'. The definition of group varies from centre to centre in the English screening programme.
b Three or more M/DHs within one DD.
An ✗ in a column indicates the feature is positive if any lesions of that type are present. A number indicates that a certain number of that feature must be present before it is considered positive. A 'visual acuity' in a column indicates that the feature is only considered positive if the visual acuity is log-MAR 0.3 or worse. EX indicates the presence of exudates and M/DH microaneurysm.

TABLE 20 Computer-assisted manual annotation (CAM) and fully automated (FA) grading strategies

Strategy	Information from manual annotations	Information from automated image analysis	Visual acuity	Patient information (age/diabetes type/male or female)
CAM	✗	✗[a]	✗	✗
FA1		✗		
FA2		✗	✗	
FA3		✗	✗	✗

a Computer intensity measurement only.
An ✗ in a column indicates the feature set was used.

© Queen's Printer and Controller of HMSO 2013. This work was produced by Olson *et al*. under the terms of a commissioning contract issued by the Secretary of State for Health. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

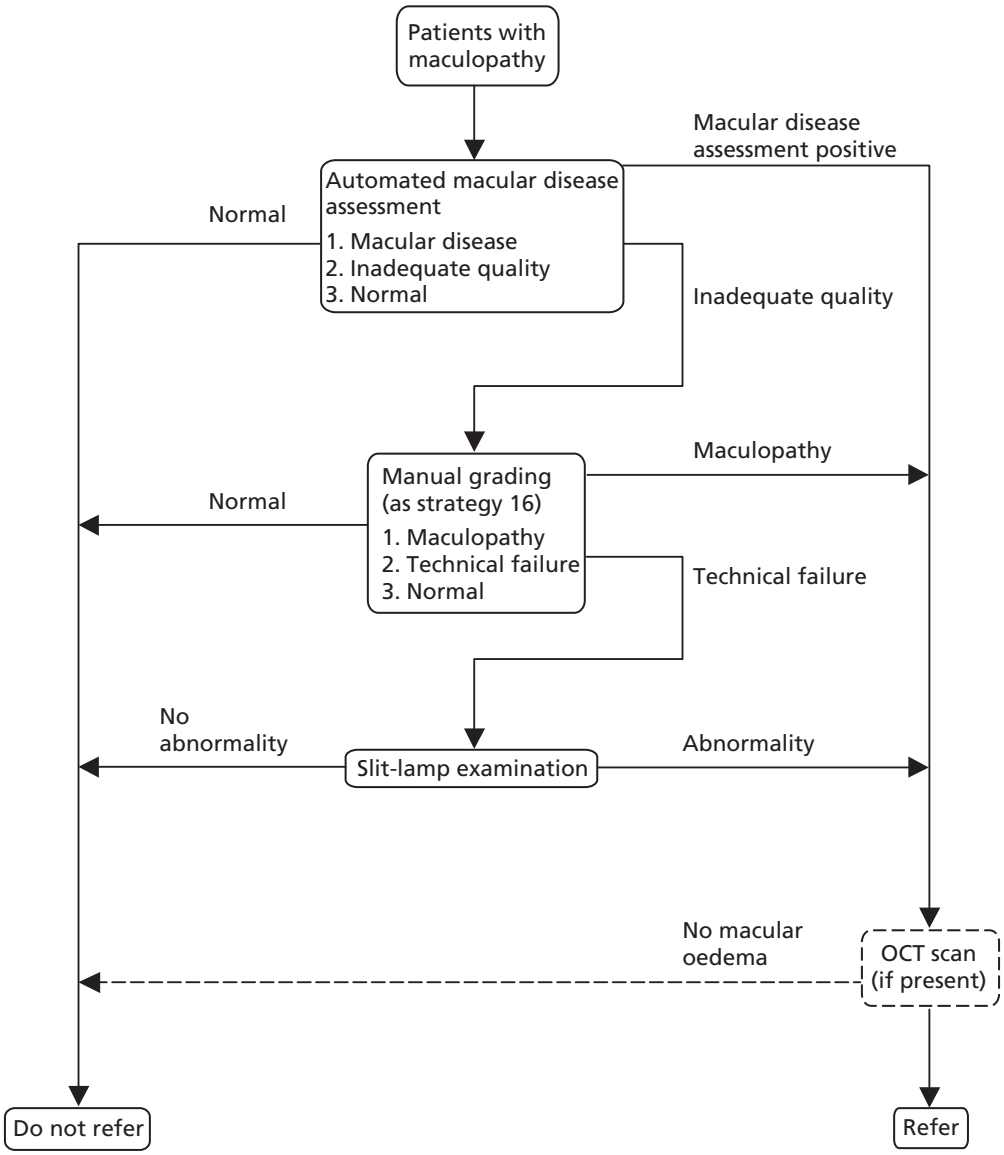


FIGURE 13 A fully automated annotation grading strategy, using manual grading to determine the outcome for technical failures from automated macular disease assessment. Manual grading strategy 16 was used because it performs best relative to the receiver operating characteristic curves shown in *Figure 16* while making a reasonable compromise of sensitivity and specificity.

Automated photograph analysis

This section describes the image analysis methods used in the computer-assisted annotation manual and fully automated annotation grading strategies.

Computer intensity measurements

It was noted that in subjects with MO the macular area often appears darker than the surrounding retina (although this must be distinguished from pupillary shadows, which are a common artefact when the pupil diameter is small). This effect is even more noticeable when imaging using infrared illumination, where the wavelengths are absorbed more strongly by the oedema. To quantify this difference in the macular intensities the following measures were calculated automatically from the retinal photographs (*Figure 14*).

1. The ratio of the mean red and mean green colour channel values within one DD.
2. The ratio of the mean red and mean green colour channel values in the annular region between one and two DD.

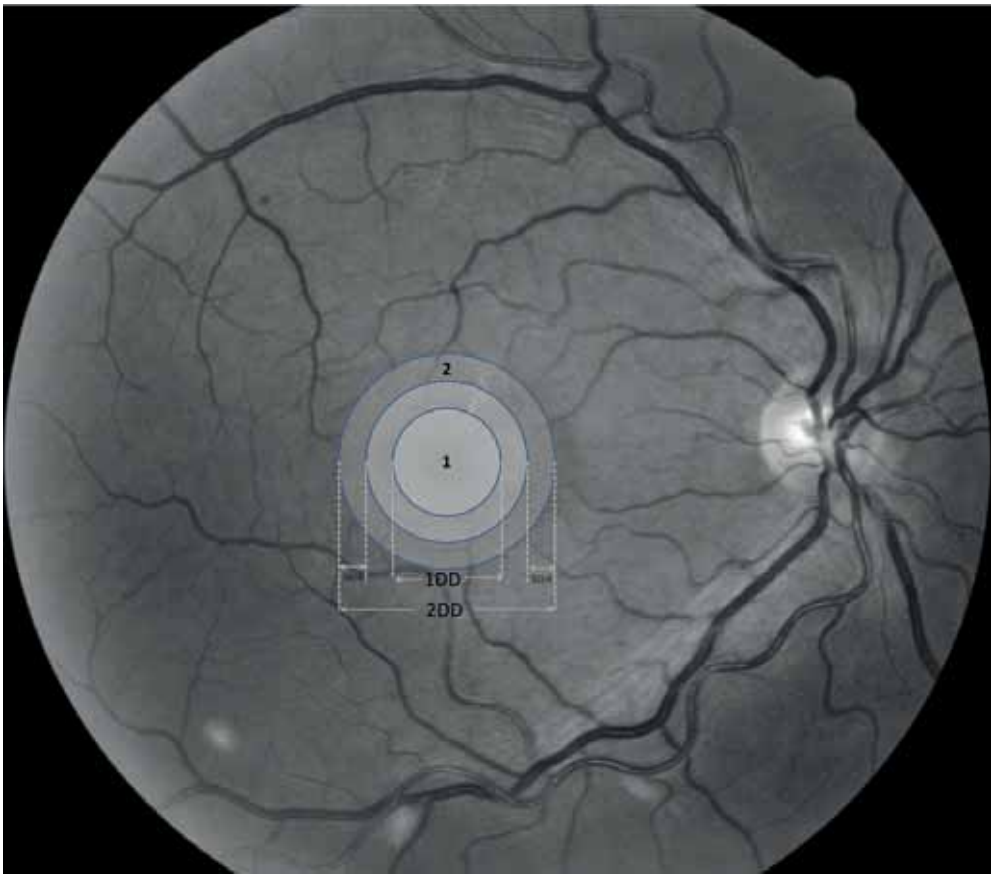


FIGURE 14 Regions of interest centred on the fovea, used to calculate the computer intensity measurements. The first region, labelled '1', is a circle with a diameter approximately equal to that of the optic disc. The second region, labelled '2', is an annular region with an inner diameter of 1.5 disc diameters and outer diameter of a 2 disc diameters.

3. The ratio of the mean green colour channel in the circular one DD region and the annular region between one and two DD.
4. As above, but using the red colour channel.

Note that these measures may be affected by several common features such as reflections from the internal limiting membrane in younger subjects (*Figure 15c*), the presence of bright exudate or drusen lesions (*Figure 15d*) or a shadow cast by a small pupil (*Figure 15e*). However, the measures may have predictive value despite these confounding factors.

Lesion detection and image quality assessment

In order to test automated grading strategies for MO detection, this project made use of software which has been validated for use in diabetic retinopathy screening[30,47–49] and described in previous technical papers.[50–54] A brief description of the automated methods is given here.

The black background was segmented and the image was corrected for uneven illumination. Next, the location of the optic disc and the fovea were determined.[53] The elliptic shape of the temporal arcades was used as a guide to make a rough estimate of the optic disc position, which was then refined using a circular Hough transform. The position of the fovea was subject to geometric constraints relative to the arcades and the disc, and was chosen at a local maximum response of a filter matched to the expected foveal darkening.

Image clarity was assessed by performing small vessel detection within a disc surrounding the detected fovea and then measuring the total vessel length. A decrease in clarity is expected to result in reduced

© Queen's Printer and Controller of HMSO 2013. This work was produced by Olson *et al*. under the terms of a commissioning contract issued by the Secretary of State for Health. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

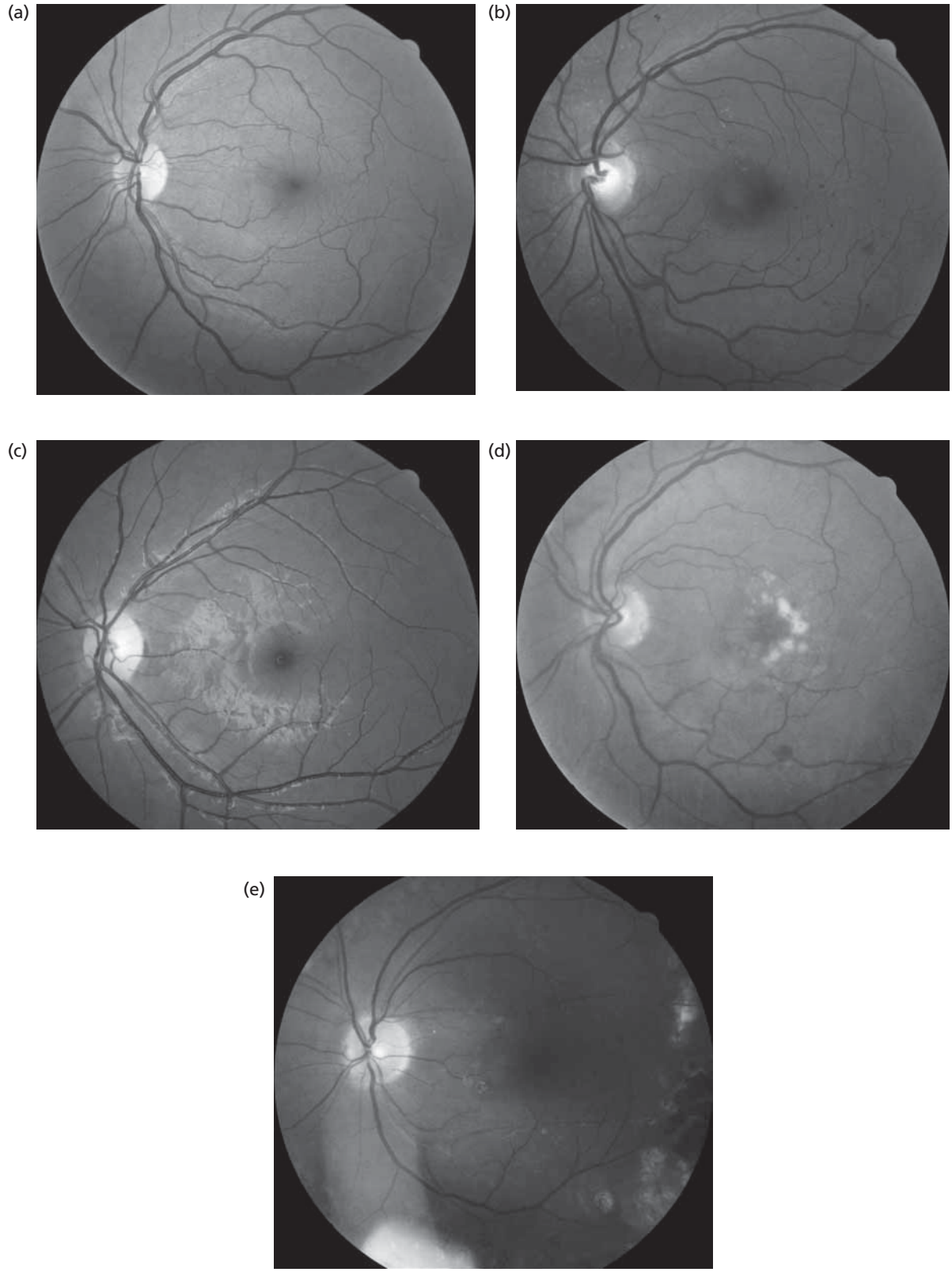


FIGURE 15 (a) Typical macular appearance with a dark region indicating the position of the centre of the fovea; (b) example retina with a darker macular region; (c) retina showing substantial reflections from the internal limiting membrane, which is often a feature in younger subjects; (d) bright and dark lesions can also affect the greyscale measurements; (e) technical problems with the image, such as pupillary shadows, will affect the greyscale measurements.

length of vessels being visible and the low between-person anatomical variation makes this a workable assessment. Images that have poor quality were classed as 'technical failures'.

Lesions (M/DHs, BHs and exudates) were then detected. A non-specific morphological filter was used to determine candidate lesions by separating dot-like objects from the linear objects such as the vasculature.[50] For M/DHs, the non-specific filter was applied at a single scale. For BHs and exudates, the non-specific filter was applied at multiple scales and the results combined.[52,54] Dark objects were detected for BHs and bright objects were detected for exudates. The second stage of lesion detection performed more detailed analysis of the candidate lesions, measuring such features as their area, contrast and, for the dark lesions, the likelihood of it lying on a vessel. For exudates, the distance to the nearest detected M/DH was also used.

For each type of lesion, a set of images in which individual lesions had been annotated was used for training. The training results and the candidate lesion features were supplied to an automated classifier, which produced a value corresponding to the likelihood of each candidate lesion being a true lesion.

Individual candidate lesion likelihoods were combined, as follows, for the candidates within circles centred on the fovea with radii of one and two DD. For M/DHs, the likelihood was thresholded, so allowing a M/DH count to be determined within one DD and two DD from the fovea. For BHs the maximum was taken of the likelihoods of candidates within the one DD circle. For exudates, the mean of the two highest likelihoods for candidates within one DD from the fovea and the mean of the three highest likelihoods for candidates within two DD from the fovea was chosen. The outputs from the automated analysis were thus:

- computer intensity measurements
- image clarity assessment
- count of M/DHs within one DD from the fovea
- count of M/DHs within two DD from the fovea
- count of M/DHs anywhere.
- likelihood of BHs within one DD from the fovea
- likelihood of BHs anywhere
- likelihood of exudates within one DD from the fovea
- likelihood of exudates within two DD from the fovea
- likelihood of exudates anywhere.

Classifiers

The computer-assisted manual and fully automated grading strategies use a classifier to determine whether or not oedema is present given a list of features, such as those above. The classifier is an example of supervised learning. It is trained using a set of data where it is known whether or not MO is present (the MO status was based on the result of their OCT scan). Once the classifier has been trained it can be used to classify previously unseen individual cases giving a probability that oedema is present. The training phase for the classifier is often very time-consuming, but it is required to be done only once.

When evaluating classifier performance it is vital to test using a completely separate set of subjects to those used in the training phase, as recognising previously seen cases is likely to artificially boost the performance. Very often the scarcity of true-positives or true-negatives means that there are not enough cases to divide into separate testing and training sets. In such cases a form of cross-validation may be used where the data are partitioned into subsets. For instance, the data could be divided into three subsets, each containing an equal number of positive cases and an equal number of negative cases. The classifier would then be trained on two of the subsets and tested on the third. In this way the performance is evaluated on the full data set. Increasing the number of subsets will give a greater confidence in the result, but will also increase the calculation time. The limiting case where each subject is a subset is known

© Queen's Printer and Controller of HMSO 2013. This work was produced by Olson *et al*. under the terms of a commissioning contract issued by the Secretary of State for Health. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

as leave-one-out testing, as the training is performed on the entire data set minus just a single test subject.

For this study a classifier known as the Random Forest classifier was used by the computer-assisted and fully automated strategies to combine the feature information and decide whether or not MO is present. This is a non-parametric classifier that makes no assumptions about the form of the feature distributions. The Random Forest classifier is an extension of classification trees. It consists of many trees that all process the same input features and then vote on which class the example belongs to. The final class is the one which achieves the most votes from all the trees in the forest. Like many non-parametric classifiers it takes much longer to train than a classifier such as a linear discriminant analysis classifier, but classification is rapid.

The features used as inputs to the Random Forest classifier are shown in *Table 21*.

Results

The sensitivities and specificities for predicting the presence of MO in at least one eye are presented in *Table 22* for each of the manual grading strategies listed in *Table 19*.

For completeness, positive and negative predictive values (percentage of those referred who have MO present in at least one eye and percentage of those not referred who do not have MO) are presented in *Table 23* for the same manual grading strategies. As the calculations of sensitivity, specificity, positive predictive value and negative predictive value are based on the same four counts of subjects in a table, very similar patterns of performance of the strategies can be seen from *Tables 22* and *23*. For example, after weighting the greater number referred under strategy 1 compared with strategy 2 gave higher sensitivity and lower specificity of strategy 1 compared with strategy 2 (see *Table 22*) and also the lower positive predictive value and higher negative predictive of strategy 1 compared with strategy 2 (see *Table 23*).

The classifiers used in the computer-assisted manual annotation and fully automated annotation grading strategies have a numerical output on a continuous scale and hence the results for these strategies are best displayed as receiver operating characteristics curves (*Figure 16*). *Figure 16* also shows the sensitivity and specificity for the manual strategies that were taken through to the economic analysis. These were manual grading strategy 1 and manual grading strategy 2, respectively, the current English and Scottish grading practices, and manual grading strategy 8 and manual grading strategy 16. Manual grading

TABLE 21 Features for use in the classifier

Features	Source
Manual lesion counts and positions (M/DH, BH, exudates)	Manual annotation
Manual lesion areas	Manual annotation
Retinopathy grade (R0 . . . R4)	Manual grading (following Scottish grading scheme)
Visual acuity (log-MAR)	Patient records
Patient information (age/diabetes type/male or female)	Patient records
Computer intensity measurements	Automated image analysis
Automatic clarity measure	Automated image analysis
Automatic M/DH counts and position (one DD, two DD, more than two DD)	Automated image analysis
Automatic BH counts and position	Automated image analysis
Automatic exudate counts and position	Automated image analysis

TABLE 22 Sensitivity and specificity for the various manual grading strategies and for the operating point chosen for automated grading strategy, FA2. The strategies (1–18) are defined in *Table 19* and the automated strategy utilises features shown in *Table 20*

Unweighted ($n = 3170$; MO $n = 243$)							Weighted ($n = 3170$[a]; MO $n = 176$[a])					
Strategy (1–18)	*n* referred	*n* not referred	Sensitivity %	95% CI	Specificity %	95% CI	*n*[a] referred	*n*[a] not referred	Sensitivity %	95% CI	Specificity %	95% CI
1	1441	1729	81.5	76.1 to 85.9	57.5	55.7 to 59.3	1122	2048	72.6	65.6 to 78.7	66.8	65.1 to 68.5
2	1450	1720	82.3	77.0 to 86.6	57.3	55.5 to 59.1	734	2436	59.5	52.1 to 66.4	79.0	77.5 to 80.4
3	1218	1952	77.8	72.1 to 82.5	64.8	63.1 to 66.6	776	2394	66.1	58.8 to 72.7	78.0	76.5 to 79.4
4	2644	526	93.8	90.1 to 96.2	17.5	16.1 to 18.9	3034	136	96.3	92.4 to 98.2	4.3	3.7 to 5.1
5	2719	451	96.3	93.1 to 98.0	15.1	13.8 to 16.4	3062	108	97.7	94.3 to 99.1	3.5	2.9 to 4.2
6	1028	2142	62.1	55.9 to 68.0	70.0	68.4 to 71.7	590	2580	50.0	42.7 to 57.3	83.2	81.9 to 84.5
7	1753	1417	88.9	84.3 to 92.3	47.5	45.7 to 49.3	1264	1906	75.3	68.4 to 81.1	62.2	60.5 to 63.9
8	2818	352	99.2	97.0 to 99.8	12.0	10.8 to 13.2	3119	51	100.0	97.9 to 100.0	1.7	1.3 to 2.2
9	2872	298	97.1	94.2 to 98.6	9.9	8.9 to 11.1	3129	41	98.6	95.6 to 99.6	1.3	0.9 to 1.8
10	2922	248	98.8	96.4 to 99.6	8.4	7.4 to 9.4	3144	26	99.7	97.3 to 100.0	0.8	0.6 to 1.2
11	1243	1927	65.4	59.3 to 71.1	63.0	61.2 to 64.7	898	2272	54.7	47.4 to 61.9	73.2	71.6 to 74.8
12	1813	1357	88.5	83.9 to 91.9	45.4	43.6 to 47.2	1312	1858	74.3	67.4 to 80.2	60.6	58.8 to 62.3
13	2951	219	99.2	97.0 to 99.8	7.4	6.5 to 8.4	3168	2	100.0	97.9 to 100.0	0.1	0.0 to 0.2
14	2845	325	99.2	97.0 to 99.8	11.0	10.0 to 12.2	3168	2	100.0	97.9 to 100.0	0.1	0.0 to 0.3
15	563	2607	48.1	41.9 to 54.4	84.8	83.4 to 86.0	490	2680	45.3	38.1 to 52.7	86.3	85.0 to 87.5
16	1603	1567	88.1	83.4 to 91.6	52.5	50.7 to 54.3	1001	2169	73.3	66.3 to 79.3	70.9	69.2 to 72.5
17	1784	1386	89.7	85.3 to 92.9	46.5	44.7 to 48.3	1285	1885	77.3	70.5 to 82.8	61.6	59.9 to 63.3
18	1515	1655	86.0	81.1 to 89.8	55.4	53.6 to 57.2	847	2323	68.4	61.2 to 74.8	75.7	74.2 to 77.2
FA2	1154	2015	82.7	77.5 to 87.0	67.4	65.7 to 69.1	921	2248	75.9	69.1 to 81.6	73.7	72.1 to 75.2

FA, fully automated.
a Counts of weighted data have been rounded to the nearest whole number.

© Queen's Printer and Controller of HMSO 2013. This work was produced by Olson *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

NIHR Journals Library www.journalslibrary.nihr.ac.uk

TABLE 23 Positive predictive value and negative predictive value for the various manual strategies and one of the automated strategies, FA2. The strategies (1–18) are defined in *Table 19* and the automated strategy utilises features shown in *Table 20*

Unweighted ($n = 3170$; MO $n = 243$)							Weighted ($n = 3170$[a]; MO $n = 176$[a])					
Strategy (1–18)	No. referred	No. not referred	PPV (%)	95% CI	NPV (%)	95% CI	No.[a] referred	No.[a] not referred	PPV (%)	95% CI	NPV (%)	95% CI
1	1441	1729	13.7	12.1 to 15.6	97.4	96.5 to 98.0	1122	2048	11.4	9.7 to 13.4	97.6	96.9 to 98.2
2	1450	1720	13.8	12.1 to 15.7	97.5	96.6 to 98.1	734	2436	14.3	12.0 to 17.0	97.1	96.3 to 97.7
3	1218	1952	15.5	13.6 to 17.7	97.2	96.4 to 97.9	776	2394	15.0	12.7 to 17.7	97.5	96.8 to 98.1
4	2644	526	8.6	7.6 to 9.8	97.1	95.3 to 98.3	3034	136	5.6	4.8 to 6.5	95.2	90.2 to 97.7
5	2719	451	8.6	7.6 to 9.7	98.0	96.3 to 98.9	3062	108	5.6	4.9 to 6.5	96.3	90.9 to 98.5
6	1028	2142	14.7	12.7 to 17.0	95.7	94.8 to 96.5	590	2580	15.0	12.3 to 18.1	96.6	95.8 to 97.2
7	1753	1417	12.3	10.9 to 13.9	98.1	97.2 to 98.7	1264	1906	10.5	8.9 to 12.3	97.7	96.9 to 98.3
8	2818	352	8.6	7.6 to 9.6	99.4	98.0 to 99.8	3119	51	5.7	4.9 to 6.5	100.0	93.0 to 100.0
9	2872	298	8.2	7.3 to 9.3	97.7	95.2 to 98.9	3129	41	5.6	4.8 to 6.4	94.2	82.6 to 98.2
10	2922	248	8.2	7.3 to 9.3	98.8	96.5 to 99.6	3144	26	5.6	4.8 to 6.5	97.8	83.4 to 99.7
11	1243	1927	12.8	11.0 to 14.8	95.6	94.6 to 96.5	898	2272	10.8	8.9 to 13.0	96.5	95.6 to 97.2
12	1813	1357	11.9	10.5 to 13.4	97.9	97.0 to 98.6	1312	1858	10.0	8.5 to 11.7	97.6	96.8 to 98.2
13	2951	219	8.2	7.2 to 9.2	99.1	96.7 to 99.7	3168	2	5.6	4.8 to 6.4	100.0	31.3 to 100.0
14	2845	325	8.5	7.5 to 9.6	99.4	97.8 to 99.8	3168	2	5.6	4.8 to 6.4	100.0	37.7 to 100.0
15	563	2607	20.8	17.6 to 24.3	95.2	94.3 to 95.9	490	2680	16.3	13.3 to 19.9	96.4	95.6 to 97.0
16	1603	1567	13.3	11.8 to 15.1	98.1	97.4 to 98.7	1001	2169	12.9	11.0 to 15.2	97.8	97.1 to 98.4
17	1784	1386	12.2	10.8 to 13.8	98.2	97.4 to 98.8	1285	1885	10.6	9.0 to 12.4	97.9	97.1 to 98.4
18	1515	1655	13.8	12.2 to 15.6	97.9	97.1 to 98.5	847	2323	14.2	12.0 to 16.8	97.6	96.9 to 98.1
FA2	1154	2015	17.4	15.3 to 19.7	97.9	97.2 to 98.5	921	2248	14.5	12.4 to 17.0	98.1	97.5 to 98.6

FA, fully automated; NPV, negative predictive value; PPV, positive predictive value.
a Counts of weighted data have been rounded to the nearest whole number.

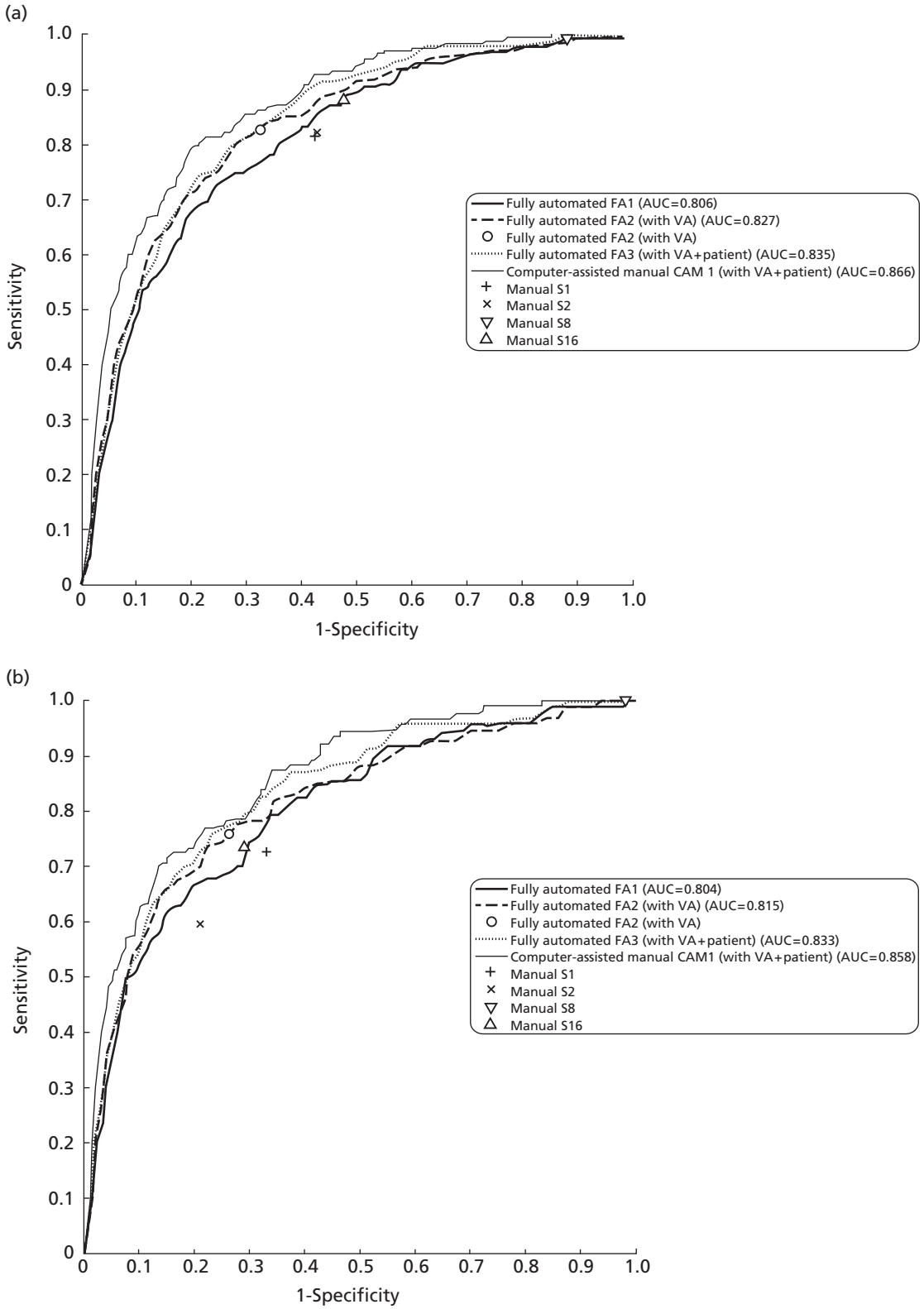


FIGURE 16 Unweighted and weighted receiver operating characteristics curves for fully automated annotation grading (FA) and computer-assisted, manual annotation grading (CAM) strategies and including the operating point chosen for FA2. The performances of the manual strategies associated with current English [strategy 1 (S1)], Scottish [strategy 2 (S2)] grading practice and of the manual grading strategies used in the economic analysis [strategy 8 (S28) and strategy 16 (S16)] are also shown. AUC, area under the receiver operating characteristics curve; VA, visual acuity.

© Queen's Printer and Controller of HMSO 2013. This work was produced by Olson *et al*. under the terms of a commissioning contract issued by the Secretary of State for Health. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

strategy 16 was chosen because it performs best relative to the receiver operating characteristics curves shown in *Figure 16* while making a reasonable compromise of sensitivity and specificity. Manual grading strategy 8 was chosen for its closeness to 100% sensitivity in order to test, in the economic analysis, a strategy involving OCT scan of all patients with photographic surrogate markers of diabetic MO.

Of the fully automated annotation grading strategies, FA2 was chosen for economic analysis since it performs better than fully automated annotation grading strategy FA1 and was similar to fully automated annotation grading strategy FA3 (though with less input information).

In order to use fully automated annotation grading strategy FA2 in the economic analysis it was necessary to choose an operating point for this strategy. The operating point was chosen to lie on a straight line between manual grading strategy 16 and the top left corner of the weighted receiver operating characteristics curve so that it achieves better sensitivity and specificity than manual grading strategy 16. At this point fully automated annotation grading strategy FA2 achieves sensitivity 75.9% and specificity 73.7%. Further details of this operating point are included in *Table 22*. It should be noted that, for the above reason, only this strategy is included in *Table 22*.

The computer-assisted manual annotation grading strategy, CAM, was not analysed further since manual annotation of lesions was considered impractical in a screening context.

Discussion and conclusions

The statistical modelling of the data collected in the study showed that MO was strongly related to the presence of lesions and was consistently higher in subjects with exudates or BHs within one DD than those with just M/DHs within one DD. Having more than two M/DHs within one DD in an eye appeared to be of particular importance in predicting the presence of MO. However, there was no evidence of a relationship between MO and the presence or count of exudates between one and two DD after adjusting for the presence, or count, of exudates within one DD.

Having worse visual acuity was associated with higher prevalence of MO.

The performance of various grading strategies for the detection of MO using retinal photographs and other risk factors for MO were explored. Owing to lower recruitment of patients with only M/DHs in the macula than occurs in the diabetic population, it was necessary to weight subjects to bring proportions to those found in an earlier sequential study.[30]

After weighting, the fully manual grading strategies that model the current grading practices in England and Scotland demonstrated different sensitivity and specificity. The higher sensitivity of the manual grading system in England, at the expense of specificity, can be explained by the referral, in England, but not in Scotland, of patients with M/DH within one DD who have visual acuity greater (worse) than log-MAR 0.3.

The modelling suggested that an ideal grading strategy would be one that takes into account the presence and count of all three types of lesion within one DD and also visual acuity. The fully manual grading strategy 16, that used M/DHs (provided that the visual acuity was worse than log-MAR 0.3), BHs and exudates within one DD, demonstrated an improved specificity by approximately 4% relative to the current grading practice in England, at similar sensitivity. It differs from the current grading practice in England, in that BHs within one DD are referred regardless of visual acuity status, and exudates between one and two DD are ignored. This is supported by the modelling that showed no evidence of a relationship between MO and the presence, or count, of exudates between one and two DD after adjusting for the presence or count of exudates within one DD.

A grading strategy using automated analysis of the macula-centred image performs better than current manual grading strategies. If automated analysis, data are combined with non-image data then further improvements are obtained, largely due to the use of visual acuity. This was supported by the modelling that showed that worse visual acuity was associated with higher prevalence of MO. As visual acuity is a standard measurement performed during diabetic retinopathy screening this additional information could be added to a fully automated strategy with little effort. Therefore, the fully automated annotation grading strategy including visual acuity (FA2) was chosen for further analysis in the next chapter.

The optimum strategy in terms of area under its receiver operating characteristic curve was the computer-assisted manual grading strategy, CAM. This uses the results of manual annotation of the individual lesions in each image. This is a time-consuming procedure and so is unlikely to be considered for routine grading practice. Therefore, this grading strategy was not taken forwards for economic analysis.

© Queen's Printer and Controller of HMSO 2013. This work was produced by Olson *et al*. under the terms of a commissioning contract issued by the Secretary of State for Health. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

Chapter 6 Health economic evaluation

Aims and objectives

The aim of this chapter is to assess the cost-effectiveness of alternative screening strategies for MO, based on the identification of different types and patterns of surrogate photographic markers. The cost-effectiveness of introducing OCT to screening programme pathways, so as to further improve the specificity of these referrals, is also assessed.

An initial analysis focuses on the cost per true case of MO appropriately detected and referred. However, in most instances improvements in the sensitivity of referral criteria come at the expense of decreased specificity. Thus, further analysis was undertaken to assess long-term cost-effectiveness, i.e. to ascertain whether or not the increased costs associated with more sensitive grading strategies, resulting from more referrals to ophthalmology clinics (appropriate and inappropriate), are worth incurring given the potential health benefits. A Markov microsimulation model was developed for this purpose.

Assessment of cost per case of macular oedema detected

Overview

In previous chapters the sensitivity and specificity of the alternative screening strategies has been established using the OCT outcome as the reference standard. There were 329 individuals in the data set where the OCT outcome was a technical failure in at least one eye, resulting in unknown oedema status at the patient level (OCT technical failure), and a further 41 individuals with two inadequate photographs. These subjects were excluded from the analysis of sensitivity/specificity at the patient level, but the expected costs associated with OCT technical failure were incorporated in the economic models comparing alternative grading strategies involving the use of OCT within the screening programme.

The short-term costs to the NHS associated with the alternative referral criteria were estimated by applying unit costs for screening and ophthalmology referral, according to the care pathway that patients in the data set would be expected to follow under the alternative strategies. For strategies entailing the use of OCT for patients within the screening programme, prior to referral, OCT costs were also applied. The individual cost estimates were derived as described below.

Following this unadjusted analysis, a further analysis was undertaken whereby the proportions of patients with different types/patterns of surrogate markers were weighted to reflect the relative frequency with which these surrogate markers would be expected to arise in a consecutive screening cohort. In order to achieve this, four patterns with retinal pathology categories were defined based on the type and location of surrogate markers present, and these were applied in a hierarchical manner such that patients were assigned to the first category for which either of their eyes met the inclusion criteria. The sensitivity and specificity of the alternative grading strategies were determined separately for patients within each of these categories. The resultant estimates were then applied in a decision tree model (*Figure 17*) where the proportions of patients within each surrogate marker category were set equal to those observed in a prior screening cohort study carried out in Grampian.[33] This approach is equivalent to that used to estimate weighted sensitivities and specificities in the previous chapter.

As the cost of a referral to the eye clinic was found to be a key driver of cost-effectiveness, and the estimate for this parameter varied substantially between Scotland and England, separate analyses were conducted for the two countries.

© Queen's Printer and Controller of HMSO 2013. This work was produced by Olson *et al*. under the terms of a commissioning contract issued by the Secretary of State for Health. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

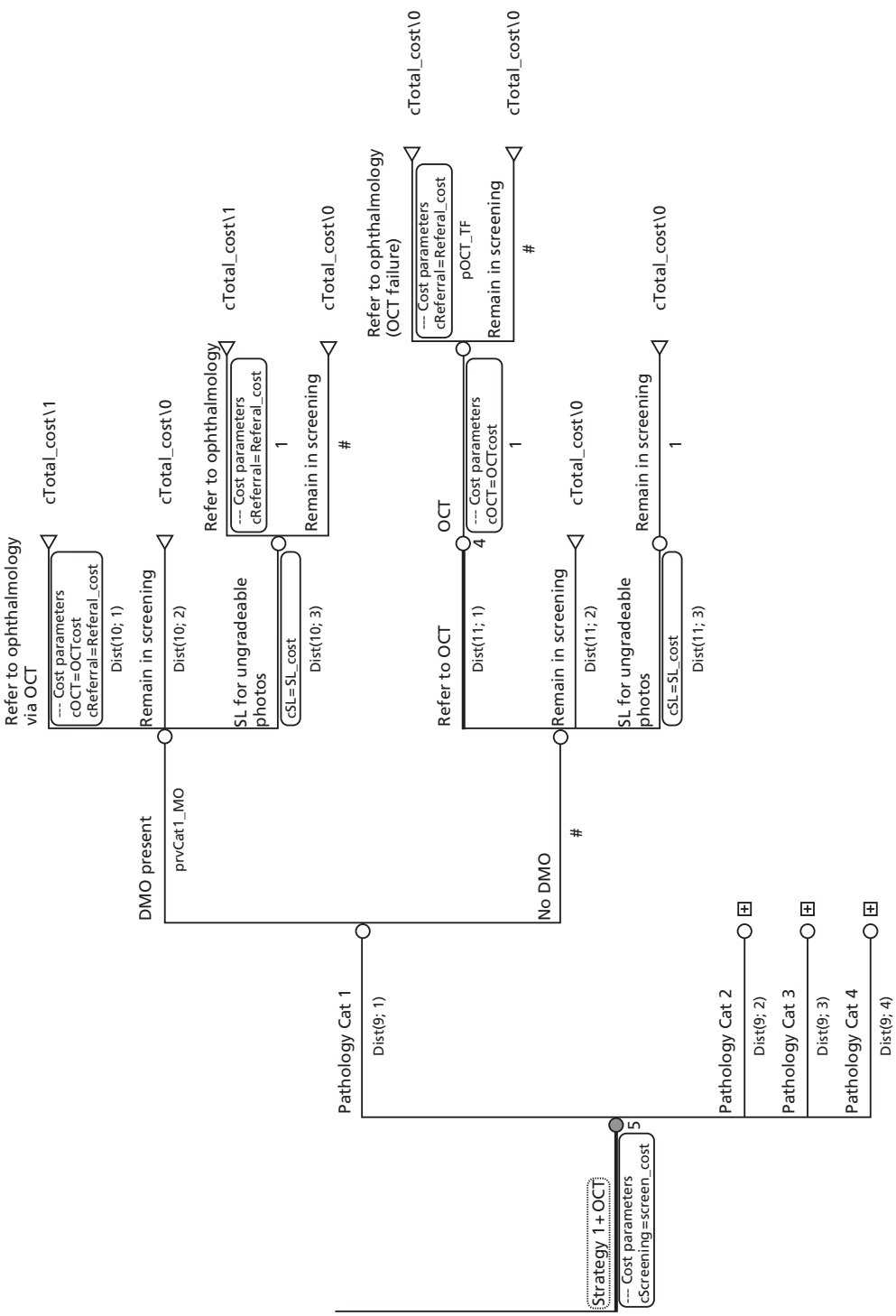


FIGURE 17 Structure of decision tree model for estimating cost per case detected (showing the arm for manual grading strategy 1 coupled with OCT). cSL, cost of slit-lamp; DMO, diabetic macular oedema; SL, slit-lamp.

Comparators

It was not feasible to consider all potential grading strategies in the economic analysis. The grading strategies chosen included the current English manual grading strategy (1) and the current Scottish manual grading strategy (2). Manual grading strategy 2 is one of the highest specificity simple manual grading strategies after weighting, whereas manual grading strategy 1 represents a moderate sensitivity, moderate specificity manual grading strategy (see *Table 23*). Based on examination of the patient-level sensitivities/specificities of alternative simple manual strategies (see *Table 23*), manual grading strategy 16 was also included as a potential alternative to manual grading strategy 1 (having similar sensitivity and higher specificity). In addition, consideration was given to the potential cost-effectiveness of utilising the fully automated annotation grading strategy with the inclusion of visual acuity (denoted FA2) running at an operating point on its receiver operating characteristic curve (see *Figure 16*), with slightly higher sensitivity and better specificity than manual grading strategy 16. Where automated grading flagged a patient's images as being of 'insufficient quality for grading', these patients were modelled to fall back to manual grading with manual grading strategy 16. Manual grading strategy 16 was chosen because of its similar sensitivity and superior specificity to manual grading strategy 1.

The impact of combining the selected manual grading strategies with OCT within the screening pathway prior to referral, to minimise false-positive referrals to eye clinics, was also assessed (see *Figure 17*). In addition, the 100% sensitive manual grading strategy 8 (any exudate or BH within one DD, or any M/DH within one DD and visual acuity $\geq$ 0.3 log-MAR) coupled with OCT prior to referral was considered.

The automated annotation grading strategy incorporating further patient characteristics was not included in the economic analysis as this was incomparable with the simple manual grading strategies (which did not utilise such information). Similarly, manual grading strategies relying on the full manual annotation of all retinal images were not included as these would be impractical for clinical practice. Thus, in total nine strategies were included in the analysis.

Costs

Costs were estimated using a variety of sources as described below. All costs were expressed in UK sterling for the 2009–10 financial year. Although discounting was not required for the cost-effectiveness analysis, the costs of capital items were annuitised over the useful lifespan of the item in question and a discount rate of 3.5% per annum was applied to account for the opportunity cost of the investment over time.[55] *Table 24* summarises all the unit costs applied in the analysis of the cost per case detected.

Screening costs

Screening costs were estimated via a survey of participating centres and included the total equipment, staff, space and overhead costs. For three of the participating screening centres these individual inputs were costed using appropriate unit costs, and the resultant total cost was divided through by the number of patients screened annually to give an estimate of the average cost per patient screened/graded (based on digital photography with manual grading).

Capital equipment was costed using 2009/10 market prices and costs were annuitised over the useful lifespan of each item as described above. Staff time was costed using published unit costs[56] [Personal Social Services Research Unit (PSSRU), 2010] reflecting the NHS salary, superannuation, national insurance, training, overhead, and building space costs attributable to different grades of staff. The costs of consumables were estimated based on programmes' reported non-staff/equipment expenditures.

Alternatively, two programmes had pre-existing charges available reflecting the cost per patient to the NHS of providing screening. These costs were applied for these centres as they were felt to accurately reflect the actual cost per patient to the NHS.

The resultant estimated cost per patient screened ranged from £32.02 to £64.56. Since the average screening costs appeared to vary by country, because of differences in screening and grading protocols

© Queen's Printer and Controller of HMSO 2013. This work was produced by Olson *et al*. under the terms of a commissioning contract issued by the Secretary of State for Health. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

TABLE 24 Unit costs applied in the cost-effectiveness analysis

Item	Unit cost (£)	Range for deterministic sensitivity analysis	Distribution parameters for probabilistic sensitivity analyses	Source
Screening programme costs				
Cost per screen (digital photography) in England	46.69	£32.50 to £64.56	Gamma	Bottom-up costing/charges
Cost per screen (digital photography) in Scotland	33.13	£32.02 to £34.23	Gamma	Bottom-up costing
Marginal cost per OCT examination (within the screening programme)	31.96	£30.00 to threshold	Gamma	Reported cost + bottom-up costing
Marginal cost of slit-lamp (within the screening programme)	27.29	£15.00 to £124.00	Gamma	Bottom-up costing
Eye clinic costs (England)				
Outpatient with slit-lamp alone	124.00	£102.00		Department of Health, 2010[57]
Outpatient with OCT only	160.00	£137.00		Department of Health, 2010[57]
Outpatient with slit-lamp + OCT	160.00	£137.00		Department of Health, 2010[57]
Outpatient with slit-lamp + fluorescein	160.00	£137.00		Department of Health, 2010[57]
Weighted average cost per initial referral	143.35	£120.81		
Eye clinic costs (Scotland)				
All diagnostic referrals	90.00			ISD, 2010[58]

ISD, Information Services Division.

and staffing structures, the average cost per patient in Scotland (£33.13) was applied in a Scottish-specific analysis, whereas the average cost per screen in English centres (£46.69) was applied in an English-specific analysis.

Automated annotation grading costs

The cost of utilising the automated annotation grading algorithm represents one of the most uncertain parameters in the model, at least within the English screening context. A version of automated disease/no disease grading has already been implemented in Scotland, and it is estimated that it is currently achieving a grading cost-saving to the screening programme [i.e. the additional cost of information technology (IT) infrastructure and hosting are being more than offset by the savings made in terms of reduced manual grading workload]. This is being achieved with automated disease/no disease grading taking place on a central server with capacity to replace 40% of level 1 (disease/no disease) manual grading episodes in Scotland. It is anticipated that the system could realise greater cost-saving to the screening programme if scaled up further (i.e. further manual grading cost-savings may be greater than the marginal costs of scaling up automation). Thus, it seems reasonable in this context to assume that using automated grading to trigger referral for those with pathology in the macula would not increase the overall image grading costs to the screening programme. Therefore, in the base-case analysis we assumed that the use of automated annotation grading would have zero impact on net grading costs. However, given the current level of uncertainty surrounding the use of automated grading in England and how it might be implemented, we conducted a threshold sensitivity analysis around the marginal cost (per patient) of

introducing automated grading for this group of patients. Of course, using automated grading in the referral pathway for MO represents just one potential use of an automated grading system within screening programmes.

Optical coherence tomography costs (within screening pathway)

The additional cost of an OCT examination for patients with suspected MO, within the screening pathway, was estimated based on the experience of two screening programmes offering OCT. The administration, clinical staff time, equipment, maintenance, space and overhead requirements were costed using the bottom-up approach described above. Total costs were divided through by the number of OCT screens generated each year at the lead centre, to estimate the OCT costs per patient screened (see *Appendix 2*). Given the uncertainty surrounding the cost of providing OCT within screening programmes (due to its currently limited use), a threshold analysis was conducted to determine the maximum cost at which OCT could be provided within screening programmes, prior to referral, while remaining cost-saving to the NHS. For the maximally sensitive manual grading strategy (manual grading strategy 8 with OCT), it was assumed that there would be a manual grading cost saving of £3.81 per patient within the English system (no need for first or second full disease grading). The savings were estimated based on a grading time of 4 minutes per patient at Agenda for Change band 5/6. The hourly costs of band 5 and band 6 graders were assumed to be similar to those of band 5 and band 6 nurses (£26 and £31 per hour, respectively).[56] A grading cost saving of £1.90 per patient was assumed for strategy 8 within the Scottish system (i.e. no need for level 2 manual grading).

Slit-lamp examination for patients classified as ungradeable

Since the application of the different strategies at the patient level resulted in different proportions of patients being classified as photographic technical failures, it was necessary to incorporate the extra costs of carrying out a slit-lamp examination for these patients within the screening programme. This was done using the same bottom-up approach as was used to estimate the cost of OCT examination within screening programmes – based on reported resource inputs required at a number of participating centres (see *Appendix 2*). The cost estimates varied according to the staff grade conducting the examination. A cost of £27.29 was applied in the base-case analysis (the average of two estimates based on an associate specialist conducting the examination, and two estimates based on band 6 or band 7 graders conducting the examination). Finally, since some centres refer patients to hospital eye services for photographic technical failures, a sensitivity analysis was undertaken whereby the payment by results (PbR) outpatient tariff[57] was applied for slit-lamp examinations occurring within the screening programme (£124).

Initial referral costs

The costs associated with initial referral to the eye clinic were based on a survey of participating centres, which ascertained the average number of clinic visits required to confirm or refute a diagnosis of MO (one visit), and the proportions of patients following alternative diagnostic pathways: (1) slit-lamp biomicroscopy alone (46%); (2) OCT alone (4%); (3) slit-lamp biomicroscopy followed by OCT (47%); or (4) slit-lamp biomicroscopy followed by fluorescein angiography (3%).

For the English-specific analysis, it was determined how alternative diagnostic procedures mapped to the Office of Population Censuses and Surveys (OPCS) procedure codes and consequently the Department of Health, Healthcare Resource Groups (HRGs). Under this approach, an OCT examination mapped to the OPCS procedure code 'Tomography evaluation of retina' (C87.3), whereas fluorescein angiography mapped to the OPCS code 'Fluorescein angiography of the eye' (C86.5). Both these codes mapped to the HRG BZ23Z (Vitreous Retinal Procedures, Category 1), the charge for which was £160 according to the PbR 2010 national tariff.[57] Initial referrals where slit-lamp alone was used to guide diagnosis were charged against the tariff for outpatient attendances [ophthalmology outpatient first attendance (£124)]. A sensitivity analysis was also conducted whereby OCT upon referral and referral without OCT was costed according to the corresponding HRGs in the NHS reference costs.[59]

© Queen's Printer and Controller of HMSO 2013. This work was produced by Olson *et al*. under the terms of a commissioning contract issued by the Secretary of State for Health. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

For Scottish programmes, the average Information Services Division, NHS National Services Scotland, specialty cost[59] for a consultant-led ophthalmology outpatient attendance was applied for all initial referrals to ophthalmology clinics (£90).

Analysis of cost per case detected

The cost per case detected was assessed by determining the expected screening, referral and treatment costs, and the expected number of true cases of MO referred under the alternative grading strategies. Initially, the analysis was conducted by applying the unadjusted sensitivity and specificity estimates in the decision tree model, with the prevalence of MO set equal to that observed in the clinical data set (7.7%). To capture the expected costs associated with OCT technical failure, for grading strategies incorporating OCT, the observed OCT technical failure rate (10.5%) was applied and those experiencing a technical failure were modelled to incur the cost of both the OCT examination within the screening programme and a subsequent referral to the eye clinic. Patients assigned the outcome of photographic technical failure by a particular strategy were modelled to receive a slit-lamp examination within the screening programme. It was assumed that these patients would receive a slit-lamp examination which would ultimately result in the correct outcome being assigned (refer or recall). Thus, a number of patients correctly identified as having MO with each strategy are in fact identified following a slit-lamp examination.

A comparison of the proportions of patients in the surrogate marker categories with the proportions observed in these categories in the prior cohort study (*Table 25*) suggested oversampling in the current study of patients with BHs or exudates within one DD and undersampling of patients with only M/DHs in the macula, or exudate(s) between one and two DD. To adjust for this in the cost-effectiveness analysis, patients were first assigned to one of the surrogate marker categories in *Table 25*. Moving through the categories in descending order they were assigned to the first category that either of their eyes satisfied. The prevalence of MO and the sensitivity/specificity of the alternative referral strategies were then estimated separately for each of the patient surrogate marker categories (*Table 26*) and applied in the decision tree model with the categories reweighted using the frequency proportions provided in *Table 25*. These frequency proportions were estimated using a subset of 1099 patients with surrogate markers arising in a consecutive cohort of 6370 individuals (with complete data) screened in Grampian.[33] They are suggestive of a maculopathy incidence rate (based on an English referable maculopathy grade) of ~6.3%. The process of frequency weighting patients in the decision model is equivalent to the design weighting procedure carried out for the statistical analysis.

Results were expressed for a cohort of 3170 patients with lesions within two DD centred on the fovea. Total screening and referral costs, true cases detected and false-positive referrals were tabulated for each strategy. Each grading strategy was compared incrementally in terms of its additional cost per extra case of MO detected in comparison with the most specific strategy. The results are presented both with and without the use of OCT prior to referral. The cost and effect of each grading strategy is also plotted graphically on the cost-effectiveness plane, with a line joining those grading strategies that represent potentially cost-effective options dependent on decision makers' willingness to pay per extra case of MO detected (cost-effectiveness frontier).

TABLE 25 Patient category proportions used in the cost-effectiveness analysis

Category	Proportion	95% CI
1. Exudate(s) within one DD	203/1099 = 0.185	0.163 to 0.209
2. BH(s) within one DD	50/1099 = 0.045	0.034 to 0.058
3. M/DH(s) (no BHs) within one DD	829/1099 = 0.754	0.729 to 0.779
4. Exudate(s) between one and two DD	17/1099 = 0.015	0.009 to 0.024

TABLE 26 Patients referred under alternative screening strategies by patient pathology category and OCT status

	1. Exudates within one DD (proportional weight = 0.1847)		2. BH(s) within one DD (proportional weight = 0.0455)		3. M/DH within one DD (proportional weight = 0.7543)		4. Exudate(s) between one and two DD (proportional weight = 0.0155)		5. No diabetic pathology (proportional weight = 0)	
Grading strategy	OCT+ve (*n* = 144)	OCT−ve (*n* = 880)	OCT+ve (*n* = 51)	OCT−ve (*n* = 372)	OCT+ve (*n* = 44)	OCT−ve (*n* = 1327)	OCT+ve (*n* = 0)	OCT−ve (*n* = 27)	OCT+ve (*n* = 4)	OCT−ve (*n* = 321)
Strategy 1	144/144 (100%)	877/880 (99.7%)	32/51 (62.7%)	100/372 (26.8%)	20/44 (45.5%)	238/1327 (17.3%)	–	24/27 (88.9%)	2/4[a] (50%)	4/321 (1.2%)
Strategy 2	144/144 (100%)	878/880 (99.8%)	51/51 (100%)	372/372 (100%)	3/44[a] (6.8%)	0/1327 (0%)	–	0/27 (0%)	2/4[a] (50%)	0/321 (0%)
Strategy 8	144/144 (100%)	879/880 (99.9%)	51/51 (100%)	372/372 (100%)	44/44 (100%)	1326/1327 (99.9%)	–	0/27 (0%)	2/4[a] (50%)	0/321 (0%)
Strategy 16	144/144 (100%)	878/880 (99.8%)	51/51 (100%)	372/372 (100%)	17/44 (38.6%)	139/1327 (10.5%)	–	0/27 (0%)	2/4[a] (50%)	0/321 (0%)
FA2	131/144 (95.1%)	488/880 (73.4%)	40/51 (84.3%)	138/371 (61.2%)	26/44 (72.7%)	254/1327 (33.0%)	–	5/27 (44%)	4/4 (100%)	52/321 (21.2%)

FA, fully automated; –ve, negative; +ve, positive.

a Cases referred through slit-lamp because of photographic failure.

Strategy 1 (current English criteria); strategy 2 (current Scottish criteria); strategy 8 (any exudates, BHs or M/DHs within less than one DD); strategy 16 (any exudates or BHs within less than one DD, or M/DH(s) with visual acuity > 0.3 log-MAR).

© Queen's Printer and Controller of HMSO 2013. This work was produced by Olson *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

Sensitivity analysis

Deterministic sensitivity analysis

Owing to uncertainty surrounding the per patient cost of providing OCT within a screening programme's pathway, a threshold analysis was undertaken to determine the maximum cost at which OCT could be provided within screening, prior to referral, while still resulting in cost-savings to the NHS. Further deterministic sensitivity analysis was undertaken to assess the sensitivity of findings to the relative frequency of different surrogate marker categories within the screening cohort, the marginal cost associated with automated grading, the cost of slit-lamp examination, and the cost of referrals to ophthalmology (i.e. applying the lower reference costs).

Probabilistic sensitivity analysis

In order to characterise the uncertainty surrounding modelled point estimates of costs and effects, arising from the joint uncertainty surrounding all input parameters, appropriate probability distributions were specified for each parameter in the model. Probabilities (sensitivities, specificities, prevalence, and frequency proportions) were specified as beta or Dirichlet distributions using the observed clinical data, whereas cost distributions for screening and OCT were specified based on the observed range of estimated values for each event. In the base-case analysis the screening cost distribution was centred on the average per patient cost estimate, and a gamma distribution was assumed. A variance parameter was selected such that the high and low screening cost estimates (see *Table 24*) were contained within the 2.5th to 97.5th percentiles of the distribution. Although referral unit costs were applied deterministically in the models, the proportion of patients receiving only a slit-lamp examination on referral to the eye clinic was allowed to vary within a uniform distribution, reflecting the reported variation in this parameter across participating centres (0–95%).

Results

Cost per case detected

Table 26 shows the number and proportion of patients in the clinical data set who would be referred/recalled by the alternative grading strategies, by surrogate marker category and the presence/absence of diabetic MO. These data were used to populate the decision tree model so that cost-effectiveness could be estimated based on the patient category frequencies observed within the current study, and also with the more realistic frequency weights from the prior cohort study.[33]

Tables 27 and *28* show the anticipated costs and consequences, from one round of screening, of using the alternative grading strategies to screen patients in the clinical data set in their unadjusted frequencies (applying English and Scottish screening and referral costs, respectively). *Tables 29* and *30* show the expected costs and outcomes for a screening cohort of the same size, but with the patient categories reweighted to reflect their expected frequency within screening programmes. *Table 29* presents the findings applying the English unit costs for screening and referral, whereas *Table 30* shows the same results applying Scottish unit costs. *Figures 18* and *19* compare all the grading strategies graphically on the cost-effectiveness plane (applying English and Scottish costs, respectively). Grading strategies falling above or behind the lines joining points on the plane represent options that are dominated (more costly and less effective than alternative options) under base-case assumptions. Those grading strategies falling on the lines offer potentially cost-effective options dependent on decision makers' willingness to pay per extra case of MO detected.

The results show that, although manual grading strategy 2 (current Scottish practice) is slightly more sensitive, less specific and more costly than the manual grading strategy 1 (current English practice), based on the unweighted analysis (see *Tables 27* and *28*), the reweighting reverses this finding such that manual

TABLE 27 Estimated screening and referral costs and true cases of MO detected under alternative referral criteria (for the 3170 subjects with usable OCT status and photographs – based on English screening and referral costs)

Strategy	Number of MO cases	Total cost (£)	MO cases detected and referred (*n*)	False-positive referrals (*n*)	Incremental cost (£)	Incremental cases	Incremental cost per case detected (£)
Strategy 2	243	358,015	200	1250	a	a	a and b
Strategy 1	243	356,806	198	1243	–1209	–2	605[b,c]
Strategy 16	243	379,743	214	1389	21,728	14	1552
FA2[d]	243	314,986	201	253	–43,029	1	Dominant
Strategy 2 + OCT	243	243,637	200	131	a	a	a and b
Strategy 1 + OCT	243	243,104	198	130	–533	–2	267[b,c]
Strategy 16 + OCT	243	252,433	214	145	8796	14	628
FA2 + OCT[d]	243	229,355	201	145	–14,282	1	Dominant
Strategy 8 + OCT	243	299,979	241	269	56,342	41	1374

FA, fully automated.
a Reference strategy.
b Strategy more costly and less effective than an alternative strategy (dominated).
c Represents a cost-saving per case missed relative to the reference strategy.
d Figures in table based on assumption that fully automated grading can be implemented at zero net increase in grading costs.

TABLE 28 Estimated screening and referral costs and true cases of MO detected under alternative referral criteria (for the 3170 subjects with usable OCT status and photographs – based on Scottish screening and referral costs)

Strategy	Number of MO cases	Total cost (£)	MO cases detected and referred (*n*)	False-positive referrals (*n*)	Incremental cost (£)	Incremental cases	Incremental cost per case detected (£)
Strategy 2	243	237,433	200	1250	a	a	a and b
Strategy 1	243	236,799	198	1243	–634	–2	317[b,c]
Strategy 16	243	251,064	214	1389	13,631	14	974
FA2[d]	243	210,171	201	253	–27,262	1	Dominant
Strategy 2 + OCT	243	182,909	200	131	a	a	a and b
Strategy 1 + OCT	243	182,592	198	130	–317	–2	158.50[b,c]
Strategy 16 + OCT	243	190,200	214	145	7291	14	520.79
FA2 + OCT[d]	243	170,229	201	145	–12,680	1	Dominant
Strategy 8 + OCT	243	235,848	241	269	52,939	41	1291.20

FA, fully automated.
a Reference strategy.
b Strategy more costly and less effective than an alternative strategy (dominated).
c Represents a cost-saving per case missed relative to the reference strategy.
d Figures in table based on assumption that fully automated grading can be implemented at zero net increase in grading costs.

© Queen's Printer and Controller of HMSO 2013. This work was produced by Olson *et al*. under the terms of a commissioning contract issued by the Secretary of State for Health. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

TABLE 29 Screening and referral cost per true case of MO detected under alternative referral criteria (for 3170 patients); adjusted for expected frequency of different patient categories and based on English screening and referral costs

Strategy	Adjusted number of MO cases	Total cost (£)	MO cases detected and referred (*n*)	False-positive referrals (*n*)	Incremental cost (£)	Incremental cases	Incremental cost per case detected (£)
Strategy 2	176	255,448	105	628	a	a	a
Strategy 1	176	310,791	128	995	55,343	23	2406[b]
Strategy 16	176	293,345	129	872	37,897	24	1579[b]
FA2[c]	176	281,465	134	786	26,017	29	897
Strategy 2 + OCT	176	197,955	105	67	a	a	a
Strategy 1 + OCT	176	218,896	128	105	20,941	23	910[b]
Strategy 16 + OCT	176	213,296	129	92	15,341	24	639[b]
FA2 + OCT[c]	176	209,688	134	82	11,733	29	405
Strategy 8 + OCT	176	305,187	176	307	107,232	71	1510

FA, fully automated.
a Reference strategy.
b Strategy more costly and less effective than an alternative strategy (dominated).
c Figures in table based on assumption that fully automated grading can be implemented at zero net increase in grading costs.

TABLE 30 Screening and referral cost per true case of MO detected under alternative referral criteria (for 3170 patients); adjusted for expected frequency of different patient categories and based on Scottish screening and referral costs

Strategy	Adjusted number of MO cases	Total cost (£)	MO cases detected and referred (*n*)	False-positive referrals (*n*)	Incremental cost (£)	Incremental cases	Incremental cost per case detected (£)
Strategy 2	176	173,246	105	628	a	a	a
Strategy 1	176	207,825	128	995	34,579	23	1503[b]
Strategy 16	176	196,882	129	872	23,636	24	985[b]
FA2[c]	176	189,231	134	786	15,985	29	551
Strategy 2 + OCT	176	145,848	105	67	a	a	a
Strategy 1 + OCT	176	163,509	128	105	17,661	23	768[b]
Strategy 16 + OCT	176	158,530	129	92	12,682	24	528[b]
FA2 + OCT[c]	176	155,148	134	82	9300	29	321
Strategy 8 + OCT	176	242,392	176	307	96,544	71	1360

FA, fully automated.
a Reference strategy.
b Strategy more costly and less effective than an alternative strategy (dominated).
c Figures in table based on assumption that fully automated grading can be implemented at zero net increase in grading costs.

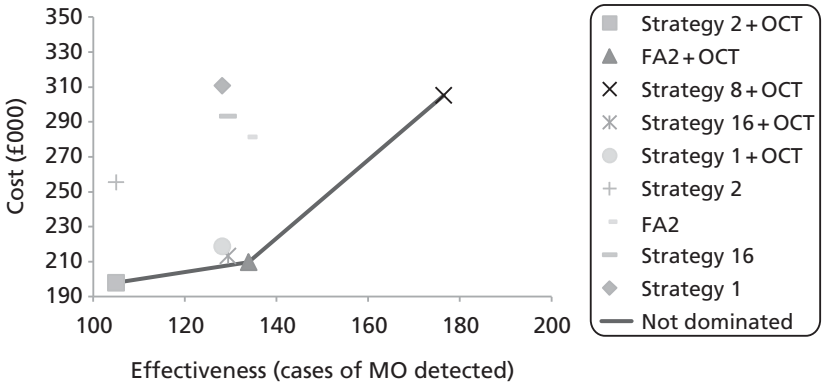


FIGURE 18 Cost-effectiveness frontier for all strategies (applying English costs).

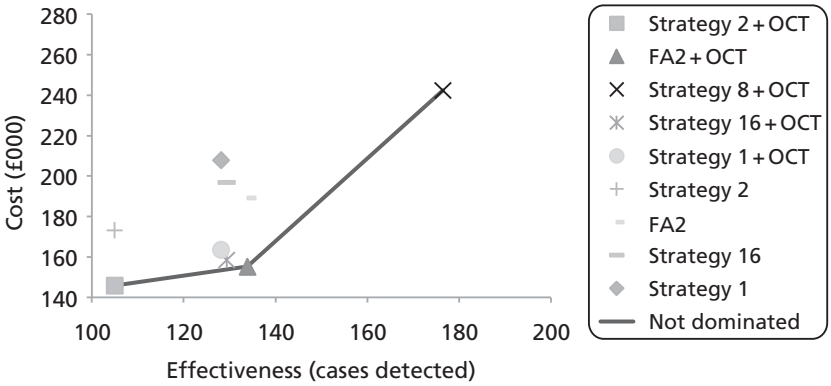


FIGURE 19 Cost-effectiveness frontier for all strategies (applying Scottish costs).

grading strategy 2 becomes substantially less sensitive, more specific and less costly than manual grading strategy 1 (see *Tables 29* and *30*). The results in *Tables 29* and *30* also suggest that manual grading strategy 16 would dominate strategy 1. The automated grading strategy (coupled with the linear classifier and fall back to manual strategy 16 in the event of automated technical failure) would also dominate manual grading strategy 1 and manual grading strategy 16 assuming it could be implemented without increasing net grading costs to a programme. A common finding from the base-case analyses is that the addition of OCT to each grading strategy (within the screening programme) prior to referral, results in a reduction in costs to the health service, with no decrement in the number of MO cases detected.

The incremental cost per extra case of MO detected varies by grading strategy. Considering the more sensitive manual grading strategy 16 compared with manual grading strategy 2 (see *Tables 29* and *30*), the additional cost per extra case detected ranges from £1579 (English costs) to £985 (Scottish costs). When the manual grading strategies are coupled with OCT, these incremental costs fall to £636 and £528, respectively. Although the automated annotation grading strategy dominates manual grading strategies 1 and 16 under the base-case assumption of zero net increase in grading costs, it costs £897 per additional case detected and referred in comparison with manual grading strategy 2 (applying English screening and referral costs). This incremental cost drops to £405 when the grading strategies are coupled with OCT prior to referral. The 100% sensitive strategy 8 (OCT for all patients with any exudate or BH less than one DD, or any M/DH less than one DD and visual acuity $\geq$ 0.3 log-MAR) costs an additional £1510 per extra case detected in comparison with the most specific grading pathway (manual grading strategy 2 + OCT) when applying English referral costs, and £1360 per extra case detected when applying Scottish referral costs. In comparison with manual grading strategy 16 plus OCT, strategy 8 costs £1955 and £1784 per extra case detected and referred when applying English and Scottish referral costs, respectively.

© Queen's Printer and Controller of HMSO 2013. This work was produced by Olson *et al*. under the terms of a commissioning contract issued by the Secretary of State for Health. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

Deterministic sensitivity analysis

Considering manual grading strategy 1 (current English practice), threshold analysis suggests that the cost of implementing OCT within the screening pathway (to reduce false-positive referrals) can rise to ~ £113 per patient before it starts increasing the overall cost per true-positive referral above the level obtained in its absence (assuming English referral costs). The corresponding threshold cost for OCT coupled with manual grading strategy 2 is ~ £110. Applying the lower Scottish referral costs, the OCT threshold costs fall to £71 and £69 for manual grading strategies 1 and 2, respectively. To put this in perspective, if an associate specialist were to perform OCT examinations within screening, and applying the same input assumptions as those used for the cost estimate based on a band 6 retinal screener (see *Appendix 2*), the corresponding cost would be ~ £64. As expected, lower referral costs also improve the cost-effectiveness of the more sensitive, less specific manual grading strategies.

Further sensitivity analysis was undertaken to determine the threshold for the net incremental grading cost associated with automated annotation grading that would result in it no longer dominating manual grading strategy 1 in terms of the cost per case detected and referred. This showed that the automated classifier would remain dominant over manual grading strategy 1 up to an additional cost of ~ £9.00 per patient graded (assuming English referral costs). The threshold cost drops to ~ £3.00 when considering the manual grading strategies coupled with OCT. Note that an alternative operating point could also be established on the automated classifier to obtain improved specificity for the same sensitivity as any of other simple manual grading strategies.

The results of further sensitivity analyses are presented in *Table 31*. In general, the ordering of grading strategies was not found to be particularly sensitive to variation in key model parameters.

Probabilistic sensitivity analysis

The probabilistic analysis characterises the joint uncertainty surrounding the estimated incremental costs and cases of MO detected with each of the alternative strategies (*Figures 20–22*). The acceptability curves indicate the probability of each grading strategy being the preferred option given different values of decision makers' maximum willingness to pay per extra case of MO detected and referred. The preferred option changes as this threshold increases.

Applying English referral costs (see *Figure 20*), manual grading strategy 2 (Scottish criteria) coupled with OCT has the highest probability of being considered cost-effective up to a 'willingness-to-pay' threshold of ~ £450 per extra case of MO detected. Above this ratio, automated annotation grading (coupled with OCT) has the higher probability of being the preferred option up to a 'willingness-to-pay' threshold of ~ £2450. Note, however, that this is dependent on the assumption that automated annotation grading can be implemented without increasing grading costs to the screening programme. Above a 'willingness-to-pay' threshold of ~ £2450, manual grading strategy 8 coupled with OCT has the highest probability of being preferred on grounds of cost-effectiveness. *Figure 21* indicates that a similar pattern of results is obtained when applying the Scottish referral costs, although the thresholds for adopting the more sensitive (less specific) manual grading strategies are slightly lower because of the lower referral costs. *Figure 22* indicates the choices between manual grading strategies when the automated grading procedures are taken out of the comparison (applying English costs). Under this scenario manual grading strategy 16 plus OCT has the higher probability of being the preferred option above a 'willingness-to-pay' threshold of ~ £800 per case detected and referred, before manual grading strategy 8 plus OCT takes over above the 'willingness-to-pay' threshold of ~ £2000 per additional case.

TABLE 31 Deterministic sensitivity analysis findings

Scenario	Strategy	Cost (£)	True-positive referrals	Incremental cost-effectiveness ratio (£)
Frequency weight for category 4 patients 0.15 rather than 0.227	Strategy 2	271,899	120	a
	Strategy 1	324,244	138	2908[b]
	Strategy 16	307,053	142	1598[b]
	FA2[a]	286,188	143	621
	Strategy 2 + OCT	205,220	120	a
	Strategy 1 + OCT	224,538	138	1073[b]
	Strategy 16 + OCT	219,450	142	647[b]
	FA2 + OCT[c]	212,407	143	312
	Strategy 8 + OCT	303,818	186	1494
Lower eye clinic referral costs (based on NHS reference costs[64] rather than PbR tariff)	Strategy 2	238,907	105	a
	Strategy 1	285,492	128	2025[b]
	Strategy 16	270,789	129	1328[b]
	FA2[c]	260,693	134	751
	Strategy 2 + OCT	194,108	105	a
	Strategy 1 + OCT	213,665	128	850[b]
	Strategy 16 + OCT	208,327	129	592[b]
	FA2 + OCT[c]	204,814	134	369
	Strategy 8 + OCT	294,278	176	1411
OCT cost per patient (within screening programme) = £50	Strategy 2	255,448	105	a
	Strategy 1	310,791	128	2406[b]
	Strategy 16	293,345	129	1579[b]
	FA2[a]	281,465	134	897
	Strategy 2 + OCT	221,300	105	a
	Strategy 1 + OCT	254,802	128	1457
	Strategy 16 + OCT	245,247	129	998[b]
	FA2 + OCT[c]	239,159	134	616[b]
	Strategy 8 + OCT	405,109	176	2589

FA, fully automated.
a Reference strategy.
b Strategy more costly and less effective than an alternative strategy (dominated).
c Figures in table based on assumption that fully automated grading can be implemented at zero net increase in grading costs.

© Queen's Printer and Controller of HMSO 2013. This work was produced by Olson *et al*. under the terms of a commissioning contract issued by the Secretary of State for Health. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

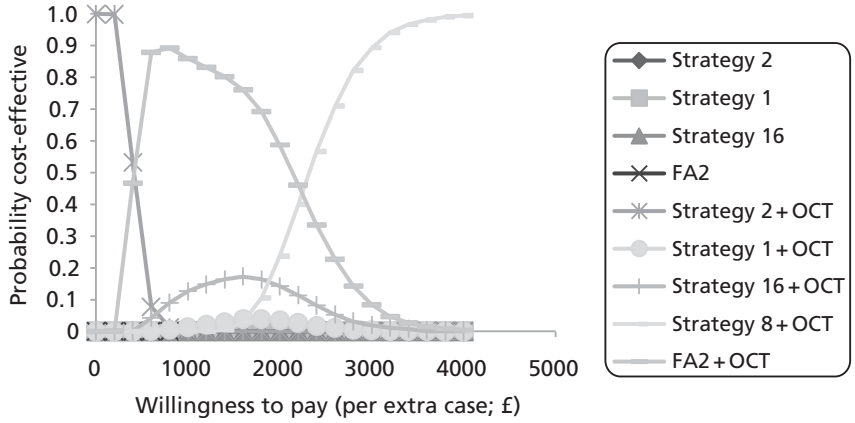


FIGURE 20 Cost-effectiveness acceptability curves for the alternative grading strategies based on English costs.

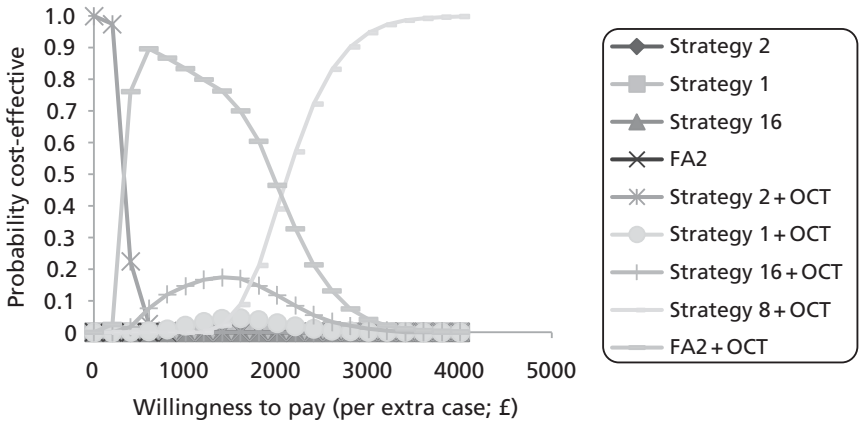


FIGURE 21 Cost-effectiveness acceptability curves for the alternative grading strategies based on Scottish costs.

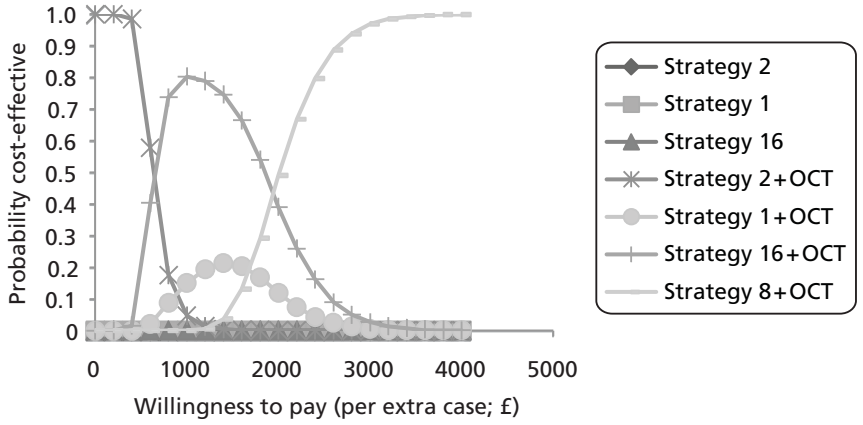


FIGURE 22 Cost-effectiveness acceptability curves excluding the automated strategies and based on English referral costs.

Assessing long-term cost-effectiveness

Model overview

In order to estimate the longer-term cost-effectiveness of the alternative referral criteria, a Markov microsimulation model was developed to simulate the progression of MO and visual loss in referred and un-referred patients over time (*Figure 23*). Patients with MO referred to the eye clinic were modelled to receive treatment and incur a lower risk of disease progression and visual impairment. Costs associated with screening, referral, treatment, ongoing monitoring and vision loss were incorporated in the

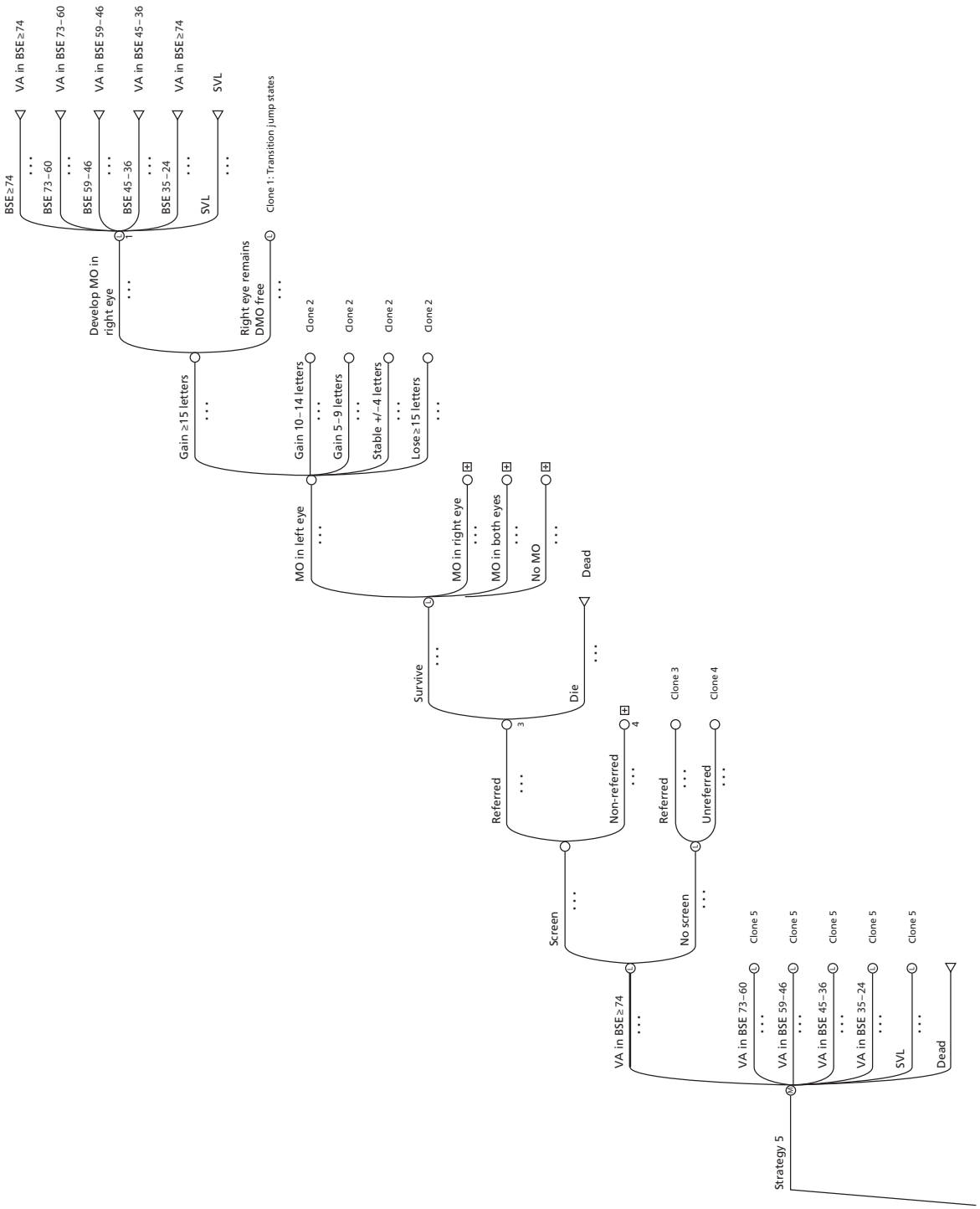


FIGURE 23 Diagrammatic representation of the model structure. BSE, better seeing eye; SVL, severe visual loss; VA, visual acuity.

© Queen's Printer and Controller of HMSO 2013. This work was produced by Olson *et al*. under the terms of a commissioning contract issued by the Secretary of State for Health. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

model – as were utility weights associated with varying degrees of visual loss – allowing costs and QALYs to be estimated over a number of years.

A microsimulation approach was used whereby individual patients with characteristics matching those of the patients in the clinical data set were randomly selected into the model one at time and simulated to progress. The modelled cohort consisted of patients with any surrogate photographic markers within one DD of the fovea, and also those with exudates between one and two DD. The relative proportions of patients with different types of macular features were adjusted to reflect their expected frequency within screening programmes (using the frequency weights in *Table 25*).

Based on the initial visual acuity in each eye, patients were assigned to one of six health states defined by their visual acuity in the better-seeing eye; visual acuity in the better-seeing eye is generally considered to be the better determinant of health-related quality of life,[60] and the health-related quality of life in patients with diabetic retinopathy is more commonly reported according to visual acuity in the better-seeing eye. These health states are represented by the nodes emanating from the Markov (strategy 5) node in *Figure 23*. The model was constructed to cycle on a 6-monthly basis (the average time interval between screening/monitoring appointments for patients with observable but non-referable maculopathy) with the tree to the right of the Markov health states representing the clinical pathways that simulated patients could follow within each cycle of the model. Tracker variables were used to update patient history and screening/referral decisions so that the pathway taken by patients in subsequent cycles of the model could be determined based on this prior information.

Patients were simulated to enter the model at an index screening visit. The adjusted sensitivity and specificity of the different strategies (see *Table 22*) were applied to patients within the alternative arms of the model. Depending on the sensitivity/specificity of the screening strategy, patients could then either be referred to the eye clinic or remain in the screening programme. The visual acuity of eyes with MO remaining in the screening programme was modelled to improve/deteriorate at the rates observed for untreated eyes in the ETDRS,[61] while eyes with MO referred to the eye clinic were assumed to receive timely treatment and improve/deteriorate at the rates observes for laser-treated eyes in the ETDRS and a more recent UK-based observational study.[61,62] The change in visual acuity for each eye with MO was modelled using 6-monthly probabilities of improvement/deterioration derived from the aforementioned studies. Eyes without MO were assumed to remain stable until the development of MO. At the end of each model cycle each patient was assigned the appropriate health state for the subsequent model cycle, based on the updated visual acuity in their better-seeing eye (furthest right nodes in *Figure 23*).

Health state utility weights obtained from the available literature[63] were used to quality adjust the time spent by patients in each visual acuity state, so as to generate QALYs. Health state utilities reflect the relative desirability of different states of health on a scale where zero represents death and one represents full health. So, for example, a year spent in the health state 'visual acuity $\geq$ 74 letters', which was assigned a utility weight of 0.83, would generate 0.83 QALYs (see *Utilities* for further details).

Costs associated with screening, referral, treatment, clinical observation and legal blindness (visual acuity $\leq$ 35 ETDRS letters) were incorporated into the model. For strategies where OCT monitoring was not available within the screening programme, referred patients with no MO were modelled to remain under 6-monthly observation in the eye clinic. Patients not referred to the eye clinic at any given screening visit were assumed to undergo screening again one year later – when they would again experience a probability of being referred to the eye clinic. In the base-case analysis it was assumed that patients with and without MO would face a constant probability of being referred each year, based on the adjusted sensitivity and specificity of the strategy in question (see *Tables 29* and *30*). As an alternative we conducted a sensitivity analysis where it was assumed that patients not referred on the index screening visit would remain non-referable (according to their surrogate markers) for 2 years and 5 years. During these time periods the visual acuity in eyes with MO was modelled to deteriorate at the rates observed for untreated eyes.[64]

At any stage in the simulation patients could die from age-specific all-cause mortality, which was modelled based on UK life tables combined with hazard ratios for all-cause mortality associated with diabetes.[65,66] Further details of the model structure, parameters and assumptions are provided opposite.

Modelled cohort

The modelled cohort consisted of patients with any surrogate photographic markers within one DD of the fovea, and also those with exudates between one and two DD. In sampling patients for entry into the model the frequency proportions from *Table 25* were applied so that simulated patients in the different surrogate marker categories would appear in the model with the same relative frequency as anticipated within screening programmes. The clinical characteristics [age, gender, type of diabetes, visual acuity (in each eye), macular surrogate marker category, retinopathy grade, and MO status (no MO/MO in right eye/MO in left eye/MO in both eyes)] of simulated patients were recorded using tracker variables, some of which were modelled to update over time (age, visual acuity, MO status).

The grading strategies

The grading strategies chosen for comparison included those examined in the cost-effectiveness analysis. This included the highly specific grading strategy 2, and the more sensitive and less specific grading strategy 1, strategy 16, and the automated annotation grading strategy FA2 (set at ~ 75.9% sensitivity and 73.7% specificity). These grading strategies were assessed both with and without the presence of OCT screening prior to referral to the eye clinic. The maximally sensitive grading strategy 8 coupled with OCT was also assessed.

Progression to macular oedema

Eyes with no MO at baseline were modelled to progress to MO using cumulative incidence data reported by Younis *et al*.[67,68] for a screening cohort in Liverpool. Younis *et al*. reported the yearly cumulative incidence of referable maculopathy, up to 6 years, in patients with no retinopathy, background retinopathy and mild pre-proliferative retinopathy in their worst eye at baseline. The reported yearly cumulative incidence rates were used to estimate the annual risks of progression to referable maculopathy conditional on survival. These annual risks were then averaged and converted to average 6-monthly risks. Finally, these probabilities were converted into probabilities of developing OCT-positive MO, based on the positive predictive value of the English referral criteria for MO (manual strategy 1) from *Table 29*. For eyes with moderate/severe retinopathy at baseline, the 3-year incidence of clinically significant MO, as reported for the placebo arm in the Protein Kinase C-DRS2 trial,[69] was used to estimate 6-monthly probabilities of progression. As an alternative approach to estimating rates of progression to MO in patients with type 2 diabetes, the proportion of patients (with varying degrees of retinopathy at baseline) receiving photocoagulation for MO at 6 years was estimated from the UK Prospective Diabetes Study[70] and used to generate 6-monthly probabilities of progression for the eyes of modelled patients (*Table 32*).

TABLE 32 Six-monthly risks of developing MO in eyes of patients without MO at baseline

Retinopathy grade in worst eye at baseline	6-monthly risk of MO (T2DM)	6-monthly risk of MO (T1DM)	6-monthly risk of treatment for MO (T2DM)
None	0.0004	0.0004	0.0007
Background	0.0032	0.0032	0.0021
Mild pre-proliferative	0.0201	0.0275	0.0183
Pre-proliferative	0.121	0.121	0.0183
Proliferative	0.121	0.121	0.0183

T1DM, type 1 diabetes mellitus; T2DM, type 2 diabetes mellitus.

© Queen's Printer and Controller of HMSO 2013. This work was produced by Olson *et al*. under the terms of a commissioning contract issued by the Secretary of State for Health. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

Based on the observed proportion of patients recruited to the clinical study with MO in both eyes (12% of those with MO in any eye), it was assumed that 12% of incident MO cases would present with both eyes affected. It was also assumed that for patients with MO in one eye, the risk of developing MO in the other eye would be increased. In a case–control study, Bhavsar and Subramanian[71] found that a prior history of 'clinically significant' MO in either eye significantly increased the risk of progression from subclinical MO to 'clinically significant' MO [unadjusted OR = 8.07 (CI 2.75 to 23.63)]. In a prospective cohort study, Varma *et al*.[72] found the incidence of 'clinically significant' MO in the second eye of patients with 'clinically significant' MO in one eye to be double the incidence of 'clinically significant' MO in a first eye. It was therefore assumed that the risk of eyes developing MO by underlying retinopathy grade (see *Table 32*) would increase similarly for simulated patients with MO already present in one eye.

A simplifying assumption of the model was that no eyes lose vision prior to the development of MO. This assumption should have limited impact on the comparison of alternative referral strategies, as any patient at risk of vision loss from proliferative retinopathy would in practice be referred to ophthalmology services regardless of whether or not MO was suspected. The rates of vision loss applied to eyes with MO in the model may also capture vision loss resulting from co-existing retinopathy.

Progression of visual impairment in eyes with macular oedema

The visual acuity of referred eyes with MO was modelled to deteriorate/improve at the rates observed for laser-treated eyes with 'clinically significant' MO, as defined by the ETDRS, whereas non-referred eyes with MO were modelled to deteriorate at the rates observed for untreated eyes with 'clinically significant' MO.[64]

Non-referred eyes

Owing to the long-established clinical practice of laser treatment for the prevention of visual loss in eyes with 'clinically significant' MO,[64] there is a lack of contemporaneous data on visual acuity outcomes for untreated eyes. Thus, non-referred eyes with MO were modelled to follow the visual acuity course observed for eyes with 'clinically significant' MO randomised to deferral of treatment in the ETDRS.[64] Although improved glucose control may reduce the rate at which patients now develop 'clinically significant' MO,[73] the definition of MO used in the current analysis maps closely to the ETDRS definition of 'clinically significant' MO, and there is no evidence available to suggest that untreated eyes with 'clinically significant' MO currently lose vision any slower than they did when ETDRS was undertaken. The main visual acuity outcome reported by the ETDRS was MVL, defined as the loss of 15 letters (three lines) or more on the ETDRS visual acuity chart.[64] Approximately 29% of eyes with 'clinically significant' MO assigned to deferral of laser photocoagulation had reached this end point by 36 months, and a subsequent report showed that ~ 40% of patients in this group had MVL at 5 years.[61] In addition, it was assumed that no untreated eyes with MO would improve by more than 10 letters (two lines) without treatment, but that a constant proportion of untreated eyes (~ 20%) would be improved by between five and nine letters (one to two lines) over follow-up.[64] These probabilities were decomposed to generate 6-month transition probabilities using the method reported by Craig and Sendi,[74] such that non-referred eyes with MO progressed as per deferred eyes with 'clinically significant' MO in the ETDRS (*Figure 24*). A simplifying assumption was that patients with deteriorating vision would not present to the eye clinic without being referred through the screening programme. This assumption works in favour of more sensitive strategies. A further assumption was that, for patients losing 15 letters or more, the average degree of visual loss would be 20 letters – this yielded a proportion of modelled eyes transiting to severe vision loss (< 24 letters) that was consistent with observations in the ETDRS. For patients with both eyes affected by MO, perfect positive correlation in visual acuity outcome was assumed between eyes.

Referred eyes

The ETDRS originally demonstrated that laser photocoagulation reduced the risk of MVL from MO by 50% at 3 years.[64] The effect was concentrated in patients with 'clinically significant' MO, in whom it was also somewhat greater (~ 12% of eyes with 'clinically significant' MO receiving immediate laser treatment had MVL at 36 months, as compared with 29% in the deferred group). At 5 years, ~ 22% of treated eyes with centre involved 'clinically significant' MO had reached this outcome. It also showed that mild improvement

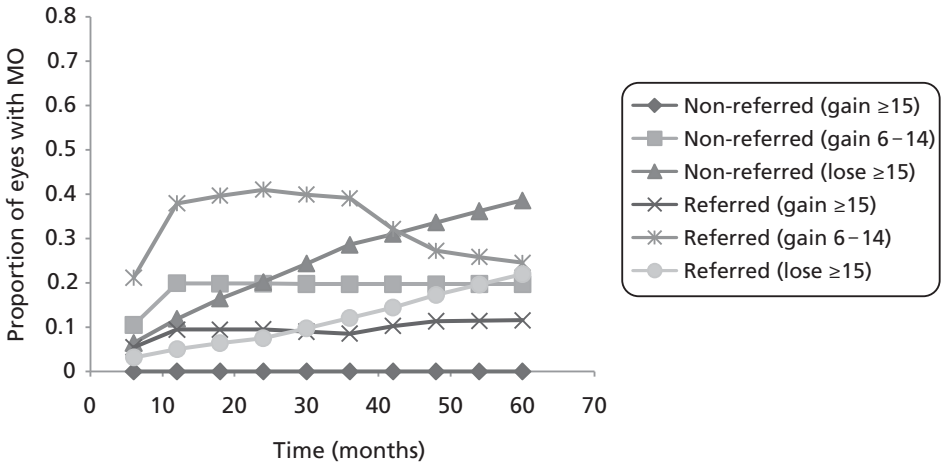


FIGURE 24 Modelled visual acuity outcomes for eyes with 'clinically significant' MO with and without referral/laser treatment.

(six letters or more) was more common in laser-treated eyes (~ 40% at 3 years), but that < 3% of treated eyes showed improvement ≥ 15 letters. However, more recent randomised controlled trials assessing the effectiveness of triamcinolone[75] and ranibizumab[76] compared with laser, for the treatment of centre involved MO, suggest better outcomes can now be achieved. Elman *et al*.[77] showed that 51% of laser-treated eyes were improved by five or more letters at 12 months, and 15% were improved by 15 letters or more. Elman *et al*.[76] also showed the proportion of laser-treated eyes demonstrating moderate improvement and deterioration (10 letters or more), compared with baseline, to be relatively stable between 12 and 24 months. The Diabetic Retinopathy Clinical Research Network trial of triamcinolone demonstrated even more favourable outcomes for laser-treated eyes at 36 months, with 62% of eyes improved by five letters or more and 26% of eyes improved by 15 letters or more. However, a recently published study of visual acuity outcomes for eyes treated with laser in routine clinic practice in the UK reported less favourable results.[68] This study of a 100 consecutive patients undergoing laser treatment for MO in a routine hospital setting found that only 28% of eyes had improved by five or more letters at 3 years and that 9% had improved by 15 letters or more (12% by 5 years).

In this study, referred eyes were modelled to deteriorate/improve at the rates observed for treated eyes in the ETDRS. However, since there is some evidence that improvements of 15 letters or more are more common with contemporary treatment protocols, the probability of eyes achieving this gain was modelled based on the study by Jyothi and Sivaprasad.[68] For modelled eyes that had lost 15 letters or more prior to referral, it was assumed that no improvement in visual acuity would be possible on initiation of treatment. This assumption and the treatment assumptions, in general, favour the more sensitive strategies.

The resultant modelled proportions of eyes with macular oedema experiencing the alternative visual acuity outcomes (over time, with and without referral) are plotted in *Figure 24*.

Mortality

The model allowed the age- and gender-specific 6-monthly mortality risk to be referenced for each simulated patient during each cycle of the model. These all-cause mortality risks for the UK general population were adjusted upwards using published age- and diabetes-specific hazard ratios for all-cause mortality.[65,66] Thus, mortality was modelled based on age, gender and type of diabetes. Although this approach may overestimate mortality somewhat for the general diabetic population, it has also been noted that the mortality risk for patients with MO is increased relative to the mortality risk for diabetic patients without MO.[78]

Costs and resource use

Costs associated with screening, OCT and initial referral were incorporated as shown in *Table 24*. In addition, costs of treatment and ongoing outpatient monitoring (for referred patients with and

© Queen's Printer and Controller of HMSO 2013. This work was produced by Olson *et al*. under the terms of a commissioning contract issued by the Secretary of State for Health. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

without MO) were incorporated in the model. All future costs were discounted at a rate of 3.5% per year, in line with the current recommendations from Her Majesty's Treasury. In relation to the capture of ongoing screening, monitoring and treatment costs, the following assumptions were applied.

Screening costs

For simulated patients not referred to the eye clinic under alternative grading strategies at any given screening appointment, the average cost of a further screening appointment was applied one year later. The impact of applying a 6-month screening interval, to reflect the fact that many patients with features in the macula are often monitored within screening programmes on a 6-monthly basis, was also assessed through sensitivity analysis. For strategies involving the use of OCT within screening programmes (prior to referral), the cost of providing photographic screening and an OCT examination was applied on a 6-monthly basis for patients sent for OCT but subsequently found not to have MO.

Follow-up up observation for eye clinic referrals without macular oedema

For strategies not utilising OCT within the screening programme, patients referred to the eye clinic without MO (not requiring treatment) were assumed to require follow-up monitoring within the hospital eye service every 6 months. This was assumed to incur the cost for an ophthalmology outpatient attendance (with ~ 50% of patients receiving an OCT examination at each visit) (see *Table 24*). For strategies utilising OCT within the screening programme, it was assumed that photo-positive patients without MO would be monitored on a 6-monthly basis with digital photography and OCT within the screening programme until the point in time when they developed MO.

Initial treatment cost for those with macular oedema

Since a survey of participating centres, and review of current National Institute for Health and Care Excellence and Scottish Medicines Consortium guidance, suggested that laser still formed the mainstay of treatment for MO in the UK, it was assumed that all patients with MO referred to ophthalmology would receive this treatment modality. This was costed using the appropriate NHS reference cost,[58] with the OPCS code 'laser photocoagulation to lesion of the retina' (C826) mapping to the HRG category BZ23Z (Retinal Vitreous Procedures-Category 1). Treatment was assumed to occur in both eyes if both eyes had MO as determined by the OCT outcome.

Ongoing treatment costs for those with macular oedema

For costing ongoing laser treatment after initial treatment, the numbers of treatments observed for patients with MO in a UK-based cohort,[68] at 3 years (1.54) and between 3 and 5 years (0.2), were used to estimate corresponding average 6-monthly costs of treatment. Patients undergoing treatment were also assumed to require 3-monthly monitoring as outpatients [£117 per visit based on the weighted average of the cost for a follow-up ophthalmology attendance (£67) and the cost of OCT in an outpatient setting (£160)]. Laser treatment was assumed to stop completely 5 years after initial treatment and also if patients entered the severe visual loss state (< 24 ETDRS letters in the better-seeing eye). We also carried out sensitivity analysis assuming that the same treatment effects could be obtained with no further laser treatment beyond 3 years.

Cost of legal blindness

In addition to the above costs, 6-monthly health and social care costs associated with legal blindness were estimated and applied to simulated patients in the model when visual acuity in the better-seeing eye fell below 35 letters (*Table 33*). These costs were taken from a study by Meads and Hyde[79] and inflated to the base costing year (2009/10).

Utilities

A review of the literature for health state utilities associated with MO revealed only one study where suitable utility weights had been elicited from a sample of the UK general population. This study reported directly elicited standard gamble utilities for a range of diabetic retinopathy health states defined in terms of best corrected visual acuity,[63] which would correlate well with visual acuity in a patients better-seeing

TABLE 33 Costs associated with legal blindness (visual acuity <35 letters in the better-seeing eye)

State costs	Unit cost (£)	Low	High	Source
Better-seeing eye < 35 letters				
Annual cost of legal blindness (first year)	6455	1375	17,100	Meads and Hyde[79]
Annual costs of legal blindness (ongoing)	6295	1325	16,800	Meads and Hyde[79]

eye. This same study also reported alternative European Quality of Life-5 Dimensions (EQ-5D) weights derived from the responses of diabetic patients with varying degrees of visual acuity. However, the sample size for the EQ-5D weights was small, especially for the lower acuity states. For this reason the direct standard gamble weights were used to adjust time spent by simulated patients in the different visual acuity states of the model (*Table 34*). We also assessed the impact of assuming that these utility weights would correlate with the visual acuity in the patient's worse-seeing eye (an assumption favouring more sensitive strategies).

Analysis

The model was analysed using Monte Carlo simulation, whereby a random number generator was used to simulate the progression of individual patients through the model one a time. Initially, the model was analysed over a 5-year period. Thereafter, further longer-term analyses were undertaken, assuming the same ongoing risks of vision loss as estimated for treated and untreated eyes in the fifth year of the model. The mean costs, years free of MVL (in either eye) and QALYs accruing to patients under the alternative referral strategies were compared to estimate incremental cost-effectiveness ratios. In order to characterise uncertainty surrounding these ratios, deterministic and probabilistic sensitivity analyses were undertaken.[80] The base-case analysis was carried out using English-specific cost data.

Probabilistic sensitivity analysis was undertaken by assigning distributions to model parameters where the data sources and approach to estimation provided sufficient information for this to be done. Beta distributions for the adjusted sensitivity and specificity estimates (of alternative grading strategies) were derived from the probabilistic simulations carried out for the analysis of the cost per case detected. Second order distributions for mean screening, OCT, referral and treatment costs were assigned by selecting variance parameters that ensured feasible low and high estimates (for each cost parameter) fell within the 2.5th and 97.5th percentile of the resultant distribution. Parameters excluded from the probabilistic sensitivity analysis included the underlying probabilities of visual loss and visual gain, and the probabilities of developing MO. The methods used to derive these probabilities precluded the estimation of uncertainty due to sampling variation, and so the impact of variation in these parameters was addressed through deterministic sensitivity analysis.

TABLE 34 Health state utilities used in the cost-effectiveness model

Visual acuity state	Standard gamble utility weight (SD)	EQ-5D utility weight (SD)
BSE ≥ 74 letters	0.83 (0.16)	0.75 (0.23)
BSE 60–73 letters	0.75 (0.20)	0.50 (0.30)
BSE 46–59 letters	0.68 (0.23)	0.68 (0.29)
BSE 36–45 letters	0.65 (0.23)[a]	0.605 (0.39)[a]
BSE 24–35 letters	0.63 (0.23)	0.53 (0.47)
BSE < 24 letters	0.58 (0.26)	0.34 (0.36)

BSE, best seeing eye.
a Inferred by interpolation between utilities for visual acuity states above and below.

© Queen's Printer and Controller of HMSO 2013. This work was produced by Olson *et al*. under the terms of a commissioning contract issued by the Secretary of State for Health. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

The results of the probabilistic sensitivity analysis are presented in the form of cost-effectiveness acceptability curves, a standard method for characterising the uncertainty surrounding estimates of cost-effectiveness. Deterministic sensitivity analysis was also used to assess the sensitivity of findings to various structural assumptions, the implications of using alternative sources of cost data and the incorporation of higher health and social care costs associated with legal blindness (see *Table 33*).

Results

Table 35 shows the expected costs, years free from MVL and QALYs for patients screened using the alternative strategies over a 5-year period. Although the strategies are noticeably different in terms of expected costs to the health service, they are harder to separate in terms of their impact on MVL and QALYs. Under the base-case analysis the differences are very small, though they slightly favour the more sensitive strategies. The introduction of OCT for all screen-positive individuals prior to referral reduces the expected costs of all strategies under the base-case costing assumptions.

Tables 36 and *37* show the impact of extending the time horizon of the analysis to 10 and 20 years, respectively. The result is a wider gap between the strategies in terms of the expected QALYs and years free from MVL, and improved cost-effectiveness of the more sensitive strategies compared with the less sensitive strategies. For example, the additional cost per QALY gained from using strategy 16 as compared with manual grading strategy 2 (current Scottish referral criteria without OCT) decreases from £882,307 (at 5 years) to £353,927 at 20 years. The corresponding additional cost per year free from MVL is £38,457 at 5 years and £22,583 at 20 years.

The longer-term cost-effectiveness results mirror the pattern observed in the short-term analysis (see *Table 29*), in that the addition of OCT results in cost-savings without reducing the health benefits. As per the short-term results, the automated strategy FA2 (with fall-back to manual strategy 16) also dominates strategies 1 and 16 under the assumption of zero net increase in grading costs associated with the introduction of automated grading. Furthermore, strategy 16 remains essentially dominant over, or very cost-effective in comparison with, strategy 1. *Figure 25* shows the cost-effectiveness frontier when considering all strategies together.

Deterministic sensitivity analysis

Table 38 presents the findings from a number of deterministic sensitivity analyses, assessing the sensitivity of findings to a number of structural and parameter assumptions in the model. Only when a number of parameters were simultaneously weighted in favour of the more sensitive grading strategies did the incremental cost per QALY estimates for these strategies begin to approach the accepted threshold range (£20,000–30,000 per QALY) in comparison with the most specific strategy (strategy 2 + OCT) (scenario 9, see *Table 38*). Considering the cost-effectiveness of fully automated annotation grading compared with manual grading strategy 16, the analysis suggests that automated grading could remain dominant up to a marginal net increase in grading cost of ~ £2.00 per patient (assuming strategies are used in combination with OCT prior to referral).

A threshold analysis was also performed to ascertain the threshold at which OCT would cease to be cost-saving when used in combination with photographic screening/grading (on a 6-monthly basis) as part of the screening pathway to monitor patients with suspected MO (rather than monitoring them on a 6-monthly basis as eye clinic outpatients). When applying English screening and referral costs, the marginal cost of providing OCT within screening could reach ~ £58 and remain cost-saving. In the Scottish context, the corresponding threshold cost for OCT drops to ~ £47. The choice between grading/screening strategies was not found to be sensitive to the cost of providing slit-lamp examinations. The results were also found to be generally insensitive to the costs associated with legal blindness and the rate of progression to MO in those with no MO at the outset.

TABLE 35 Expected costs, years free from MVL and QALYs per patient over a 5-year period using alternative screening strategies (based on a simulated cohort of 100,000 patients with macular pathology)

Strategy	Cost (mean) (£)	Incremental cost (£)	Years free from MVL (mean)	Incremental years free from MVL	Incremental cost per year free of MVL (£)	QALYs (mean)	Incremental QALYs	Incremental cost per QALY (£)
Strategy 2	727	a	4.2424	a	a	3.5331	a	a
Strategy 1	869	142	4.2452	0.0028	50,779[b]	3.5333	0.0001	1,286,593[b]
Strategy 16	829	102	4.2450	0.0027	38,457[b]	3.5333	0.0001	882,307[b]
FA2[c]	800	73	4.2454	0.0030	24,265	3.5333	0.0001	583,654
Strategy 2 + OCT	607	a	4.2424	a	a	3.5331	a	a
Strategy 1 + OCT	706	99	4.2452	0.0028	35,402[b]	3.5333	0.0001	896,991[b]
Strategy 16 + OCT	679	72	4.2450	0.0027	27,146[b]	3.5333	0.0001	622,805[b]
FA2 + OCT[c]	659	52	4.2454	0.0030	17,285	3.5333	0.0001	415,754
Strategy 8 + OCT	882	275	4.2494	0.0070	39,165	3.5335	0.0003	882,981

FA, fully automated.
a Reference strategy.
b Strategy more costly and less effective than an alternative strategy (dominated).
c Figures in table based on assumption that fully automated grading can be implemented at zero net increase in grading costs.

© Queen's Printer and Controller of HMSO 2013. This work was produced by Olson *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

TABLE 36 Expected costs, years free from MVL and QALYs per patient over a 10-year period using alternative screening strategies (based on a simulated cohort of 100,000 patients with macular pathology)

Strategy	Cost (mean) (£)	Incremental cost (£)	Years free from MVL (mean)	Incremental years free from MVL	Incremental cost per year free of MVL (£)	QALYs (mean)	Incremental QALYs	Incremental cost per QALY (£)
Strategy 2	1385	a	7.1537	a	a	5.9989	a	a
Strategy 1	1580	195	7.1592	0.0055	35,638[b]	5.9992	0.0003	671,828[b]
Strategy 16	1528	143	7.1588	0.0051	28,024[b]	5.9992	0.0003	558,528[b]
FA2[c]	1486	101	7.1596	0.0059	17,136	5.9992	0.0003	336,170
Strategy 2 + OCT	1158	a	7.1537	a	a	5.9989	a	a
Strategy 1 + OCT	1298	140	7.1592	0.0055	25,586[b]	5.9992	0.0003	482,338[b]
Strategy 16 + OCT	1261	103	7.1588	0.0051	20,185[b]	5.9992	0.0003	402,297[b]
FA2 + OCT[c]	1230	72	7.1596	0.0059	12,216	5.9992	0.0003	239,646
Strategy 8 + OCT	1465	307	7.1655	0.0117	26,179	5.9995	0.0005	594,178

FA, fully automated.
a Reference strategy.
b Strategy more costly and less effective than an alternative strategy (dominated).
c Figures in table based on assumption that fully automated grading can be implemented at zero net increase in grading costs.

TABLE 37 Expected costs, years free from MVL and QALYs per patient over a 20-year period using alternative screening strategies (based on a simulated cohort of 100,000 patients with macular pathology)

Strategy	Cost (mean) (£)	Incremental cost (£)	Years free from MVL (mean)	Incremental years free from MVL	Incremental cost per year free of MVL (£)	QALYs (mean)	Incremental QALYs	Incremental cost per QALY (£)
Strategy 2	2164	a	10.2631	a	a	8.7029	a	a
Strategy 1	2374	210	10.2703	0.0072	29,170[b]	8.7033	0.0004	473,005[b]
Strategy 16	2320	156	10.2700	0.0069	22,583[b]	8.7033	0.0004	353,927[b]
FA2[c]	2277	113	10.2709	0.0078	14,399	8.7034	0.0005	222,210
Strategy 2 + OCT	1814	a	10.2631	a	a	8.7029	a	a
Strategy 1 + OCT	1965	151	10.2703	0.0072	20,975[b]	8.7033	0.0004	340,113[b]
Strategy 16 + OCT	1925	111	10.2700	0.0069	16,069[b]	8.7033	0.0004	251,833[b]
FA2 OCT[c]	1894	80	10.2709	0.0078	10,194	8.7034	0.0005	157,317
Strategy 8 + OCT	2109	295	10.2788	0.0157	18,832	8.7038	0.0009	329,497

a Reference strategy.
b Strategy more costly and less effective than an alternative strategy (dominated).
c Figures in table based on assumption that fully automated grading can be implemented at zero net increase in grading costs.

© Queen's Printer and Controller of HMSO 2013. This work was produced by Olson *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

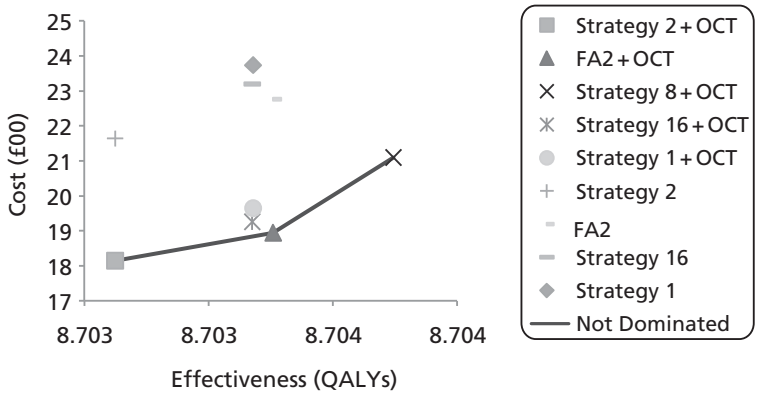


FIGURE 25 Cost-effectiveness frontier for all strategies.

Probabilistic sensitivity analysis

Figures 26 and *27* summarise the results of the probabilistic sensitivity analysis, applying base-case structural assumptions over a 20-year time horizon. The analysis accounts for the joint uncertainty surrounding screening, referral and treatment costs, as well as the uncertainty surrounding the sensitivity/specificity of the alternative screening strategies, the relative frequencies of different pathology categories (see *Table 25*), and the utility weights applied to the visual acuity states. Since the methods for estimating the probabilities of progression to visual loss and MO did not enable the estimation of appropriate portability distributions, the probabilistic sensitivity analysis did not incorporate the uncertainty surrounding these parameters.

The figures demonstrate how the probability of the more sensitive grading strategies being cost-effective increases as societies 'willingness to pay' per QALY or year free of MVL increases. With QALYs as the unit of health outcome (see *Figure 26*), manual grading strategy 2 with OCT (the most specific and least costly strategy) retains the highest probability of being the preferred option up to a 'willingness to pay' threshold of ~ £240,000, at which point the automated grading strategy (coupled with OCT) takes over. However, with years free from MVL as the unit of outcome, this ceiling ratio drops to ~ £12,000 (see *Figure 27*). These threshold values assume that the cost of implementing automated grading is offset by cost-savings resulting from reductions in manual grading time.

Discussion

The cost-effectiveness modelling has demonstrated how costs to the health service increase as more sensitive referral strategies are adopted over more specific screening strategies.

A useful finding is that given the relatively high costs associated with referrals to hospital outpatient eye clinics, the incorporation of OCT within screening pathways, targeted towards those with suspected MO, could reduce costs to the NHS without reducing the number of OCT-positive cases of MO detected and referred.

When looking at the grading strategies based on observing the presence of surrogate photographic markers, the addition of OCT prior to referral reduced the overall costs of detection and referral without impacting on the number of OCT-positive MO cases detected.

The OCT costs were based on the experience of the lead centre, where a band 6 screener carries out the OCT examinations under the supervision of a consultant ophthalmologist. However, deterministic sensitive analysis suggests that, if used in combination with retinal photography (on a 6-monthly basis)

TABLE 38 Deterministic sensitivity analysis conducted over a 20-year time horizon

Scenario	Strategy	Cost (mean) (£)	Years free from MVL (mean)	QALYs (mean)	Incremental cost per year free of MVL (£)	Incremental cost per QALY (£)
1. Application of Scottish screening, referral and treatment costs	Strategy 2	1643	10.2631	8.7029	a	a
	Strategy 1	1805	10.2703	8.7033	22,503[b]	364,890[b]
	Strategy 16	1763	10.2700	8.7033	17,372[b]	272,251[b]
	FA2[c]	1730	10.2709	8.7034	11,086	171,082
	Strategy 2 + OCT	1446	10.2631	8.7029	a	a
	Strategy 1 + OCT	1572	10.2703	8.7033	17,502[b]	283,803[b]
	Strategy 16 + OCT	1539	10.2700	8.7033	13,463[b]	210,995[b]
	FA2 + OCT[c]	1512	10.2709	8.7034	8410	129,787
	Strategy 8 + OCT	1719	10.2788	8.7038	17,427	304,924
2. Cases missed by strategies at index screening visit also missed in subsequent screening visit (3-year delay)	Strategy 2	2155	10.2552	8.7025	a	a
	Strategy 1	2368	10.2632	8.7030	26,369[b]	465,247[b]
	Strategy 16	2313	10.2628	8.7030	20,636[b]	346,475[b]
	FA2[c]	2271	10.2636	8.7030	13,691	219,383
	Strategy 2 + OCT	1806	10.2552	8.7025	a	a
	Strategy 1 + OCT	1959	10.2632	8.7030	18,941[b]	334,192[b]
	Strategy 16 + OCT	1919	10.2628	8.7030	14,759[b]	247,796[b]
	FA2 + OCT[c]	1888	10.2636	8.7030	9678	155,081
	Strategy 8 + OCT	2109	10.2788	8.7038	12,844	234,722
3. Acuity in the worse-seeing eye determines visual acuity state and quality-of-life pay-off	Strategy 2	2269	10.2631	8.6722	a	a
	Strategy 1	2479	10.2703	8.6731	29,170[b]	247,056[b]
	Strategy 16	2425	10.2700	8.6731	22,583[b]	179,967[b]
	FA2[c]	2381	10.2709	8.6731	14,272	122,064
	Strategy 2 + OCT	1920	10.2631	8.6722	a	a
	Strategy 1 + OCT	2070	10.2703	8.6731	20,836[b]	176,469[b]
	Strategy 16 + OCT	2030	10.2700	8.6731	15,924[b]	126,900[b]
	FA2 + OCT[c]	1999	10.2709	8.6731	10,067	86,099
	Strategy 8 + OCT	2212	10.2788	8.6739	18,640	174,949
4. Differential discounting (3.5% for costs, 1.5% per annum for benefits)	Strategy 2	2164	11.8596	10.0772	a	a
	Strategy 1	2374	11.8677	10.0777	25,676[b]	409,370[b]

continued

© Queen's Printer and Controller of HMSO 2013. This work was produced by Olson *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

TABLE 38 Deterministic sensitivity analysis conducted over a 20-year time horizon (*continued*)

Scenario	Strategy	Cost (mean) (£)	Years free from MVL (mean)	QALYs (mean)	Incremental cost per year free of MVL (£)	Incremental cost per QALY (£)
	Strategy 16	2320	11.8674	10.0777	19,837[b]	306,293[b]
	FA2[c]	2277	11.8685	10.0778	12,620	191,675
	Strategy 2 + OCT	1814	11.8596	10.0772	a	a
	Strategy 1 + OCT	1965	11.8677	10.0777	18,462[b]	294,357[b]
	Strategy 16 + OCT	1925	11.8674	10.0777	14,115[b]	217,939[b]
	FA2 + OCT[c]	1894	11.8685	10.0778	8934	135,699
	Strategy 8 + OCT	2109	11.8772	10.0782	16,742	286,099
5. Laser treatment initiated only in those with MO with visual acuity ≥ 0.3 log-MAR (≤ 70 ETDRS letters), benefit remains unaffected	Strategy 2	2123	10.2631	8.7029	a	a
	Strategy 1	2333	10.2703	8.7033	29,170[b]	473,005[b]
	Strategy 16	2278	10.2700	8.7033	22,438[b]	351,658[b]
	FA2[c]	2235	10.2709	8.7034	14,272	220,244
	Strategy 2 + OCT	1774	10.2631	8.7029	a	a
	Strategy 1 + OCT	1924	10.2703	8.7033	20,836[b]	337,861[b]
	Strategy 16 + OCT	1884	10.2700	8.7033	15,924[b]	249,564[b]
	FA2 + OCT[c]	1853	10.2709	8.7034	10,067	155,351
	Strategy 8 + OCT	2067	10.2788	8.7038	18,704	327,263
6. Zero treatment costs beyond 3 years post diagnosis	Strategy 2	2031	10.2631	8.7029	a	a
	Strategy 1	2238	10.2703	8.7033	28,753[b]	466,248[b]
	Strategy 16	2183	10.2700	8.7033	22,004[b]	344,852[b]
	FA2[c]	2140	10.2709	8.7034	13,890	214,345
	Strategy 2 + OCT	1682	10.2631	8.7029	a	a
	Strategy 1 + OCT	1829	10.2703	8.7033	20,419[b]	331,103[b]
	Strategy 16 + OCT	1789	10.2700	8.7033	15,490[b]	242,758[b]
	FA2 + OCT[c]	1758	10.2709	8.7034	9685	149,451
	Strategy 8 + OCT	1969	10.2788	8.7038	18,321	320,562
7. Assume 40-letter loss for all patients experiencing MVL	Strategy 2	2306	10.2631	8.6916	a	a
	Strategy 1	2507	10.2703	8.6923	27,920[b]	299,223[b]
	Strategy 16	2453	10.2700	8.6923	21,280[b]	223,383[b]
	FA2[c]	2409	10.2709	8.6924	13,125	130,176

continued

TABLE 38 Deterministic sensitivity analysis conducted over a 20-year time horizon (*continued*)

Scenario	Strategy	Cost (mean) (£)	Years free from MVL (mean)	QALYs (mean)	Incremental cost per year free of MVL (£)	Incremental cost per QALY (£)
	Strategy 2 + OCT	1956	10.2631	8.6916	a	a
	Strategy 1 + OCT	2098	10.2703	8.6923	19,724[b]	211,391[b]
	Strategy 16 + OCT	2059	10.2700	8.6923	14,911[b]	156,520[b]
	FA2 + OCT[c]	2027	10.2709	8.6924	9047	89,733
	Strategy 8 + OCT	2237	10.2788	8.6930	17,938	197,632
8. All treatment benefits maintained beyond 5 years (i.e. no further visual loss in those treated successfully)	Strategy 2	2163	10.3409	8.7067	a	a
	Strategy 1	2373	10.3526	8.7073	17,953	321,117
	Strategy 16	2319	10.3519	8.7073	14,073	245,183
	FA2[c]	2275	10.3532	8.7074	9056	158,228
	Strategy 2 + OCT	1814	10.3409	8.7067	a	a
	Strategy 1 + OCT	1964	10.3526	8.7073	12,823	229,369
	Strategy 16 + OCT	1925	10.3519	8.7073	10,013	174,457
	FA2 + OCT[c]	1893	10.3532	8.7074	6388	111,607
	Strategy 8 + OCT	2107	10.3680	8.7081	10,776	209,912
9. Simultaneous application of scenarios 2, 3 and 6; combined with application of a 40-letter visual acuity loss for those suffering MVL	Strategy 2	2613	10.2552	8.6511	a	a
	Strategy 1	2802	10.2632	8.6526	23,398[b]	133,098[b]
	Strategy 16	2747	10.2628	8.6525	17,501[b]	97,443[b]
	FA2[c]	2702	10.2636	8.6526	10,504	60,766
	Strategy 2 + OCT	2264	10.2552	8.6511	a	a
	Strategy 1 + OCT	2393	10.2632	8.6526	15,970[b]	90,844[b]
	Strategy 16 + OCT	2353	10.2628	8.6525	11,624[b]	64,719[b]
	FA2 + OCT[c]	2319	10.2636	8.6526	6492	37,552
	Strategy 8 + OCT	2500	10.2788	8.6552	10,004	58,093

a Reference strategy.
b Strategy more costly and less effective than an alternative strategy (dominated).
c Figures in table based on assumption that fully automated grading can be implemented at zero net increase in grading costs.

© Queen's Printer and Controller of HMSO 2013. This work was produced by Olson *et al*. under the terms of a commissioning contract issued by the Secretary of State for Health. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

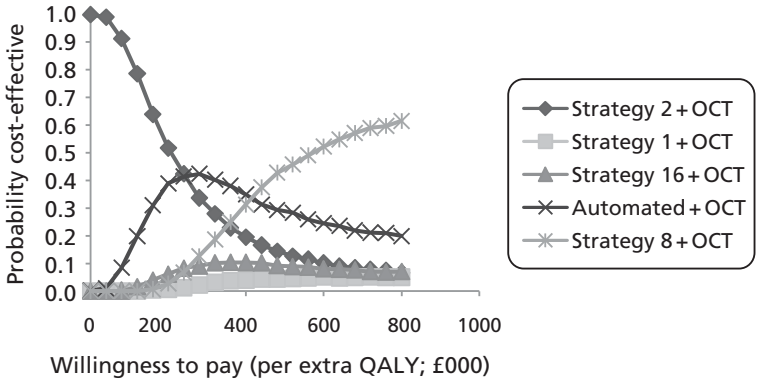


FIGURE 26 Cost-effectiveness acceptability curves for the alternative strategies based on a 20-year time horizon and using QALYs as the measure of effect.

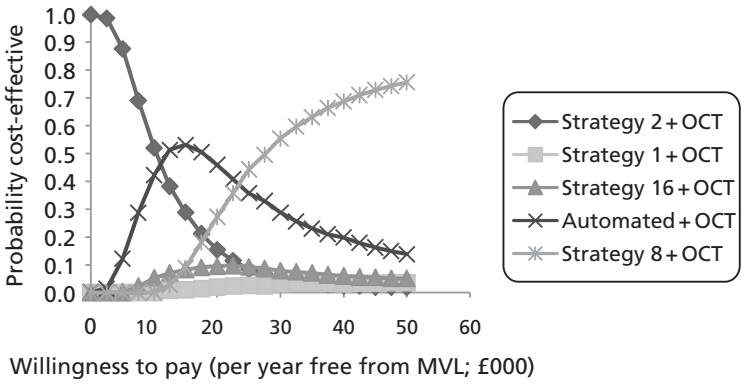


FIGURE 27 Cost-effectiveness acceptability curves for the alternative grading strategies based on a 20-year time horizon and using years free from MVL as the measure of effect.

to monitor patients within screening, the use of OCT could remain cost saving up to an incremental cost of ~ £58 per patient.

In the Scottish context, this threshold cost drops to ~ £47. To put this in perspective, if an associate specialist were to perform OCT examinations within screening programmes, assuming the same time input required by the band 6 retinal screener (see *Appendix 2*), the cost would be ~ £64.

A Markov microsimulation model was developed to assess the longer-term cost-effectiveness of moving from more specific to more sensitive and more costly grading strategies (e.g. manual grading strategy 2 + OCT, to manual grading strategy 1 + OCT, to manual grading strategy 16 + OCT, to automated grading + OCT, to manual grading strategy 8 + OCT).

This analysis showed that while costs to health service increase substantially with the more sensitive/less specific strategies, the associated increase in cases detected does not translate into large gains in health outcomes over time.

The reason for the small differences in years free of MVL and QALYs between the strategies is down to a number of factors.

First, the adjusted prevalence of MO in the modelled cohort was relatively low in the first place, so only a small proportion of the cohort was initially modelled to be at risk of visual loss (5.6%).

Second, although the relative difference in the adjusted sensitivity of the strategies appeared quite substantial, the repetitive nature of screening diminishes the importance of this difference since any

patients missed in one round of screening have a chance of being picked up during the next round prior to any visual loss occurring (i.e. only 12% of non-referred patients with MO were modelled to progress to MVL within 1 year, compared with 5% of referred/treated patients as per the ETDRS).[61,64]

The expected QALYs for patients screened under the alternative grading strategies were particularly insensitive to differences in grading strategy sensitivity. This is because the QALYs were determined by visual acuity in the better-seeing eye of simulated patients.

Since the majority of patients missed by less sensitive grading strategies had MO in only one eye, and most often their worse-seeing eye, any loss of vision in this eye did not impact on the patient's health-related quality of life in the base-case scenario.

Despite the only very small differences in health outcomes between grading strategies, increases in grading strategy sensitivity (with accompanying reductions in specificity) did result in substantial cost increases to the health service over time, as a result of more subjects being sent for OCT or referred to eye clinics for ongoing monitoring or treatment.

To assess the sensitivity of findings to the base-case parameter values and structural assumptions, extensive deterministic sensitivity analysis was undertaken. As a best case scenario for the more sensitive grading strategies, patients' visual acuity state and health-related quality of life status were assumed to be driven entirely by acuity in the worse-seeing eye, and additional costs of treatment were assumed to cease beyond 3 years (for the same treatment benefit).

On top of this, it was assumed that cases missed by less sensitive grading strategies would continue to be missed in subsequent rounds of screening for a period of 3 years (and be at increased risk of MVL in the interim) and that those suffering MVL would lose on average 40 ETDRS letters. Under this relatively unlikely scenario, the incremental cost per QALY estimates for the more sensitive grading strategies remained unfavourable when compared with thresholds used to judge cost-effectiveness in the UK (i.e. £20,000–30,000 per QALY).

These results would suggest that, given the repetitive nature of screening and the relatively slow progression of visual loss from MO, the adoption of more sensitive/less specific screening strategies over more specific/less sensitive grading strategies is unlikely to be considered cost-effective.

Conversely, the results would suggest that if significant improvements in specificity can be achieved without reducing sensitivity – for example, by adopting the more complex automated annotated grading strategy based on automated image grading and pattern detection – such an approach could offer a more cost-effective alternative than current grading strategies based on simple manual algorithms. This, of course, will depend on the cost at which such a system can be implemented within different screening contexts relative to the manual grading cost-savings that might be realised as a result.

To give an example, the results presented here would suggest that the automated annotation grading strategy might continue to dominate grading strategy 16 up to a net increase in grading costs of ~ £2.00 per patient with suspected MO (with OCT also being used prior to referral). However, more detailed consideration would need to be given to how such a pathway would be implemented and how much it would cost in different local contexts.

Strengths and limitations

In order to assess longer-term cost-effectiveness, we created a flexible model whereby the individual eyes of each patient were modelled separately. This enabled us to assess the impact of different assumptions about how loss of acuity in one eye or both eyes affects the health-related quality of life of patients (i.e. either assuming the patient's better-seeing eye determines health-related quality of life or assuming the worse-seeing eye determines health-related quality of life).

© Queen's Printer and Controller of HMSO 2013. This work was produced by Olson *et al*. under the terms of a commissioning contract issued by the Secretary of State for Health. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

We tried to use the best available evidence on the effects of laser treatment compared with no treatment for those patients with MO, in order to assess the benefits of improved detection and referral. In doing so, we updated older effect profiles for laser treatment to accommodate the observation of improved outcomes being associated with its use in more contemporary studies.[68]

A lack of contemporary data on progression of MO in untreated patients required the assumption that patients without treatment would progress at the rate observed in the ETDRS.[64]

A weakness of the modelling related to the sample selection approach and the subsequent need to rely on surrogate marker category frequency weights (obtained from a separate cohort study) to estimate the likely sensitivity/specificity of the alternative options within routine screening. However, the overall findings and conclusions were generally insensitive to the weighting process.

A further potential weakness was the assumption that all patients with MO would be treated with laser, rather than intravitreal anti-vascular endothelial growth factor injections. This decision was taken based on current National Institute for Health and Care Excellence guidance[81] as well as a survey of participating centres which suggested that laser still formed the mainstay of treatment throughout the study period.

Our model may be updated in the future to incorporate antivascular endothelial growth factor treatment. It is possible that this could improve the cost-effectiveness of more sensitive referral strategies by improving outcomes for additional patients detected and referred.

However, by the same token it might also improve outcomes for patients with MO missed by less sensitive strategies but picked up at a subsequent screening visit (or presenting with clinical symptoms).

Comparison with other studies

We are not aware of any other studies which have specifically assessed the cost-effectiveness of alternative approaches to screening for diabetic MO. Many studies have assessed the cost-effectiveness of screening for diabetic retinopathy[82–85] in which detection of MO was one of the modelled benefits.

In comparison with these previous screening models, our model provides more flexibility for assessing the impact on quality of life of MVL in one or both eyes; some previous studies applied more crude annual risks of progression from MO directly to irreversible legal blindness (sometimes applying reported risks per eye with MO at the level of the patient).

Conclusions

The findings of the economic modelling would suggest that given the relatively low prevalence of MO in patients with surrogate photographic markers, coupled with the repetitive nature of screening, the adoption of more sensitive grading strategies is unlikely to be cost-effective unless this can be achieved for no (or only a very small) decrease in specificity.

A further important finding from the economic analysis is that OCT within screening pathways offers potential.

Chapter 7 Discussion

Reason for study

The risk of retinopathy increases with disease duration.[1] Type 2 diabetes is often diagnosed several years after onset and almost 40% of people with type 2 diabetes are found to have retinopathy at diagnosis. This is potentially sight-threatening in between 4% and 8% of cases. Sixty per cent of people with type 2 diabetes will have retinopathy 20 years from onset.

There are two main mechanisms responsible for visual loss in diabetic retinopathy. The first of these is proliferative retinopathy and the development of new blood vessels. The second is MO, the build-up of fluid involving the area of the retina associated with best vision, the macula.

Although proliferative disease is most likely to lead to serious vision loss, MO is more common and is the leading cause of MVL in people with diabetes. However, laser treatment is only moderately effective, at best, for MO.

At the time of the introduction of national programmes for screening for diabetic retinopathy in the UK, screening for diabetic MO was not deemed to meet strict World Health Organization guidelines for screening.

Diabetic retinopathy at the macula (maculopathy) is more common than proliferative diabetic retinopathy so retinal screening programmes have had to develop pathways to address this, although it is acknowledged that the evidence base for doing so is limited.

The current recommended method of retinal screening for diabetic retinopathy is digital fundus photography, a two-dimensional technology that cannot detect MO directly. Current photographic grading schemes rely on a combination of surrogate markers, chosen by expert consensus, to infer the presence of MO.

These manual grading schemes use combinations of features of retinopathy including M/DHs, BHs, exudates and also visual acuity. At present, there is no consensus among the four nations as to which, and how many, features should be used to infer the presence of diabetic MO.

Clinical experience from the Grampian Retinal Screening Programme and Liverpool Screening Programme suggests that only 12–14% of these patients have evidence of MO, when examined by ophthalmologists using slit-lamp biomicroscopy.

Aim of the study

The primary aim of the study was to determine the best photographic surrogate markers for detecting potentially sight-threatening diabetic MO within English and Scottish national screening programmes. Specifically we wished to:

(a) investigate whether or not particular distributions and combinations of lesions (M/DHs, BHs and exudates), assessed manually or automatically, were more specific surrogate markers of MO than current practice, using OCT as the reference standard
(b) assess the costs and consequences of using alternative distributions and combinations of these lesions to screen for MO, using either automated or manual detection

© Queen's Printer and Controller of HMSO 2013. This work was produced by Olson *et al*. under the terms of a commissioning contract issued by the Secretary of State for Health. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

(c) model the long-term cost and quality-of-life implications of using alternative distributions and combinations of surrogate markers to screen for MO.

Once the study was under way, several screening programmes were found to be using OCT as part of the screening process to reduce false-positive referrals to the hospital eye service.

Consequently, we added a further aim to assess the costs and consequences of using OCT within retinal screening programmes in addition to improving the photographic surrogate markers, as this would affect how photographic markers would be used in future.

Study design

This was a multicentre, prospective, observational cohort study. We recruited 3540 subjects attending one of seven diabetic retinal screening programmes who had features of diabetic retinopathy visible within the macular region.

The specific diabetic retinopathy features of interest as surrogate markers for MO were M/DHs, BHs and exudates. All subjects were recruited and imaged at the participating centres in Aberdeen, Birmingham, Dundee, Dunfermline, Edinburgh, Liverpool and Oxford.

All subjects had their best corrected visual acuity assessed followed by a single retinal photograph of each eye. An OCT scan was acquired within 4 weeks of the retinal photograph.

The retinal photographs, OCT images and relevant demographic study data were uploaded using a customised web-based tool developed for the study.

All the retinal photographs were annotated for features of retinopathy by a research nurse who had 3 years' experience working as a retinal grader.

Software was developed to enable the research nurse to annotate key retinal landmarks and all lesions associated with maculopathy (M/DHs, BHs and exudates) as well as non-diabetic features with similar appearance which could confound the analysis (flame haemorrhages, drusen and CWS).

Recruitment

A total of 3540 patients were recruited from the seven participating centres, of whom 370 were excluded as either the retinal photographs or OCT images were not of adequate quality to meet the inclusion criteria in the study. The median age of the remaining 3170 subjects was 60 years. Of these subjects, 60.7% were male and 84.5% were Caucasian. There was a preponderance of people with type 2 diabetes (77.4%) compared with type 1 diabetes.

Of the 3170 patients, 243 (7.7%) were confirmed as having MO after review of the OCT images. The prevalence of MO differed greatly by centre, ranging from 3.7% to 12.2%. The prevalence of MO also differed between scanners, ranging from 4.5% to 11.8%.

Macular oedema and photographic features of retinopathy

The detection of MO was strongly related to the presence of lesions and was consistently higher in subjects with exudates or BHs (without exudates) within one DD than in those with just M/DHs within one DD.

The detection of MO was also strongly related to increasing counts of these three types of lesions. This was true of all three types of lesion within one DD (adjusted for each other) including M/DHs.

There was no evidence of a relationship between MO and the presence or count of exudates between one and two DD.

Having more than two M/DHs in an eye appeared to be of particular importance in predicting the presence of MO.

Although it would normally be impractical in manual screening to count individual lesions, it would be possible to identify an image with more than two M/DHs.

Macular oedema and patient characteristics

Having poor visual acuity was associated with increased detection of MO. Other patient characteristics associated with MO in univariate analyses were older age, Caucasian ethnicity and having type 2 diabetes rather than type 1 diabetes.

These apparent relationships between higher detection of MO with older age, poor visual acuity and type 2 diabetes were unlikely to have been entirely independent of each other.

Development of grading strategies

In order to evaluate the best combination of photographic features that would help predict the presence of MO on OCT, a number of grading strategies were explored including those currently used by screening programmes.

In addition to features of retinopathy, the inclusion of those patient characteristics that were found to be statistically significant in a univariate analysis were modelled.

We compared the performance of three broad approaches, namely manual grading strategies, computer-assisted manual annotation grading strategies and fully automated annotation grading strategies.

Because the proportion of patients recruited who had M/DHs only in the macula was lower than in the typical diabetic population attending a screening programme, it was necessary to weight subjects to bring proportions in line with those found in an earlier prospective cohort study.

The clinical effectiveness of manual grading strategies

The effectiveness of manual grading strategies is demonstrated by the trade-off between sensitivity and specificity.

The English grading system, which uses the largest number of features, is the most sensitive but the least specific, whereas the Scottish grading system is the most specific and the least sensitive.

A number of new manual grading strategies were devised with the aim of improving on existing approaches.

One of the manual grading strategies, labelled strategy 16, demonstrated an improved specificity by approximately 4% relative to the current grading practice in England, for a similar sensitivity. It differs from the current grading practice in England in that BHs within one DD are referred regardless of visual acuity status, and exudates between one DD and two DD are not included in the selection criteria. This grading strategy would appear to make grading easier and could improve appropriateness of referrals.

The current manual grading strategies that do not include visual acuity or the presence of M/DHs could be improved by taking these into account.

Current manual grading strategies that include exudates between one and two DD could be improved by removing this aspect.

© Queen's Printer and Controller of HMSO 2013. This work was produced by Olson *et al*. under the terms of a commissioning contract issued by the Secretary of State for Health. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

An ideal grading strategy would able to take into account the presence and count of all three types of lesion within one DD and also visual acuity.

The clinical effectiveness of computer-assisted manual annotation grading strategies

The optimum grading strategy, in terms of area under its receiver operating characteristic curve, was the computer-assisted manual annotation grading strategy, CAM. This uses the results of manual annotation of the individual lesions in each image.

This is a time-consuming procedure and so is unlikely to be considered for routine screening practice. Therefore, this manual annotation grading strategy was not taken forward for economic analysis.

The clinical effectiveness of fully automated annotation grading strategies

Fully automated annotation grading strategies were developed using previously described algorithms for detection of lesions of diabetic retinopathy.

The outputs of the lesion detectors were combined using an automated classifier which was trained to detect MO. Optionally, visual acuity and other non-image data were also input to the classifier.

A grading strategy using automated analysis of the macula-centred image performs slightly better than current manual grading strategies.

If automated annotation grading data are combined with non-image data, then further improvements are obtained; a large proportion of this improvement is due to visual acuity alone. Because visual acuity is a standard measurement performed during diabetic retinopathy screening, this additional information could be added to a fully automated annotation grading strategy with little effort.

Cost-effectiveness

The economic analysis was conducted in two phases. An initial analysis focused on the cost per true case of MO appropriately detected and referred, both before and after adjusting for the expected frequency of different surrogate marker categories within the screening cohort.

This process relied on the estimated sensitivity/specificity of the alternative grading strategies obtained from the clinical data (see *Table 23*), combined with estimates of screening and referral costs obtained from a variety of sources (see *Table 24*).

In addition to estimating the cost-effectiveness of alternative simple manual grading strategies, the potential cost-effectiveness of using fully automated annotation grading, in combination with information on visual acuity within a linear classifier, was assessed.

Furthermore, we also assessed the cost-effectiveness of introducing OCT within the screening pathway, so as to further improve the specificity of referrals for suspected MO.

Following this initial analysis, a Markov microsimulation model was developed and used to assess the likely impact of the same grading/screening strategies on longer-term costs and health outcomes. This process relied on epidemiological, cost and health state utility data derived from the published literature.

Results of the cost-effectiveness analysis for manual grading strategies including quality-adjusted life-years and long-term costs (aims b and c, Chapter 2*)*

The cost-effectiveness modelling clearly demonstrates how costs to the health service increase as more sensitive manual grading strategies are adopted over more specific grading strategies.

Despite the only very small differences in health outcomes between manual grading strategies, increases in sensitivity (with accompanying reductions in specificity) did result in substantial increases in costs to the health service over time, as a result of more subjects being sent for OCT or referred to eye clinics for ongoing monitoring or treatment.

While the relative difference in the adjusted sensitivity of the manual grading strategies appeared quite substantial, the repetitive nature of screening diminishes the importance of this difference since any patients missed in one round of screening have a chance of being picked up during the next round prior to any visual loss occurring.

Even assuming very favourable parameters for the more sensitive manual grading strategies, the incremental cost per QALY estimates for the more sensitive grading strategies remained unfavourable when compared with thresholds used to judge cost-effectiveness in the UK (i.e. £20,000–30,000 per QALY).

Cost-effectiveness of manual grading strategies compared with a fully automated annotation grading strategy (aims b and b, Chapter 2*)*

The results would suggest that if significant improvements in specificity can be achieved without reducing sensitivity – for example, by adopting the more complex classifier based on automated image grading and pattern detection – such an approach could offer a more cost-effective alternative than current manual grading strategies based on simple manual algorithms.

Costs and consequences of using optical coherence tomography within retinal screening programmes in addition to improving the photographic surrogate markers

A valuable finding is that given the relatively high costs associated with referrals to hospital outpatient eye clinics, the incorporation of OCT within the screening pathway, targeted towards those with suspected MO, could potentially reduce costs to the NHS without reducing the number of OCT-positive cases of MO detected and referred.

When looking at those grading strategies based on observing the presence of surrogate photographic markers, the addition of OCT prior to referral, as part of the screening pathway, to each of these grading strategies reduced the overall costs of detection and referral without impacting on the number of OCT-positive MO cases detected.

The OCT costs were based on the experience of the lead centre, where a band 6 screener carries out the OCT examinations under the supervision of a consultant ophthalmologist.

However, deterministic sensitive analysis suggests that if used in combination with retinal photography (on a 6-monthly basis) to monitor patients within screening, the use of OCT could remain cost-saving up to an incremental cost of ~ £58 per patient. In the Scottish context, this threshold cost drops to ~ £47.

Strengths

This was a multicentre study reflecting everyday clinical practice throughout the UK. This allowed for all patients attending retinal screening programmes to be invited if they met the inclusion criteria.

© Queen's Printer and Controller of HMSO 2013. This work was produced by Olson *et al*. under the terms of a commissioning contract issued by the Secretary of State for Health. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

To avoid intercentre variation, all OCT operators were required to be accredited before submitting data for the study. Operators submitted a portfolio that included the images collected using the OCT scanner they would be using for the study.

The web-based system allowed for the centralisation of retinal image and OCT grading using quality-assured graders.

The database of annotated images allowed us to explore in detail the relationships between the various photographic features of retinopathy and the presence of MO seen with OCT.

We were also able to compare the clinical effectiveness and cost-effectiveness of the various existing manual grading strategies on a single data set, thus highlighting their individual strengths and weakness.

We built on existing automated algorithms for assessing diabetic retinopathy. We were able to incorporate the automated detections of features along with patient characteristics in a novel automated annotation grading algorithm to optimise screening for MO.

The economic analysis included a survey of costs and pathways of implementation in the participating centres. The results of the study can therefore be applied across England and Scotland.

Limitations

Since its commercial launch in 1996, OCT technology has been continuously evolving. Improvements have resulted in improved signal-to-noise ratio and axial and lateral resolution, and shortened acquisition times.

A number of scanners are now available, but there is poor agreement between retinal thickness measurements made on the different scanners. It was decided to include a variety of scanners in order to benefit from the higher-quality images from the newer devices and to make the results more generally applicable.

In order to include results from different OCT scanners, experiments were undertaken before the study began to estimate the differences in thickness between the new scanners to allow corrections to be made.

The wide variation in MO detection between centres is most likely due to differences in the sensitivity of the OCT scanner than true variations in prevalence.

The centres with the latest high-resolution scanners all detect a higher proportion of their subjects as having MO than centres relying on older technology.

In centres with older scanners a proportion of the true cases of MO will be graded as having no MO. Such cases will affect the estimated sensitivities and specificities of the different strategies.

Three hundred and seventy people were excluded from the final analysis because an adequate quality image could not be obtained either by retinal photography or by using OCT.

This was related to the type of scanner used and some patient-related characteristics such as poor visual acuity and older age.

In the recruited cohort there was an under-representation of those with just M/DHs in the macula when compared with a typical screening cohort. This reflected a variation in current practice between centres with regard to how M/DHs in the macula region are handled by the screening programme.

We were able to adjust for this by weighting the data using information from an earlier reference graded prospective cohort study. However, the overall findings and conclusions were generally insensitive to the weighting process.

There was a lack of contemporary data on the progression of MO in untreated patients for economic modelling. This required the assumption that patients without treatment would progress at the rate observed in the ETDRS.[64]

We tried to use the best available evidence on the effects of laser treatment compared with no treatment for those patients with MO, in order to assess the benefits of improved detection and referral.

In doing so, we updated older effect profiles for laser treatment to accommodate the observation of improved outcomes being associated with its use in more contemporary studies.[68]

A further potential weakness was the assumption that all patients with MO would be treated with laser, rather than intravitreal antivascular endothelial growth factor injections.

This decision was taken based on current National Institute for Health and Care Excellence guidance,[81] as well as a survey of participating centres which suggested that laser still forms the mainstay of treatment.

Implications for health care

Compared with all current manual grading schemes, for the same sensitivity, a fully automated strategy, using the automated detection of patterns of photographic surrogate markers, achieves a higher specificity for detecting MO in people with diabetes, especially if visual acuity is included in the automated strategy.

Overall costs to the health service are likely to increase if more sensitive referral strategies are adopted over more specific screening strategies for MO, for only very small gains in QALYs.

The addition of OCT to each screening strategy, prior to referral, results in a potential for a reduction in costs to the health service with no decrement in the number of MO cases detected.

Retinal thickness measurements utilised by different scanners should be standardised.

Recommendations for research

Research should be undertaken into the most cost-effective method of screening for diabetic MO that takes into consideration a treatment strategy where the timing of intervention is based on a patient's perception of their visual disability.

This would also require research to be undertaken on the progression of MO in untreated patients and its impact on quality of life.

© Queen's Printer and Controller of HMSO 2013. This work was produced by Olson *et al*. under the terms of a commissioning contract issued by the Secretary of State for Health. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

Acknowledgements

This project could not have been successfully completed without the co-operation of all the centres. Much of the workload of acquiring the data has been carried out by the retinal screeners, supported by the clinical staff. To them we are very grateful.

Our particular thanks go to Julie Raeper, a Senior Retinal Screener at Aberdeen Royal Infirmary, who manually graded and annotated all the images.

A large part of the success of the project lies in the widespread use of web-based technology – both to ensure that data acquisition was standardised and easy to do, and to facilitate the reading of the images.

Keith Goatman took the major responsibility for this. He also took a lead role in the development of the analysis software with the support of Alan Fleming, on whose image analysis software this part of the project was based.

Thanks must go to those who regularly made the journey to Aberdeen to sit on the Trials Steering Committee and take part in the Investigators' Meetings.

In particular we would like to thank those members not involved directly in the study: Mr Stephen Graham, our patient representative, Professor Alex Elliott, Dr A Manivannan and Mrs Alison Farrow. Their independent views were very useful contributions to our discussions. Thanks are also owed to the chair of the Trials Steering Committee for guiding the committee, initially Dr Caroline Styles and latterly Dr Rod Harvey.

Contributions of authors

Dr John Olson was the principal investigator and had overall responsibility, co-ordination and supervision, with special responsibility for clinical issues and the design of the study.

Professor Peter Sharp was involved in project management, interpretation of the findings and contributed to the writing of the report.

Dr Keith Goatman managed the project, overseeing the recruitment of patients, and developing the software. He wrote the early drafts of the report and contributed to the final report.

Dr Gordon Prescott was responsible for the statistical analysis of the data and contributed *Chapter 4* and part of *Chapter 5*.

Dr Graham Scotland was responsible for the health economic analysis and contributed *Chapter 6*.

Dr Alan Fleming produced the initial software for analysing retinal images for features of diabetic retinopathy. He worked with Dr Goatman on the development of the software for assessing the surrogate photographic markers of MO and contributed to the writing of *Chapter 5*.

Dr Sam Philip worked with Dr Olson in the clinical interpretation of the results and in the writing of several of the chapters of the report. The weightings used to adjust the data to reflect current practice were derived from his earlier work at Aberdeen.

© Queen's Printer and Controller of HMSO 2013. This work was produced by Olson *et al*. under the terms of a commissioning contract issued by the Secretary of State for Health. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

Dr Cynthia Santiago, **Dr Shyamanga Borooah**, **Professor Deborah Broadbent**, **Professor Victor Chong**, **Professor Paul Dodson**, **Professor Simon Harding**, **Professor Graham Leese**, **Dr Caroline Styles**, **Dr Ken Swa** and **Ms Helen Wharton** were responsible for providing patient data from the collaborating centres, contributing to the interpretation of the findings and reading the final report.

Ethical permission

The main approval was obtained from the North of Scotland Research Ethics Committee (REC reference: 07/S0801/107). Date of approval: 17 December 2007.

References

1. Goatman KA. A reference standard for the measurement of macular oedema. *Br J Ophthalmol* 2006;**90**:1197–202. http://dx.doi.org/10.1136/bjo.2006.095885
2. Facey K, Cummins E, Macpherson K, Morris A, Reay L, Slattery J. *Organisation of services for diabetic retinopathy screening. Health Technology Assessment report 1*. Glasgow: Health Technology Board for Scotland; 2002.
3. Bhagat N, Grigorian RA, Tutela A, Zarbin MA. Diabetic macular oedema: pathogenesis and treatment. *Surv Ophthalmol* 2009;**54**:1–32.
4. Kinyoun J, Barton F, Fisher M, Hubbard L, Aiello L, Ferris F. Detection of diabetic macular oedema. Ophthalmoscopy versus photography – Early Treatment Diabetic Retinopathy Study Report Number 5. The ETDRS Research Group. *Ophthalmology* 1989;**96**:746–50; discussion 750–1.
5. Bresnick GH, Mukamel DB, Dickinson JC, Cole DR. A screening approach to the surveillance of patients with diabetes for the presence of vision-threatening retinopathy. *Ophthalmology* 2000;**107**:19–24. http://dx.doi.org/10.1016/S0161-6420(99)00010-X
6. Sanchez-Tocino H, Alvarez-Vidal A, Maldonado MJ, Moreno-Montanes J, Garcia-Layana A. Retinal thickness study with optical coherence tomography in patients with diabetes. *Invest Ophthalmol Vis Sci* 2002;**43**:1588–94.
7. Massin P, Girach A, Erginay A, Gaudric A. Optical coherence tomography: a key to the future management of patients with diabetic macular oedema. *Acta Ophthalmol Scand* 2006;**84**:466–74. http://dx.doi.org/10.1111/j.1600-0420.2006.00694.x
8. Hee MR, Puliafito C, Duker JS, Reichel E, Coker JG, Wilkins JR, *et al.* Topography of diabetic macular oedema with optical coherence tomography. *Ophthalmology* 1998;**105**:360–70.
9. Strom C. Comparison of objective retinal thickness analysis and subjective stereo fundus photography in diabetic macular oedema. *Invest Ophthalmol Vis Sci* 2004;**45**:1450–5.
10. Stanga PE, Downes SM, Ahuja RM, Chong NH, Antcliff R, Reck AC, *et al.* Comparison of optical coherence tomography and fluorescein angiography in assessing macular oedema in retinal dystrophies: preliminary results. *Int Ophthalmol* 2001;**23**:321–5.
11. Chew EY, Ferris FL, Csaky KG, Murphy RP, Agrón E, Thompson DJS, *et al.* The long-term effects of laser photocoagulation treatment in patients with diabetic retinopathy: the early treatment diabetic retinopathy follow-up study. *Ophthalmology* 2003;**110**:1683–9.
12. Aiello LP, Wong JS. Role of vascular endothelial growth factor in diabetic vascular complications. *Kidney Int Suppl* 2000;**77**:S113–19. http://dx.doi.org/10.1046/j.1523-1755.2000.07718.x
13. Aiello LP, Edwards AR, Beck RW, Bressler NM, Davis MD, Ferris F, *et al.* Factors associated with improvement and worsening of visual acuity 2 years after focal/grid photocoagulation for diabetic macular oedema. *Ophthalmology* 2010;**117**:946–53.
14. Bonini-Filho MA, Jorge R, Barbosa JC, Calucci D, Cardillo JA, Costa RA. Intravitreal injection versus sub-Tenon's infusion of triamcinolone acetonide for refractory diabetic macular oedema: a randomised clinical trial. *Invest Ophthalmol Vis Sci* 2005;**46**:3845–9.
15. Yilmaz T, Weaver CD, Gallagher MJ, Cordero-Coma M, Cervantes-Castaneda RA, Klisovic D, *et al.* Intravitreal triamcinolone acetonide injection for treatment of refractory diabetic macular oedema: a systematic review. *Ophthalmology* 2009;**116**:902–11; quiz 912–13.

© Queen's Printer and Controller of HMSO 2013. This work was produced by Olson *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

16. Cunningham ET, Adamis AP, Altaweel M, Aiello LP, Bressler NM, D'Amico DJ, *et al.* A phase II randomised double-masked trial of pegaptanib, an anti-vascular endothelial growth factor aptamer, for diabetic macular oedema. *Ophthalmology* 2005;**112**:1747–57.

17. Michaelides M, Kaines A, Hamilton RD, Fraser-Bell S, Rajendram R, Quhill F, *et al.* A prospective randomised trial of intravitreal bevacizumab or laser therapy in the management of diabetic macular oedema (BOLT study) 12-month data: report 2. *Ophthalmology* 2010;**117**:1078–86.e2.

18. Bolz M, Ritter M, Schneider M, Simader C, Scholda C, Schmidt-Erfurth U. A systematic correlation of angiography and high-resolution optical coherence tomography in diabetic macular oedema. *Ophthalmology* 2009;**116**:66–72.

19. Massin P, Bandello F, Garweg JG, Hansen LL, Harding SP, Larsen M, *et al.* Safety and Efficacy of Ranibizumab in Diabetic Macular Oedema (RESOLVE Study). *Diabetes Care* 2010;**33**:2399–405.

20. Nguyen QD, Shah SM, Heier JS, Do DV, Lim J, Boyer D, *et al.* Primary end point (six months) results of the ranibizumab for oedema of the macula in diabetes (READ-2) study. *Ophthalmology* 2009;**116**:2175–81.e1.

21. Nguyen QD, Shah SM, Khwaja A, Channa R, Hatef E, Do DV, *et al.* Two-year outcomes of the ranibizumab for oedema of the macula in diabetes (READ-2) study. *Ophthalmology* 2010;**117**:2146–51.

22. Mitchell P, Bandello F, Schmidt-Erfurth U, Lang GE, Massin P, Schlingemann RO, *et al.* The RESTORE study: ranibizumab monotherapy or combined with laser versus laser monotherapy for diabetic macular oedema. *Ophthalmology* 2011;**118**:615–25.

23. Wilson JMG, Jungner G. *Principles and practice of screening for disease*. Geneva: World Health Organization; 1968.

24. Flach AJ. The incidence, pathogenesis and treatment of cystoid macular oedema following cataract surgery. *Trans Am Ophthalmol Soc* 1998;**96**:557–634.

25. Vestgaard M, Ringholm L, Laugesen CS, Rasmussen KL, Damm P, Mathiesen ER. Pregnancy-induced sight-threatening diabetic retinopathy in women with type 1 diabetes. *Diabetic Med* 2010;**27**:431–5. http://dx.doi.org/10.1111/j.1464-5491.2010.02958.x

26. Rossing K, Jacobsen P, Hommel E, Mathiesen E, Svenningsen A, Rossing P, *et al.* Pregnancy and progression of diabetic nephropathy. *Diabetologia* 2002;**45**:36–41. http://dx.doi.org/10.1007/s125-002-8242-4

27. Rasmussen KL, Laugesen CS, Ringholm L, Vestgaard M, Damm P, Mathiesen ER. Progression of diabetic retinopathy during pregnancy in women with type 2 diabetes. *Diabetologia* 2010;**53**:1076–83. http://dx.doi.org/10.1007/s00125-010-1697-9

28. Early Treatment for Diabetic Retinopathy Study Research Group. Early Treatment for Diabetic Retinopathy Study. Photocoagulation for diabetic macular oedema. *Arch Ophthalmol* 1985;**103**:1796–806.

29. Browning DJ, Apte RS, Bressler SB, Chalam KV, Danis RP, Davis MD, *et al.* Association of the extent of diabetic macular oedema as assessed by optical coherence tomography with visual acuity and retinal outcome variables. *Retina* 2009;**29**:300–5.

30. Philip S, Fleming AD, Goatman KA, Fonseca S, McNamee P, Scotland GS, *et al.* The efficacy of automated 'disease/no disease' grading for diabetic retinopathy in a systematic screening programme. *Br J Ophthalmol* 2007;**91**:1512–17. http://dx.doi.org/10.1136/bjo.2007.119453

31. Holladay JT. Proper method for calculating average visual acuity. *J Refract Surg* 1997;**13**:388–91.

32. Polito A, Del Borrello M, Isola M, Zemella N, Bandello F. Repeatability and reproducibility of fast macular thickness mapping with stratus optical coherence tomography. *Arch Ophthalmol* 2005;**123**:1330–7. http://dx.doi.org/10.1001/archopht.123.10.1330

33. Paunescu L. Reproducibility of nerve fibre thickness, macular thickness, and optic nerve head measurements using StratusOCT. *Invest Ophthalmol Vis Sci* 2004;**45**:1716–24.

34. Carlton J, Karnon J, Smith KJ, Marr J. The clinical effectiveness and cost-effectiveness of screening programmes for amblyopia and strabismus in children up to the age of 4–5 years: a systematic review and economic evaluation. *Health Technol Assess* 2008;**12**(25).

35. Scanlon PH, Foy C, Chen FK. Visual acuity measurement and ocular co-morbidity in diabetic retinopathy screening. *Br J Ophthalmol* 2008;**92**:775–8. http://dx.doi.org/10.1136/bjo.2007.128561

36. Liazos E, Broadbent DM, Beare N, Kumar N. Spontaneous resolution of diabetic macular oedema after discontinuation of thiazolidenediones. *Diabet Med* 2008;**25**:860–2. http://dx.doi.org/10.1111/j.1464-5491.2008.02491.x

37. Ambrosius WT, Danis RP, Goff DC, Greven CM, Gerstein HC, Cohen RM, *et al.* Lack of association between thiazolidinediones and macular oedema in type 2 diabetes: the ACCORD eye substudy. *Arch Ophthalmol* 2010;**128**:312–18.

38. Scottish Diabetic Retinal Screening Collaborative. Scottish Diabetic Retinopathy Grading Scheme 2007. 2007. URL: www.ndrs.scot.nhs.uk/ClinGrp/Docs/Grading Scheme 2007 v1.1.pdf (accessed 13 April 2013).

39. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* 1986;**1**:307–10. http://dx.doi.org/10.1016/j.ijnurstu.2009.10.001

40. Bland JM, Altman DG. Measuring agreement in method comparison studies. *Stat Methods Med Res* 1999;**8**:135–60. http://dx.doi.org/10.1191/096228099673819272

41. Shrout PE, Fleiss JL. Intraclass correlations: uses in assessing rater reliability. *Psychol Bull* 1979;**86**:420–8. http://dx.doi.org/10.1037//0033-2909.86.2.420

42. Brown H, Prescott R. *Applied mixed models in medicine (statistics in practice).* Hoboken, NJ: John Wiley & Sons; 2006.

43. Kish L. *Survey sampling*. New York, NY: John Wiley & Sons; 1965. pp. 424–6.

44. Bethlehem J. Weighting. In Lavrakas PJ, editor. *The encyclopaedia of survey research methods*. vol. 2. New York, NY: Sage; 2008. pp. 957–60.

45. Sapsford R. *Survey research*. London: Sage; 1999. pp. 31–2.

46. Warwick D, Lininger CA. *The sample survey: theory and practice*. New York, NY: McGraw-Hill; 1975. pp. 273–4.

47. Fleming AD, Goatman KA, Philip S, Prescott GJ, Sharp PF, Olson JA. Automated grading for diabetic retinopathy: a large-scale audit using arbitration by clinical experts. *Br J Ophthalmol* 2010;**94**:1606–10. http://dx.doi.org/10.1136/bjo.2009.176784

48. Fleming AD, Goatman KA, Philip S, Williams GJ, Prescott GJ, Scotland GS, *et al.* The role of haemorrhage and exudate detection in automated grading of diabetic retinopathy. *Br J Ophthalmol* 2010;**94**:706–11. http://dx.doi.org/10.1136/bjo.2008.149807

49. Goatman K, Charnley A, Webster L, Nussey S. Assessment of automated disease detection in diabetic retinopathy screening using two-field photography. *PLoS ONE* 2011;**6**:e27524. http://dx.doi.org/10.1371/journal.pone.0027524

© Queen's Printer and Controller of HMSO 2013. This work was produced by Olson *et al*. under the terms of a commissioning contract issued by the Secretary of State for Health. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

50. Fleming A, Philip S, Goatman K, Olson J, Sharp P. Automated microaneurysm detection using local contrast normalisation and local vessel detection. *IEEE Trans Med Imaging* 2006;**25**:1223–32.

51. Fleming AD, Philip S, Goatman KA, Olson JA, Sharp PF. Automated assessment of diabetic retinal image quality based on clarity and field definition. *Invest Ophthalmol Vis Sci* 2006;**47**:1120–5. http://dx.doi.org/10.1167/iovs.05-1155

52. Fleming AD, Philip S, Goatman KA, Olson JA, Sharp PF. Automated detection of exudates for diabetic retinopathy screening. *Phys Med Biol* 2007;**52**:331–45. http://dx.doi.org/10.1088/0031-9155/52/24/012

53. Fleming A, Goatman K, Philip S, Olson J, Sharp P. Automatic detection of retinal anatomy to assist diabetic retinopathy screening. *Phys Med Biol* 2007;**52**:331–45. http://dx.doi.org/10.1088/0031-9155/52/2/002

54. Fleming A, Goatman K, Williams G, Philip S, Sharp P, Olson J. Automated detection of blot haemorrhages as a sign of referable retinopathy. In McKenna S, Hoey J, editors. *Proceedings of 12th Annual Conference on Medical Image Understanding and Analysis*. Dundee; British Machine Vision Association; 2008. pp. 235–9.

55. HM Treasury. *Green book guidance on discounting*. URL: www.hm-treasury.gov.uk/green_book_guidance_discounting.htm (accessed 1 May 2013).

56. Personal Social Services Research Unit (PPSRU). *Unit costs of health and social care 2010*. Canterbury: PSSRU, University of Kent; 2010.

57. Department of Health. 2010. URL: http://webarchive.nationalarchives.gov.uk/+/www.dh.gov.uk/en/Managingyourorganisation/Financeandplanning/NHSFinancialReforms/index.htm (accessed 1 May 2013).

58. Information Services Division (ISD) Scotland. *Specialty costs*. 2010. URL: www.isdscotland.org/Health-Topics/Finance/Costs/File-Listings-2010.asp (accessed 1 May 2013).

59. Department of Health. *NHS reference costs 2009–2010*. URL: www.gov.uk/government/publications/nhs-reference-costs-2009-2010 (accessed 1 May 2013).

60. National Institute for Health and Care Excellence (NICE). *Macular oedema (diabetic): ranibizumab: final appraisal determination document*. London: NICE; 2011.

61. Early photocoagulation for diabetic retinopathy. ETDRS report number 9. Early Treatment Diabetic Retinopathy Study Research Group. *Ophthalmology* 1991;**98**(Suppl. 5):766–85.

62. Jyothi S, Sivaprasad S. Five-year visual outcome following laser photocoagulation of diabetic macular oedema. *Eye* 2011;**25**:850–9. http://dx.doi.org/10.1038/eye.2011.102

63. Lloyd A, Nafees B, Gavriel S, Rousculp MD, Boye KS, Ahmad A. Health utility values associated with diabetic retinopathy. *Diabet Med* 2008;**25**:618–24. http://dx.doi.org/10.1111/j.1464-5491.2008.02430.x

64. Early Treatment Diabetic Retinopathy Study research group. Early Treatment Diabetic Retinopathy Study (ETDRS). Photocoagulation for diabetic macular oedema. Early Treatment Diabetic Retinopathy Study report number 1. *Arch Ophthalmol* 1985;**103**:1796–806.

65. Mulnier HE, Seaman HE, Raleigh VS, Soedamah-Muthu SS, Colhoun HM, Lawrenson RA. Mortality in people with type 2 diabetes in the UK. *Diabet Med* 2006;**23**:516–21. http://dx.doi.org/10.1111/j.1464-5491.2006.01838.x

66. Soedamah-Muthu SS, Fuller JH, Mulnier HE, Raleigh VS, Lawrenson RA, Colhoun HM. All-cause mortality rates in patients with type 1 diabetes mellitus compared with a non-diabetic population from the UK general practice research database, 1992–1999. *Diabetologia* 2006;**49**:660–6. http://dx.doi.org/10.1007/s00125-005-0120-4

67. Younis N, Broadbent DM, Harding SP, Vora JP. Incidence of sight-threatening retinopathy in type 1 diabetes in a systematic screening programme. *Diabet Med* 2003;**20**:758–65. http://dx.doi.org/10.1046/j.1464-5491.2003.01035.x

68. Younis N, Broadbent DM, Vora JP, Harding SP. Incidence of sight-threatening retinopathy in patients with type 2 diabetes in the Liverpool Diabetic Eye Study: a cohort study. *Lancet* 2003;**361**:195–200. http://dx.doi.org/10.1016/S0140-6736(03)12267-2

69. PKC-DRS2 Group, Aiello LP, Davis MD, Girach A, Kles KA, Milton RC, *et al.* Effect of ruboxistaurin on visual loss in patients with diabetic retinopathy. *Ophthalmology* 2006;**113**:2221–30.

70. Kohner EM, Stratton IM, Aldington SJ, Holman RR, Matthews DR, Group UKPDS (IKPDS). Relationship between the severity of retinopathy and progression to photocoagulation in patients with type 2 diabetes mellitus in the UKPDS (UKPDS 52). *Diabet Med* 2001;**18**:178–84.

71. Bhavsar KV, Subramanian ML. Risk factors for progression of subclinical diabetic macular oedema. *Br J Ophthalmol* 2011;**95**:671–4. http://dx.doi.org/10.1136/bjo.2010.182337

72. Varma R, Choudhury F, Klein R, Chung J, Torres M, Azen SP, *et al.* Four-year incidence and progression of diabetic retinopathy and macular oedema: the Los Angeles Latino Eye Study. *Am J Ophthalmol* 2010;**149**:752–3.

73. Wong TY, Mwamburi M, Klein R, Larsen M, Flynn H, Hernandez-Medina M, *et al.* Rates of progression in diabetic retinopathy during different time periods: a systematic review and meta-analysis. *Diabetes Care* 2009;**32**:2307–13. http://dx.doi.org/10.2337/dc090615

74. Craig BA, Sendi PP. Estimation of the transition matrix of a discrete-time Markov chain. *Health Econ* 2002;**11**:33–42. http://dx.doi.org/10.1002/hec.654

75. Network DRCR. A randomised trial comparing intravitreal triamcinolone acetonide and focal/grid photocoagulation for diabetic macular oedema. *Ophthalmology* 2008;**115**:1447–9; 1449.e1–10.

76. Network DRCR, Elman MJ, Aiello LP, Beck RW, Bressler NM, Bressler SB, *et al.* Randomised trial evaluating ranibizumab plus prompt or deferred laser or triamcinolone plus prompt laser for diabetic macular oedema. *Ophthalmology* 2010;**117**:1064–77.e35.

77. Elman MJ, Bressler NM, Qin H, Beck RW, Ferris FL, Friedman SM, *et al.* Expanded 2-year follow-up of ranibizumab plus prompt or deferred laser or triamcinolone plus prompt laser for diabetic macular oedema. *Ophthalmology* 2011;**118**:609–14.

78. Hirai FE, Knudtson MD, Klein BE, Klein R. Clinically significant macular oedema and survival in type 1 and type 2 diabetes. *Am J Ophthalmol* 2008;**145**:700–6.

79. Meads C, Hyde C. What is the cost of blindness? *Br J Ophthalmol* 2003;**87**:1201–4. http://dx.doi.org/10.1136/bjo.87.10.1201

80. Briggs AH. Handling uncertainty in cost-effectiveness models. *Pharmacoeconomics* 2000;**17**:479–500. http://dx.doi.org/10.2165/00019053-200017050-00006

81. National Institute for Health and Care Excellence (NICE). *Ranibizumab for the treatment of diabetic macular oedema*. URL: http://guidance.nice.org.uk/TA237 (accessed 1 May 2013).

82. Vijan S, Hofer TP, Hayward RA. Cost-utility analysis of screening intervals for diabetic retinopathy in patients with type 2 diabetes mellitus. *JAMA* 2000;**283**:889–96. http://dx.doi.org/10.1001/jama.283.7.889

83. Javitt JC, Aiello LP, Chiang Y, Ferris FL, 3rd, Canner JK, Greenfield S. Preventative eye care in people with diabetes is cost-saving to the federal government. Implications for health-care reform. *Diabetes Care* 1994;**17**:909–17. http://dx.doi.org/10.2337/diacare.17.8.909

© Queen's Printer and Controller of HMSO 2013. This work was produced by Olson *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

84. Davies R, Roderick P, Canning C, Brailsford S. The evaluation of screening policies for diabetic retinopathy using simulation. *Diabet Med* 2002;**19**:762–70. http://dx.doi.org/10.1046/j.1464-5491.2002.00773.x

85. James M, Turner DA, Broadbent DM, Vora J, Harding SP. Cost effectiveness analysis of screening for sight threatening diabetic eye disease. *BMJ* 2000;**320**:1627–31. http://dx.doi.org/10.1136/bmj.320.7250.1627

86. Fercher AF. Optical coherence tomography – development, principles, applications. *Zeitschrift für medizinische Physik* 2010;**20**:251–76. http://dx.doi.org/10.1016/j.zemedi.2009.11.002

87. Virgili G, Menchini F, Dimastrogiovanni AF, Rapizzi E, Menchini U, Bandello F, *et al.* Optical coherence tomography versus stereoscopic fundus photography or biomicroscopy for diagnosing diabetic macular oedema: a systematic review. *Invest Ophthalmol Vis Sci* 2007;**48**:4963–73.

88. Huang D, Swanson EA, Lin CP, Schuman JS, Stinson WG, Chang W, *et al.* Optical coherence tomography. *Science* 1991;**254**:1178–81. http://dx.doi.org/10.1126/science.1957169

89. Kinyoun J, Barton F, Fisher M, Hubbard L, Aiello L, Ferris F. Detection of diabetic macular edema. Ophthalmoscopy versus photography – early treatment diabetic retinopathy study report number 5. The ETDRS Research Group. *Ophthalmology* 1989;**96**:746–50; discussion 750–1.

90. Davis MD, Gangnon RE, Lee L-Y, Hubbard LD, Klein BEK, Klein R, *et al.* The Age-Related Eye Disease Study severity scale for age-related macular degeneration: AREDS Report No. 17. *Arch Ophthalmol* 2005;**123**:1484–98.

91. Zeimer R, Shahidi M, Mori M, Zou S, Asrani S. A new method for rapid mapping of the retinal thickness at the posterior pole. *Invest Ophthalmol Vis Sci* 1996;**37**:1994–2001.

Appendix 1 Techniques for measuring retinal thickness

Macular oedema has traditionally been assessed clinically using a combination of slit-lamp biomicroscopy, stereo photography and stereo fluorescein angiography. However, these techniques have a number of limitations that make them unsuitable for this study. Foremost is that they offer only qualitative assessments, which are also relatively insensitive to changes in retinal thickness. Furthermore, slit-lamp examination does not provide a pictorial record and, together with stereo photography, is known to be biased by the presence or absence of exudates. Angiography is a sensitive test for leakage, but not for thickening. Best corrected visual acuity has also been used as a surrogate indication of thickening, but is neither sensitive nor specific. Four imaging techniques that offer objective, quantitative measures of macular thickening are described below. *Table 39* lists the axial and lateral resolutions of the four methods.

Ultrasound

Ultrasound frequencies between 10 MHz and 20 MHz are the most useful for retinal imaging. Frequencies as high as 50 MHz have been used for high-resolution imaging of the anterior segment, but these have insufficient penetration for retinal imaging. Interfaces between materials with different acoustic properties generate strong echoes, whereas materials that scatter the ultrasound beam return weaker echoes. Fluid-filled structures, such as the vitreous, or cysts, neither reflect nor scatter ultrasound. Therefore, there is strong contrast between MO and normal retinal tissue.

The two-dimensional cross-sectional ultrasound image is composed of a series of A-scans. These are single axial profiles recording the strength of echoes from different tissue depths. One advantage of ultrasound imaging is that it is not affected by optical opacities, such as cataract or vitreous haemorrhage. However, it has relatively poor resolution (approximately 200 µm axially and laterally) and is an invasive technique, requiring contact with the eye.

Optical coherence tomography

Optical coherence tomography[86] is a popular, rapid and non-invasive technique for cross-sectional retinal imaging that has proved convenient for longitudinal studies and as a trial outcome measure.[1,87] It is often referred to as an optical analogue of ultrasound imaging, measuring backscattered light rather than sound. However, unlike ultrasound, OCT is a non-invasive, non-contact technique. Nevertheless, much OCT terminology has been borrowed from ultrasound imaging, such as the 'A-scan' (a signal vs. depth profile) and 'B-scan' (a collection of A-scans giving a two-dimensional cross-section). The first commercial OCT scanner was launched in 1996 by Carl Zeiss Meditec. It used a super-luminescent diode light source and time domain interferometry based on a moving mirror and a Michelson interferometer.[88] In 2001 Zeiss released their most recent time domain OCT system, the Stratus OCT. Although faster than its predecessors it still only acquires 400 A-scans/second, which limits the number of cross-sections that can be acquired before eye movement is a problem. The latest generation of OCT scanners are based on spectral domain techniques, which dispense with the moving mirror and are thereby able to increase the acquisition rate by up to two orders of magnitude. Three such scanners, the Zeiss Cirrus OCT, Topcon 3D OCT-1000 and Heidelberg Spectralis, were used in this study.

Prior to OCT, retinal cross-sections were possible only using ultrasound imaging. The introduction of OCT improved the axial resolution of cross-sections by two orders of magnitude, allowing structures to be seen in vivo that were formerly visible only by histological examination (*Figure 28*). Retinal thickness measurements are calculated from the cross-sectional images using software to segment the inner and outer limiting boundaries of the retina, although there is disagreement about which boundary best

© Queen's Printer and Controller of HMSO 2013. This work was produced by Olson *et al*. under the terms of a commissioning contract issued by the Secretary of State for Health. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

TABLE 39 Summary of axial and lateral resolution of retinal thickness techniques. Note that the axial sampling used by the Zeiss Stratus OCT results in a much higher thickness map lateral resolution at the centre of the region than at the periphery. Modern spectral domain scanners more commonly sample the area rectilinearly and have constant resolution across the map

Method	Axial resolution (μm)	Lateral resolution (μm)
Ultrasound (B mode)	150–200	250–500
OCT (cross-section)	2–15	10–20
OCT (Stratus map)	2–15	10 (centre)–1500 (edge)
Retinal Thickness Analyser	50	380
Heidelberg cSLO	150–300	10–20

cSLO, Confocal Scanning Laser Ophthalmoscope.

FIGURE 28 Optical coherence tomography retinal cross-section.

represents the lower limit of the retina. Although thickness measurements are calculated for every A-scan line, these are usually combined to form a nine-region thickness map centred on the fovea, as shown in *Figure 29*. This map layout was first used by the ETDRS[89] and later by the Age-Related Eye Disease Study.[90]

Retinal Thickness Analyser

The Retinal Thickness Analyser (RTA) was launched in 2000 by Talia Technology Ltd, based on research at Johns Hopkins University.[91] It consists of an integrated fundus camera and laser system for non-invasive retinal thickness measurement.

The thickness measurement is based on the same principle as the slip-lamp biomicroscope. A green (543 mm) helium–neon laser is projected as a slit, 3 mm in length and approximately 15 μm wide. The slit is not projected perpendicularly onto the retinal surface, but at an oblique angle of 15 degrees, so that light reflected from different depths results in a lateral displacement of the reflection: each 100 μm change in depth results in a 27 μm lateral shift. These lateral displacements are recorded using a charge-coupled device (image sensor; CCD) camera and automatic measurements used to generate a thickness map.

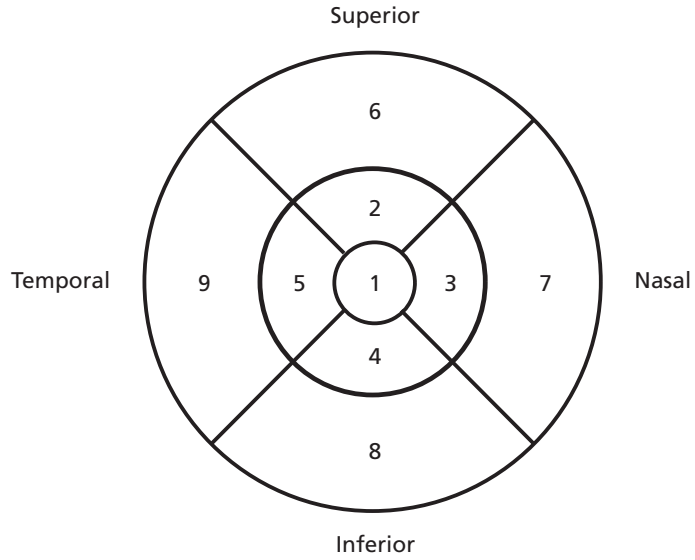


FIGURE 29 Early Treatment Diabetic Retinopathy Study/Age-Related Eye Disease Study regions. The inner circle has a diameter of 1 mm, the middle circle a diameter of 3 mm and the outer circle a diameter of 6 mm. Left eyes are reflected left-to-right such that the temporal region is always to the left and the nasal region to the right.

Confocal scanning laser ophthalmoscope

The scanning laser ophthalmoscope produces an image by rapidly scanning a laser spot across the retina, both horizontally and vertically. Any laser wavelength can be used, but red or infrared light is usually chosen for three-dimensional imaging, as the longer wavelengths are able to penetrate deeper into the retina. A three-dimensional image may be acquired by placing a confocal aperture in front of the detector. This allows only light from a given depth range to be detected. A three-dimensional volume is formed from a series of two-dimensional images at different depths, as selected by the confocal aperture position. Although the confocal scanning laser ophthalmoscope has good lateral resolution (at least as good as a fundus camera), the axial resolution is poor at around 150–300 µm, limiting the accuracy of the thickness measurements and showing little or no anatomical features in the axial direction.

© Queen's Printer and Controller of HMSO 2013. This work was produced by Olson *et al*. under the terms of a commissioning contract issued by the Secretary of State for Health. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

Appendix 2 Cost estimates for optical coherence tomography and slit-lamp examination within the screening programme

The following tables indicate the bottom-up cost calculations used to estimate the cost of OCT examinations within screening programmes.

Equipment

OCT machine	Unit cost (£)	Useful lifespan	EAC (£)	Annual throughput generated in screening programme	Cost per patient screened (£)	Source/notes
Scenario 1	36,000	10	4329	937	4.62	Centre information
Scenario 2	50,000	10	6012	937	6.42	Centre information
Annual service contract			3403	937	3.63	Centre information

EAC, equivalent annual cost.

Staff costs

Task	Band	Patients screened per day clinic	Working hours per day	Time per patient (minutes)	Unit cost per hour (£)	Cost per patient (£)	Source/notes
Administration time	3			5	12.86	1.07	Centre-reported data; PSSRU, 2010[56]
Examination	6	24	7.5	18.75	31	9.69	Centre-reported data; PSSRU, 2010[56]
Associated grading costs	6			5	31	2.58	Centre-reported data; PSSRU, 2010[56]
Additional L3 grader costs	Consultant			5	127	10.58	Centre-reported data; PSSRU, 2010[56]
Total						23.93	

L3, level 3.

Consumables	
Letter and postage	£0.30
Total cost per OCT exam (scenario 1)	£31.96
Total cost per OCT exam (scenario 2)	£33.64

© Queen's Printer and Controller of HMSO 2013. This work was produced by Olson *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

Slit-lamp examination costs

The following tables indicate the bottom-up cost calculations used to estimate the cost of providing slit-lamp examinations within screening programmes.

Equipment costs

Slit-lamp	Units	Cost (£)	Lifespan	EAC (£)	Annual throughput	Cost per patient (£)	Source/notes
Scenario 1	1	11,250	10	1353	3780	0.36	Centre reported; assumes shared
Scenario 2	1	15,000	10	1804	450	4.01	Centre reported; assumes exclusive to screening programme
Scenario 3	2	22,500	10	2705	1307	2.07	Centre reported; assumes exclusive to screening programme
Scenario 4	1	15,000	10	1804	800	2.25	Centre-reported data

EAC, equivalent annual cost.

Staff costs

Scenario	Clinic time (hours)	Clinic throughput	Time per patient (minutes)	Examiner band	Unit cost per hour (£)	Cost per patient (£)	Source/notes
Scenario 1	4	13	0.308	Associate specialist	90	27.69	Incorporates oncosts, overheads and treatment space; PSSRU, 2010[56]
Scenario 2	3.5	8	0.438	Associate specialist	90	39.38	Incorporates oncosts, overheads and treatment space; PSSRU, 2010[56]
Scenario 3	3.5	12	0.292	Band 6	31	9.04	Incorporates oncosts and overheads; PSSRU, 2010[56]
Scenario 3	3.5	12	0.292	Band 7	36	10.50	Incorporates oncosts and overheads; PSSRU, 2010[56]
Average scenario 3						9.77	Incorporates oncosts and overheads; PSSRU, 2010[56]
Scenario 4	3.5	10	0.350	Band 7	36	12.60	Incorporates oncosts and overheads; PSSRU, 2010[56]

Treatment space (scenarios 3 and 4)

Examination space (m^2)	New build cost per m^2 (£)	Total cost (£)	Lifespan (years)	EAC (£)	Use for slit-lamp	Annual attibutable cost to slit-lamp examinations (£)	Annual maintenance	Annual throughput	Cost per patient (£)
25	3000	75,000	60	3007	0.20	601	200	984	0.81

EAC, equivalent annual cost.

Notes: Data from NHS Grampian Estates Department; assumes space used for slit-lamp 1 day a week, 41 weeks per year, 24 patients per day.

Administration costs (all scenarios)

Administrative staff time	Unit cost (£)
Estimated cost per patient assuming similar to OCT clinic administration costs	1.07

Consumables

Item	Unit cost (£)
Tropicamide or phenylephrine 2.5% (drops)	0.98
Letter and postage	0.30

Total cost per slit-lamp examinations

Scenario	Total cost (£)
Scenario 1	30.40
Scenario 2	45.73
Scenario 3	15.01
Scenario 4	18.02
Average	**27.29**

© Queen's Printer and Controller of HMSO 2013. This work was produced by Olson *et al*. under the terms of a commissioning contract issued by the Secretary of State for Health. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

Appendix 3 Demographics and statistical modelling – additional data

Presence of different types of lesion at the seven centres

Lesion information was missing for three left eyes and two right eyes, so the denominator for left eyes was 3167 and for right eyes was 3168. Counts of subjects and eyes with MO and features by centre in each eye within one DD unless otherwise specified. Subjects can have several different lesions or none in each eye. Percentages have not been presented for very small counts.

Mutually exclusive groups of lesions by centre

Mutually exclusive (non-overlapping) groups of features were identified so that every eye could be classified into one group and one group only. The groups for comparison were (1) M/DH only (not BH or exudates); (2) BH, but not exudates (BH only + BH and M/DH only); and (3) exudates (regardless of what else). Prior to the recruitment these lesions were expected to be present in 11.3%, 1.4% and 3.5% of scanned images, but 69.8%, 8.6% and 21.6% of images with some lesions.

The percentages of left eyes in the groups: no lesions within one DD, M/DHs only within one DD, BHs only or BHs with M/DHs within one DD, exudates only or with M/DHs or BHs within one DD, other lesions within one DD were 28.1%, 40.3%, 8.4%, 20.4% and 2.8%, respectively. Very similar percentages were found in the right eyes: 27.9%, 41.7%, 8.7%, 18.9% and 2.8%. These percentages, particularly those for M/DHs only, did not match those in the paragraph above, and this was one reason why weighting of the subjects according to the features of their worst eye was considered necessary.

Note that in *Tables 43* and *45* there are some very large coefficients and upper ends of the 95% CIs for more than two BHs and for unknown visual acuity. This is a consequence of there being very few subjects with these characteristics, giving very wide CIs, and several of them having MO, inflating the estimates of ORs. These estimates are very volatile and should not be treated as a true reflection of the relationships that would be found in population of screened people with lesions.

© Queen's Printer and Controller of HMSO 2013. This work was produced by Olson *et al*. under the terms of a commissioning contract issued by the Secretary of State for Health. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

TABLE 40 Subjects and eyes with MO and features by centre in left and right eyes within one DD unless otherwise specified. Subjects can have several different lesions or none in an eye

Centre	MO+ve (%)	MO−ve (%)	BH (%)	CWS (%)	Drusen (%)	Exudate (%)	Flame haemorrhages (%)	M/DH (%)	Exudate, one to two DD (%)
Left eye									
Aberdeen	66 (7.3)	843 (92.7)	161 (17.7)	5	60 (6.6)	270 (29.7)	28 (3.1)	609 (67.0)	252 (27.7)
Birmingham	18 (2.1)	824 (97.9)	69 (8.2)	5	79 (9.4)	97 (11.5)	22 (2.6)	563 (66.9)	78 (9.3)
Dundee	20 (7.9)	234 (92.1)	32 (12.7)	1	8 (3.2)	37 (14.7)	4	126 (50.2)	38 (15.1)
Edinburgh	6 (2.8)	212 (97.2)	34 (15.6)	1	15 (6.9)	83 (38.1)	5	152 (69.7)	53 (24.3)
Liverpool	7 (1.7)	409 (98.3)	23 (5.5)	1	15 (3.6)	50 (12.0)	3	196 (47.1)	31 (7.5)
Dunfermline	3 (1.9)	156 (98.1)	8 (5.0)	0	9 (5.7)	15 (9.4)	1	89 (56.0)	12 (7.5)
Oxford	24 (6.5)	348 (93.5)	57 (15.3)	2	47 (12.6)	95 (25.5)	9	260 (69.9)	105 (28.2)
Total	144 (4.5)	3026 (95.5)	384 (12.1)	15 (0.5)	23 (7.4)	647 (20.4)	72 (2.3)	1995 (63.0)	569 (18.0)
Right eye									
Aberdeen	54 (5.9)	855 (94.1)	143 (15.7)	5	62 (6.8)	244 (26.8)	21 (2.3)	579 (63.7)	240 (26.4)
Birmingham	23 (2.7)	819 (97.3)	73 (8.7)	2	98 (11.7)	103 (12.2)	19 (2.3)	579 (68.8)	85 (10.1)
Dundee	15 (5.9)	239 (94.1)	32 (12.6)	2	12 (4.7)	38 (15.0)	6	144 (56.7)	38 (15.0)
Edinburgh	8 (3.7)	210 (96.3)	40 (18.3)	1	10 (4.6)	62 (28.4)	3	163 (74.8)	40 (18.3)
Liverpool	7 (1.7)	409 (98.3)	27 (6.5)	2	18 (4.3)	50 (12.0)	8	201 (48.4)	37 (8.9)
Dunfermline	5 (3.1)	154 (96.9)	12 (7.5)	0	8 (5.0)	19 (11.9)	2	96 (60.4)	15 (9.4)
Oxford	23 (6.2)	349 (93.8)	61 (16.4)	2	44 (11.8)	82 (22.0)	10 (2.7)	270 (72.6)	87 (23.4)
Total	135 (4.3)	3035 (95.7)	388 (12.2)	14 (0.4)	252 (8.0)	598 (18.9)	69 (2.2)	2032 (64.1)	542 (17.1)

−ve, negative; +ve, positive.

TABLE 41 Exclusive groups of features for comparison are (1) M/DH only (no BHs or exudates); (2) BH, but no exudates (BH only or BH+M/DH only); and (3) exudates (regardless of what else)

Centre		No lesions within one DD	M/DH only within one DD	BH or BH + M/DH within one DD	Exudate + any within one DD	Other within one DD	Total
Left eye							
Aberdeen	Count	193	306	115	270	25	909
	% within centre	21.2	33.7	12.7	29.7	2.8	100.0
Birmingham	Count	225	441	51	97	28	842
	% within centre	26.7	52.4	6.1	11.5	3.3	100.0
Dundee	Count	99	85	24	37	6	251
	% within centre	39.4	33.9	9.6	14.7	2.4	100.0
Edinburgh	Count	42	68	23	83	2	218
	% within centre	19.3	31.2	10.6	38.1	0.9	100.0
Liverpool	Count	200	145	15	50	6	416
	% within centre	48.1	34.9	3.6	12.0	1.4	100.0
Dunfermline	Count	54	76	7	15	7	159
	% within centre	34.0	47.8	4.4	9.4	4.4%	100.0
Oxford	Count	76	154	31	95	16	372
	% within centre	20.4	41.4	8.3	25.5	4.3	100.0
Total	Count	889	1275	266	647	90	3167
	% within centre	28.1	40.3	8.4	20.4	2.8	100.0
Right eye							
Aberdeen	Count	239	314	92	244	20	909
	% within centre	26.3	34.5	10.1	26.8	2.2	100.0
Birmingham	Count	206	450	53	103	29	841
	% within centre	24.5	53.5	6.3	12.2	3.4	100.0
Dundee	Count	94	88	28	38	6	254
	% within centre	37.0	34.6	11.0	15.0	2.4	100.0
Edinburgh	Count	35	88	30	62	3	218
	% within centre	16.1	40.4	13.8	28.4	1.4	100.0
Liverpool	Count	177	153	20	50	15	415
	% within centre	42.7	36.9	4.8	12.0	3.6	100.0

continued

© Queen's Printer and Controller of HMSO 2013. This work was produced by Olson *et al*. under the terms of a commissioning contract issued by the Secretary of State for Health. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

TABLE 41 Exclusive groups of features for comparison are (1) M/DH only (no BHs or exudates); (2) BH, but no exudates (BH only or BH+M/DH only); and (3) exudates (regardless of what else) (*continued*)

Centre		No lesions within one DD	M/DH only within one DD	BH or BH + M/DH within one DD	Exudate + any within one DD	Other within one DD	Total
Dunfermline	Count	54	72	9	19	5	159
	% within centre	34.0	45.3	5.7	11.9	3.1	100.0
Oxford	Count	80	155	45	82	10	372
	% within centre	21.5	41.7	12.1	22.0	2.7	100.0
Total	Count	885	1320	277	598	88	3168
	% within centre	27.9	41.7	8.7	18.9	2.8	100.0

TABLE 42 Relationship of individual lesions to right eye MO (OR; analysis unweighted)

Right eye												
Feature category (mutually exclusive)	Unadjusted OR[a]	Lower 95% CI	Upper 95% CI	Adjusted OR[b]	Lower 95% CI	Upper 95% CI	Adjusted OR[c]	Lower 95% CI	Upper 95% CI	Adjusted OR[d]	Lower 95% CI	Upper 95% CI
M/DH only	1											
BH not exudate	5.2	3.1	8.8	4.8	2.9	8.1	4.4	2.6	7.5	3.8	2.2	6.6
Exudate + any	5.2	3.4	8.1	4.9	3.1	7.6	6.2	3.9	9.7	6.0	3.7	9.6
Count of lesions of each type												
M/DH	1.23	1.19	1.27	1.16	1.12	1.21	1.18	1.13	1.22	1.19	1.14	1.24
BH	2.71	2.25	3.25	2.03	1.67	2.46	2.00	1.63	2.45	1.83	1.47	2.27
Exudate	1.15	1.12	1.18	1.10	1.07	1.14	1.12	1.09	1.16	1.12	1.09	1.16
Exudate within one to two DD	1.06	1.04	1.08									

a Unadjusted results.
b Adjusting for centre.
c Adjusting for centre, gender, age, glitazone and diabetes (type 1, type 2 and secondary/unknown).
d Adjusting for centre, gender, age, glitazone, diabetes (type 1, type 2 and secondary/unknown) and visual acuity.

© Queen's Printer and Controller of HMSO 2013. This work was produced by Olson *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

NIHR Journals Library www.journalslibrary.nihr.ac.uk

TABLE 43 Relationship of individual lesions within one DD to right eye MO (OR; analysis unweighted)

Right eye									
Feature category within a DD (mutually exclusive)	Unadjusted OR[a]	Lower 95% CI	Upper 95% CI	Adjusted OR[b]	Lower 95% CI	Upper 95% CI	Adjusted OR[c]	Lower 95% CI	Upper 95% CI
Zero M/DHs	1.00			1.00			1.00		
One M/DH	1.42	0.70	2.91	1.22	0.57	2.59	1.35	0.63	2.88
Two M/DHs	1.28	0.55	2.99	1.20	0.50	2.89	1.46	0.60	3.55
More than two M/DHs	7.38	4.37	12.46	4.15	2.32	7.42	5.07	2.80	9.18
Zero BHs	1.00			1.00			1.00		
One BH	2.88	1.77	4.71	2.09	1.24	3.55	1.90	1.11	3.25
Two BHs	11.43	6.00	21.77	4.24	2.01	8.95	3.75	1.73	8.15
More than two BHs	40.65	21.13	78.20	15.34	7.14	32.94	16.18	7.28	35.94
Zero exudates	1.00			1.00			1.00		
One exudate	0.94	0.29	3.02	0.86	0.26	2.87	0.98	0.29	3.32
Two exudates	2.50	1.12	5.58	1.89	0.75	4.75	2.08	0.80	5.38
More than two exudates	6.77	4.67	9.81	4.34	2.82	6.69	5.73	3.62	9.06
Visual acuity better[d]	1.00			1.00			1.00		
Visual acuity worse[d]	6.66	4.62	9.60	5.20	3.39	7.98	4.06	2.60	6.33
Visual acuity unknown/missing	13.18	4.10	42.35	28.56	6.15	132.77	20.12	4.03	100.42

a Unadjusted results.
b Adjusting for centre, other features and visual acuity.
c Adjusting for centre, gender, age, glitazone, diabetes (type 1, type 2 and secondary/unknown) and visual acuity. There were too few with amblyopia to adjust for this.
d Visual acuity better (log-MAR < 0.3; Snellen 6/9.5 or better), visual acuity worse (log-MAR ≥ 0.3; Snellen 6/12 or worse).

TABLE 44 Relationship of individual lesions to right eye MO (OR; analysis weighted)

Right eye												
Feature category (mutually exclusive)	**Unadjusted OR[a]**	**Lower 95% CI**	**Upper 95% CI**	**Adjusted OR[b]**	**Lower 95% CI**	**Upper 95% CI**	**Adjusted OR[c]**	**Lower 95% CI**	**Upper 95% CI**	**Adjusted OR[d]**	**Lower 95% CI**	**Upper 95% CI**
M/DH	1.0			1.0			1.0			1.0		
BH not exudate	5.7	2.9	11.3	5.1	2.5	10.1	4.7	2.3	9.5	4.2	2.0	8.6
Exudate + any	5.6	3.5	8.8	5.1	3.2	8.2	6.3	3.9	10.3	6.1	3.7	10.1
Count of lesions of each type												
M/DH	1.27	1.22	1.32	1.22	1.17	1.28	1.23	1.17	1.29	1.23	1.17	1.29
BH	3.13	2.42	4.05	2.06	1.56	2.72	2.04	1.53	2.72	1.89	1.40	2.55
Exudate	1.17	1.13	1.22	1.10	1.06	1.15	1.12	1.07	1.17	1.12	1.07	1.17
Exudate with one to two DD[e]	1.07	1.05	1.09									

a Unadjusted results.
b Adjusting for centre.
c Adjusting for centre, and the other features.
d Adjusting for centre, gender, age, glitazone, diabetes (type 1, type 2 and secondary/unknown) and visual acuity. There were too few with amblyopia to adjust for this.
e Counts of exudates between one and two DD is non-significant after adjusting for the presence of exudates within one DD and so is dropped from subsequent models.

© Queen's Printer and Controller of HMSO 2013. This work was produced by Olson *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

TABLE 45 Relationship of individual lesions within one DD to right eye MO (OR; analysis weighted)

Right eye									
Feature category (mutually exclusive)	Unadjusted OR[a]	Lower 95% CI	Upper 95% CI	Adjusted OR[b]	Lower 95% CI	Upper 95% CI	Adjusted OR[c]	Lower 95% CI	Upper 95% CI
Zero M/DHs	1.00			1.00			1.00		
One M/DH	1.75	0.74	4.13	1.72	0.71	4.16	1.80	0.74	4.39
Two M/DHs	2.22	0.89	5.53	2.38	0.93	6.10	2.58	1.00	6.64
More than two M/DH	7.68	3.74	15.77	5.18	2.42	11.09	5.91	2.74	12.73
Zero BHs	1.00			1.00			1.00		
One BH	3.65	1.81	7.35	2.02	0.94	4.33	1.80	0.83	3.91
Two BHs	14.71	5.72	37.86	4.13	1.39	12.23	3.58	1.15	11.10
More than two BH	60.61	24.12	152.30	18.99	6.50	55.49	21.33	7.02	64.82
Zero exudates	1.00			1.00			1.00		
One exudate	1.24	0.27	5.77	0.87	0.18	4.25	0.99	0.20	4.92
Two exudates	3.31	1.16	9.43	1.98	0.60	6.53	2.21	0.65	7.52
More than two exudates	8.96	5.68	14.15	4.65	2.73	7.94	5.99	3.42	10.47
Visual acuity better[d]	1.00			1.00			1.00		
Visual acuity worse[d]	5.85	3.77	9.09	4.19	2.53	6.94	3.28	1.94	5.54
Visual acuity unknown/missing	6.50	1.29	32.77	12.02	1.32	109.53	8.18	0.78	85.39

a Unadjusted results.
b Adjusting for centre and visual acuity.
c Adjusting for centre, gender, age, glitazone, diabetes (type 1, type 2 and secondary/unknown) and visual acuity. There were too few with amblyopia to adjust for this.
d Visual acuity better (log-MAR < 0.3; Snellen 6/9.5 or better), visual acuity worse (log-MAR ≥ 0.3; Snellen 6/12 or worse).

Appendix 4 Protocol

Version 2.1 (4th June 2009).

1. Study summary

Full title
Improving the value of screening for diabetic macular oedema using surrogate photographic markers.

Short title
Screening for diabetic MO using surrogate photographic markers.

Official website
http://www.abdn.ac.uk/ismo/

Funding
£432,174 from the National Institute for Health Research (NIHR) Health Technology Assessment (HTA) programme. Grant reference 06/402/49.

Ethical approval
Main REC: North of Scotland Research Ethics Committee.

REC ref: 07/S0801/107.

Date of approval: 17/12/07.

Important dates
Start date: 1st May 2008.

50% target recruitment point: 1st May 2009.

100% target recruitment point: 31st October 2009.

End date: 30th April 2010.

Draft final report due: 14th May 2010.

Collaborators
Aberdeen University School of Medicine, NHS Grampian, NHS Tayside, NHS Lothian and Borders, Royal Liverpool and Broadgreen University Hospital NHS Trust, Queen Margaret Hospital, Dunfermline, Oxford Radcliffe Hospitals NHS Trust, Heart of England NHS Trust, Norfolk and Norwich University Hospitals NHS Trust.

Study objectives
The purpose of this study is to determine the best method for detecting sight-threatening macular oedema using photographic surrogate markers for people with diabetes in the context of national screening programmes.

© Queen's Printer and Controller of HMSO 2013. This work was produced by Olson *et al*. under the terms of a commissioning contract issued by the Secretary of State for Health. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

Background

Macular oedema is associated with several conditions which cause irreversible vision loss, including diabetic retinopathy. It may be classified according to the fluid distribution: diffuse oedema is a general thickening of the central retina caused by either extensive capillary dilation or capillary closure, while focal oedema is centred on specific vascular abnormalities, such as microaneurysms. Accumulated fluid defocuses the image on the retina, reducing visual acuity. If oedema persists, the increased pressure may lead to irreparable photoreceptor damage or retinal detachment.

Since retinal thickening is not visible directly on the retinal photographs used by screening programmes, people are referred to ophthalmology clinics on the basis of a range of surrogate photographic markers, such as exudates within a certain distance of the foveal centre. Evidence from the Early Treatment of Diabetic Retinopathy Study suggests that exudates may be a sensitive marker of macular oedema (Bresnick 2000). However, as the ETDRS excluded patients with mild retinopathy in the absence of exudate, the results are not applicable to screening programmes in the United Kingdom, where 60% of patients have no visible signs of retinopathy and over 30% have only mild retinopathy. Evidence from the Grampian Retinal Screening Programme suggests that only 12% of patients with surrogate markers referred to an ophthalmologist have indications of macular oedema when examined by slit-lamp biomicroscopy. Similarly, a retrospective analysis from Liverpool, including 257 patients referred from the screening programme to the ophthalmology clinic between December 2001 and June 2002, found that only 14% had evidence of macular oedema (unpublished data).

Macular oedema has traditionally been assessed clinically using a combination of slit-lamp biomicroscopy, stereo photography and stereo fluorescein angiography. However, these techniques have a number of limitations. Foremost is that they are only qualitative assessments, which are relatively insensitive to thickness changes. Furthermore, slit-lamp examination does not provide a pictorial record and, together with stereo photography, is known to be biased by the presence or absence of exudates. Although the angiogram is a sensitive test for leakage, the assessment of thickening is very subjective. Best corrected visual acuity has also been used as a surrogate indication of thickening, but is neither sensitive nor specific, being affected by several factors besides macular thickness.

Study design

A total of 4000 patients with photographic signs of maculopathy (exudates within two disc diameter radius, blots or dot haemorrhages within one disc diameter radius) shall be recruited from the seven study centres. Each subject will have photography and optical coherence tomography on both eyes where possible. 10% of patients are expected to have ungradeable images and will be unsuitable for the study.

All relevant lesions visible in the colour photographs will be annotated by the research nurse in Aberdeen using computer-assisted software. The optical coherence tomography images will be analysed both quantitatively and qualitatively by the research nurse. Software will be developed to analyse the distribution of retinal lesions in order to find the significance of the photographic marker patterns on the likelihood of clinically significant macular oedema.

Proposed outcome measures

A sample of patients attending diabetic retinopathy-screening programmes will receive an optical coherence tomography examination in addition to the standard digital photograph. An expert grader will assess all digital images independently for the presence of different surrogate photographic markers. The sensitivity and specificity of standard referral criteria (based on the presence of surrogate markers) will be assessed using optical coherence tomography as the reference standard.

Although manual grading will always have a role in the assessment of retinal images, the scale of the current diabetes epidemic means that automation will have to play a key role if the Liverpool Declaration's target of offering annual systematic screening to at least 80% of the estimated 35 million people with diabetes in Europe is to be achieved. Furthermore, given the low specificity of current referral criteria, it is important to investigate whether specificity can be improved using more complex patterns of surrogate markers, using either computer-assisted annotation or fully automated computer analysis, so that scarce ophthalmology resources are used most effectively.

Derived sensitivity/specificity estimates will be used to assess the costs and consequences (i.e. the number of appropriate/inappropriate ophthalmology referrals) of using alternative surrogate markers, and patterns of surrogate markers, for the detection of macular oedema, using manual, computer-assisted annotation and fully automated detection systems.

It is important to look at the implications for health care delivery of introducing this technology into systematic screening programmes for diabetic retinopathy. We will therefore model long-term costs and outcomes (visual loss and quality adjusted life years) of the alternative screening strategies using epidemiological literature and available cost estimates.

Sample size calculation

If 33333 patients are screened across five centres, 4000 patients would be expected to have surrogate markers (12%) of whom 400 would be expected to have macular oedema. If there are 4000 patients with any markers, the power for detecting a change in the referral specificity of at least 20% will be much higher than 80%.

However, a new diagnostic test (using a more specific combination of features) which has increased specificity may also have decreased the sensitivity. Therefore, since there are fewer true cases with macula oedema than true controls, the crucial issue for the power of this study is whether or not a small change in sensitivity can be detected.

With 400 cases of macula oedema there will be 80% power to detect a difference in sensitivity of 3% (99% vs. 96%) between two diagnostic tests (any markers vs. a specific combination of markers) when the percentage of true cases where the diagnostic tests disagree is expected to be 5%. A McNemar's test of equality of paired proportions has been used with a 0.05 significance level.

Research governance

Research activities at each of the participating centres will be carried out in accordance with the Department of Health's Research Governance Framework for Health and Social Care. The project will be registered with the Institute of Applied Health Sciences (University of Aberdeen) and each appropriate NHS trust. Copies of the protocol and records of all important documents will be cross referenced and filed so that they are available at any given time. Records relating to all procedures carried out will be kept so that the research process is clearly understandable and repeatable. NHS Grampian has agreed to act as the sponsors for the study.

A Trial Steering Committee has been set up. The chair, Dr Caroline Styles, is a consultant ophthalmologist in Fife with experience of retinal screening (she was later replaced by Dr Rod Harvey). Professor Alex Elliot, a medical physicist from Glasgow, brings medical image processing experience. Mr Steve Graham is the committee patient representative. Ms Alison Farrow is a retinal photographer in Aberdeen (she was later replaced by Dr A Manivannan). Finally Dr John Olson and Professor Peter Sharp are also members.

© Queen's Printer and Controller of HMSO 2013. This work was produced by Olson *et al*. under the terms of a commissioning contract issued by the Secretary of State for Health. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

2. Centre contact details

Centre	Principal investigator(s)	Other contacts
University of Aberdeen and NHS Grampian	Dr John Olson	Dr Keith Goatman
Heart of England NHS Trust	Professor Paul Dodson	Ms Jane Pitt
NHS Tayside	Professor Graham Leese Dr John Ellis	
NHS Lothian	Dr Ken Swa	Dr Shyamanga Borooah
Royal Liverpool & Broadgreen University Hospitals Trust	Professor Simon Harding Professor Deborah Broadbent	Dr Yalin Zheng
Oxford Radcliffe Hospitals NHS Trust	Professor Victor Chong	Ms Sue Beatty
Queen Margaret Hospital, Dunfermline	Dr Caroline Styles	

Dr John Olson is the project Chief Investigator (john.olson@nhs.net). For general enquiries about the study contact Dr Keith Goatman (k.a.goatman@abdn.ac.uk).

Dr Caroline Styles (caroline.styles@nhs.net) is the chair of the Trial Steering Committee (later replaced by Dr Rod Harvey).

3. Patient selection and recruitment

Each centre shall collect data from 800 patients fulfilling the inclusion criteria below.

Inclusion Criteria

Age 18 or older.

A trained grader confirms that the fundus photograph of at least one eye shows diabetic eye disease including any of:

Microaneurysms/dot haemorrhages within one disc diameter radius of the macula.

Blot haemorrhages within one disc diameter radius of the macula.

Exudates within two disc diameter radius of the macula.

Able and willing to provide signed informed consent.

Exclusion Criteria

The patient has had macular or pan-retinal laser treatment [clarification to protocol added version 2.0].

The patient has had an intraocular injection [added in protocol version 2.0].

The patient is pregnant (since oedema may be present independent of gestational diabetes) [Added in protocol version 2.0].

Intraocular surgery within one year of enrolment.

The screening fundus photograph is a technical failure (i.e. the macula vessels are not clearly visible and/or the field of view does not include a region of diameter two disc diameter radius s about the centre of the macula).

Contraindications to pupillary dilation/intolerance or hypersensitivity to mydriatics in either eye if dilation is necessitated.

How patients are recruited will depend on the local set-up at each centre. In Aberdeen patients with exudates or blots within 1 disc diameter radius (grade code M2) are routinely referred from the screening service to see an ophthalmologist. These patients are now routinely offered optical coherence tomography scans and therefore only need to be consented to allow their data to be used in the study. Patients with exudates within 1 disc diameter radius and 2 disc diameter radius (grade code M1) will need to be invited for an optical coherence tomography scan they would not normally receive, so will need to be consented both for the imaging and use of data. Finally, patients with microaneurysms/dot haemorrhages within 1 disc diameter radius do not trigger a specific grade in the Scottish grading system and will have to be manually chosen during screening grading. As for the M1 category, these patients will have to be invited for an optical coherence tomography they would not normally receive, so will need to be consented both for imaging and use of their data.

4. Retinal photography, visual acuity and Glitazone

Visual acuity

Log-MAR (or Snellen if Log-MAR not available) pin-hole or best-corrected visual acuity measurement for both eyes. [Best-corrected added version 2.1]

Record if the patient has a long-standing, non-diabetic reason for a low visual acuity measurement, for instance due to amblyopia. This must be recorded on the web upload form. [Added version 2.1]

Photography

Single views of both eyes:

Approximately 45 degree field of view.

Approximately macula centred.

Colour digital photograph, at least 3 megapixels (ideally less than 7 megapixels and, if using JPEG compression, set for highest quality).

Adequate field of view (should show 2 disc diameter radius around centre of macula).

Adequate clarity (i.e. adequate to see macular microaneurysms, if present). If there are problems obtaining a retinal photograph it is very likely to be difficult obtaining an adequate optical coherence tomography scan.

Mydriasis only necessary if pupil size too small for imaging. Camera small pupil facility acceptable providing 2 disc diameter radius visible around macula centre.

Note that the maximum allowed time between photography and optical coherence tomography is four weeks to reduce errors due to lesion changes between photograph and optical coherence tomography.

Example images showing adequate and inadequate clarity and field of view are available from the website. (http://www.abdn.ac.uk/ismo/).

Glitazone [Added version 2.1]

The glitazone class of drugs appears to be associated with diabetic macular oedema [1]. If the patient has used any of the following drugs within the past six months this must be recorded on the web upload form:

© Queen's Printer and Controller of HMSO 2013. This work was produced by Olson *et al*. under the terms of a commissioning contract issued by the Secretary of State for Health. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

Rosiglitazone: Avandia, Avandamet.

Pioglitazone: Actos Competact.

[1] Fong DS and Contreras R. Glitazone use associated with diabetic macular edema. *American Journal of Ophthalmology* 2009;**147**:583–586.

5. Optical Coherence Tomography

Optical coherence tomography scanner choice

When the project was first proposed every centre had use of the Zeiss Stratus optical coherence tomography scanner (often known as OCT3). Since the funding was granted a number of new instruments based on a spectral domain method have become available. Initial tests on the Zeiss Cirrus, Topcon OCT1000 and Heidelberg Spectralis show these new scanners have some advantages for this study. Hence centres will not be required to use the Stratus but instead encouraged to use the best available equipment.

In order to reduce variation due to different manufacturers, models and operators:

Only accredited operators shall acquire scans.

It is recommended that, where possible, the same scanner is used for the duration of the study. If more than one scanner is used the scanner used must be indicated. If a scanner must be replaced it may be replaced by any model.

Dates of services and faults should be recorded in the study folder.

Optical Coherence Tomography protocol

Two types of data are required from the optical coherence tomography scanner:

A thickness map divided into nine Early Treatment Diabetic Retinopathy Study regions centred on the centre of the macula and with a diameter of 6 mm.

Orthogonal B-mode 6-mm cross-sections centred on the centre of the macula.

Specific instructions for different optical coherence tomography models are given in appendix A to the protocol.

Accreditation process

As in other multicentre imaging studies, to avoid inter-centre imaging variation all optical coherence tomography operators are required to be accredited before submitting data for the study. Operators should submit a simple portfolio of images (described below) collected using the optical coherence tomography scanner intended for the study. Images will be checked for macula position, adequate image quality and differences between thicknesses in the nine Early Treatment Diabetic Retinopathy Study regions of the repeat scans. The accreditation images are as follows:

(a) Normal eyes

Repeat macula maps of the same normal eye as per scanner model protocol above.

(b) Macular oedema

Repeat macula maps of the same eye showing obvious macular oedema (i.e. central thickness at least 300 microns).

For scanners (such as the Stratus) which do not automatically generate high-resolution orthogonal cross-sections ("cross hair scans") also perform the cross-hair scan.

Accreditation Upload procedure

Go to the study website: http://www.abdn.ac.uk/ismo/

Login using your supplied username and password. The admin functions will appear once you are logged in. Select "Accreditation" from the menu and follow the instructions for uploading images.

6. Data transfer

Photographs and optical coherence tomography scans are transferred to Aberdeen using the web-based upload service at:

http://www.abdn.ac.uk/ismo/

Log in using your supplied username and password and the admin functions will appear. Select "upload data" and follow the on-screen instructions.

Note: the server automatically generates an anonymous subject ID. If you enter the subject name and hospital ID number these will print a front sheet for your study folder, but will not be transmitted to Aberdeen.

Data protection and personal information

The personal patient information you enter is not transferred to the webserver, it is just used to produce the print out.

However, patient information may appear on the photographs or optical coherence tomography images which are transferred to Aberdeen. Although the University of Aberdeen webserver has been registered to allow the storage of personal information, personal information is not required for this study we do not wish to keep it, and have not sought ethical permission to do so. **Therefore <u>all</u> personal information will be automatically removed from images that are uploaded to Aberdeen.** It is therefore vital that if centres wish to be able to cross reference data from their own clinical records that they keep a local copy of the anonymous study ID and real patient ID.

Image grading

(a) Colour photographs

All fundus photographs will be assessed in Aberdeen by a trained reference grader. An experienced grader will train the reference grader in the identification of the features of diabetic retinopathy and in the use of the software. The trainer will quality assure the annotations of 100 images as part of the training phase. The following features will be noted:

Presence of microaneurysms/dot haemorrhages and blots within 1 disc diameter radius.

Presence of exudates within 1 disc diameter radius and 2 disc diameter radius.

All the above lesions will be annotated on images, using computer-assisted annotation software, to enable pattern analysis of the lesion distributions.

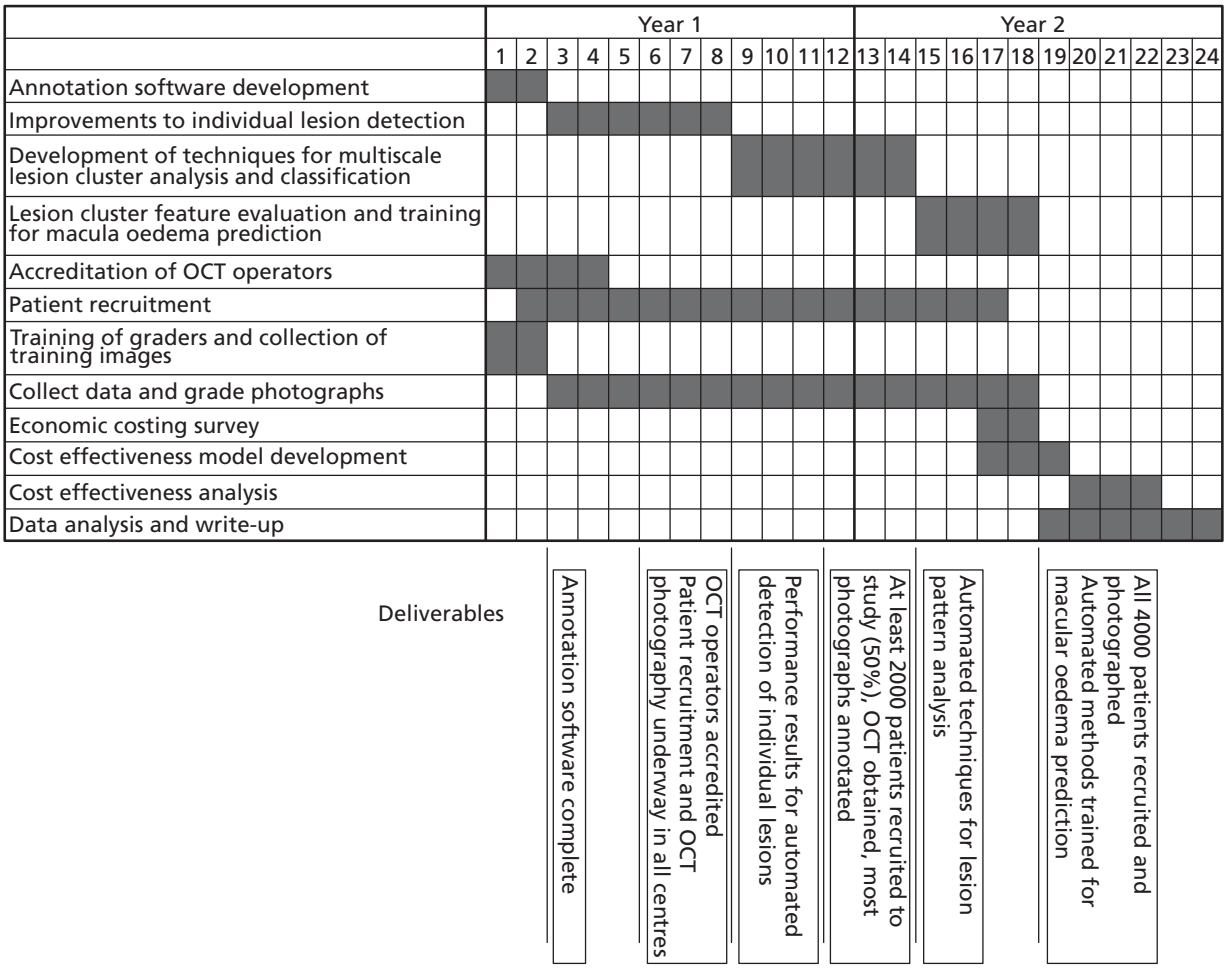
© Queen's Printer and Controller of HMSO 2013. This work was produced by Olson *et al*. under the terms of a commissioning contract issued by the Secretary of State for Health. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

(b) Optical Coherence Tomography data

All the optical coherence tomography images will be graded quantitatively and qualitatively for area, amount and site of retinal thickening based on the generated thickness maps. Optical coherence tomography studies suggest that oedema may be reliably detected by slit-lamp biomicroscopy if it has a retinal thickness of at least 300 µm (Hee 1998). This agrees with the Early Treatment Diabetic Retinopathy Study, which used a reference thickness of 250 µm for "clinically significant" macular oedema.

In this study, macular oedema will be defined as an optical coherence tomography retinal thickness of 300 µm or greater (relative to Zeiss Stratus measurements) in any map region within one disc diameter radius of the centre of the fovea (i.e. in any of the five central Early Treatment Diabetic Retinopathy Study map regions).

8. Project timetable



Appendix A: Specific Optical Coherence Tomography scanner instructions

(a) Zeiss Stratus Optical Coherence Tomography

Data acquisition

For each eye, use the fast macular map protocol to obtain six intersecting radial lines with a length of 6mm. Total acquisition time is approximately two seconds.

Scans will be considered acceptable if: (1) the standard deviation of the central foveal measurements from the six cross-sections is less than 10% of the mean central measurement, and (2) the recorded signal strength is at least 4, and (3) there are no warnings about low analysis confidence, missing data or high variance, and (4) there are no visible boundary tracking errors.

Data export

Two pieces of evidence are required to confirm the presence of macular oedema: (1) the thickness map giving the Early Treatment Diabetic Retinopathy Study region thicknesses in microns and (2) a cross-sectional image showing the thickening is due to dark, fluid-filled spaces.

Thickness maps: Generate thickness maps by selecting the best scans for the left and right eyes and clicking on the "analysis" tab and selecting "Retinal Thickness/Volume (OU)". Export the result as a PDF document using the "Print to PDF file" option. An example is shown in figure Stratus 1.

Does the horizontal cross-section image included on the thickness map (a) pass through the region of greatest thickening and (b) clearly show the thickening? If not, export an additional cross-sectional image for one or both eyes as below:

Cross-sectional image(s): The PDF report includes a low-resolution/low-quality image of the horizontal cross-section for each eye. If one or both images either does not pass through the region of thickening, or is insufficient quality to show the oedema, a higher quality image of the cross-section which best shows the thickening should be exported. Select one eye at a time. Click on the analysis tab and click on the "Scan Selection" button at the bottom of the analysis section. The six cross-sections are shown on the left of the screen (see figure Stratus 2). Select the one which best shows the thickening and click on the "Export JPEG" button to save the cross-section as a JPEG image (see figure Stratus 3). Repeat for the other eye if necessary.

(b) Zeiss Cirrus Optical Coherence Tomography

Data acquisition

The Cirrus has two macular mapping protocols. The first (512 × 128) acquires 128 horizontal cross-sections, each containing 512 A-scans. The second (200 × 200) acquires 200 horizontal cross-sections, each containing 200 A-scans. We recommend use of the 200 × 200 protocol as it produces equally high-quality maps in a slightly shorter time; less time for patient movement, increasing the probability of obtaining a successful scan. High-resolution orthogonal cross-sections are also acquired automatically during the scan.

Scans will be considered acceptable if: (1) the signal strength is at least 5/10, and (2) there are no obvious signs of movement, and (3) there are no visible boundary tracing errors, and (4) the fixation point is within 250 microns of the centre of the foveal pit (where visible).

Data export

Two pieces of evidence are required to confirm the presence of macular oedema: (1) the thickness map values and (2) dark fluid-filled spaces on the cross-sectional image(s). If the orthogonal cross-sections do

© Queen's Printer and Controller of HMSO 2013. This work was produced by Olson *et al*. under the terms of a commissioning contract issued by the Secretary of State for Health. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

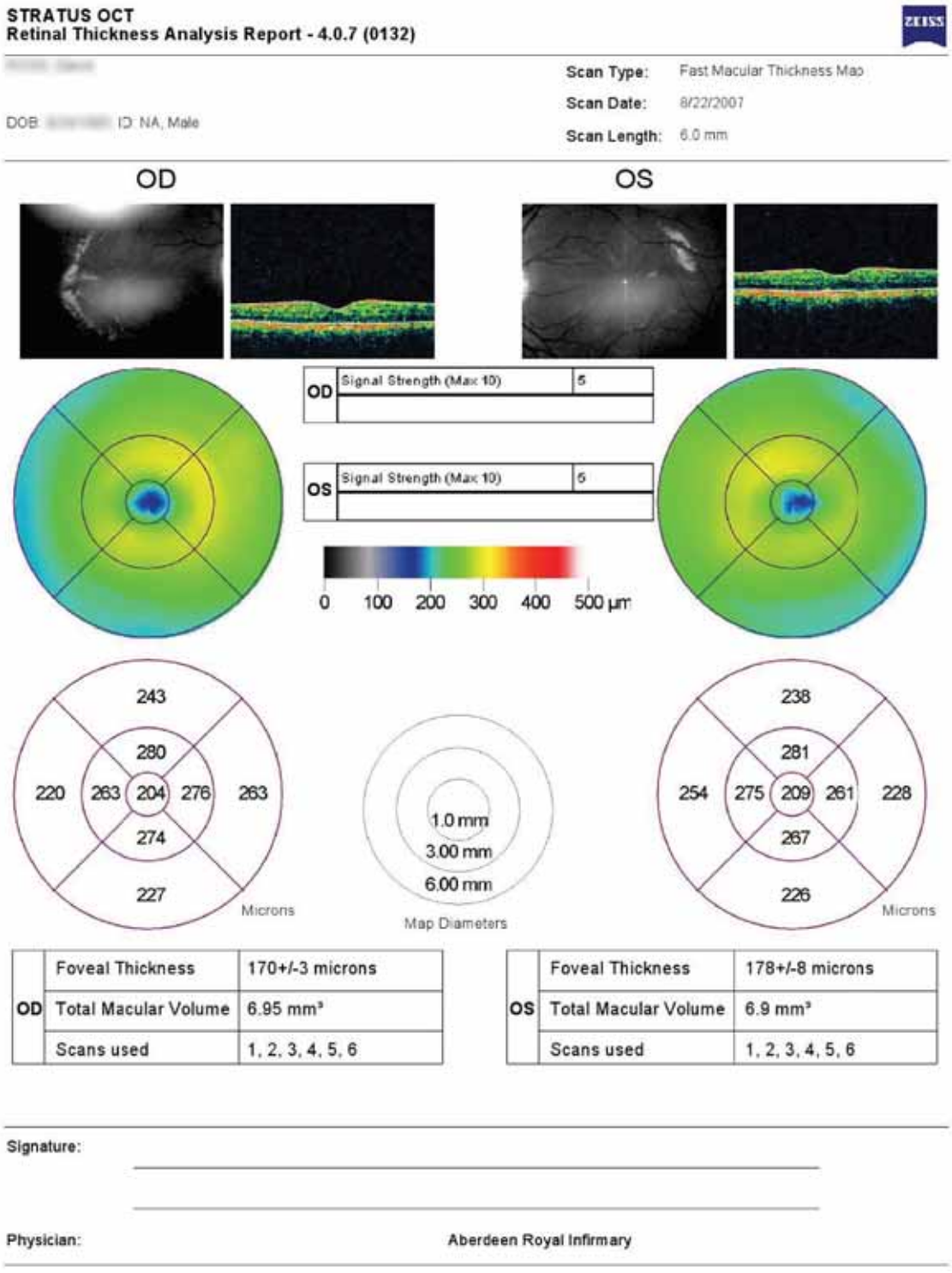


FIGURE Stratus 1 – example thickness map (PDF file).

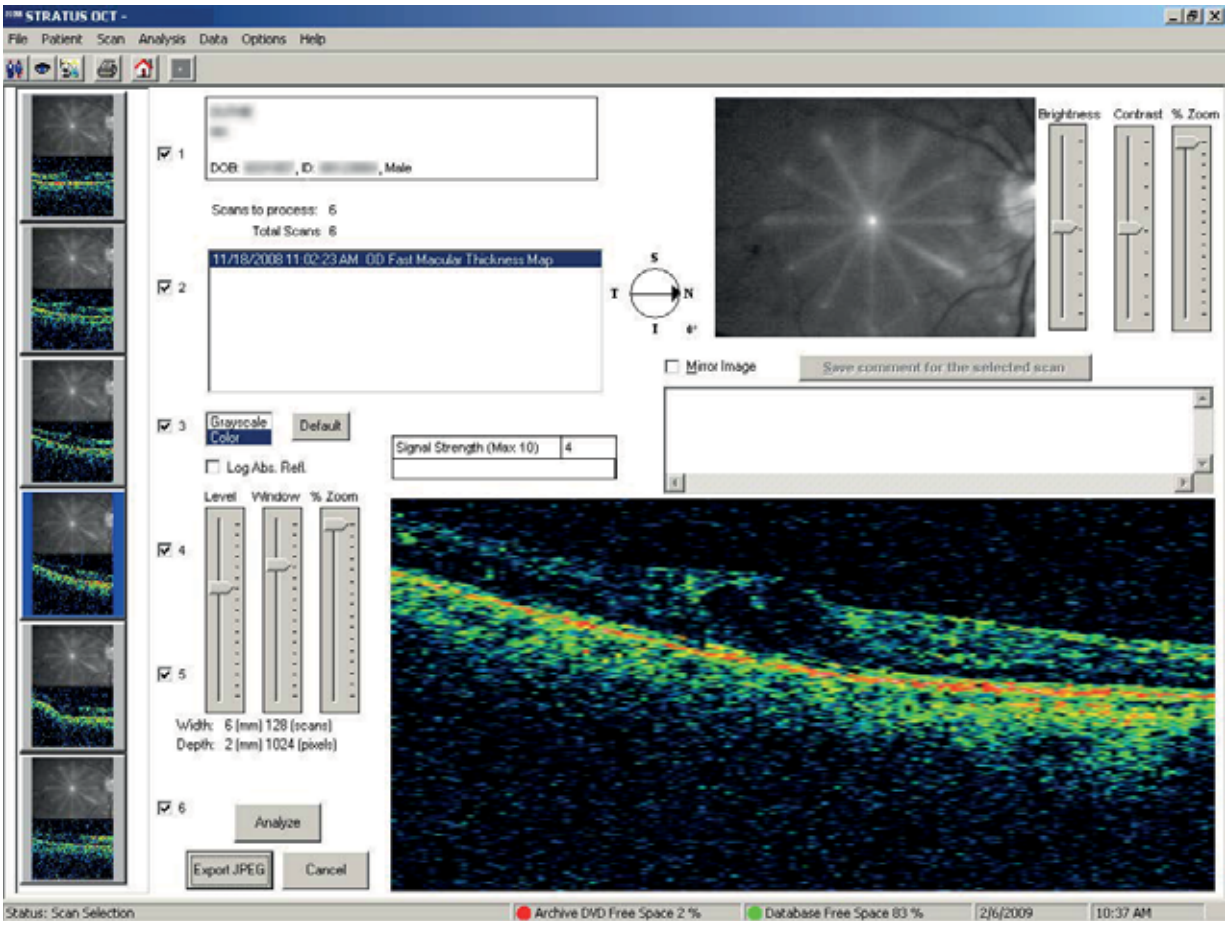


FIGURE Stratus 2 – Cross-section selection for export.

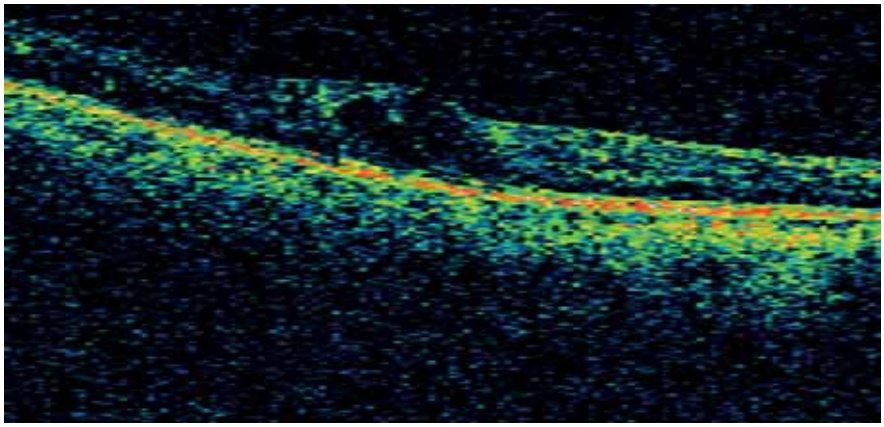


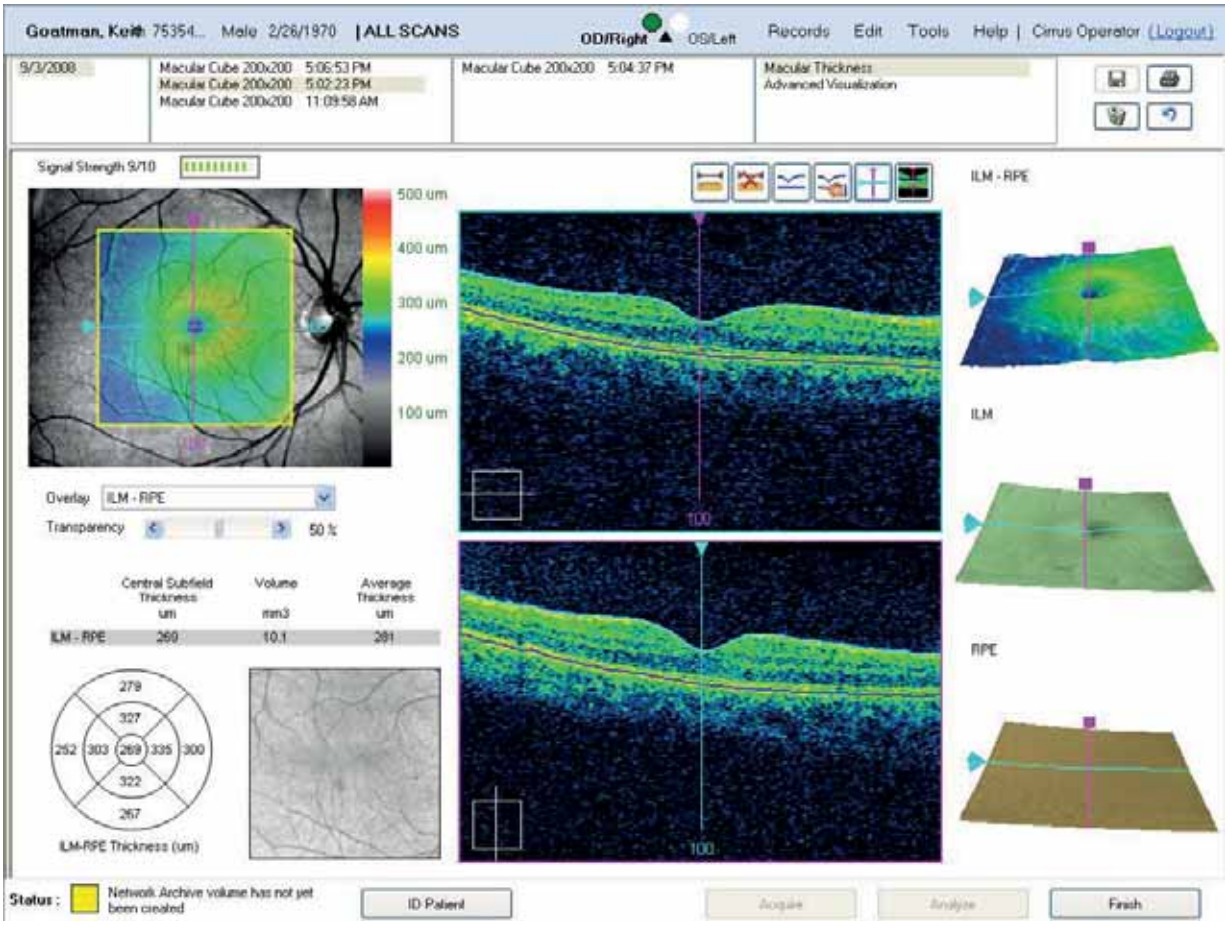
FIGURE Stratus 3 – Exported cross-section (JPEG file).

not intersect the region of thickening then a second image is required with the cross hairs moved to pass through the thickened region.

To export the data select print and choose the option to save the printout as a TIFF image (note the quality of the PDF image is too poor for the study).

Upload the optical coherence tomography images and colour photographs using the web-upload system at: http://www.abdn.ac.uk/ismo/

© Queen's Printer and Controller of HMSO 2013. This work was produced by Olson *et al*. under the terms of a commissioning contract issued by the Secretary of State for Health. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.



(c) Topcon OCT-1000 Mark I

The Topcon OCT-1000 is essentially a non-mydriatic fundus camera with an optical coherence tomography spectrometer attached. As well as the spectral domain optical coherence tomography data it also acquires 3MP colour photographs which are ideal for use in this study.

Important note about volume scaling

The OCT-1000 has an option to change the relative horizontal and vertical scale of the B-mode slice display. Topcon call this the "volume scale", and it has options of 1 : 1, 1 : 2 or 1 : 3. You do not need to change this setting for the study but you should inform Aberdeen of your preferred local setting. The value may be viewed/changed by selecting Tools from the main menu bar, Options => Volume scale.

Data capture

Use the "3D scan" acquisition mode.

Select the 6 × 6 mm rectangular scan area and 512 × 128 lines (128 lines of 512 A-scans per line).

Use the camera macular-centred (M) fixation point.

Once an acceptable scan has been acquired save the data.

Colour photograph export

The colour photograph is exported by selecting "Export" from the main menu bar and selecting "Fullsize fundus". Select PNG filetype (the jpeg quality is too poor for the study) and choose a filename, ideally including the anonymous patient ID (e.g. for Aberdeen A0123_photo_left.png).

Optical Coherence Tomography data export

Three pieces of information are required for the study:

The nine region Early Treatment Diabetic Retinopathy Study thickness map.

The central horizontal B-mode section (slice 64).

The horizontal B-mode section which best includes the severest thickening within the Early Treatment Diabetic Retinopathy Study region, if different from (b) above.

These are obtained from the captured data as follows:

Analyse data: Make sure the raw study data has been analysed (this estimates the region boundaries on the 128 B-mode sections).

The default view shows B-mode slice 64 on the left hand side and the colour photograph on the right.

Display Early Treatment Diabetic Retinopathy Study grid: Click on the Early Treatment Diabetic Retinopathy Study icon by the colour photograph to overlay thickness values on the colour photograph (see figure Topcon example 1).

Optional reposition: In images where the foveal pit is obvious, place a cursor on the deepest part of the pit and check that this corresponds to the centre of the Early Treatment Diabetic Retinopathy Study grid. If the difference is more than approximately 200 microns (for reference the central Early Treatment Diabetic Retinopathy Study region is 1000 microns diameter) reposition the Early Treatment Diabetic Retinopathy Study grid using the "reposition" option in the grid menu to the left of the colour photograph. This is done by dragging the grid while holding down the left mouse button. Note that moving the grid will mean part of one or more of the outer regions will be outside the captured data (indicated by the green box). If the fixation error is too large and the grid is moved too far then N/A will appear in place of a thickness measurement. This will be counted as a technical failure and should be repeated. However, the N/A designation appears too conservative. For example, figure Topcon example 1 shows a small corrective shift of approximately 130 microns. Figure Topcon example 2 shows a much larger correction shift of approximately 1000 microns, where 55% of the area of the outer inferior region is outside the collected optical coherence tomography image. This is not acceptable and consequently images requiring a corrective shift of more than 500 microns will be considered a technical failure and should be repeated.

Central section screen-shot: take a screenshot showing the central horizontal section (slice 64) and the colour photograph with the Early Treatment Diabetic Retinopathy Study grid overlaid. Select Export from the main menu bar followed by "Screenshot". Save the screenshot as a PNG format image, ideally using a filename convention which includes the patient anonymous ID. Figure Topcon 1 shows an example screenshot.

Maximum thickness screen-shot: if the central slice does not include the area of thickening, change the displayed section by scrolling the mouse wheel until the section with the thickening is displayed and repeat the screenshot as above. *If the central section adequately shows the thickening then this image is not required.*

© Queen's Printer and Controller of HMSO 2013. This work was produced by Olson *et al*. under the terms of a commissioning contract issued by the Secretary of State for Health. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

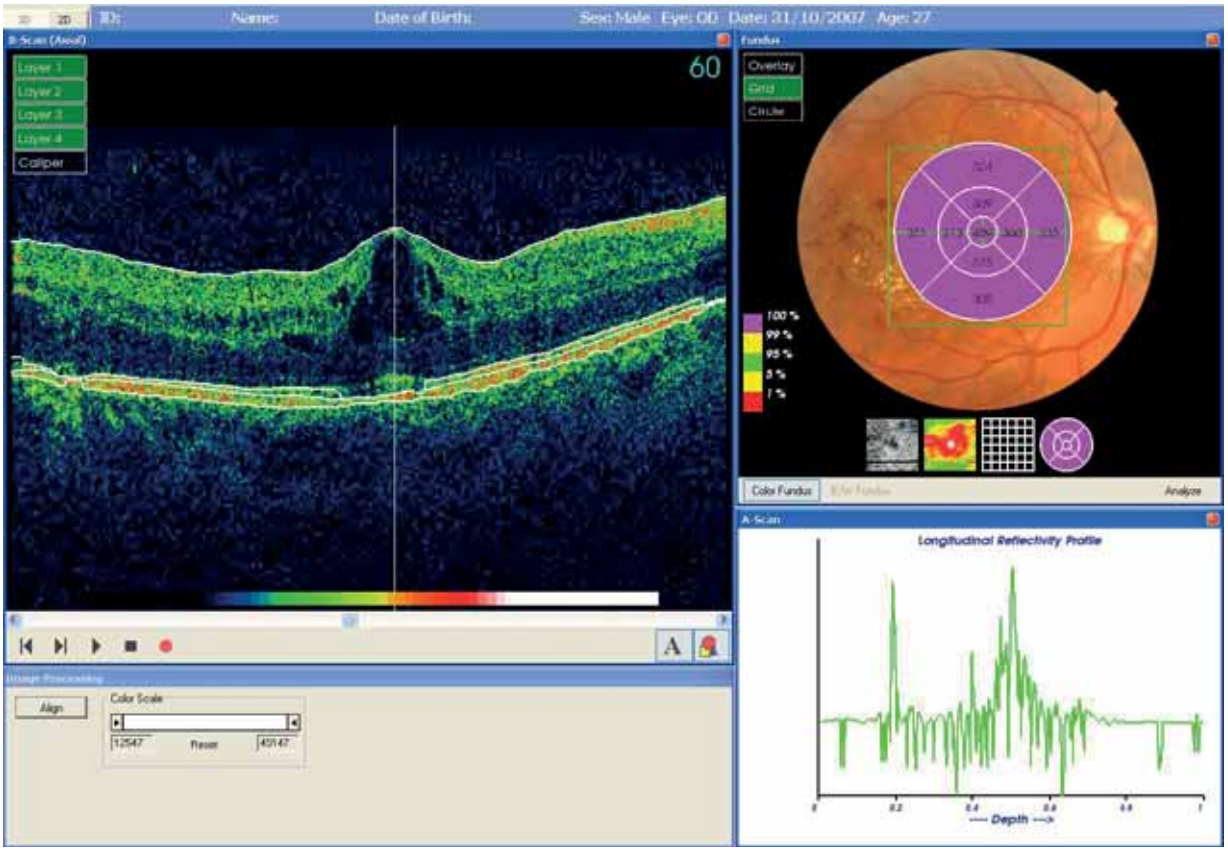


FIGURE Topcon example 1 showing B-mode section 60 on the left hand side and the colour photograph with Early Treatment Diabetic Retinopathy Study region thicknesses on the right.

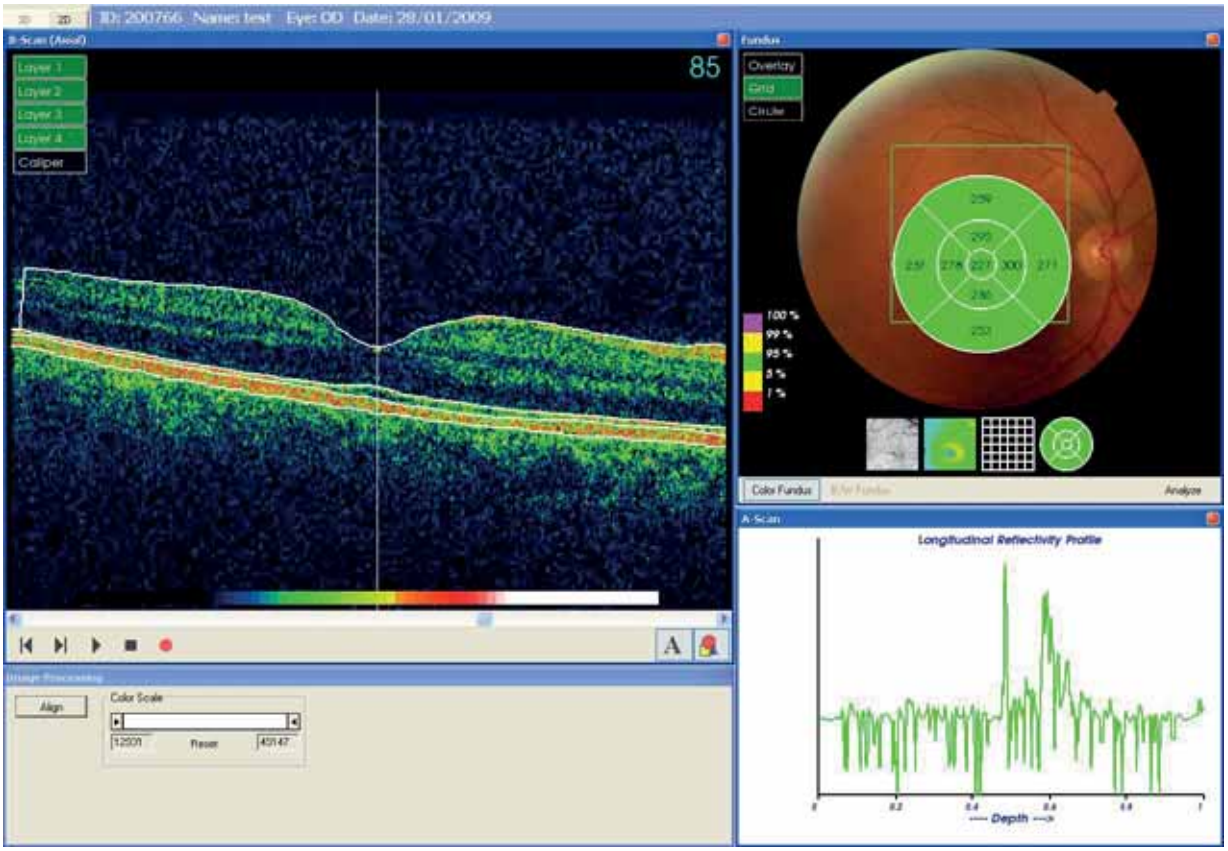


FIGURE Topcon example 2 showing an excessive reposition of the Early Treatment Diabetic Retinopathy Study grid.

For each eye there will be one colour photograph and one or two optical coherence tomography images to upload using the web-based system at http://www.abdn.ac.uk/ismo/

(d) Heidelberg Spectralis

The Heidelberg Spectralis is currently the fastest optical coherence tomography scanner on the market, acquiring 40,000 A-scan lines per second. It has two features not found on any other optical coherence tomography scanner: eye tracking and multiple lines acquisition to reduce noise (ART). It also has more user adjustable parameters than any other scanner.

The default macular thickness protocol acquires 19 horizontal sections with horizontal and vertical field of views of 20 and 15 degrees respectively. By default each line is acquired multiple times (with ART enabled). Note: This default protocol is not suitable for the study, as it does not cover the required 6 mm diameter region about the centre of the macula.

Data acquisition

The field of view should be centred on the macula.

Change the horizontal field of view (using the right cursor key) from 20 degrees to 30 degrees. The current field of view and number of sections is displayed at the bottom of the screen in acquisition mode.

Change the vertical field of view (using the UP cursor key) from 15 degrees to 25 degrees (which also increases the number of slides from 19 to 31).

The number of ART repeat scans can be set between 2 and 100. A setting of 3 is adequate for this study.

When the correct field of view is obtained lock the position using the large black circular button on the touch pad.

Signal strength should be greater than 15/40 (image quality bar blue rather than red).

Data analysis and export

Three pieces of information are required for the study:

The nine region Early Treatment Diabetic Retinopathy Study thickness map.

The horizontal B-mode section through the centre of the macula.

The horizontal B-mode section which includes the severest thickening within the Early Treatment Diabetic Retinopathy Study region. If this is the same slice as (b) above another image is not required.

(a) Nine region thickness map

For the nine region thickness map select concentric circle diameters of "1, 3, 6 mm Early Treatment Diabetic Retinopathy Study" (see figure Spectralis 2).

If the foveal pit is visible and there is an obvious fixation error the Early Treatment Diabetic Retinopathy Study grid may be moved using the mouse so as to be centred on the foveal pit. If the require shift is too large one or more of the outer regions will go outside the field of view and the thickness value will be blank. These are considered a technical failure and should be repeated.

Export the screenshot image in PNG format (Please do not use JPEG. Even using the highest quality setting the images are saved with the colour information at half the resolution of the brightness information, which blurs coloured features with sharp edges, such as the text.)

© Queen's Printer and Controller of HMSO 2013. This work was produced by Olson *et al*. under the terms of a commissioning contract issued by the Secretary of State for Health. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

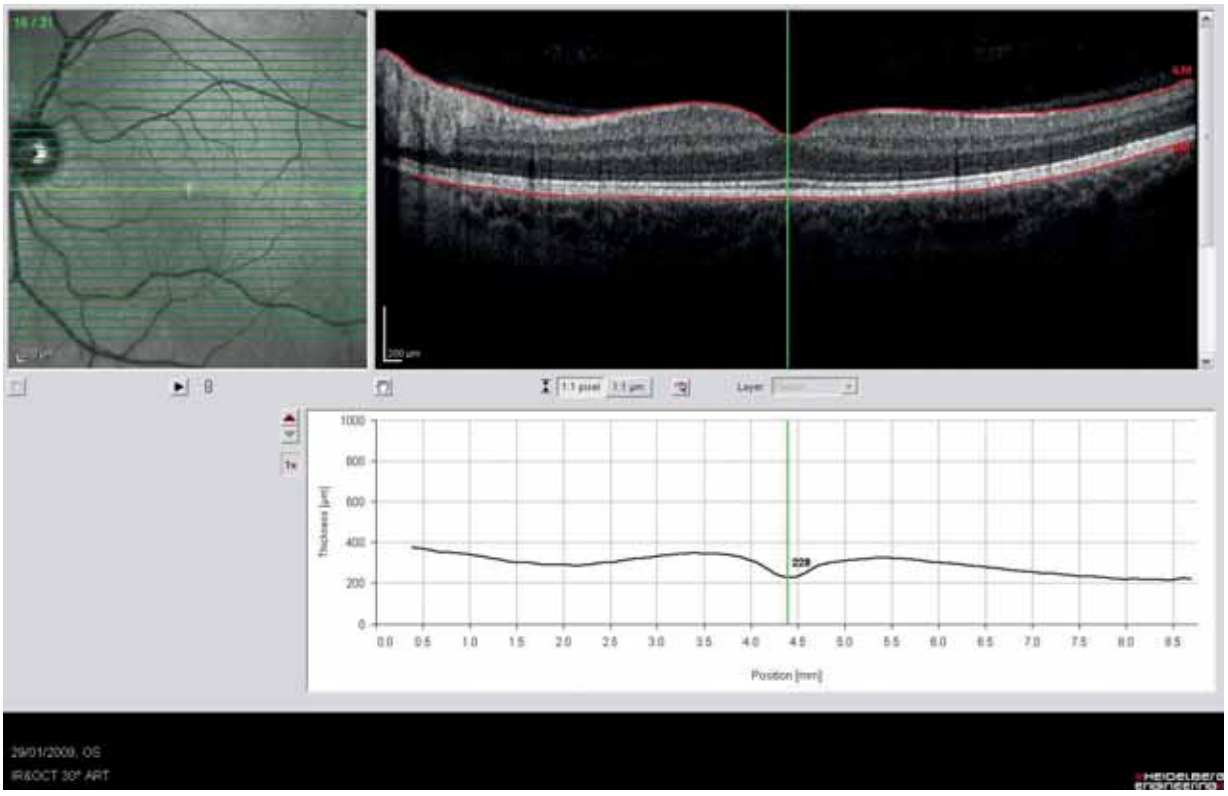


FIGURE Spectralis 1 Example screen print of horizontal section.

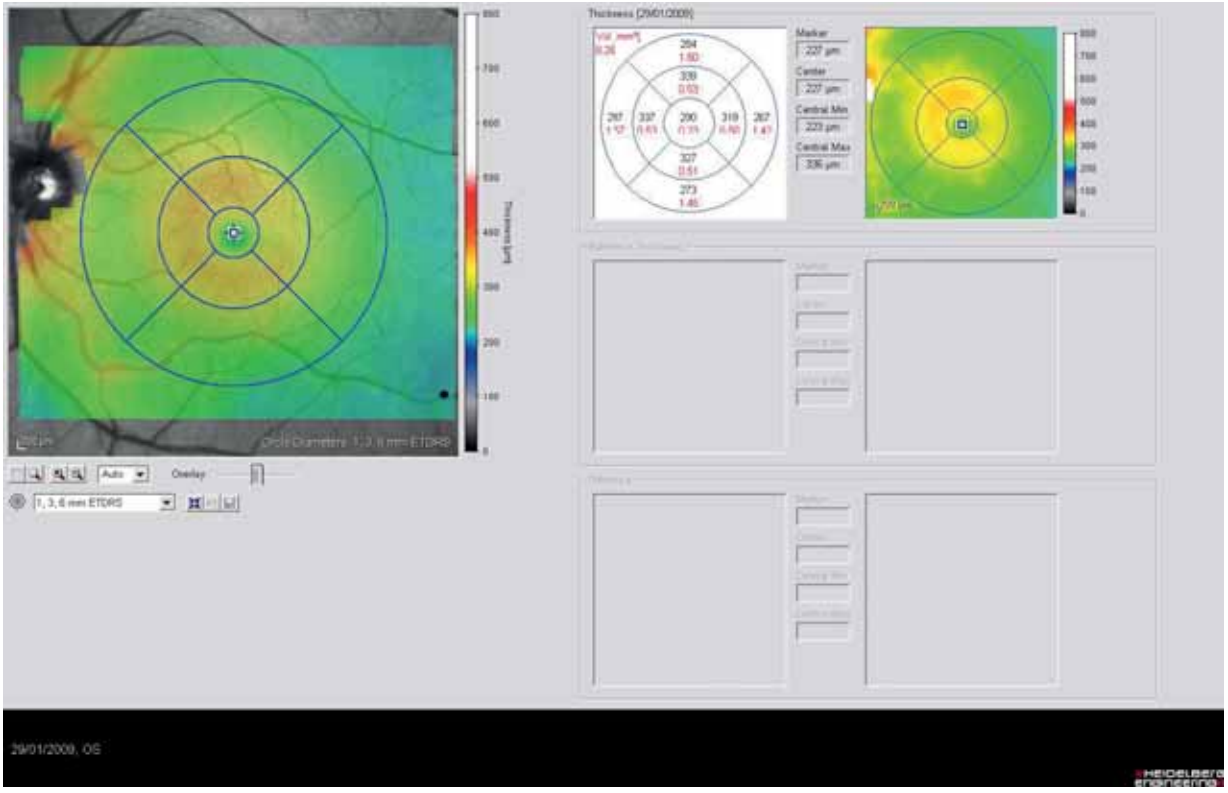


FIGURE Spectralis 2 Example screen print showing macular thickness map.

(b) and (c) Horizontal cross-section

Produce a screenshot of the horizontal slice centred on the macula (see figure Spectralis 1) and save in PNG format.

Optionally, if the region of greatest thickening is not visible on the central slice, select the slice which shows the greatest thickening and produce a screenshot as above.

Upload the data using the web-based system at: http://www.abdn.ac.uk/ismo/

For each eye there will be one colour photograph and two or three (where the thickening is not visible on the central section) optical coherence tomography images to upload.

© Queen's Printer and Controller of HMSO 2013. This work was produced by Olson *et al*. under the terms of a commissioning contract issued by the Secretary of State for Health. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

EME
HS&DR
HTA
PGfAR
PHR

Part of the NIHR Journals Library
www.journalslibrary.nihr.ac.uk

This report presents independent research funded by the National Institute for Health Research (NIHR). The views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR or the Department of Health

Published by the NIHR Journals Library