



This is a peer-reviewed, final published version of the following document and is licensed under Creative Commons: Attribution 4.0 license:

Lois, Noemi, Cook, Jonathan, Wang, Ariel, Aldington, Stephen J, Mistry, Hema, Maredza, Mandy, McAuley, Danny, Aslam, Tariq, Bailey, Clare, Chong, Victor, Ghanchi, Faruque, Scanlon, Peter H ORCID logoORCID: <https://orcid.org/0000-0001-8513-710X>, Sivaprasad, Sobha, Steel, David, Styles, Caroline, Azuara-Blanco, Augusto, Prior, Lindsay and Waugh, Norman (2021) Multimodal imaging interpreted by graders to detect re-activation of diabetic eye disease in previously treated patients: the EMERALD diagnostic accuracy study. Health Technology Assessment, 25 (32). pp. 1-104. doi:10.3310/hta25320

Official URL: <http://doi.org/10.3310/hta25320>

DOI: <http://dx.doi.org/10.3310/hta25320>

EPrint URI: <https://eprints.glos.ac.uk/id/eprint/9774>

Disclaimer

The University of Gloucestershire has obtained warranties from all depositors as to their title in the material deposited and as to their right to deposit such material.

The University of Gloucestershire makes no representation or warranties of commercial utility, title, or fitness for a particular purpose or any other warranty, express or implied in respect of any material deposited.

The University of Gloucestershire makes no representation that the use of the materials will not infringe any patent, copyright, trademark or other property or proprietary rights.

The University of Gloucestershire accepts no liability for any infringement of intellectual property rights in any material deposited but will remove such material from public view pending investigation in the event of an allegation of any such infringement.

PLEASE SCROLL DOWN FOR TEXT.

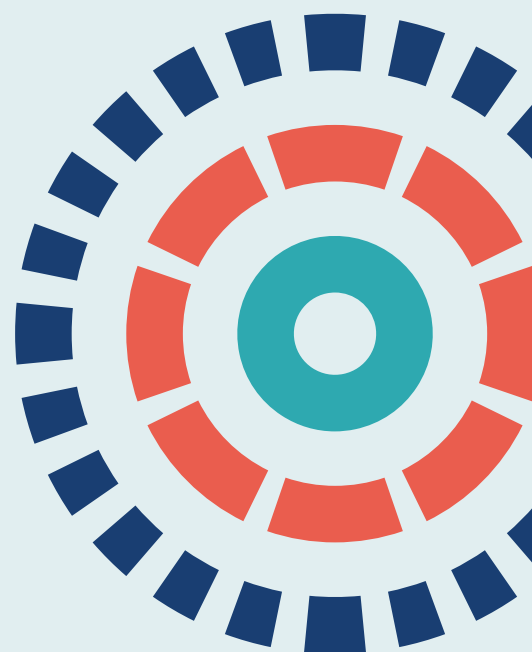
Health Technology Assessment

Volume 25 • Issue 32 • May 2021

ISSN 1366-5278

Multimodal imaging interpreted by graders to detect re-activation of diabetic eye disease in previously treated patients: the EMERALD diagnostic accuracy study

Noemi Lois, Jonathan Cook, Ariel Wang, Stephen Aldington, Hema Mistry, Mandy Maredza, Danny McAuley, Tariq Aslam, Clare Bailey, Victor Chong, Faruque Ghanchi, Peter Scanlon, Sobha Sivaprasad, David Steel, Caroline Styles, Augusto Azuara-Blanco, Lindsay Prior and Norman Waugh on behalf of the EMERALD Study Group



Multimodal imaging interpreted by graders to detect re-activation of diabetic eye disease in previously treated patients: the EMERALD diagnostic accuracy study

Noemi Lois^{1*}, Jonathan Cook², Ariel Wang²,
Stephen Aldington³, Hema Mistry⁴,
Mandy Maredza⁴, Danny McAuley^{1,5}, Tariq Aslam⁶,
Clare Bailey⁷, Victor Chong⁸, Faruque Ghanchi⁹,
Peter Scanlon³, Sobha Sivaprasad¹⁰, David Steel^{11,12},
Caroline Styles¹³, Augusto Azuara-Blanco¹⁴,
Lindsay Prior¹⁴ and Norman Waugh⁴ on behalf of the
EMERALD Study Group

¹The Wellcome-Wolfson Institute for Experimental Medicine, Queen's University Belfast, Belfast, UK

²Centre for Statistics in Medicine, University of Oxford, Oxford, UK

³Gloucestershire Hospitals NHS Foundation Trust, Gloucester, UK

⁴Warwick Medical School, University of Warwick, Coventry, UK

⁵The Regional Intensive Care Unit, Royal Victoria Hospital, Belfast, UK

⁶The Manchester Academic Health Science Centre, Manchester Royal Eye Hospital and Division of Pharmacy and Optometry, School of Health Sciences, Faculty of Biology, Medicine and Health, University of Manchester, Manchester, UK

⁷Bristol Eye Hospital, Bristol, UK

⁸Royal Free Hospital NHS Foundation Trust, London, UK

⁹Bradford Teaching Hospitals NHS Foundation Trust, Bradford Royal Infirmary, Bradford, UK

¹⁰National Institute for Health Research Moorfields Biomedical Research Centre, London, UK

¹¹Sunderland Eye Infirmary, Sunderland, UK

¹²Institute of Genetic Medicine, Newcastle University, Newcastle upon Tyne, UK

¹³Queen Margaret Hospital, Dunfermline, UK

¹⁴Centre for Public Health, Queen's University Belfast, Belfast, UK

*Corresponding author

Declared competing interests of authors: Jonathan Cook was a member of the National Institute for Health Research (NIHR) Health Technology Assessment (HTA) Efficient Trial Designs Board (between 2014 and 2016), NIHR HTA End of Life Care and Add-on Studies Board member (between 2014 and 2016) and member of a NIHR Research for Patient Benefit Programme regional advisory committee (South Central/South East and Central) between 2015 and 2019. Stephen Aldington reports grants

from the NIHR HTA programme for other work during the conduct of the study. Danny McAuley is a member of the Medical Research Council (MRC) and NIHR Efficacy and Mechanism Evaluation (EME) committees and is the Director of the MRC and NIHR EME programmes. Tariq Aslam reports grants from Bayer (Leverkusen, Germany) and Théa Pharmaceuticals Limited (Newcastle-under-Lyme, UK). Tariq Aslam also reports payment for lectures from Bayer, Théa Pharmaceuticals Limited, Novartis International AG (Basel, Switzerland) and Allergan (Dublin, Ireland); and advisory board fees from Novartis International AG, Bayer, Théa Pharmaceuticals, Allergan and Alimera Sciences, Inc. (Alpharette, GA, USA). Clare Bailey sat on advisory boards for Novartis International AG, Bayer and Roche Holding AG (Basel, Switzerland), and reports lecture fees from Bayer, Novartis International AG, Roche Holding AG, Allergan and Alimera Sciences, Inc. Victor Chong is a part-time employee of Boehringer Ingelheim International GmbH (Ingelheim am Rhein, Germany); the study was performed outwith Boehringer Ingelheim, and Boehringer Ingelheim did not endorse the study or its result. Faruque Ganchi reports consulting fees or paid advisory work from Alimera Sciences, Inc., Allergan, Bayer, Novartis International AG and Roche Holding AG; lecture fees or honoraria: Alimera Sciences, Inc., Allergan, Bayer, Novartis International AG and Roche Holding AG; payment for lectures from Alimera Sciences, Inc., Allergan, Bayer and Novartis International AG; and reports expenses from Allergan, Bayer, Boehringer Ingelheim, Chengdu List Pharmaceutical Co., Ltd, Novartis International AG, PanOptica (Mount Arlington, NJ, USA) and Roche Holding AG. Peter Scanlon reports grants and support to the institution from the NIHR HTA programme (project reference number 13/142/04). He also reports money to the institution from Gloucestershire Hospitals NHS Foundation Trust for employment, grants to the institution outside the submitted work from Bayer, Allergan, Boehringer Ingelheim and Novartis International AG, and grants for being a co-applicant on Innovate UK Grant Application. He also reports payment for various lectures and expenses. Sobha Sivaprasad is a member of HTA committee, received research grants from and attended advisory boards of Novartis International AG, Bayer, Allergan, Gyroscope (London, UK), Roche, Oxurion (Leuven, Belgium), Apellis Pharmaceuticals (Waltham, MA, USA), Boehringer Ingelheim, Heidelberg Engineering Inc. (Franklin, MA, USA) and Optos (Dunfermline, UK). David Steel has been a consultant to Gyroscope, Roche and Alcon (Geneva, Switzerland), and has received research funding from Bayer and Alcon. Augusto Azuara-Blanco is a member of the NIHR HTA Prioritisation Committee.

Published May 2021

DOI: 10.3310/hta25320

This report should be referenced as follows:

Lois N, Cook J, Wang A, Aldington S, Mistry H, Maredza M, *et al.* Multimodal imaging interpreted by graders to detect re-activation of diabetic eye disease in previously treated patients: the EMERALD diagnostic accuracy study. *Health Technol Assess* 2021;**25**(32).

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

ISSN 1366-5278 (Print)

ISSN 2046-4924 (Online)

Impact factor: 3.370

Health Technology Assessment is indexed in MEDLINE, CINAHL, EMBASE, the Cochrane Library and Clarivate Analytics Science Citation Index.

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

Editorial contact: journals.library@nihr.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

Health Technology Assessment (HTA) research is undertaken where some evidence already exists to show that a technology can be effective and this needs to be compared to the current standard intervention to see which works best. Research can evaluate any intervention used in the treatment, prevention or diagnosis of disease, provided the study outcomes lead to findings that have the potential to be of direct benefit to NHS patients. Technologies in this context mean any method used to promote health; prevent and treat disease; and improve rehabilitation or long-term care. They are not confined to new drugs and include any intervention used in the treatment, prevention or diagnosis of disease.

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.

This report

The research reported in this issue of the journal was funded by the HTA programme as project number 15/42/08. The contractual start date was in April 2017. The draft report began editorial review in March 2020 and was accepted for publication in June 2020. 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 and Social Care. 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 and Social Care.

© Queen's Printer and Controller of HMSO 2021. This work was produced by Lois *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. 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).

NIHR Journals Library Editor-in-Chief

Professor Ken Stein Professor of Public Health, University of Exeter Medical School, UK

NIHR Journals Library Editors

Professor John Powell Chair of HTA and EME Editorial Board and Editor-in-Chief of HTA and EME journals. Consultant Clinical Adviser, National Institute for Health and Care Excellence (NICE), UK, and Professor of Digital Health Care, Nuffield Department of Primary Care Health Sciences, University of Oxford, UK

Professor Andrée Le May Chair of NIHR Journals Library Editorial Group (HS&DR, PGfAR, PHR journals) and Editor-in-Chief of HS&DR, PGfAR, PHR journals

Professor Matthias Beck Professor of Management, Cork University Business School, Department of Management and Marketing, University College Cork, Ireland

Dr Tessa Crilly Director, Crystal Blue Consulting Ltd, UK

Dr Eugenia Cronin Senior Scientific Advisor, Wessex Institute, UK

Dr Peter Davidson Consultant Advisor, Wessex Institute, University of Southampton, UK

Ms Tara Lamont Senior Scientific Adviser (Evidence Use), Wessex Institute, University of Southampton, UK

Dr Catriona McDaid Senior Research Fellow, York Trials Unit, Department of Health Sciences, University of York, UK

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

Professor Geoffrey Meads Emeritus Professor of Wellbeing Research, University of Winchester, 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 Professor of Child Health Research, UCL Great Ormond Street Institute of Child Health, UK

Professor Jonathan Ross Professor of Sexual Health and HIV, University Hospital Birmingham, UK

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

Professor Ken Stein Professor of Public Health, University of Exeter Medical School, UK

Professor Jim Thornton Professor of Obstetrics and Gynaecology, Faculty of Medicine and Health Sciences, University of Nottingham, UK

Please visit the website for a list of editors: www.journalslibrary.nihr.ac.uk/about/editors

Editorial contact: journals.library@nihr.ac.uk

Abstract

Multimodal imaging interpreted by graders to detect re-activation of diabetic eye disease in previously treated patients: the EMERALD diagnostic accuracy study

Noemi Lois^{1*}, Jonathan Cook², Ariel Wang², Stephen Aldington³, Hema Mistry⁴, Mandy Maredza⁴, Danny McAuley^{1,5}, Tariq Aslam⁶, Clare Bailey⁷, Victor Chong⁸, Faruque Ghanchi⁹, Peter Scanlon³, Sobha Sivaprasad¹⁰, David Steel^{11,12}, Caroline Styles¹³, Augusto Azuara-Blanco¹⁴, Lindsay Prior¹⁴ and Norman Waugh⁴ on behalf of the EMERALD Study Group

¹The Wellcome-Wolfson Institute for Experimental Medicine, Queen's University Belfast, Belfast, UK

²Centre for Statistics in Medicine, University of Oxford, Oxford, UK

³Gloucestershire Hospitals NHS Foundation Trust, Gloucester, UK

⁴Warwick Medical School, University of Warwick, Coventry, UK

⁵The Regional Intensive Care Unit, Royal Victoria Hospital, Belfast, UK

⁶The Manchester Academic Health Science Centre, Manchester Royal Eye Hospital and Division of Pharmacy and Optometry, School of Health Sciences, Faculty of Biology, Medicine and Health, University of Manchester, Manchester, UK

⁷Bristol Eye Hospital, Bristol, UK

⁸Royal Free Hospital NHS Foundation Trust, London, UK

⁹Bradford Teaching Hospitals NHS Foundation Trust, Bradford Royal Infirmary, Bradford, UK

¹⁰National Institute for Health Research Moorfields Biomedical Research Centre, London, UK

¹¹Sunderland Eye Infirmary, Sunderland, UK

¹²Institute of Genetic Medicine, Newcastle University, Newcastle upon Tyne, UK

¹³Queen Margaret Hospital, Dunfermline, UK

¹⁴Centre for Public Health, Queen's University Belfast, Belfast, UK

*Corresponding author n.lois@qub.ac.uk

Background: Owing to the increasing prevalence of diabetes, the workload related to diabetic macular oedema and proliferative diabetic retinopathy is rising, making it difficult for hospital eye services to meet demands.

Objective: The objective was to evaluate the diagnostic performance, cost-effectiveness and acceptability of a new pathway using multimodal imaging interpreted by ophthalmic graders to detect reactivation of diabetic macular oedema/proliferative diabetic retinopathy in previously treated patients.

Design: This was a prospective, case-referent, cross-sectional diagnostic study.

Setting: The setting was ophthalmic clinics in 13 NHS hospitals.

Participants: Adults with type 1 or type 2 diabetes with previously successfully treated diabetic macular oedema/proliferative diabetic retinopathy in one/both eyes in whom, at the time of enrolment, diabetic macular oedema/proliferative diabetic retinopathy could be active or inactive.

Methods: For the ophthalmic grader pathway, review of the spectral domain optical coherence tomography scans to detect diabetic macular oedema, and seven-field Early Treatment Diabetic Retinopathy Study/ultra-wide field fundus images to detect proliferative diabetic retinopathy, by trained ophthalmic graders. For the current standard care pathway (reference standard), ophthalmologists examined patients face to face by slit-lamp biomicroscopy for proliferative diabetic retinopathy and, in addition, spectral domain optical coherence tomography imaging for diabetic macular oedema.

Outcome measures: The primary outcome measure was sensitivity of the ophthalmic grader pathway to detect active diabetic macular oedema/proliferative diabetic retinopathy. The secondary outcomes were specificity, agreement between pathways, cost-consequences, acceptability and the proportion of patients requiring subsequent ophthalmologist assessment, unable to undergo imaging and with inadequate quality images/indeterminate findings. It was assumed for the main analysis that all patients in whom graders diagnosed active disease or were 'unsure' or images were 'ungradable' required examination by an ophthalmologist.

Results: Eligible participants with active and inactive diabetic macular oedema (152 and 120 participants, respectively) and active and inactive proliferative diabetic retinopathy (111 and 170 participants, respectively) were recruited. Under the main analysis, graders had a sensitivity of 97% (142/147) (95% confidence interval 92% to 99%) and specificity of 31% (35/113) (95% confidence interval 23% to 40%) to detect diabetic macular oedema. For proliferative diabetic retinopathy, graders had a similar sensitivity and specificity using seven-field Early Treatment Diabetic Retinopathy Study [sensitivity 85% (87/102), 95% confidence interval 77% to 91%; specificity 48% (77/160), 95% confidence interval 41% to 56%] or ultra-wide field imaging [sensitivity 83% (87/105), 95% confidence interval 75% to 89%; specificity 54% (86/160), 95% confidence interval 46% to 61%]. Participants attending focus groups expressed preference for face-to-face evaluations by ophthalmologists. In the ophthalmologists' absence, patients voiced the need for immediate feedback following grader's assessments, maintaining periodic evaluations by ophthalmologists. Graders and ophthalmologists were supportive of the new pathway. When compared with the reference standard (current standard pathway), the new grader pathway could save £1390 per 100 patients in the review of people with diabetic macular oedema and, depending on the imaging modality used, between £461 and £1189 per 100 patients in the review of people with proliferative diabetic retinopathy.

Conclusions: For people with diabetic macular oedema, the ophthalmic grader pathway appears safe and cost saving. The sensitivity of the new pathway to detect active proliferative diabetic retinopathy was lower, but may still be considered acceptable for patients with proliferative diabetic retinopathy previously treated with laser. Suggestions from focus group discussions should be taken into consideration if the new pathway is introduced to ensure its acceptability to users.

Limitations: Lack of fundus fluorescein angiography to confirm diagnosis of active proliferative diabetic retinopathy.

Future work: Could refinement of the new pathway increase its sensitivity to detect proliferative diabetic retinopathy? Could artificial intelligence be used for automated reading of images in this previously treated population?

Trial registration: Current Controlled Trials ISRCTN10856638 and ClinicalTrials.gov NCT03490318.

Funding: This project was funded by the National Institute for Health Research (NIHR) Health Technology Assessment programme and will be published in full in *Health Technology Assessment* Vol. 25, No. 32. See the NIHR Journals Library website for further project information.

Contents

List of tables	xiii
List of figures	xvii
List of boxes	xix
List of abbreviations	xxi
Plain English summary	xxiii
Scientific summary	xxv
Chapter 1 Background	1
Diabetic macular oedema	1
Proliferative diabetic retinopathy	2
The burden of diabetic retinopathy, diabetic macular oedema and proliferative diabetic retinopathy and NHS pressures to meet service demands	7
Chapter 2 Diagnostic accuracy	9
Methods	9
Setting	9
Population	9
Pathways evaluated	9
Outcome measures	10
Selection of ophthalmic graders and training	10
Patient flow in EMERALD	11
Schedule of assessments for EMERALD	12
Masking	13
Data collection and quality checks	13
Sample size	13
Statistical analyses	13
Eligible participants for the EMERALD ophthalmic grader's pathway	14
Eye-level classification	15
Participant-level classification	17
Definition of grader's assessments of optical coherence tomography, OPTOS and ETDRS	18
Diagnostic performance of the EMERALD ophthalmic graders	18
Diagnostic performance analysis methods	19
Sensitivity analyses of diagnostic performance	19
Secondary analyses of diagnostic accuracy	19
Results	21
Participant characteristics	21
Sample description	21
Diagnostic performance of the imaging tests	25
Diagnostic performance for diabetic macular oedema patients eligible for the new ophthalmic grader pathway	25
Results: main analysis for diabetic macular oedema	28
Results: sensitivity analysis for diabetic macular oedema	28
Sensitivity analysis 1	28

<i>Sensitivity analysis 2, 3 and 6</i>	28
<i>Diagnostic performance for proliferative diabetic retinopathy patients eligible for the new ophthalmic grader pathway</i>	30
Results: main analysis for proliferative diabetic retinopathy	37
Results: sensitivity analysis for proliferative diabetic retinopathy	37
<i>Sensitivity analysis 1</i>	37
<i>Sensitivity analyses 2, 4, 5 and 6</i>	37
Results: additional analysis for proliferative diabetic retinopathy	37
<i>Additional analyses 1–3</i>	37
<i>Paired comparisons of OPTOS and ETDRS assessment of proliferative diabetic retinopathy</i>	38
<i>Secondary sensitivity analyses on combined diseases (diabetic macular oedema and proliferative diabetic retinopathy) for all patients</i>	39
Discussion	42
Chapter 3 Focus group research	45
The aim of the focus group work	45
Methods	45
<i>Sampling and recruitment strategy</i>	45
<i>What happened in the focus groups?</i>	46
<i>Focus groups for professionals</i>	46
<i>Analysing the data</i>	47
Findings	49
<i>Focus group phase 1: the journey through the clinic</i>	49
<i>Focus group phase 2: the virtual clinic</i>	52
<i>Focus group phase 3: the ranking exercise</i>	55
<i>Responses of consultants</i>	55
<i>Responses of photographers and graders</i>	56
Discussion	57
Summary	58
Conclusion and implications	59
Chapter 4 Cost-consequences analysis	61
Introduction	61
Methods	61
Outcome measures	63
Results	64
<i>Costs</i>	64
<i>Diabetic macular oedema: comparison of the ophthalmic grader pathway with current standard of care</i>	66
<i>Proliferative diabetic retinopathy: comparison of grader pathway with current standard of care</i>	67
<i>Proliferative diabetic retinopathy: comparison with enhanced reference standard</i>	69
<i>Results for EQ-5D-5L, NEI VFQ-25 and VisQoL measures</i>	70
Summary/discussion	73
<i>Diabetic macular oedema</i>	73
<i>Proliferative diabetic retinopathy</i>	73
<i>Modelling of cost-effectiveness</i>	74
Chapter 5 General discussion	77
Statement of principal findings	77
Strengths and limitations	77
Implementation of the new care pathway: matters to consider	78
Cost-consequences analysis	79

The potential for future automated grading of images	79
Imaging modalities	80
Virtual clinics	81
Conclusion	82
Recommendations for future research	82
Acknowledgements	85
References	91
Appendix 1 Additional tables from <i>Chapter 2</i> diagnostic accuracy	97
Appendix 2 Dealing with a small number of participants	103

List of tables

TABLE 1 List of analyses	15
TABLE 2 List of the planned principal analyses (main and sensitivity analyses of diagnostic accuracy)	16
TABLE 3 List of unplanned additional principal analyses of diagnostic accuracy	17
TABLE 4 Combination of eye-level 'active' status to person-level 'active' status	18
TABLE 5 List of secondary analyses and additional post-hoc analysis of diagnostic accuracy: all patients	20
TABLE 6 Demographic characteristics of participants	23
TABLE 7 The DMO classification based on slit-lamp biomicroscopic examination by an ophthalmologist with access to SD-OCT images (reference standard)	23
TABLE 8 The PDR classification based on slit-lamp biomicroscopic examination by an ophthalmologist (reference standard)	24
TABLE 9 The DMO grading data: graders assessment based on OCT images compared with reference standard	24
TABLE 10 The PDR grading data: graders and ophthalmologists assessment based on UWF OPTOS and seven-field ETDRS images compared with reference standard	26
TABLE 11 Missing images at person level	27
TABLE 12 Diabetic macular oedema classifications: eligible patients for new pathway ($n = 272$)	27
TABLE 13 Diagnostic performance for DMO patients	30
TABLE 14 Proliferative diabetic retinopathy classifications: eligible patients for the new pathway ($n = 281$)	31
TABLE 15 Diagnostic performance for PDR patients	34
TABLE 16 Paired comparisons in sensitivity and specificity between G-OPTOS and G-ETDRS (main analysis)	38
TABLE 17 Paired comparisons in sensitivity and specificity between grader OPTOS and ETDRS (SENA1)	39
TABLE 18 Patient-level classifications: all patients ($n = 397$)	40
TABLE 19 Secondary analyses at patient level: all patients	41

TABLE 20 Proposed maximum variation sampling strategy for FGs in the UK: number of groups in each category by country and age group	45
TABLE 21 Number of patients who were invited, gave consent and attended FG meetings at five UK study sites	47
TABLE 22 Ranking of preferences	47
TABLE 23 Staff costs (2019 prices)	62
TABLE 24 Unit costs of equipment (2019 prices)	63
TABLE 25 Time (in minutes) taken to complete specified activities per patient by ophthalmologists and ophthalmic graders	64
TABLE 26 Alternative costs for staff time	65
TABLE 27 The costs of procedures	66
TABLE 28 Relative costs of follow-up of DMO	67
TABLE 29 Cost comparison: grader pathway and ophthalmologist standard care	68
TABLE 30 Cost comparison: grader pathway and detection of eyes needing treatment	68
TABLE 31 Cost comparison: grader pathway and detection of eyes with pre-retinal of vitreous haemorrhage	69
TABLE 32 Cost and consequences comparison using ERS	69
TABLE 33 Costs comparison of the reading of images by graders and ophthalmologists for 100 patients	70
TABLE 34 Mean EQ-5D-5L utility and vision-specific outcome scores by status of DMO (active vs. inactive)	71
TABLE 35 Mean EQ-5D-5L utility and vision-specific outcome scores by status of PDR (active vs. inactive)	71
TABLE 36 Demographics of participants: active diseases based on ophthalmologist slit-lamp examination	97
TABLE 37 Findings identified by ophthalmologists (when undertaking slit-lamp biomicroscopy = reference standard) and recorded as potential factors that could lead to inadequate image quality	97
TABLE 38 Other findings/comorbidities identified by the ophthalmologist when undertaking slit-lamp biomicroscopy	98
TABLE 39 Reason for missing eye-level assessments	99
TABLE 40 PDR test results using OPTOS or using ETDRS vs. reference standard or enhanced standard	99

TABLE 41	Secondary analyses on randomly selected eyes: all patients – SECA1A	100
TABLE 42	Secondary analyses at eye level: all patients	100
TABLE 43	Recruitment proportion by different time points	101

List of figures

FIGURE 1 Spectral domain-optical coherence tomography in DMO	2
FIGURE 2 Spectral domain-optical coherence tomography in DMO with CRT of > 400 µm	3
FIGURE 3 Spectral domain-optical coherence tomography following resolution of DMO	3
FIGURE 4 Fundus images of an eye with HRC PDR and a healthy eye	4
FIGURE 5 Fundus image of an eye treated with laser panretinal photocoagulation and with regressed PDR	6
FIGURE 6 Standard fundus image in a patient with DMO	6
FIGURE 7 Scheme demonstrating the seven-field ETDRS fields	6
FIGURE 8 Ultra-wide field fundus images	7
FIGURE 9 Study flow diagram	22
FIGURE 10 Flow diagram: diagnosis as determined by the ophthalmic graders for DMO patients – main and SENA1	29
FIGURE 11 Flow diagram: diagnosis as determined by the ophthalmic graders for PDR patients comparing OPTOS and ETDRS – main and SENA1	33
FIGURE 12 Issues discussed by male participants in focus group 5	48
FIGURE 13 Mean NEI VFQ-25 scores among patients eligible for the new ophthalmic grader pathway categorised by DMO status (active vs. inactive)	71
FIGURE 14 Proportion of DMO patients split by EQ-5D-5L dimensions and levels	72
FIGURE 15 Proportion of PDR patients split by EQ-5D-5L dimensions and levels	72
FIGURE 16 Current standard of care and new proposed pathway for people with previously treated and stable DMO and PDR	82
FIGURE 17 Barriers to increasing physical activity identified in 14 FGs	103
FIGURE 18 Activities for a regenerated park	103
FIGURE 19 Number of new (additional) causes of flu referenced in sequential interviews (<i>n</i> = 54)	104
FIGURE 20 Consequences of traumatic brain injury as identified in 32 interviews with traumatic brain injury outpatients	104

List of boxes

BOX 1 How *Figure 12* was constructed

48

List of abbreviations

AI	artificial intelligence	NICTU	Northern Ireland Clinical Trials Unit
anti-VEGF	anti-vascular endothelial growth factor	NIHR	National Institute for Health Research
CARF	Central Administrative Research Facility	NPDR	non-proliferative diabetic retinopathy
CI	confidence interval	NVD	new vessels in the disc
CRF	case report form	NVE	new vessels elsewhere
CRT	central retinal thickness	OCT	optical coherence tomography
DMO	diabetic macular oedema	O-FTF	ophthalmologist face-to-face examination
DR	diabetic retinopathy	PDR	proliferative diabetic retinopathy
EMERALD	Effectiveness of Multimodal imaging for the Evaluation of Retinal oedema And new vessels in Diabetic retinopathy	PI	principal investigator
EQ-5D-5L	EuroQol-5 Dimensions, five-level version	PPI	patient and public involvement
ERS	enhanced reference standard	PRP	panretinal photocoagulation
ETDRS	Early Treatment Diabetic Retinopathy Study	PSSRU	Personal and Social Services Research Unit
FG	focus group	SD-OCT	Spectral Domain Optical Coherence Tomography
FN	false negative	SENA	sensitivity analysis
FP	false positive	TN	true negative
HTA	Health Technology Assessment	TP	true positive
NEI VFQ-25	National Eye Institute Visual Function-25	TRD	tractional retinal detachment
		VisQoL	Vision and Quality of Life

Plain English summary

More and more people are developing diabetes. Diabetic macular oedema and proliferative diabetic retinopathy are complications of diabetes, which could cause blindness. Thus, people with diabetic macular oedema and proliferative diabetic retinopathy need to be treated in a timely manner and reviewed in clinic for life.

The population in the world is ageing. As a result, there are more people with eye diseases. There are also more treatments now for people with eye diseases. The workload in hospitals is increasing, making it difficult for the NHS to cope with the demand. There are not enough ophthalmologists (eye doctors) to look after patients. Delayed appointments and treatment mean that patients may lose sight.

The goal of EMERALD (Effectiveness of Multimodal imaging for the Evaluation of Retinal oedema And new vessels in Diabetic retinopathy) was to see if patients with treated and stable diabetic macular oedema or proliferative diabetic retinopathy could be followed by 'ophthalmic graders', who are not doctors but are trained to diagnose diabetic macular oedema and proliferative diabetic retinopathy. In EMERALD, trained ophthalmic graders examined photographs of the back of the eye of people with diabetic macular oedema and proliferative diabetic retinopathy. They checked if diabetic macular oedema and proliferative diabetic retinopathy remain inactive. If so, patients could continue follow-up with the ophthalmic graders. If diabetic macular oedema or proliferative diabetic retinopathy were active, graders would immediately refer patients to ophthalmologists.

EMERALD found that graders were excellent at detecting diabetic macular oedema, and this could give ophthalmologists time to see other patients. Graders were not quite as good at detecting active proliferative diabetic retinopathy. However, considering that patients had already had treatment, this may still be safe. Patients participating in focus group discussions mentioned that they would prefer to see ophthalmologists, so they could ask questions about their eye condition. If this was not possible, they would like to have immediate results from graders and still see the ophthalmologist from time to time.

Scientific summary

Background

Diabetic retinopathy is a leading cause of sight loss in people of working age. Patients with diabetic retinopathy may lose sight from diabetic macular oedema and/or proliferative diabetic retinopathy. In diabetic macular oedema fluid accumulates at the macula, the retinal area responsible for central sight, with subsequent central visual loss (e.g. that required for reading). In proliferative diabetic retinopathy abnormal 'new vessels' grow in the retina and may rupture causing a vitreous haemorrhage or scarring that could lead to a tractional retinal detachment. Vitreous haemorrhage and tractional retinal detachment cause loss of central and peripheral vision. Owing to increasing numbers of people with diabetes, it is expected that the burden of diabetic retinopathy will continue to rise, despite improvements in glycaemic control and screening for retinopathy having reduced the risk of advanced retinopathy.

The estimated prevalence of diabetic macular oedema and proliferative diabetic retinopathy is similar, at $\approx 7\%$. Considering the prevalence of diabetes in UK (≈ 3.9 million in 2019), a minimum of 273,000 people have diabetic macular oedema and proliferative diabetic retinopathy in the UK.

Diabetic macular oedema is treated with macular laser photocoagulation (when the central retinal thickness, measured by spectral domain optical coherence tomography, is $< 400 \mu\text{m}$) or intravitreal injections of antivascular endothelial growth factor therapies (when the central retinal thickness is $\geq 400 \mu\text{m}$). Patients should be followed up every 3–4 months after laser treatment. After antivascular endothelial growth factors, patients should usually be followed up monthly during the first year of treatment and every 1–3 months thereafter, as diabetic macular oedema can recur.

Proliferative diabetic retinopathy is currently treated with laser panretinal photocoagulation. After treatment, patients are followed up at 4- to 6-month intervals for life, as proliferative diabetic retinopathy can recur and vitreous haemorrhage/tractional retinal detachment could still occur. A high proportion of patients followed up in Hospital Eye Services have treated and inactive proliferative diabetic retinopathy.

Currently in the NHS, ophthalmologists assess patients during follow-up visits. At each visit, patients with diabetic macular oedema/proliferative diabetic retinopathy receive a visual acuity test, often undertaken by a nurse; Spectral Domain Optical Coherence Tomography obtained by a photographer/imaging technician; and fundus examination by slit-lamp biomicroscopy by an ophthalmologist. Based on slit-lamp biomicroscopy and Spectral Domain Optical Coherence Tomography, the ophthalmologist determines whether or not diabetic macular oedema is present; based on slit-lamp biomicroscopy the ophthalmologist determines if there is active proliferative diabetic retinopathy. Spectral Domain Optical Coherence Tomography is non-invasive, safe and fast, obtaining scans of the macula. Fundus (retinal) photographs or fundus fluorescein angiography are not routinely carried out to determine the activity of proliferative diabetic retinopathy, but they are used in selected patients. Fundus fluorescein angiography requires injection of a dye into a peripheral vein. Images are taken as the dye circulates in the retina. Standard cameras used to obtain fundus photographs and fundus fluorescein angiography cannot image the retinal periphery, but newer ultra-wide field imaging captures nearly the entire retina in a single image.

The large number of people with diabetic macular oedema/proliferative diabetic retinopathy and the need for patients to be followed up at short intervals is making it difficult for the NHS to cope, especially because of a shortage of ophthalmologists. Difficulties will increase given the increasing prevalence of diabetes. Identifying new ways to increase NHS capacity/efficiency without compromising quality of care is essential. EMERALD (Effectiveness of Multimodal imaging for the Evaluation of Retinal oedema And new vessels in Diabetic retinopathy) was planned with this in mind.

Objective

The objective was to determine whether or not patients with successfully treated diabetic macular oedema/proliferative diabetic retinopathy could be followed through a new care pathway involving multimodal retinal imaging and image assessment by trained ophthalmic graders. Diagnostic accuracy, cost-effectiveness and acceptability of this new pathway to patients and health-care professionals were evaluated against the current standard of care.

Methods

Design

This was a prospective, case-referent, cross-sectional diagnostic study.

Setting

This was carried out in specialist Hospital Eye Services.

Participants

Adults with diabetes, with previously successfully treated diabetic macular oedema/proliferative diabetic retinopathy in one/both eyes; at the time of enrolment, diabetic macular oedema/proliferative diabetic retinopathy could be active or inactive.

Clinical pathways assessed

New pathway: multimodal imaging (Spectral Domain Optical Coherence Tomography to detect diabetic macular oedema; seven-field Early Treatment Diabetic Retinopathy Study and ultra-wide field fundus images to detect proliferative diabetic retinopathy) with subsequent review by trained, tested and certified ophthalmic graders.

Standard care pathway: ophthalmologist examining patients in clinic as per current standard practice (for diabetic macular oedema slit-lamp biomicroscopy examination and Spectral Domain Optical Coherence Tomography; for proliferative diabetic retinopathy slit-lamp biomicroscopy examination).

Outcomes

The primary outcome was sensitivity of the new pathway to detect active diabetic macular oedema/proliferative diabetic retinopathy.

The secondary outcomes were specificity, concordance, cost-effectiveness, acceptability of the new pathway to patients and health-care professionals, proportions of patients requiring subsequent assessment by ophthalmologist, unable to undergo imaging and with images of inadequate quality for interpretation.

EMERALD patient flow

Patients with previously treated diabetic macular oedema/proliferative diabetic retinopathy were identified from clinical records, electronic databases or in clinic. At their review appointment, an ophthalmologist confirmed patient eligibility, obtained informed consent and determined whether or not active/inactive diabetic macular oedema/proliferative diabetic retinopathy was present (reference standard). Visual acuity, Spectral Domain Optical Coherence Tomography and fundus examination were carried out as per routine standard practice. In some participating sites patients were evaluated in 'research' clinics, and in others they were evaluated in usual NHS clinics.

Non-stereoscopic seven-field Early Treatment Diabetic Retinopathy Study and ultra-wide angle fundus images were obtained, anonymised, uploaded to a central facility and allocated randomly to ophthalmic graders. Graders did not grade images from their own centre (to ensure masking to the reference standard).

Graders did not grade both seven-field Early Treatment Diabetic Retinopathy Study and ultra-wide fundus images from the same patient (to prevent the grading of one technology influencing the grading of the other). Graders judged whether there was active/inactive diabetic macular oedema/proliferative diabetic retinopathy or if they were uncertain.

Given the possibility of new vessels not being seen by the ophthalmologist on slit-lamp biomicroscopy but detected on photographs, EMERALD also evaluated an 'enhanced' reference standard for proliferative diabetic retinopathy consisting of the reference standard supplemented by evaluation of seven-field Early Treatment Diabetic Retinopathy Study and ultra-wide field fundus images reviewed by an ophthalmologist expert in diabetic retinopathy. If active proliferative diabetic retinopathy was detected by one of these three methods, it was considered that the enhanced reference standard identified active proliferative diabetic retinopathy. Seven-field Early Treatment Diabetic Retinopathy Study and ultra-wide field images of the same participant were reviewed by different ophthalmologists, who did not grade images from their own centre.

Focus groups

To determine acceptability of the new pathway to patients and health professionals, focus group discussions were undertaken.

Sample size and statistical analysis

The sample size was based on the number of patients with reactivated (active) diabetic macular oedema and proliferative diabetic retinopathy which would enable sensitivity to be tested against a pre-specified target level of 80%, considered the minimum acceptable level for the ophthalmic grader's pathway. A lower specificity was thought acceptable; a target of 65% was used to confirm sufficiency of the sample size for assessing specificity. To detect sensitivity of the new pathway with 80% and 90% power (10% and 12% higher than the 80% minimal target set) required 89 participants with each diabetic macular oedema and PDR that had reactivated, with two-sided 5% significance level. Ninety-three participants whose disease had not reactivated would enable a specificity of 80% to be detected with 90% power. A 95% confidence interval for photographer sensitivity and specificity would have a confidence interval (Wilson method) with a width of 10–20%, depending on the observed level. Allowing for 10% missing/indeterminate results, 104 individuals with each active and inactive diabetic macular oedema and active and inactive proliferative diabetic retinopathy were required.

Separate analyses were planned for diabetic macular oedema and proliferative diabetic retinopathy. Participants were categorised as having active or inactive diabetic macular oedema/proliferative diabetic retinopathy according to the reference standard, at the person level. Those with previously successfully treated diabetic macular oedema/proliferative diabetic retinopathy constitute 'eligible' participants for each analysis (diabetic macular oedema/proliferative diabetic retinopathy) for the new pathway. This person-based assessment reflects the consequences of the clinical decision. The diagnostic performance of the new pathway was quantified against the reference standard. Reflecting how the new pathway would function in practice, 'unsure', 'ungradable' and 'active' classifications required 'referral' and examination by an ophthalmologist under the main analyses.

Planned sensitivity analyses included (1) assessment of the impact of 'unsure' and 'ungradable' on the diagnostic performance of the ophthalmic grader; (2) using the ophthalmologist's decision to do further treatment, rather than presence of active disease; (3) detection of more severe disease (central-involving diabetic macular oedema, pre-retinal or vitreous haemorrhage in proliferative diabetic retinopathy); (4) diagnostic performance within routine NHS clinics (vs. 'research' clinics); and (5) for proliferative diabetic retinopathy only, diagnostic performance of the ophthalmic grader against the 'enhanced' reference standard. The impact of using ultra-wide field versus seven-field Early Treatment Diabetic Retinopathy Study images on the diagnostic performance of the new pathway was assessed under the principal analyses for proliferative diabetic retinopathy using both reference standard and enhanced reference standard. Additional analyses were carried out in the proliferative

diabetic retinopathy group to aid understanding findings from pre-planned analyses. Agreement between proliferative diabetic retinopathy assessment methods was quantified.

Secondary analyses included evaluation of eye level data; analysis including all patients (with or without diabetic macular oedema/proliferative diabetic retinopathy); assessment of the overall referral (diabetic macular oedema and proliferative diabetic retinopathy); and use of visual acuity as a proxy to detect active disease.

Analyses were carried out using Stata® version 15 (StataCorp LP, College Station, TX, USA). A statistical analysis plan was agreed and made accessible on the EMERALD website and the EMERALD protocol published prior to data analysis.

Health economic evaluation

Costs of ophthalmic grader and standard pathways were prospectively obtained, including collection of time costs for each procedure. It was hypothesised that the new pathway would have the same sensitivity as the standard care pathway but at lower cost, making the analysis a cost-consequences one, including assessment of ophthalmologist time released by the new pathway. Diabetic macular oedema and proliferative diabetic retinopathy were assessed separately. If there was marginal loss in sensitivity in the new pathway, a cost-effectiveness analysis was planned in which the disutility of the visual impact of the marginal loss would be assessed against costs saved. If there was an unacceptably low sensitivity for proliferative diabetic retinopathy, no modelling would be done.

Results

Three-hundred and ninety-seven participants, 272 eligible with diabetic macular oedema (152 active diabetic macular oedema and 120 inactive diabetic macular oedema) and 281 eligible with proliferative diabetic retinopathy (111 active proliferative diabetic retinopathy and 170 inactive proliferative diabetic retinopathy) were recruited. Most eligible participants with diabetic macular oedema were white ($n = 240$, 88%), male ($n = 175$, 64%) and over half were aged ≥ 60 years ($n = 159$, 58%). Most eligible participants with proliferative diabetic retinopathy were white ($n = 234$, 83%), male ($n = 185$, 66%) and slightly less than half ($n = 133$, 47%) were aged ≥ 60 years.

Under the main analysis of diabetic macular oedema (grader referring patient to ophthalmologists due to presence of active diabetic macular oedema or unsure or ungradable), graders had a sensitivity of 97% (142/147; 95% confidence interval 92% to 99%) with a specificity of 31% (35/113; 95% confidence interval 23% to 40%) when compared with the reference standard. Similar results were obtained for analysis evaluating people with diabetic macular oedema requiring further treatment (sensitivity 95%; 81/85, 95% confidence interval 89% to 98%; specificity 21%; 36/175, 95% confidence interval 15% to 27%) and those with central-involving diabetic macular oedema (sensitivity 94%; 121/129, 95% confidence interval 88% to 97%; specificity 56%; 72/128, 95% confidence interval 48% to 65%). Results were similar to those of the main analysis when only referral for active diabetic macular oedema was considered (i.e. excluding the 'unsure' and 'ungradable') and when patients were assessed in NHS clinics (vs. 'research' clinics).

Under the main analysis of proliferative diabetic retinopathy (grader referring patients to ophthalmologists due to presence of active proliferative diabetic retinopathy or unsure or ungradable), graders had similar sensitivity and specificity whether they used seven-field Early Treatment Diabetic Retinopathy Study (sensitivity 85%; 87/102, 95% confidence interval 77% to 91%; specificity 48%; 77/160, 95% CI 41% to 56%) or ultra-wide field (sensitivity 83%; 87/105, 95% confidence interval 75% to 89%; specificity 54%; 86/160, 95% CI 46% to 61%) images. Sensitivity and specificity were similar when grading patients

requiring further treatment (for seven-field Early Treatment Diabetic Retinopathy Study images: sensitivity of 88%; 74/84, 95% confidence interval 79% to 93%; specificity 46%; 82/178, 95% confidence interval 39% to 53%; for ultra-wide field images sensitivity of 86%; 77/90, 95% confidence interval 77% to 91%; specificity 52%; 91/175, 95% confidence interval 45% to 59%). Sensitivity and specificity of the graders to detect more severe disease (proliferative diabetic retinopathy with pre-retinal or vitreous haemorrhage) were slightly higher when using ultra-wide field imaging (sensitivity 87%; 62/71, 95% confidence interval 78% to 93%; specificity 49%; 95/193, 95% confidence interval 42% to 56%; for seven-field Early Treatment Diabetic Retinopathy Study sensitivity 80%; 53/66, 95% confidence interval 69% to 88%; specificity 40%; 79/196, 95% confidence interval 34% to 47%). Results against the enhanced reference standard were similar to those against the reference standard (for seven-field Early Treatment Diabetic Retinopathy Study images, sensitivity of 82%; 111/135, 95% confidence interval 75% to 88%; specificity 54%; 68/127, 95% confidence interval 45% to 62%; for ultra-wide field images sensitivity 80%; 110/138, 95% confidence interval 72% to 86%; specificity 60%; 76/127, 95% confidence interval 51% to 68%). Findings were similar whether patients were assessed in NHS or 'research' clinics. Sensitivity and specificity, however, were lower when considering referrals due to active proliferative diabetic retinopathy only (i.e. excluding 'unsure' and 'ungradable').

Thirty-six participants attended ten focus groups in Northern Ireland ($n = 4$), Scotland ($n = 2$) and England ($n = 4$). Participants preferred face-to-face evaluations by ophthalmologists, where information about their eye condition could be received and anxieties assuaged. In the absence of ophthalmologists, participants voiced the need for immediate results from the grader's reading of images. Patients wanted periodic evaluation by ophthalmologists, even if at longer intervals. Patients are uncertain of the professional identity, training and performance of graders. Graders and ophthalmologists were supportive of the new pathway, but graders expressed caution about their ability to answer potential questions from patients unrelated to the activity of their disease.

The cost-consequences analysis, in monetary terms, showed that for diabetic macular oedema, where sensitivity was very good, the cost-difference (savings) for the grader's pathway would be £1390 per 100 patients. For proliferative diabetic retinopathy, if sensitivity was considered acceptable, the cost would be reduced by £461 for seven-field Early Treatment Diabetic Retinopathy Study images and by £1889 for ultra-wide field images, per 100 patients. The difference arises because ultra-wide images require less time to be obtained and read than seven-field Early Treatment Diabetic Retinopathy Study.

Conclusions

The sensitivity of the new grader's pathway to determine diabetic macular oedema was 94% or above in all analyses suggesting that, for diabetic macular oedema, the new ophthalmic grader pathway would be safe. The sensitivity to determine proliferative diabetic retinopathy was 80% or above in all planned analyses with one exception (referrals for active proliferative diabetic retinopathy only, excluding 'unsure' and 'ungradable'). This level of sensitivity, although potentially less than ideal, may be acceptable for patients previously treated with laser panretinal photocoagulation, for whom this new pathway is being proposed. Where waiting targets are not achieved and people with serious eye conditions are waiting longer than acceptable, this new pathway would be reasonable and justified. Ultra-wide field imaging had slightly higher sensitivity to detect proliferative diabetic retinopathy with pre-retinal/vitreous haemorrhage (i.e. high-risk) and was less costly than seven-field Early Treatment Diabetic Retinopathy Study, thus, would be the preferred option. Despite the relatively low specificity of the new ophthalmic grader pathway, this pathway would save ophthalmologists' time that could then be redirected to a more timely evaluation and treatment of patients requiring urgent care, which would likely be sight-saving for many patients. Recommendations from the focus groups should be followed if the new pathway is to be introduced to ensure acceptability to users.

Recommendations for future research

For proliferative diabetic retinopathy, a pilot study could be run prior to its widespread implementation in the NHS, to further ensure its safety. Measures to further improve the sensitivity of this pathway (e.g. providing continuous feedback to graders, selecting patients entering the pathway, enhancing resolution of screens used to view images) could be introduced and evaluated.

Could artificial intelligence be used for automated reading of images in this previously treated population?

Trial registration

This trial is registered as ISRCTN10856638 and ClinicalTrials.gov NCT03490318.

Funding

This project was funded by the National Institute for Health Research (NIHR) Health Technology Assessment programme and will be published in full in *Health Technology Assessment* Vol. 25, No. 32. See the NIHR Journals Library website for further project information.

Chapter 1 Background

Parts of this report have been reproduced with permission from our published protocol: Lois *et al.*¹ This is an open access article distributed in accordance with the Creative Commons Attribution 4.0 Unported (CC BY 4.0) license, which permits others to copy, redistribute, remix, transform and build upon this work for any purpose, provided the original work is properly cited, a link to the licence is given, and indication of whether changes were made. See: <https://creativecommons.org/licenses/by/4.0/>. The text includes minor additions and formatting changes to the original text.

Parts of this report have been reproduced from Lois *et al.*² This is an open access article distributed in accordance with the terms of the Creative Commons Attribution 4.0 (CC BY-NC-ND 4.0) license, which permits others to copy and redistribute the material in any medium or format, provided the original work is properly cited. See: <https://creativecommons.org/licenses/by/4.0/>

Diabetic retinopathy (DR) is the most common microvascular complication of diabetes and a leading cause of sight loss in people of working age.³ Patients with DR may lose their vision as a result of the development of diabetic macular oedema (DMO) and/or proliferative diabetic retinopathy (PDR), which are the major complications of DR. DMO and PDR can happen in the same person and even in the same eye.

Diabetic macular oedema

Diabetic macular oedema is caused by the accumulation of fluid, often accompanied by lipid (and occasionally blood), in the central part of the retina, the macula, which is responsible for the generation of detailed central vision. Thus, DMO results in people progressively losing central sight and experiencing difficulties reading, recognising faces and doing any work for which detailed central vision is required (*Figure 1*).

In 2010, the prevalence of DMO in England was estimated to be 7% of all people with diabetes.⁴ An individual participant data meta-analysis that included 22,896 individuals from 35 studies conducted in Asia, Australia, Europe and the USA found a very similar estimate of prevalence of DMO, with an overall age-standardised prevalence of 6.8%.⁵ Based on these estimates and considering the prevalence of diabetes in the UK,⁶ it can be presumed that there are a minimum of 273,000 people in the UK affected by DMO.

In the UK, treatments for DMO include focal or grid macular laser, eye injections of anti-vascular endothelial growth factor (anti-VEGF) therapies and eye injections of steroids. All treatments aim to dry the macula and improve or maintain vision. Macular laser is advised and offered when the central retinal thickness (CRT), measured by means of spectral domain optical coherence tomography (SD-OCT) is < 400 µm, which may be considered mild DMO. Patients in whom there is drop-out of perifoveal capillaries contiguous to the area of leakage causing DMO, as determined by fundus fluorescein angiography, are not good candidates for macular laser. Patients with more severe DMO (CRT on SD-OCT of ≥ 400 µm) are eligible to receive anti-VEGFs; currently, the anti-VEGFs approved by the NHS are ranibizumab (Lucentis®; Novartis Pharmaceuticals UK Ltd, London, UK) and aflibercept (Eylea®; Bayer plc, Reading, UK) (*Figure 2*).^{7,8} Bevacizumab (Avastin®; Genentech Inc., South San Francisco, CA, USA) is also available off label. Intraocular steroids are reserved for patients with DMO who do not respond to macular laser or anti-VEGFs and are pseudophakic (i.e. have had their cataracts removed).⁹

Macular laser is delivered in a single session. Re-treatments may be required to clear DMO and are usually given at no earlier than 3- to 4-month intervals. Anti-VEGF injections are given monthly until the macula is dry (i.e. DMO clears). Once patients have been treated successfully (i.e. DMO is no longer present), long-term follow-up is required, as DMO may recur. Typically, once dried, patients are followed up every 3–4 months following laser treatment for DMO; patients are followed up monthly initially and every 1–3 months thereafter following treatment with anti-VEGFs.¹⁰ Follow-up continues for the rest of the patient's life.

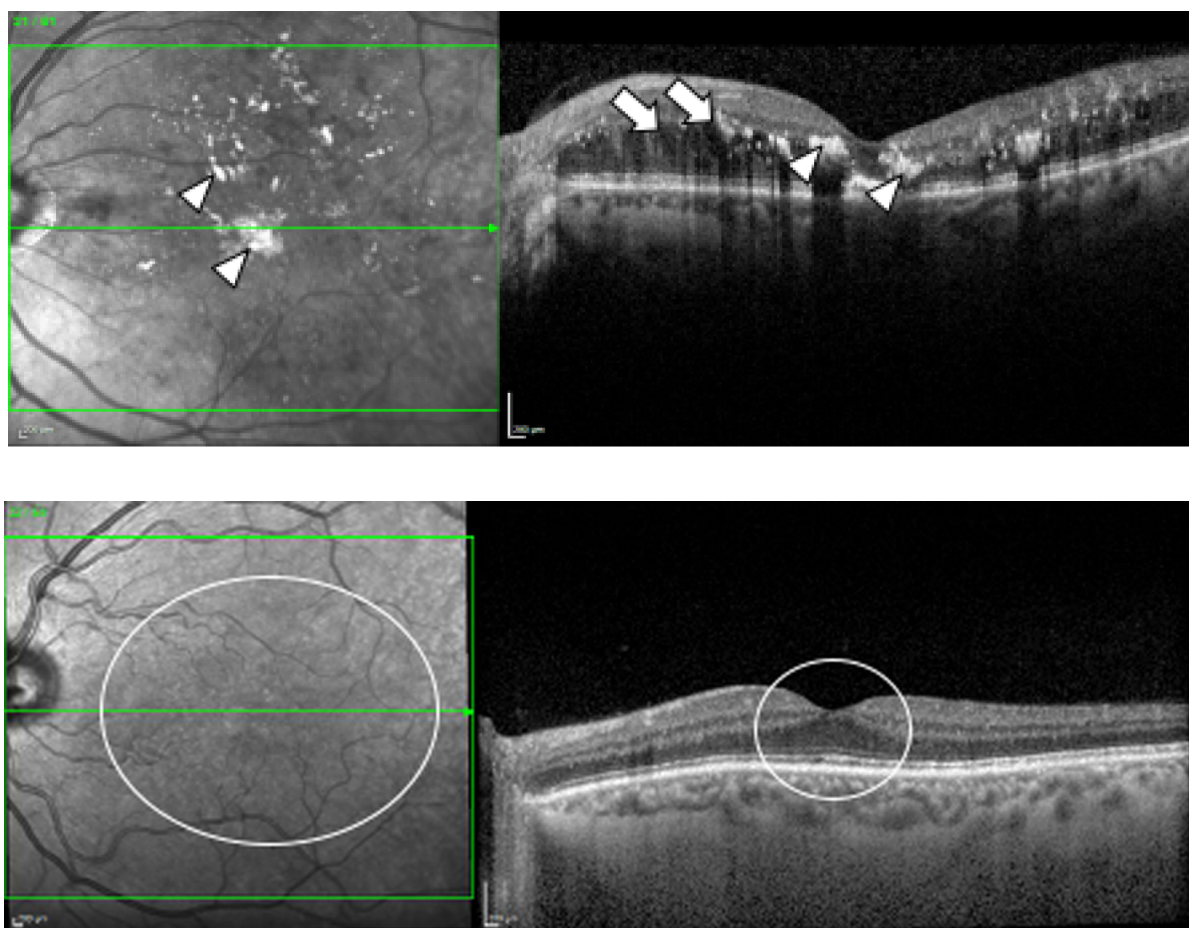


FIGURE 1 Spectral domain-optical coherence tomography in DMO. Fundus image (left) and spectral-domain optical coherence tomography scan (right) of the left eye of an individual with diabetes and DMO (top) and, for comparison, of someone with diabetes but no DMO (bottom). Lipid (the so-called 'hard exudates') (arrowheads; hard exudates appear as white areas on the images) and fluid (arrows; fluid appears as black areas on the image) are present in DMO (top). The normal structure of the macula (the central area of the retina, marked approximately with the circle on the bottom image, left) and of the fovea (the central area of the macula, marked approximately with a circle in the bottom image, right) is observed in the eye with no DMO (bottom).

Currently in the NHS, patients with previously treated DMO are evaluated in clinic during follow-up appointments with ophthalmologists who have expertise in retinal diseases, even if DMO is no longer present. At each visit, patients receive a visual acuity test, which is most often undertaken by a nurse or a visual acuity technician, before having a SD-OCT scan, obtained by a photographer or by an imaging technician, and then being seen by an ophthalmologist. The ophthalmologist checks the back of the eye using slit-lamp biomicroscopy and, most often, a non-contact fundus lens, and determines, also using SD-OCT scans, whether DMO is present or absent (*Figure 3*). SD-OCT is a non-invasive, user-friendly, fast and safe imaging technology that obtains scans of the back of the eye. SD-OCT allows the CRT to be measured (which is often increased when DMO is present) and visualises fluid in the retina, which is the hallmark of DMO (see *Figures 1* and *2*). SD-OCT has been extensively used in clinical trials and clinical practice to determine the presence of DMO, select treatment and monitor the patient's response to treatment.¹¹⁻¹⁶

Proliferative diabetic retinopathy

In PDR, abnormal, newly formed blood vessels ('new vessels') grow on the optic nerve head [the so-called new vessels in the disc (NVD)] and/or on the surface of the retina [the so-called new vessels elsewhere (NVE)] and towards the inside of the eye (i.e. the vitreous cavity) (*Figure 4*). These blood vessels may bleed, causing what is called a vitreous haemorrhage. They can also lead to the formation of scarring

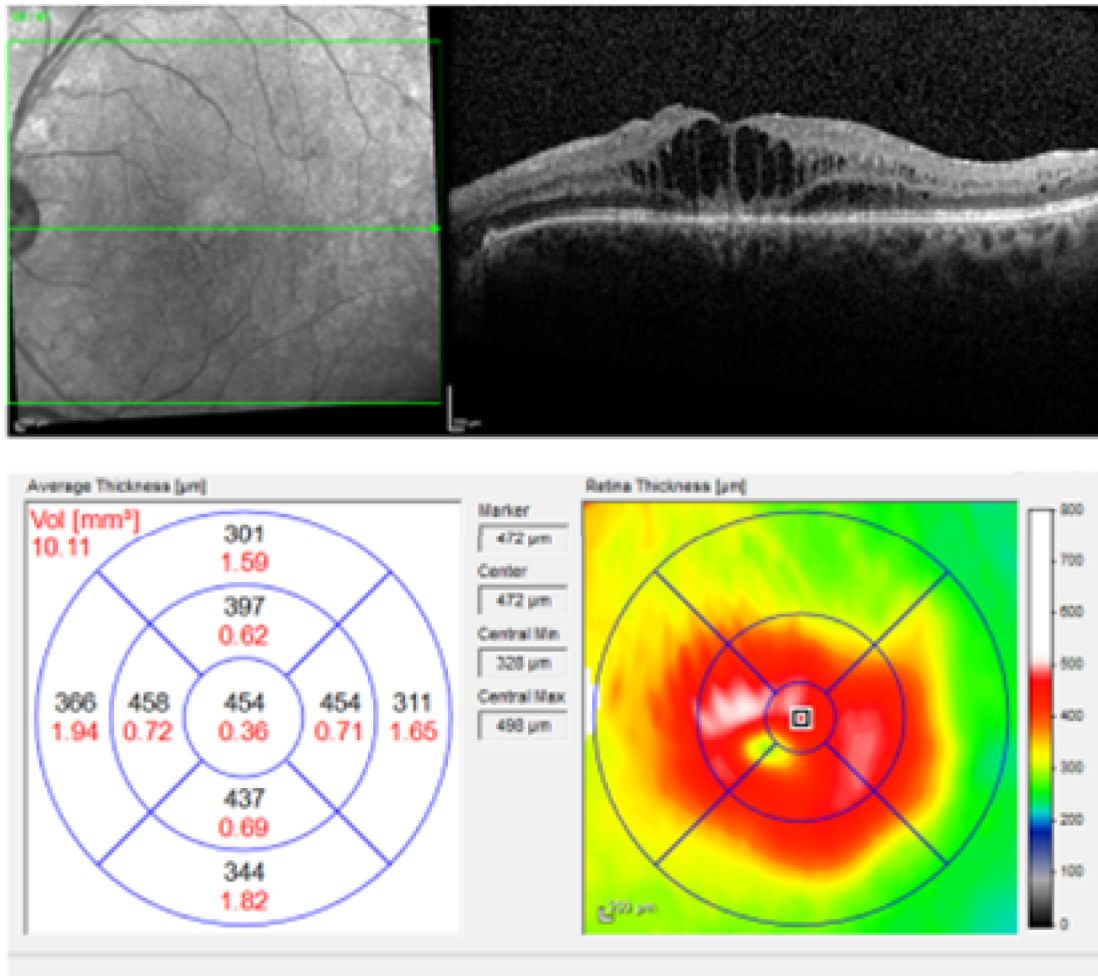


FIGURE 2 Spectral domain-optical coherence tomography in DMO with CRT of $> 400 \mu\text{m}$. Fundus image (top left) and SD-OCT scan (top right) of the left eye of a patient with severe DMO, with $> 400 \mu\text{m}$ CRT ($454 \mu\text{m}$ in the central 1 mm field, bottom left). Thickening of the retina is shown in red in the colour map (bottom right).

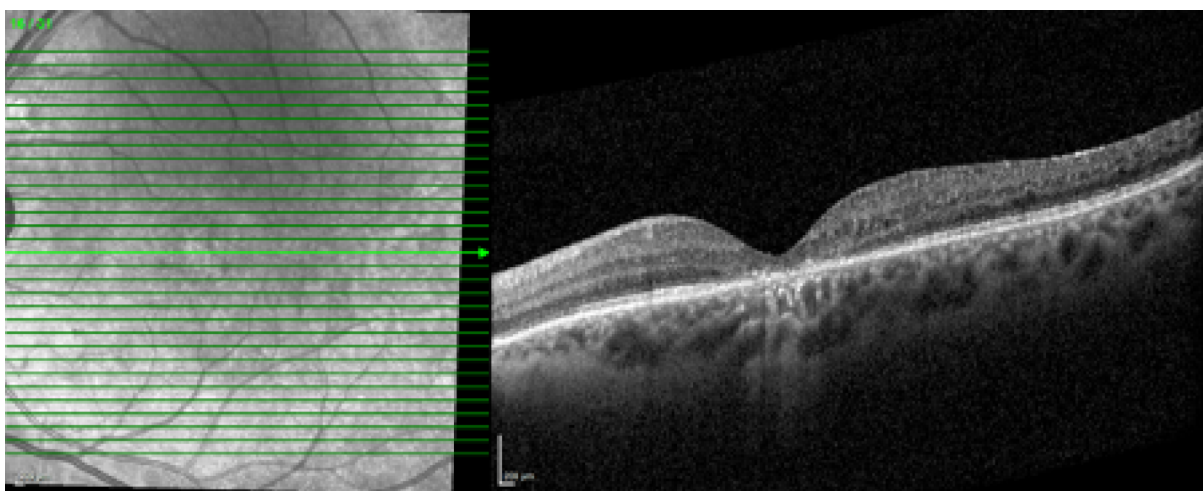


FIGURE 3 Spectral domain-optical coherence tomography following resolution of DMO. Fundus image (left) and SD-OCT scan (right) of the left eye of the patient shown in Figure 2 following anti-VEGF treatment. The DMO has resolved (no DMO present).

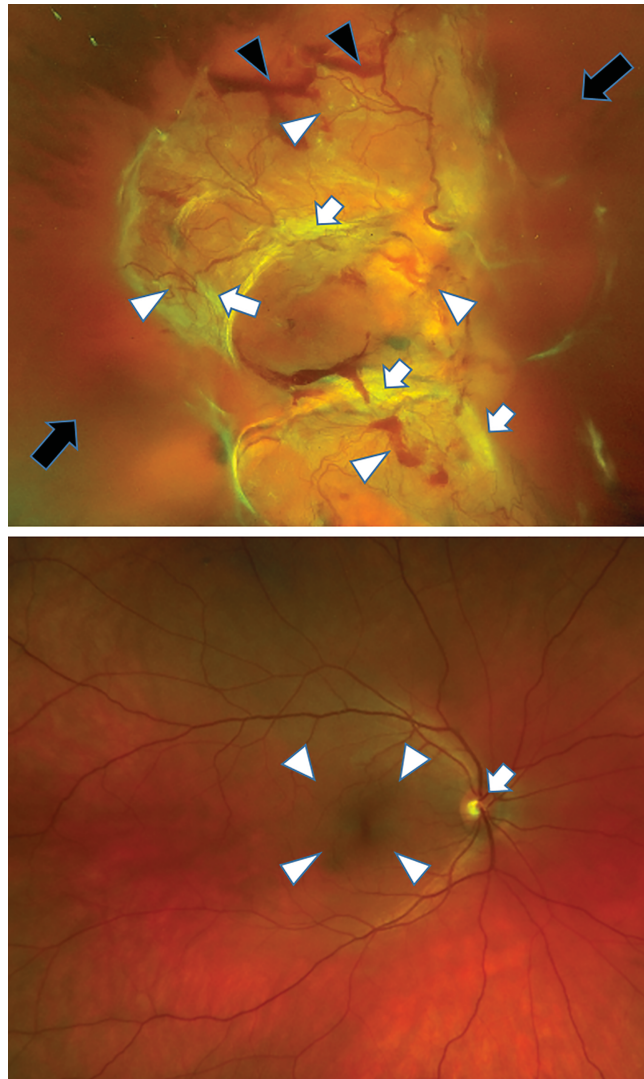


FIGURE 4 Fundus images of an eye with HRC PDR and a healthy eye. Fundus image of the right eye (top) of a patient with PDR and very large new vessels (white arrowheads) and scarring (white bands seen on the image, white arrows), pre-retinal haemorrhages (black arrows) and vitreous haemorrhage (black arrowheads) obscuring the view of the retina. For comparison, a fundus image of a healthy eye is shown (bottom). The optic nerve (white arrow) and macular area (white arrowhead) are healthy.

tissue that could then contract and pull on the retina and detach it from the wall of the eye, causing a tractional retinal detachment (TRD). Both vitreous haemorrhage and TRD can lead to loss of not only central but also peripheral vision. Once it occurs, vitreous haemorrhage may clear spontaneously or may require surgery (vitrectomy). If TRD affects or threatens the macula it will require surgery, which is often very challenging. Furthermore, visual outcomes following TRD repair may be disappointing.

The estimated prevalence of PDR in the individual participant data meta-analysis referred to above was 6.96%.⁵ Based on the prevalence of diabetes in the UK,⁶ it could be assumed that there are $\approx 271,440$ people in the UK affected by PDR.

The Early Treatment Diabetic Retinopathy Study (ETDRS)¹⁷ demonstrated the value of scatter laser panretinal photocoagulation (PRP) for the treatment of PDR. Following the results of the ETDRS, immediate treatment is recommended for all patients with type 1 or type 2 diabetes when high-risk characteristics are present, including NVD of a certain size (more than one-third to one-quarter of the disc area) and/or any new vessels if pre-retinal haemorrhage or vitreous haemorrhage are detected.¹⁷ Treatment at earlier stages could be considered in selected cases (e.g. poor attendance at clinics) and

especially in people with type 2 diabetes. The cost-effectiveness of applying PRP at earlier stages than the presence of high-risk characteristics (i.e. when there is PDR but not reaching high-risk characteristics and when there is severe non-proliferative disease and before new vessels develop) is unclear, and cannot be recommended for all patients at present.¹⁸ PRP is delivered as an outpatient procedure, under topical anaesthetic or sub-Tenon's (local) anaesthesia. Most often, treatment is completed in two sessions. Laser PRP is usually undertaken following ETDRS guidelines. For some time, an argon laser was the only laser type used to perform PRP, but recently other types of laser, including the so-called 'pattern' laser, have been introduced. However, there is no evidence to date of the efficacy and safety of any lasers other than the argon laser, and of any laser protocols/strategies other than the ETDRS.¹⁹ In this regard, it is important to note that a recent, large, well-conducted randomised controlled trial (RCT) comparing laser PRP with anti-VEGF therapies for the treatment of PDR found that, in the PRP arm, eyes receiving 'pattern' PRP were at a higher risk of worsening PDR [60% vs. 39%; hazard ratio (HR) 2.04, 99% confidence interval (CI) 1.02 to 4.08; $p = 0.008$] than those receiving conventional single-spot PRP laser.²⁰

The National Institute for Health and Care Excellence (NICE) is expected to appraise anti-VEGFs for the treatment of PDR in the next year or two. Anti-VEGFs are not currently used to treat PDR in the UK.

Once scatter laser PRP is complete, patients require follow-up (*Figure 5*). Initially, immediately after treatment, patients are followed up at a shorter time interval (e.g. 4–8 weeks) to ensure that regression of the disease occurs. However, once the disease is stabilised and regression occurs, patients are often followed up every 6–8 months or even at longer intervals. Follow-up is required as new vessels in PDR could return and vitreous haemorrhage and TRD could still ensue. Closer follow-up should be considered if PRP is performed using a 'pattern' laser, given the information provided above.

Currently in the NHS, most patients with treated and stable PDR are followed up in clinic by ophthalmologists. In fact, it has been shown that a high proportion of all patients with DR reviewed in hospital eye services have treated and stable PDR.²¹ At each follow-up visit, patients receive a visual acuity test, which is most often undertaken by a nurse or a visual acuity technician, followed by a SD-OCT scan, which is obtained by a photographer/imaging technician, to rule out the presence of concomitant DMO; they are then seen by an ophthalmologist, an expert on retinal diseases. The ophthalmologist examines the patient by slit-lamp biomicroscopy using, most often, a non-contact fundus lens, and determines whether there is active PDR or the disease remains stable. Fundus photographs (photographs of the retina) are not routinely obtained in clinic to determine whether or not active PDR is present, although they may be taken in selected cases.

If required, fundus photographs can be obtained with standard cameras that provide 30- to 45-degree images to allow imaging of the central part of the retina, including the macula (*Figure 6*). To visualise other areas of the retina using these standard cameras, the patients are asked to look up, left, down and right so that images of the superior, nasal, inferior and temporal retina can be obtained, for example, when imaging a right eye. However, even when this is carried out, the peripheral retina cannot be imaged.

To date, most of the research studies on DR have used seven-field ETDRS imaging to document the status of the retina. Seven-field ETDRS imaging consists of seven fields that are located in the mid-peripheral retina as follows (*Figure 7*):

- field 1 – centred at the optic nerve head
- field 2 – centred at the fovea
- field 3 – temporal to the fovea
- field 4 – superotemporal
- field 5 – inferotemporal
- field 6 – superonasal
- field 7 – inferonasal.

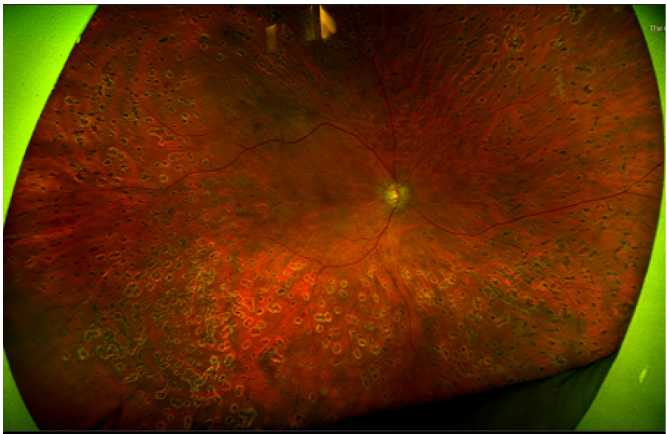


FIGURE 5 Fundus image of an eye treated with laser panretinal photocoagulation and with regressed PDR. Ultra-wide field fundus image of the right eye of a patient with PDR following laser panretinal photocoagulation. Laser scars (small, black areas surrounding by clearer 'halos') throughout the retina are seen.



FIGURE 6 Standard fundus image in a patient with DMO. Standard fundus photograph of the left eye of a patient with DR and macular oedema. With a single image, only the central retinal area is visualised.

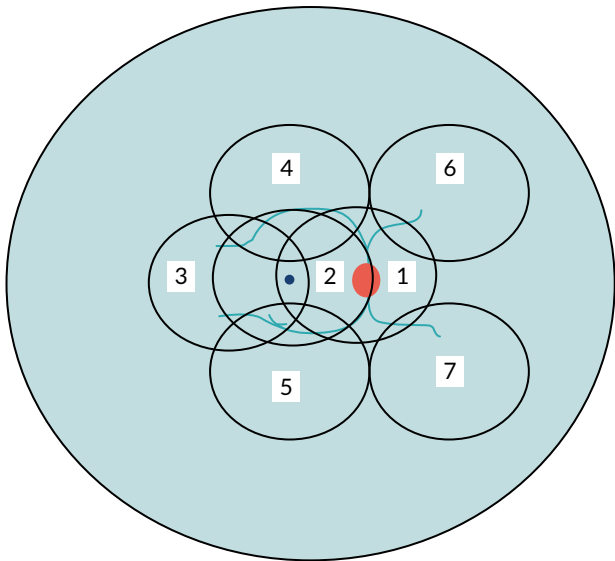


FIGURE 7 Scheme demonstrating the seven-field ETDRS fields.

Seven-field imaging takes some time to be obtained and, similarly, to be evaluated; for this reason, it is impractical and not used in routine clinics.

Furthermore, as shown in *Figure 7*, much of the retina remains uncovered (i.e. not photographed) when seven-field ETDRS fundus images are used. More recently, ultra-wide field imaging has become available. The first ultra-wide field fundus technology developed, which remains the widest field imaging system in existence, is produced by Optos (Dunfermline, UK). Using a scanning laser ophthalmoscope, a view of nearly the entire retina (200 degrees) can be obtained in a single image (*Figure 8*). To visualise the far peripheral superior and inferior retina as well, patients are asked to look up and down. With the three images (centre, superior and inferior) the whole retina is covered. Ultra-wide field imaging allows identification of DR lesions in the peripheral retina that would not be imaged if seven-field ETDRS was used instead; this could have important prognostic implications.²²

The burden of diabetic retinopathy, diabetic macular oedema and proliferative diabetic retinopathy and NHS pressures to meet service demands

Diabetes is on the rise. The estimated prevalence of diabetes in the UK was 3.9 million in 2019, over 100,000 more than in 2018.⁶ If nothing changes, 5 million people in the UK will have diabetes by 2025.⁶ The increasing number of people with diabetes will translate to a larger number of individuals with DMO and PDR. Given that patients are required to return to clinics at short intervals to receive treatment (until their disease has been controlled), and that these visits are required for the rest of

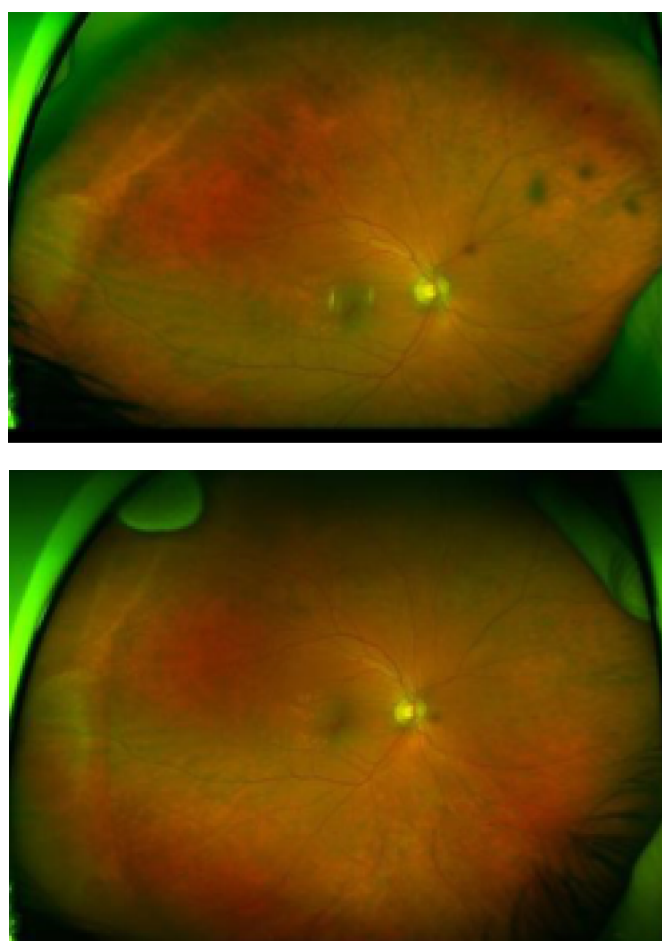


FIGURE 8 Ultra-wide field fundus images. Full visualisation of the superior (top), central (centre) and inferior (bottom) retina. (*continued*)

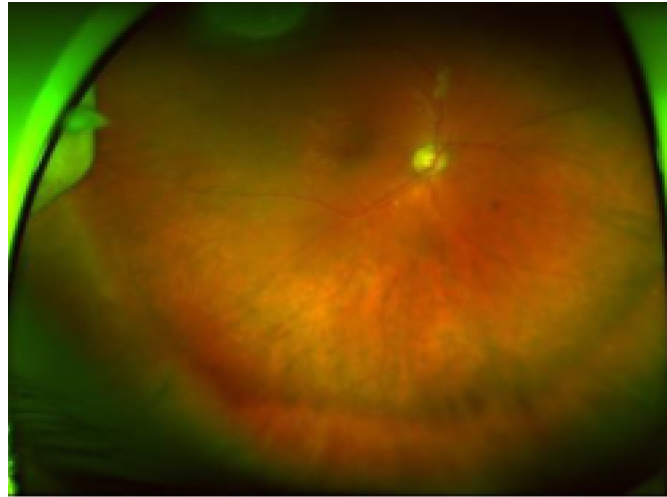


FIGURE 8 Ultra-wide field fundus images. Full visualisation of the superior (top), central (centre) and inferior (bottom) retina.

their lives, the already excessive workload in hospital eye services related to DMO/PDR is expected to worsen. This poses major problems for the ability of ophthalmic clinics to evaluate and treat patients in time, especially because of the shortage of ophthalmologists. Delaying appointments may subsequently lead to poorer visual outcomes for patients; such delays are widespread in the UK, as recently highlighted by the Royal College of Ophthalmologists.²³ Hospital eye services are further stretched by the fact that anti-VEGFs have been introduced not only for DMO but also for other conditions, including age-related macular degeneration and retinal vein occlusion, and frequent evaluation and treatment is required for these groups as well. Thus, it is imperative that new ways to increase efficiency and capacity in the NHS are identified and, if safe and acceptable, implemented.

With the above in mind, EMERALD (Effectiveness of Multimodal imaging for the Evaluation of Retinal oedema And new vesseLs in Diabetic retinopathy) was conceived with the purpose of determining whether or not patients who had previously received treatment for DMO and/or PDR and in whom treatment had been successful [i.e. DMO cleared and PDR became quiescent (inactive)] could be followed up with multimodal retinal imaging and review of these images by a trained ophthalmic grader, rather than by ophthalmologists, as is currently standard practice.

Chapter 2 Diagnostic accuracy

The purpose of EMERALD was to determine the diagnostic accuracy, acceptability to patients and health-care professionals, and cost-effectiveness of a new care pathway, the ophthalmic grader pathway, when compared with the current standard of care (standard care pathway) for the surveillance of people with previously successfully treated DMO and/or PDR.

Methods

EMERALD was designed as a case-referent, cross-sectional diagnostic study with both sampling (selection) of patients and data collection carried out prospectively.²⁴ This approach provides both a cost-efficient study design and a low risk of bias in terms of diagnostic accuracy.²⁵

At the time of study conception, a patient and public involvement (PPI) group was established. EMERALD was presented to the PPI group and the plans for the study discussed. The PPI group confirmed the research question was important and that the tests proposed were adequate and feasible to patients. The PPI group provided essential input to all patient-related materials elaborated for the study (patient information sheet and consent form). The PPI group will also be actively involved in the dissemination of the results.

Setting

EMERALD was conducted in 13 specialised hospital eye services in the UK. Participating sites were in England ($n = 11$), Scotland ($n = 1$) and Northern Ireland ($n = 1$). All participating sites had extensive experience in the management of people with DR, DMO and PDR.

Population

Patients with DR and DMO and/or PDR were eligible for EMERALD if they met the following inclusion criteria:

- Adults (aged ≥ 18 years) with type 1 or type 2 diabetes and with previously successfully treated DMO and/or PDR in one or both eyes, and in whom, at the time of enrolment in the study, DMO and/or PDR was active or inactive. Patients could be recruited into the study only once.

Active DMO was defined as a CRT of $> 300 \mu\text{m}$ on SD-OCT and/or by the presence of intraretinal/subretinal fluid on SD-OCT owing to DMO. Inactive DMO was defined as the lack of intraretinal/subretinal fluid at the macula on SD-OCT. Active PDR was defined as the presence of subhyaloid/vitreous haemorrhage and/or active new vessels (new vessels with a lack of fibrosis on them), whether in the disc (NVD) or elsewhere in the retina (NVE). Inactive PDR was defined as the lack of preretinal/vitreous haemorrhage and the lack of active NVD or NVE.

The exclusion criteria used were as follows:

- patients unable to provide informed consent
- patients unable to read, speak or understand English.

Pathways evaluated

Ophthalmic grader pathway

The new care pathway tested, the ophthalmic grader pathway, consisted of reviewing SD-OCT images to detect DMO and the review of seven-field ETDRS/ultra-wide field fundus images to detect PDR; this was carried out by trained, tested and certified ophthalmic graders (see *Selection of ophthalmic*

graders and training). Following evaluation of these images, graders determined whether there was active DMO/PDR, inactive DMO/PDR or whether they were unsure or unable to grade (ungradable), in which case patients would be referred to an ophthalmologist for assessment. If there was no DMO and/or inactive PDR only, the grader would arrange a review appointment for the patient in the ophthalmic grader pathway at a pre-determined interval.

Standard-of-care pathway

The standard-of-care pathway considered in EMERALD was the current standard of care, as follows:

- for people with DMO – ophthalmologist evaluating patients in clinic by slit-lamp biomicroscopy and with access to SD-OCT scans
- for people with PDR – ophthalmologist evaluating patients in clinic by slit-lamp biomicroscopy.

For the purpose of EMERALD, and owing to ethics considerations, all patients were first evaluated through the standard care pathway, as explained in *Patient flow in EMERALD*.

Enhanced reference standard for proliferative diabetic retinopathy

It is possible that ophthalmologists could miss new blood vessels when evaluating patients by slit-lamp biomicroscopy. To account for this, EMERALD also evaluated an ‘enhanced’ reference standard for PDR that consisted of the reference standard (i.e. evaluation of the patient by slit-lamp biomicroscopy by an ophthalmologist) supplemented by the evaluation of seven-field ETDRS and ultra-wide field fundus images, which were reviewed by an ophthalmologist who was an expert in DR. If active PDR was detected in one of these three evaluations (slit-lamp biomicroscopy, seven-field ETDRS fundus images and ultra-wide field fundus images), it was considered that there was active PDR based on the enhanced reference standard (ERS). To obtain masking to the reference standard, ophthalmologists did not grade images from their own centre. Seven-field ETDRS and ultra-wide field images of the same participant were not reviewed by the same ophthalmologist to avoid the reading of one imaging technology influencing the reading of the other.

Outcome measures

Primary outcome

The primary outcome was the sensitivity of the new pathway (ophthalmic grader pathway) in detecting active DMO/PDR, using the standard care pathway as the reference standard.

Secondary outcomes

The secondary outcomes were:

- specificity, concordance (agreement) between the new pathway (ophthalmic grader pathway) and the standard care pathway, and positive and negative likelihood ratios
- cost-effectiveness
- acceptability
- proportion of patients requiring subsequent full clinical assessment
- proportion of patients unable to undergo imaging, with inadequate quality images or indeterminate findings.

Selection of ophthalmic graders and training

Currently, ophthalmic photographers/imaging technicians obtain images and interpret them routinely, but make no decisions on the care of patients. In ophthalmic services, there are also ophthalmic graders that have been trained to interpret findings on fundus images for the purpose of undertaking DR screening.

For EMERALD, the ophthalmic graders at each participating site were selected as follows. To start, local principal investigators (PIs) provided names of individuals that they believed had experience of obtaining and/or grading images of patients with DMO and PDR; these individuals were approached to confirm their interest and willingness to participate in EMERALD. Some ophthalmic graders who were selected for EMERALD were already involved in the grading of images for DR screening.

Graders identified by the PIs were asked to fill in questionnaires detailing their experience of imaging/grading DMO/PDR, their experience of recognising features of DMO/PDR and whether or not they felt confident identifying DMO on SD-OCT and new vessels on fundus images. Graders who stated that they did not have experience of imaging/grading DMO/PDR and those stating that they could not recognise features of DMO/PDR were not invited to take part in EMERALD as graders.

Formal training was then provided to potential EMERALD graders; during training sessions, features of active/inactive DMO/PDR were reviewed and discussed and extensive clinical examples were presented. This training was provided to all EMERALD ophthalmic graders prior to the initiation of the study. Thus, prior to the initiation of the study, a 2-day face-to-face course was provided to potential graders. This was followed by two additional half-day training sessions.

A web-based teaching module with examples of DMO/PDR was prepared so that graders could consolidate their knowledge. Clear guidelines on when patients would need to be referred for an assessment by an ophthalmologist were also given.

To ensure that the graders who were selected had the required level of experience to undertake the task of grading images, all potential graders were required to take a test involving the reading of optical coherence tomographies (OCTs) and fundus images. Only graders who reached a minimum of 80% of correct answers in detecting the presence of DMO (active DMO) and active PDR, when present, were invited to act as graders. Graders were allowed to undergo further training and take the test a second time, but if the minimum number of correct answers was not reached at this second test they were not selected to be graders for EMERALD and did not carry out this role for the study.

Patient flow in EMERALD

Patients with previously successfully treated DMO/PDR were identified from clinical records, electronic databases or in clinic. Verbal and written information about the study was given to all potential participants and their questions, if any, were answered. At their review appointment, an ophthalmologist confirmed patient eligibility, obtained informed consent and determined whether or not DMO was present and whether there was active or inactive PDR, setting the reference standard. All participants underwent visual acuity testing, SD-OCT scans and fundus examination, as undertaken per routine standard practice. In some participating sites, 159 (40%) patients were evaluated in 'research' clinics, and in others 238 (60%) patients were evaluated fully within usual NHS clinics.

Once the reference standard was set, non-stereoscopic seven-field ETDRS and ultra-wide angle fundus images were obtained. These images were then coded (identifiers removed), uploaded to a central facility [the Central Administrative Research Facility (CARF), Queen's University, Belfast] and allocated randomly to the EMERALD ophthalmic graders. Graders did not grade images from their own centre (to ensure masking to the reference standard) and did not grade seven-field ETDRS/ultra-wide fundus images from the same patient to prevent the grading of one technology influencing the grading of the other. For each patient/imaging modality, graders judged whether there was active/inactive DMO and/or active/inactive PDR, or if they were uncertain about this. Graders also determined, based on their findings, whether patients needed referral to an ophthalmologist for review/treatment or they could continue their care in the ophthalmic grader pathway.

Schedule of assessments for EMERALD

Case report forms (CRFs) were specifically designed to collect all of the information required for the purpose of the EMERALD study. The following information was obtained during the standard care pathway and recorded:

- medical and ophthalmic history, including previous treatments for DMO/PDR
- best-corrected visual acuity
- ophthalmologists' findings on slit-lamp biomicroscopy, including:
 - presence of anterior segment neovascularisation and possible determinants of poor-quality images, including media opacity and small pupillary size
 - presence/absence of DMO and/or active/inactive PDR
 - presence/absence and location of active/inactive NVD and NVE in the retina and/or pre-retinal haemorrhage/vitreous haemorrhage
 - presence of any other co-existent eye disease (e.g. glaucoma)
 - proposed plan for the patient (observation or treatment).

While in clinic, patients completed the EuroQol-5 Dimensions, five-level version (EQ-5D-5L), National Eye Institute Visual Function-25 (NEI VFQ-25) and Vision and Quality of Life (VisQoL) questionnaires. Those willing to participate in the focus group (FG) discussions were asked for their consent and informed that they might be contacted at a later date for this purpose.

Patients underwent seven-field ETDRS and ultra-wide field imaging. Images were obtained following the guidelines provided in the standard operational procedures, as set in the EMERALD study manual. Once obtained, fundus photographs and SD-OCT images were anonymised and transferred to the CARF of Queens' University Belfast, where they were uploaded to an electronic website developed for the study and randomly allocated and made accessible to the EMERALD ophthalmic graders and ophthalmologists determining the ERS for PDR. Images (SD-OCTs for graders and seven-field ETDRS and ultra-wide field fundus images for graders and ophthalmologists) were viewed in an ophthalmic viewer platform (Ophthalsuite, BlueWorks, Coimbra, Portugal) and graded by ophthalmic graders and ophthalmologists.

Ophthalmic graders determined and recorded in the appropriate CRF:

- whether there was active/inactive DMO/PDR or if they were unsure or unable to grade presence of the disease (ungradable)
- the presence/absence, location and activity of new vessels or whether they were unsure or unable to grade presence of the new vessels (ungradable)
- the presence/absence and degree of completeness (partial or complete) of previous laser PRP
- the presence/absence of pre-retinal haemorrhage and vitreous haemorrhage
- the presence and type of other abnormalities, if observed
- whether the patient could continue surveillance in the ophthalmic grader pathway or if an assessment by an ophthalmologist was required and the reasons why.

Similarly, ophthalmologists who were determining the ERS determined and recorded the following information in the appropriate CRF:

- whether there was active/inactive PDR or if they were unsure or unable to grade presence of the disease (ungradable)
- presence/absence, location and activity of new vessels or whether they were unsure or unable to grade presence of the new vessels (ungradable)
- presence/absence and degree of completeness (partial or complete) of previous laser PRP
- presence/absence of pre-retinal haemorrhage and vitreous haemorrhage
- presence and type of other abnormalities, if observed
- whether or not the patient could continue surveillance in the ophthalmic grader pathway or if an assessment by an ophthalmologist was required and the reasons why.

Masking

Ophthalmic graders and ophthalmologists evaluating images for the ERS were masked to the results of the reference standard. To ensure this, graders and ophthalmologists did not interpret images from patients recruited in their own centres; they did not have access to the results of the reference standard. Ophthalmologists undertaking the standard-of-care evaluation were also masked to the findings/decisions made by the ophthalmic graders (who reviewed the images at a later date). Patients were also masked to the findings/decisions made by the ophthalmic graders (these were not available at the time of the clinical visit for the study).

Data collection and quality checks

As stated above, CRFs were specifically prepared for the purpose of EMERALD and used to collect study data. Monitoring was undertaken during the study to check the accuracy of entries in CRFs against source documents, adherence to protocol and procedures and adherence to the International Conference of Harmonisation Good Clinical Practice guidelines²⁶ and regulatory requirements. Monitoring visits were carried out by a monitor from the Northern Ireland Clinical Trials Unit (NICTU). Training on the protocol and study procedures was provided by the chief investigator and the NICTU to research staff at all participating sites prior to the initiation of the study.

Study data were transferred from CRFs to a web-based clinical trial database, which was elaborated for the purpose of EMERALD, by NICTU personnel and processed electronically. Data quality-control checks were carried out by a data manager to ensure accuracy. Data errors were documented in quality control reports and corrective actions implemented. Data validation and discrepancy reports were generated following data entry to identify discrepancies, such as data out of range, inconsistencies or protocol deviations, based on the data validation checks programmed in the clinical trial database.

Sample size

The sample size was determined by setting a target number of people with reactivated (active) DMO and PDR required to enable sensitivity to be tested against a prespecified target level of 80%. The required sample size was calculated using formula T1 from Obuschowski²⁷ in Microsoft Excel® (Microsoft Corporation, Redmond, WA, USA); this was a Wald test-based calculation. This target level was considered to be the minimum acceptable for the new pathway (ophthalmic grader pathway) to be clinically viable. A lower specificity was considered acceptable; a target level of 65% was used to confirm sufficiency of the sample size to assess specificity. Eighty-nine participants with DMO/PDR that had reactivated (active DMO/PDR) was sufficient to detect if the sensitivity of the new pathway is 10% and 12% higher than the 80% minimal target set with 80% and 90% power, respectively, at the two-sided 5% significance level.²⁸ Ninety-three participants who have not reactivated would enable detection with a specificity 15% higher than the 65% target with 90% power. A 95% confidence interval for the sensitivity and specificity of the ophthalmic grader pathway would have a confidence interval (Wilson method) with a width of 10–20% depending on the observed level.²⁹ Allowing for 10% of missing/indeterminate results, 104 individuals who have reactivated and 104 who have not reactivated were required (208 individuals for each DMO and PDR), which led to a need for a maximum of 416 participants in the study. Because some participants had both DMO and PDR, and they contributed to both DMO and PDR targets, the number of patients required was smaller than 416.

Statistical analyses

Overview of principal analyses

The statistical analyses were conducted in accordance with the study protocol¹ and the statistical analysis plan, which was signed off and made publicly available on the EMERALD website (www.nictu.hscni.net/emerald-trial/; accessed March 2020) prior to data analyses. Additional analyses conducted beyond those planned for in the statistical analysis plan are identified as such in this chapter (as well as in *Chapter 4*). To address the primary objective, the principal analyses were carried out with DMO and PDR patients assessed in two separate sets of analyses at the person level, one for each disease (see *Table 2*). Analyses were focused

on the participant being considered eligible for the new pathway by virtue of having previously successfully treated DMO/PDR in one or both eyes ('Patients eligible for new pathway'). The main analysis was according to the ophthalmic graders decision to refer to an ophthalmologist (irrespective of the reason, i.e. active disease or if the grader was unsure or unable to grade; referred to throughout this chapter as 'referral' = 'active' + 'unsure' + 'ungradable'); this would be what would occur if the ophthalmic grader pathway were to be implemented. The reference standard was standard care: slit-lamp biomicroscopic examination by an ophthalmologist [referred to throughout this chapter as ophthalmologists face-to-face (O-FTF)] with access to SD-OCT images (referred to as O-FTF + OCT) for DMO, and slit-lamp biomicroscopic examination by ophthalmologist (O-FTF) for PDR to detect the presence of active disease (active DMO or active PDR) in either eye. For PDR only, an ERS was used for some analyses. This consisted of the combined findings of ophthalmologist slit-lamp biomicroscopic examination (O-FTF), ultra-wide field fundus images (referred to throughout this chapter as O-OPTOS) and seven-field ETDRS fundus images (referred to throughout this chapter as O-ETDRS). Positive detection of active PDR constituted a positive ERS result. The diagnostic performance of the new ophthalmic grader pathway was quantified and compared with that of the standard care pathway. For PDR, there were two sets of results: one using ultra-wide field fundus imaging-based assessment (referred to as OPTOS) and one using the seven-field ETDRS-based assessment (referred to as ETDRS). The impact of using either OPTOS or ETDRS on the diagnostic performance of the new pathway was formally compared.

Table 1 summarises all of the analyses undertaken in EMERALD. *Tables 2* and *3* summarise the planned principal analyses (main and sensitivity analysis) and the unplanned additional principal analyses undertaken.

Eligible participants for the EMERALD ophthalmic grader's pathway

Eligible participants contributed to the principal analyses under the new pathway for DMO (see *Table 2*) if they had at least one eye with DMO that was eligible (i.e. previously successfully treated). Similarly, eligible participants contributed to the principal analyses under the new pathway for PDR (see *Tables 2* and *3*) if they had at least one eye with PDR that was eligible (i.e. previously successfully treated).

Participants in EMERALD, however, could have one or both eyes with 'active DMO' or 'active PDR' ('recurrent', 'de novo' or 'persistent'; see below for definitions), or 'inactive DMO' or 'inactive PDR', or 'no DMO' or 'no PDR' according to the diagnosis established at the standard care pathway.

Definitions

'Active, recurrent DMO/PDR' refers to previously successfully treated DMO/PDR that was present at the time of the EMERALD examination.

'Active, de novo DMO/PDR' refers to DMO/PDR never present before but observed at the time of the EMERALD examination.

'Active, persistent DMO/PDR' refers to DMO/active PDR that was present before, never cleared (e.g. unsuccessfully treated) and was present at the time of the EMERALD examination.

'Inactive DMO/PDR' refers to inactive disease at the time of the EMERALD examination.

Note that some eyes with inactive DMO/PDR would have been previously successfully treated and would be considered eligible eyes; others, however, may not have had previously successfully treated disease, but at the time the patient attends the EMERALD visit their disease may be inactive and, thus, would be considered ineligible eyes (see *Eye-level classification*).

'No DMO/PDR' refers to DMO or PDR never present and not detected at the time of the EMERALD examination.

TABLE 1 List of analyses

Disease(s)	Image(s)	Population	Level of analysis	Analysis ^a	Location in report		
					Figure	Table number(s)	Appendix 1
DMO	SD-OCT	Eligible patients for new pathway	Person	Main analysis SENA1 SENA2 SENA3 SENA6	Figure 10	Tables 2, 7, 9, 12 and 13	
PDR	UWF OPTOS and seven-field ETDRS	Eligible patients for new pathway	Person	Main analysis SENA1 SENA2 SENA4 SENA5 SENA6 Additional 1 Additional 2 Additional 3	Figure 11 Figure 11 Figure 11	Tables 2, 8, 10, 14, 15, 16 and 17 Tables 2, 8, 10, 14 and 15 Tables 3, 8, 10, 14 and 15	
Combined diseases	SD-OCT and UWF OPTOS/ seven-field ETDRS	All patients	Person	SECA2A Additional 4 SECA2B SECA2C	N/A	Tables 5, 18 and 19	
DMO	SD-OCT	All patients	Eye	SECA1A SECA1B	N/A	Table 5	Table 41 Table 42
PDR	UWF OPTOS/ seven-field ETDRS	All patients	Eye	SECA1A SECA1C	N/A	Table 5	Table 41 Table 42

N/A, not applicable; SECA, secondary analysis; SENA, sensitivity analysis; UWF, ultra-wide field fundus images.
^a Details of each of these analyses are presented in Table 2.

Eye-level classification

The eye-level classification was based on the following six scenarios:

- ‘Scenario 1 – Previous Active, successfully treated, Today Active’ (corresponding to ‘recurrence’) – also defined as ‘Eligible eye’ for the analysis of new pathway.
- ‘Scenario 2 – Previous Active, successfully treated, Today Inactive’ (corresponding to ‘inactive’) – also defined as ‘Eligible eye’ for the analysis of new pathway.
- ‘Scenario 3 – Previous Active, not successfully treated, Today Active’ (corresponding to ‘persistence’) – also defined as ‘Ineligible eye’ for the analysis of new pathway.
- ‘Scenario 4 – Never present, Today Active’ (corresponding to ‘de novo’, i.e. newly diagnosed) – also defined as ‘Ineligible eye’ for the analysis of new pathway.
- ‘Scenario 5 – Previous Active, not successfully treated, Today Inactive’ – (corresponding to ‘inactive’) also defined as ‘Ineligible eye’ for the analysis of new pathway.
- ‘Scenario 6 – DMO/PDR never present, Today No DMO/PDR’ (i.e. no disease) – also defined as ‘Ineligible eye’ for the analysis of new pathway.

TABLE 2 List of the planned principal analyses (main and sensitivity analyses of diagnostic accuracy)^a

Analysis name	Level of analysis	DMO		PDR	
		Index test positive	Reference standard	Index test positive	Reference standard
Main	Person	OCT-based ophthalmic grader referral^b for either eye	O-FTF + OCT assessment of active DMO in either eye	OPTOS-based ophthalmic grader referral for either eye ETDRS-based ophthalmic grader referral for either eye	O-FTF assessment of active PDR in either eye
SENA1	Person	OCT-based ophthalmic grader identification of active disease in either eye	O-FTF + OCT assessment of active DMO in either eye	OPTOS-based ophthalmic grader identification of active disease in either eye ETDRS-based ophthalmic grader identification of active disease in either eye	O-FTF assessment of active PDR in either eye
SENA2	Person	OCT-based ophthalmic grader referral for either eye	O-FTF + OCT assessment of active DMO in either eye requiring treatment	OPTOS-based ophthalmic grader referral for either eye ETDRS-based ophthalmic grader referral for either eye	O-FTF assessment of active PDR in either eye requiring treatment
SENA3	Person	OCT-based ophthalmic grader identification of central-involving DMO in either eye	O-FTF + OCT assessment of central-involving DMO in either eye	N/A	N/A
SENA4	Person	N/A	N/A	OPTOS-based ophthalmic grader referral for either eye ETDRS-based ophthalmic grader referral for either eye	O-FTF assessment of active PDR with pre-retinal or vitreous haemorrhage in either eye
SENA5	Person	N/A	N/A	OPTOS-based ophthalmic grader referral for either eye ETDRS-based ophthalmic grader referral for either eye	Enhanced standard
SENA6	Person	OCT-based ophthalmic grader referral for either eye (participants assessed in routine clinic setting only)	O-FTF + OCT assessment of active DMO in either eye (participants assessed in routine clinic setting only)	OPTOS-based ophthalmic grader referral for either eye (participants assessed in routine clinic setting only) ETDRS-based ophthalmic grader referral for either eye (participants assessed in routine clinic setting only)	O-FTF assessment of active PDR in either eye (participants assessed in routine clinic setting only)

O-FTF + OCT, reference standard for DMO (i.e. standard care); O-FTF, reference standard for PDR (i.e. standard care); N/A, not applicable; SENA, sensitivity analysis.

a Principal analyses were focused on 'eligible patients for the new pathway' and referred to the analysis carried out using data from the eligible eye(s) of each patient included in EMERALD. [Note that all patients included in EMERALD were eligible for the purpose of evaluating one or other disease or both (DMO or PDR); they may have contributed with data from 'eligible eyes' if they have had DMO, PDR or both, in one or both eyes that were previously successfully treated, becoming stable at one point following treatment – see *Methods*.]

b Grader referral for DMO or PDR = 'active' + 'unsure' + 'ungradable'.

Bold highlights the differences between analyses.

TABLE 3 List of unplanned additional principal analyses of diagnostic accuracy^a

Analysis name	Level of analysis	DMO		PDR	
		Index test positive	Reference standard	Index test positive	Reference standard
Additional 1	Person	N/A	N/A	OPTOS-based ophthalmologist identification of active disease in either eye ETDRS-based ophthalmologist identification of active disease in either eye	O-FTF assessment of active PDR in either eye
Additional 2	Person	N/A	N/A	OPTOS-based ophthalmologist identification of active disease in either eye ETDRS-based ophthalmologist identification of active disease in either eye	O-FTF assessment of active PDR with pre-retinal or vitreous haemorrhage in either eye
Additional 3	Person	N/A	N/A	OPTOS-based ophthalmic grader referral^b for either eye ETDRS-based ophthalmic grader referral for either eye	O-FTF assessment + ophthalmologist OPTOS images of active PDR in either eye O-FTF assessment + ophthalmologist ETDRS images of active PDR in either eye

N/A, not applicable.

a Three additional analyses on PDR that were not listed in the statistical analysis plan were conducted in the light of the main PDR results to better understand the content and potential value of the two imaging modalities along with the reliability of the O-FTF reference standard for PDR.

b Grader referral for PDR = 'active' + 'unsure' + 'ungradable'.

Bold highlights the differences between analyses.

Participant-level classification

All participants who were included in the EMERALD study were at least eligible for the new DMO or the new PDR pathway, or eligible for both the DMO and the PDR pathways. Each eye for each EMERALD participant was assessed with regard to the status of PDR and DMO, irrespective of whether or not they were 'eligible' for both cohorts. What data were included varies according to the analysis (see *Tables 2 and 3* for a full list). The statement 'eligible participants' for the new pathway refers to analyses from data obtained from participants contributing with one or both eligible eyes (one or both eyes had to have previously successfully-treated DMO or PDR, respectively) (Scenario 1 or 2 in one or both eyes).

To be eligible for the new DMO pathway (see *Table 2*), participants needed to have at least one eye with previously successfully treated DMO (i.e. 'eligible'). At the time of the EMERALD examination, if DMO was present in one or both eyes they were classed as having active DMO ('recurrent DMO'), and if neither eye had DMO then participants were classed as having inactive DMO. Inactive DMO includes patients who have one previously successfully treated eye in which DMO is inactive at the time of examination and who have no DMO in the other eye. It also includes those for whom one or both eyes had been previously successfully treated for DMO, but both eyes were inactive at the time of examination. Participants who did not have an eye that had been previously successfully treated were ineligible for the new DMO pathway.

To be eligible for the new PDR pathway (see *Tables 2 and 3*), participants needed to have at least one eye with previously successfully treated PDR (i.e. 'eligible'). At the time of the EMERALD examination, if PDR was active in one or both eyes they were classed as having active PDR ('recurrent PDR'), and if neither eye had active PDR then participants were classed as having inactive PDR. Inactive PDR includes patients who had one previously successfully treated eye in which PDR is inactive at the time of examination and who had no PDR in the other eye. It also includes those in whom one or both

eyes had been previously successfully treated for PDR but both eyes were inactive at the time of examination. Participants who did not have an eye that had been previously successfully treated were ineligible for the new PDR pathway.

Definition of grader's assessments of optical coherence tomography, OPTOS and ETDRS

A grader's decision on whether or not to refer the patient, an inherently person-level decision, was used, rather than an assessment of the single eye. The basis of the referral decision was the grader's assessment of the corresponding disease (DMO/PDR). Furthermore, to more closely reflect how the new pathway would function in practice, a patient whom graders classified as having DMO or PDR within the eye under the category 'unsure' or 'ungradable' were considered alongside those classified as 'active', as both would be anticipated to require further examination by an ophthalmologist under the main analysis (referral = 'active' + 'unsure' + 'ungradable'). How the ophthalmic graders decision was tested varied according to the analysis (see *Tables 2 and 3* for a full list). *Table 4* shows the graders decision of person-level active status based on the combination of eye-level results. Only those samples in which there was an overall classification available were included in the sensitivity analyses.

Diagnostic performance of the EMERALD ophthalmic graders

The diagnostic performance of the EMERALD ophthalmic graders was assessed against the standard care reference standard to determine:

1. sensitivity (the proportion of patients determined by the ophthalmologist to suffer from active DMO/PDR who had been correctly referred/identified by EMERALD ophthalmic graders)
2. specificity (the proportion of patients determined by the ophthalmologist to suffer from inactive DMO/PDR who had been correctly referred/identified by EMERALD ophthalmic graders)
3. overall agreement (a measure of how well the ophthalmic graders' assessment agrees with the ophthalmologist assessment)
4. positive likelihood ratio (the probability that a patient with active DMO/PDR is correctly assessed by the ophthalmic graders, divided by the probability that a patient with inactive DMO/PDR is incorrectly assessed by the ophthalmic graders as being active)
5. negative likelihood ratio (the probability that a patient with active DMO/PDR is incorrectly assessed by the ophthalmic graders as being active, divided by the probability that a patient with inactive DMO/PDR is correctly assessed).

Agreement between assessment methods was also quantified, where appropriate.

TABLE 4 Combination of eye-level 'active' status to person-level 'active' status

Eye-level active status	Person-level active status	
	Main	SENA1/SENA3
'Both' eyes 'Active' or 'either' eye 'Active', no matter the status of the other eye	Referral ^a	Active
'Both' eyes 'Unsure' or 'Ungradable'	Referral ^a	Not Active
One eye 'Unsure' or 'Ungradable', while the other eye 'No' disease, 'Ungradable' or 'Unsure'	Referral ^a	Not Active
One eye 'Unsure' or 'Ungradable', while the other eye 'Missing'	Referral ^a	Missing
'Both' eyes 'No disease' or 'Inactive disease'	Not Active	Not Active
One eye 'No disease' while the other eye 'Inactive disease' and neither eye 'Missing'	Not Active	Not Active
'Both' eyes 'Missing', and none of the eyes would be referral	Missing	Missing
One eye 'No disease' or 'Inactive disease', while the other eye 'Missing'	Missing	Missing

SENA, sensitivity analysis.

For PDR, the enhanced reference was the reference standard for PDR (O-FTF) + ophthalmologist OPTOS + ophthalmologist ETDRS. If a patient was gradable by any of the three tests, he/she was considered as gradable and, hence, the one available result was used.

^a Grader referral for DMO or PDR = 'active' + 'unsure' + 'ungradable'.

Diagnostic performance analysis methods

For all diagnostic accuracy analyses the sensitivity, specificity, and positive and negative likelihood ratios were calculated (with appropriate 95% CIs). The difference in sensitivity and specificity between OPTOS and ETDRS fundus images, as assessed by the ophthalmic graders, was compared with corresponding 95% CIs produced using Newcombe's method for paired data³⁰ and McNemar's test.³¹

Sensitivity analyses of diagnostic performance

In addition to the main analysis for DMO and PDR, a number of sensitivity analyses were conducted (listed in Table 2).

Sensitivity analysis (SENA)1 assessed the impact of the 'unsure' and 'ungradable' test classifications and of the ophthalmic grader's grade on the diagnostic performance by defining the index test positive as definite assessment of the active disease (in contrast to allowing referral for 'unsure' or 'ungradable'). The reference standards of the O-FTF + OCT assessment for DMO and the O-FTF for PDR of those present with active disease were used.

SENA2 assessed the ophthalmic graders' referral assessments against the reference standard of those requiring treatment for both DMO and PDR. For DMO only, SENA3 focused on diagnostic performance for assessment among patients considered to possess a form of DMO that would be more important to recognise earlier, central-involving DMO (in contrast to non-central-involving DMO against the reference standard of O-FTF + OCT assessment). For PDR only, SENA4 assessed the diagnostic performance of the ophthalmic grader against the reference standard of O-FTF assessment to detect active PDR with pre-retinal or vitreous haemorrhage (i.e. high risk PDR). For PDR only, SENA5 assessed the diagnostic performance of the ophthalmic grader against the ERS (O-FTF assessment supplemented by ophthalmologist evaluation using OPTOS and ETDRS images) to detect active PDR. A further sensitivity analysis (SENA6) assessed diagnostic accuracy for a subset of participants only, who were assessed in a 'typical' NHS clinic setting, in contrast to a research clinic, but otherwise under the same conditions as the main analysis.

For PDR only, three additional post hoc analyses (see Table 3) that were not listed in the statistical analysis plan were also conducted in the light of the main PDR results, to further understand the content and potential value of the two imaging modalities evaluated (ultra-wide field and seven-field fundus images) along with the reliability of the O-FTF reference standard for PDR. 'Additional 1' compared the ophthalmologists' identification of active PDR by OPTOS or by ETDRS in either eye with the reference standard of O-FTF assessment for active PDR. 'Additional 2' compared the ophthalmologists' identification of active PDR by OPTOS or by ETDRS in either eye with the reference standard of O-FTF assessment for patients with more severe disease (i.e. PDR with pre-retinal or vitreous haemorrhage). 'Additional 3' compared the ophthalmic graders' referral for either eye by OPTOS or by ETDRS with the reference standard of O-FTF assessment + ophthalmologist OPTOS assessment of active PDR in either eye or O-FTF assessment + ophthalmologist ETDRS assessment of active PDR in either eye.

Secondary analyses of diagnostic accuracy

In addition to the principal analyses that were conducted separately at the person level for DMO and PDR, two distinct secondary analyses of diagnostic accuracy were planned (Table 5). Both of these secondary analyses focused on the entire EMERALD participant population ('All participants'), whether or not they were eligible for a disease (DMO/PDR) pathway. First, a limited set of eye-level analyses were carried out using the positive identification of active disease at the eye level [secondary analysis (SECA) 1A–C]. A random eye was selected when two eyes were eligible, with ophthalmic grader eye-specific assessment against the reference standard. Second, person-level analyses were carried out for the overall referral status of a patient, irrespective of whether it was active DMO or PDR that required referral (SECA2A–C). SECA2A utilised an ophthalmic grader referral assessment for either disease; for DMO, OCT assessment was used, and for PDR the OPTOS and ETDRS assessment were used in turn (i.e. two sets of results). The reference standards of O-FTF + OCT assessment for DMO and O-FTF

TABLE 5 List of secondary analyses and additional post-hoc analysis of diagnostic accuracy: all patients^a

Analysis name	Level of analysis	DMO index test positive	DMO reference standard	PDR index test(s) positive	PDR reference standard
SECA1A	Eye	OCT-based ophthalmic grader assessment of active DMO in the randomly selected eye	O-FTF + OCT assessment of active DMO in the randomly selected eye	OPTOS-based ophthalmic grader assessment of active PDR in randomly selected eye ETDRS-based ophthalmic grader assessment of active PDR in randomly selected eye	O-FTF assessment of active PDR in the randomly selected eye
SECA1B	Eye	OCT-based ophthalmic grader identification of central-involving DMO in either eye	O-FTF + OCT assessment of central-involving DMO in either eye		
SECA1C	Eye	N/A	N/A	OPTOS-based ophthalmic grader referral ETDRS-based ophthalmic grader referral	Enhanced standard
Combined DMO/PDR test positive				Combined DMO/PDR reference standard	
SECA2A	Person	Ophthalmic grader referral^c based on OCT for DMO and either OPTOS-based ophthalmic grader referral for PDR or ETDRS-based ophthalmic grader referral for PDR (referral for either disease will be considered a referral for the combined test)		O-FTF + OCT assessment of active DMO and O-FTF assessment of active PDR in either eye	
Additional 4 ^b	Person	Ophthalmic grader assessment of active DMO based on OCT and either OPTOS-based ophthalmic grader assessment of active PDR or ETDRS-based ophthalmic grader assessment of active PDR		O-FTF + OCT assessment of active DMO and O-FTF assessment of active PDR in either eye	
SECA2B	Person	Ophthalmic grader referral based on OCT for DMO and either OPTOS-based ophthalmic grader referral or ETDRS-based ophthalmic grader referral for PDR, and visual acuity of < 6/12 (or ETDRS equivalent letter) (referral for either disease, or owing to visual acuity, will be considered a referral for the combined test)		O-FTF + OCT assessment of active DMO and O-FTF assessment of active PDR in either eye	
SECA2C	Person	Ophthalmic grader referral based on OCT for DMO and either OPTOS-based ophthalmic grader referral or ETDRS-based ophthalmic grader referral for PDR, and visual acuity of < 6/12 (or ETDRS equivalent letter) (referral for either disease, or owing to visual acuity, will be considered a referral for the combined test)		O-FTF + OCT assessment of active DMO and O-FTF assessment of active PDR in either eye requiring treatment	

N/A, not applicable.

a The 'All patients' population will include in the analysis those with 'persistent' disease, 'de novo' disease, or 'no disease' in one eye, irrespective of the disease status of the other eye, as well as those who were eligible for the new ophthalmic grader pathway (i.e. those who had previously successfully treated DMO/PDR).

b One additional analysis on PDR that was not listed in the statistical analysis plan was conducted to further understand the content and potential implementation method of the new pathway.

c Grader referral for DMO or PDR = 'active' + 'unsure' + 'ungradable'.

Notes

O-FTF + OCT is the reference standard for DMO (i.e. standard care); O-FTF is the reference standard for PDR (i.e. standard care). Bold highlights the differences between analyses.

assessment for PDR for those present with active disease were used (either disease constituting active disease at the person level). SECA2B was the same as SECA2A, except that a visual acuity of less than 6 out of 12 would also be considered a valid reason for referral. SECA2C was the same as SECA2A, except the reference standard was an ophthalmologist's assessment of the presence or absence of active disease that required treatment. Diagnostic performance for both secondary analyses was assessed using the same outcomes and methods as for the principal analyses.

One additional post hoc analysis ('Additional 4') that was not listed in the statistical analysis plan was also conducted to further understand the content and the potential implementation method of the new pathway. The test results for Additional 4 were graders' assessment of active disease(s) in either eye rather than graders' referral [which would include active disease(s), unsure or ungradable] in either eye (SECA2A). Both SECA2A and Additional 4 would be of interest to how the new pathway would function in clinical practice.

Results

Participant characteristics

A flow diagram is provided (*Figure 9*) that shows the flow of participants and the O-FTF assessment reference standard for each disease. In total, 490 patients were screened, of whom 401 (82%) consented and were recruited. A total of 397 patients were eligible to participate in the study and four patients were ineligible. Five eyes could not be assessed and were excluded at eye-level assessment, but were included at patient-level assessment (assessed the other eye only): one right eye could not be assessed owing to blindness; four left eyes could not be assessed owing to blindness ($n = 2$), total retinal detachment ($n = 1$) and artificial eye ($n = 1$).

Sample description

Baseline characteristics of participants

The median age of the 397 patients included in EMERALD was 60 (range 18–103) years. Nearly two-thirds (65%) of the patients were male and the majority (86%) were of white ethnicity. In total, 317 (80%) patients had a diagnosis of DMO and of these 182 (57%) were from the older (≥ 60 years) age group. A total of 287 (72%) patients had a diagnosis of PDR and of these 136 (47%) were from the older age group (*Table 6*). The participants' characteristics by active diseases status are summarised in *Appendix 1, Table 36*.

Reference standard diagnosis characteristics

Tables 7 and 8 display the EMERALD population at eye level and at person level based on the six patient classification scenarios for DMO and for PDR, respectively. The reference standard was determined at both the eye and person level. Of the 397 patients included in the analyses, 272 (80%) were eligible for DMO in the new pathway and these patients were used in the principal person-level diagnostic accuracy analyses for DMO. Of these, 152 (56%) had DMO at least in one eye that was previously successfully treated and was active (i.e. had recurrence of DMO) when evaluated in the EMERALD study visit (i.e. when entering the study). A total of 120 (44%) had DMO in at least one eye that had been previously successfully treated and remained inactive (i.e. had no DMO present) when evaluated in the EMERALD study visit (i.e. when entering the study) (see *Figure 9* and *Table 7*). In total, 281 (71%) of the 397 patients were eligible for PDR in the new pathway and these patients were used in the principal person-level diagnostic accuracy analyses for PDR. Of these, 111 (40%) had PDR in at least one eye that was previously successfully treated and was active (i.e. had recurrence of active PDR) when evaluated in the EMERALD study visit (i.e. when entering the study). A total of 170 (60%) patients had PDR in at least one eye that had been previously successfully treated and remained inactive when evaluated in the EMERALD study visit (i.e. when entering the study).

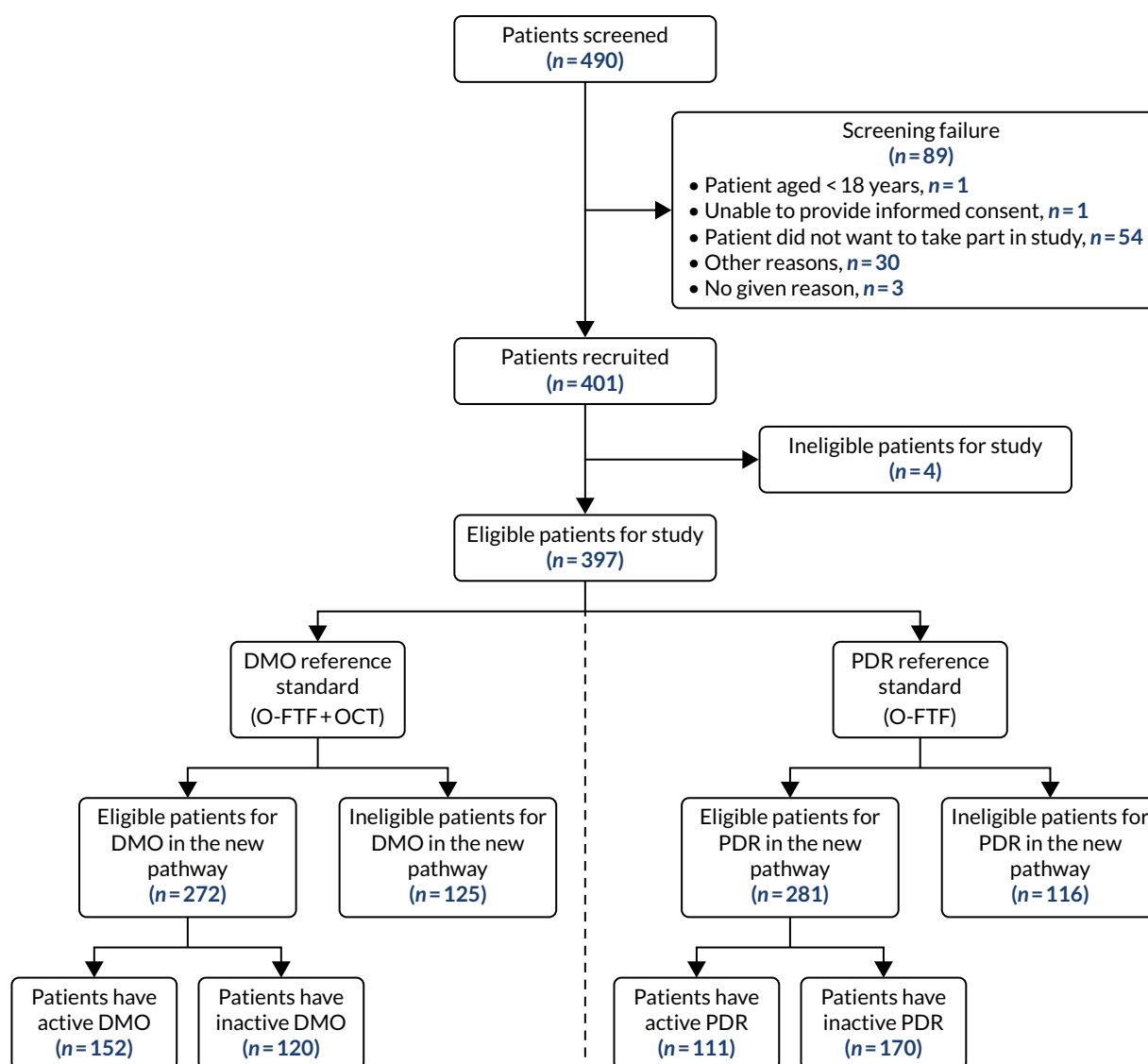


FIGURE 9 Study flow diagram.

(see Figure 9 and Table 8). In total, 272 (69%) out of the 397 participants had active DMO or active PDR, and 49 (12%) had active DMO and active PDR according to the ophthalmologists assessments (see Appendix 1, Table 36).

Eye characteristics and eye comorbidities of participants

Appendix 1, Tables 37 and 38, describes other findings (abnormalities) that were identified by ophthalmologists when undertaking slit-lamp biomicroscopy and were recorded as potential factors that could lead to inadequate image quality, as well as comorbidities noted. Comorbidities were uncommon in EMERALD participants. The most common comorbidities were epiretinal membrane and glaucoma (reported in approximately 3% of eyes) (see Appendix 1, Table 38).

Grading results diagnosis characteristics

Table 9 displays the OCT-based ophthalmic graders' grading results at eye level against the reference standard for assessing DMO. OCT images were available for 385 (97%) right eyes and 380 (97%) left eyes. Of these, only a very small number, six (2%) right eyes and five (1%) left eyes, were considered ungradable. In a similarly small number of eyes, the grader was unsure if DMO was present [13 (3%) right eyes and 12 (3%) left eyes]. A total of 198 (51%) right eyes and 188 (49%) left eyes were identified as having DMO, and there were 168 (44%) right eyes and 175 (46%) left eyes with no DMO.

TABLE 6 Demographic characteristics of participants

Characteristic	Patients with DMO (N = 317), n (%)	Eligible for DMO in the new pathway (N = 272), n (%)	Patients with PDR (N = 287), n (%)	Eligible for PDR in the new pathway (N = 281), n (%)	Total (N = 397), n (%)
Sex					
Male	205 (65)	175 (64)	187 (65)	185 (66)	257 (65)
Female	112 (35)	97 (36)	100 (35)	96 (34)	140 (35)
Age (years)					
18–59	135 (43)	113 (42)	151 (53)	148 (53)	188 (47)
≥ 60	182 (57)	159 (58)	136 (47)	133 (47)	209 (53)
Ethnic origin					
White	274 (86)	240 (88)	240 (84)	234 (83)	340 (86)
Black	20 (6)	17 (6)	19 (7)	19 (7)	26 (7)
Asian	16 (5)	11 (4)	20 (7)	20 (7)	22 (6)
Middle Eastern	3 (1)	1 (< 1)	5 (2)	5 (2)	5 (1)
Other	4 (1)	3 (1)	3 (1)	3 (1)	4 (1)

TABLE 7 The DMO classification based on slit-lamp biomicroscopic examination by an ophthalmologist with access to SD-OCT images (reference standard)

DMO status		Eye level		Person level
		Right eye (N = 396), ^a n (%)	Left eye (N = 393), ^a n (%)	Person (either eye) (N = 397), n (%)
Active	Previously successfully treated (scenario 1)	97 (24)	92 (23)	152 (38)
	Previously unsuccessfully treated (scenario 3)	30 (8)	17 (4)	–
	Newly diagnosed (scenario 4)	15 (4)	11 (3)	–
Inactive	Previously successfully treated (scenario 2)	113 (29)	119 (30)	120 (30)
	Previously unsuccessfully treated (scenario 5)	19 (5)	23 (6)	–
No disease	Scenario 6	121 (31)	127 (32)	–
Unclassifiable ^b		1 (< 1)	4 (1)	–
	Owing to no view of fundus: cataract	1 (< 1)	–	–
	Owing to no view of fundus: haemorrhage	–	3 (1)	–
	Owing to no view of fundus	–	1 (< 1)	–
Ineligible for new DMO pathway ^c		–	–	125 (31)

a Five eyes in total could not be assessed. One right eye could not be assessed owing to blindness; four left eyes could not be assessed owing to blindness (n = 2), total retinal detachment (n = 1) and artificial eye (n = 1).

b One participant had unclassifiable DMO in their left eye owing to no view of fundus (vitreous haemorrhage present and right eye ineligible for new DMO pathway based on O-FTF + OCT assessment): this patient was classified as ineligible for the new DMO pathway at the person level. This participant was eligible for the new PDR pathway. The other four participants who had an unclassifiable eye were eligible for the new DMO pathway.

c Ineligible participants for the new DMO pathway referred to those with both eyes ineligible for the new DMO pathway, which included both eyes not having been previously successfully treated (Scenarios 3 and 5), de novo disease (Scenario 4) or no disease (Scenario 6), with the exception of one participant (listed above) who had one eye DMO unclassifiable and the other eye ineligible for the new DMO pathway.

TABLE 8 The PDR classification based on slit-lamp biomicroscopic examination by an ophthalmologist (reference standard)

PDR status		Eye level		Person level
		Right eye (N = 396), ^a n (%)	Left eye (N = 393), ^a n (%)	Person (either eye) (N = 397), ^a n (%)
Active	Previously successfully treated (scenario 1)	72 (18)	71 (18)	111 (28)
	Previously unsuccessfully treated (scenario 3)	4 (1)	4 (1)	–
	Newly diagnosed (scenario 4)	4 (1)	4 (1)	–
Inactive	Previously successfully treated (scenario 2)	180 (45)	171 (44)	170 (43)
	Previously unsuccessfully treated (scenario 5)	0 (0)	1 (< 1)	–
No disease	(Scenario 6)	134 (34)	140 (36)	–
Unclassifiable ^b		2 (1)	2 (1)	–
	Owing to no view of fundus: cataract	1 (< 1)	–	–
	Owing to no view of fundus: corneal graft	–	1 (< 1)	–
	Owing to no view of fundus: reason not specified	1 (< 1)	1 (< 1)	–
Ineligible for new PDR pathway ^c		–	–	116 (29)

a Five eyes in total could not be assessed. One right eye could not be assessed owing to blindness; four left eyes could not be assessed owing to blindness (n = 2), total retinal detachment (n = 1) and artificial eye (n = 1).

b All four participants who had an unclassifiable eye were eligible for the new PDR pathway.

c Ineligible participants for the new PDR pathway referred to those with both eyes ineligible for the new PDR pathway, which including both eyes not been previously successfully treated (Scenarios 3 and 5), de novo disease (Scenario 4) or no disease (Scenario 6).

TABLE 9 The DMO grading data: graders assessment based on OCT images compared with reference standard

DMO grading	Graders OCT (G-OCT)		Reference standard (O-FTF + OCT)	
	Right eye ^a (n = 396)	Left eye ^a (n = 393)	Right eye ^a (n = 396)	Left eye ^a (n = 393)
Total DMO present (active)	198	188	142	120
No DMO	168	175	253	269
Unsure if DMO present	13	12	–	–
Ungradeable	6	5	1	4
No images	11	13	–	–

G-OCT, grader assessment based on SD-OCT images; O-FTF + OCT, reference standard for DMO.

a Five eyes in total could not be assessed: one right eye could not be assessed owing to blindness; four left eyes could not be assessed owing to blindness (n = 2), total retinal detachment (n = 1) and artificial eye (n = 1).

Shaded cells refer to the reference standard.

Table 10 displays the UWF OPTOS and seven-field ETDRS-based ophthalmic graders' and ophthalmologists' grading results for PDR patients at eye level against the reference standard. Seven-field ETDRS images were available for 376 (95%) right eyes and 362 (92%) left eyes. Of these, a small number [19 (5%) right eyes and 27 (7%) left eyes] were considered ungradable. In a similarly small number of eyes, the grader was unsure if PDR was present [20 (5%) right eyes and 17 (5%) left eyes]. A total of 88 (23%) right eyes and 83 (23%) left eyes were identified as having active PDR based on seven-field ETDRS images. There was a small number of eyes in which PDR was identified, but the grader stated that they were unsure if there was active disease or not [12 (3%) right eyes and 14 (4%) left eyes]. A total of 71 (42%) right eyes and 65 (40%) left eyes were classified by the graders as having inactive PDR, and 166 (44%) right eyes and 156 (43%) left eyes were classified by the graders as having no PDR. UWF OPTOS images were available for 379 (96%) right eyes and 374 (95%) left eyes. Of these, a small number [16 (4%) right eyes and 21 (6%) left eyes] were considered ungradable. In a similarly small number of eyes, the grader was unsure if PDR was present [20 (5%) right eyes and 21 (6%) left eyes]. A total of 85 (22%) right eyes and 73 (20%) left eyes were identified as having active PDR based on UWF OPTOS images. There was a small number of eyes in which PDR was identified but the grader was unsure if there was active disease or not [11 (3%) right eyes and 16 (4%) left eyes]. A total of 95 (25%) right eyes and 81 (22%) left eyes were identified as having inactive PDR, and 152 (40%) right eyes and 162 (43%) left eyes were identified as having no PDR, based on UWF OPTOS images.

Missing test images

Table 11 displays the reasons for missing test images in EMERALD at the person level. Of the 397 patients included in the analyses, 345 (87%) had SD-OCT, seven-field ETDRS and ultra-wide field fundus images for each eye available for assessment. A total of 52 patients had at least one imaging assessment missing. Of these, 17 (33%) would be considered as not assessable by images and, therefore, the patient would need to be referred to clinical practice (e.g. media opacity). In total, 35 (67%) patients would be considered as missing for the EMERALD study and would also be considered as missing in practice (e.g. patient does not attend for the taking of the image). The details of eye-level missing images by assessment are summarised in Appendix 1, Table 39.

Diagnostic performance of the imaging tests

The results of the diagnostic performance of the EMERALD ophthalmic graders are presented in the following three sections:

1. Diagnostic performance for DMO patients eligible for the new ophthalmic grader pathway.
2. Diagnostic performance for PDR patients eligible for the new ophthalmic grader pathway.
3. Secondary sensitivity analyses on combined diseases (DMO and PDR) for all patients.

Diagnostic performance for diabetic macular oedema patients eligible for the new ophthalmic grader pathway

Summary of ophthalmic graders' grading results for diabetic macular oedema

Table 12 displays the OCT-based ophthalmic graders' grading results at the person level against the reference standard. The graders' decision on whether or not to refer the patient, an inherently person-level decision, was used rather than an assessment of the single eye. Furthermore, to address how the new pathway would function in practice, graders' assessment results of DMO of 'unsure' and 'ungradable' were considered alongside those classified as 'active' to be referred to an ophthalmic clinic for ophthalmological assessment. Of the 272 DMO patients who were eligible for the new pathway, 152 (56%) were identified by the ophthalmologists' slit-lamp biomicroscopic examination with access to SD-OCT images (O-FTF + OCT is the reference standard) as having DMO that was active at the time of recruitment, whereas 120 (44%) were identified by the O-FTF + OCT reference standard as having DMO that was inactive at the time of recruitment.

TABLE 10 The PDR grading data: graders and ophthalmologists assessment based on UWF OPTOS and seven-field ETDRS images compared with reference standard

PDR grading	Graders seven-field (G-ETDRS)		Ophthalmologist seven-field (O-ETDRS)		Graders UWF (G-OPTOS)		Ophthalmologist UWF (O-OPTOS)		Reference standard (O-FTF)	
	Right eye ^a (n = 396)	Left eye ^a (n = 393)	Right eye ^a (n = 396)	Left eye ^a (n = 393)	Right eye ^a (n = 396)	Left eye ^a (n = 393)	Right eye ^a (n = 396)	Left eye ^a (n = 393)	Right eye ^a (n = 396)	Left eye ^a (n = 393)
Total PDR present	171	162	185	185	191	170	206	213		
Active	88	83	58	60	85	73	70	75	80	79
Inactive	71	65	98	106	95	81	117	121	180	172
Unsure if active	12	14	29	19	11	16	19	17	–	–
No PDR	166	156	150	137	152	162	143	125	134	140
Unsure if PDR present	20	17	30	31	20	21	25	23	–	–
Ungradable	19	27	9	10	16	21	4	12	2	2
No images	20	31	22	30	17	19	18	20	–	–

G-ETDRS, grader assessment based on seven-field ETDRS fundus images; G-OPTOS, grader assessment based on ultra-wide field fundus images using OPTOS; O-ETDRS, ophthalmologist assessment based on seven-field ETDRS fundus images; O-FTF, reference standard for PDR; O-OPTOS, ophthalmologist assessment based on ultra-wide field fundus images using OPTOS.

a Five eyes in total could not be assessed. One right eye could not be assessed owing to blindness; four left eyes could not be assessed owing to blindness (n = 2), total retinal detachment (n = 1) and artificial eye (n = 1).

Shaded cells refer to the reference standard.

TABLE 11 Missing images at person level

	Number of participants (n = 397)
All images are available (all three ^a imaging modalities for both eyes ^b)	345
Missing at least one imaging assessment	52
Image missing: patient would need referral	17
No view of fundus: vitreous haemorrhage	2
No view of fundus: corneal graft	1
Poor pupillary dilatation	4
Patient unable to co-operate with imaging	1
Patient unable to sit at the camera	2
Media opacity	5
Patient could not stand the light	1
Unable to see fundus owing to pathology	1
Image missing: missing	35
Patient does not attend for the taking of the image	9
Only one eye image captured ^c	5
Images not recorded correctly	11
Other unknown reason	10

a OCT, ETDRS and OPTOS.

b Five eyes in total could not be imaged. For these five patients, only three imaging assessment per patients were taken.

c At the initiation of the study and for a short period until an amendment was implemented, images were being obtained in 'eligible' eyes only.

TABLE 12 Diabetic macular oedema classifications: eligible patients for new pathway (n = 272)

Classification	Reference standard (O-FTF + OCT)			G-OCT		
	Active DMO ^a	Active DMO requiring treatment ^b	Central-involving active DMO ^c	Referral ^d for DMO ^e	Active DMO ^f	Central-involving active DMO ^g
Positive	152	85	132	220	209	177
Negative	120	187	139	40	50	80
Missing/cannot classify	0	0	1	12	13	15

G-OCT, grader assessment based on SD-OCT images; O-FTF + OCT, reference standard for DMO.

a Ophthalmologist slit-lamp biomicroscopic examination with access to OCT images identifying active DMO in either eye.

b Ophthalmologist slit-lamp biomicroscopic examination with access to OCT images identifying active DMO in either eye requiring treatment.

c Ophthalmologist slit-lamp biomicroscopic examination with access to OCT images identifying central-involving active DMO in either eye.

d Grader referral for DMO: 'active' + 'unsure' + 'ungradable'.

e Graders' referral of DMO based on OCT images, which includes 'DMO present', 'Unsure if DMO present' and 'Ungradable'. In practice, patients under these three categories would be referred to an ophthalmic clinic.

f Grader identification of active DMO (active DMO only, 'unsure' and 'ungradable' excluded) in either eye based on OCT images.

g Grader identification of central-involving active DMO in either eye based on OCT images.

Shaded cells refer to the reference standard.

Compared with this reference standard, 220 (81%) out of the 272 patients under graders' assessment (G-OCT) were referred for further ophthalmological examination; of these, 209 (209/220; 95%) were referred because graders (G-OCT) identified active DMO (i.e. active DMO present in their opinion) and 11 (11/220; 5%) because graders were unsure or considered OCT images ungradable.

The flow of study participants who followed the new pathway of graders' assessment for DMO based on OCT is shown in *Figure 10*. Of the 397 EMERALD participants, four (1%) had missing OCT images for both eyes and were excluded from the analyses. Eleven (3%) participants were excluded from the main analysis and 12 (3%) were excluded from SENA1 because of missing OCT images for one eye, which led to no overall graders assessment result at the patient level. Of the 382 participants who had valid grading results from the graders to be included in the main analysis, 281 (74%) were identified by the graders as referring for DMO (i.e. having active DMO or unsure or ungradable), of whom 220 (78%) were eligible for the new DMO pathway. In total, 101 (26%) of the 382 participants were identified by the graders as not having active DMO (including no DMO and inactive DMO), of whom 40 (40%) were eligible for the new DMO pathway. Of the 381 participants who had valid grading results from the graders to be included in SENA1, 262 (69%) were identified by the graders as having active DMO (i.e. excluding unsure and ungradable), of whom 209 (80%) were eligible for the new DMO pathway. In total, 119 (31%) out of the 381 participants were identified by the graders as not having active DMO, of whom 50 (42%) were eligible for the new DMO pathway. The diagnostic performance for the main analysis and sensitivity analyses for DMO patients are given in *Table 13*. Results from the main analysis and SENA1 are also presented in the flow diagram (see *Figure 10*). Owing to missing OCT images that led to no overall graders assessment result, the grader's diagnostic performance under main analysis was tested with a slightly smaller number of referral for DMO [disease present, $n = 147$; true positive (TP), $n = 142$; false negative (FN), $n = 5$] than the O-FTF-OCT (ophthalmologist face-to-face examination with access to optical coherence tomography scans) examination of active DMO ($n = 152$), and a slightly smaller number of not active DMO [disease absent, $n = 113$; false positive (FP), $n = 78$; true negative (TN), $n = 35$] than the O-FTF-OCT examination of not active DMO ($n = 120$). The grader's diagnostic performance under SENA1 was also tested with a slightly smaller number of active DMO (disease present, $n = 146$; TP, $n = 139$; FN, $n = 7$) than the O-FTF-OCT examination of active DMO ($n = 152$), and a slightly smaller number of not active DMO (disease absent, $n = 113$; FP, $n = 70$; TN, $n = 43$) than the O-FTF-OCT examination of not active DMO ($n = 120$).

Results: main analysis for diabetic macular oedema

The main analysis tested the ophthalmic graders' decision of referring patients to ophthalmologists because of the presence of active DMO, or if they were unsure or unable to grade the images (ungradable) in either eye based on the OCT images. The reference standard was the ophthalmologist slit-lamp biomicroscopic examination with access to OCT images identifying active DMO in either eye. Under this main analysis of DMO, graders were found to have a sensitivity of 97% (95% CI 92% to 99%) with a specificity of 31% (95% CI 23% to 40%) when compared with the reference standard.

Results: sensitivity analysis for diabetic macular oedema

Sensitivity analysis 1

Sensitivity analysis 1 assessed the impact of the 'unsure' and 'ungradable' ophthalmic graders' grading results on the diagnostic performance, by defining the index test positive as definite assessment of active DMO (in contrast to allowing referral for unsure or ungradable DMO). Under this analysis, graders were found to have a sensitivity of 95% (95% CI 90% to 98%) and a specificity of 38% (95% CI 30% to 47%).

Sensitivity analysis 2, 3 and 6

Similar results were obtained for the analysis evaluating people with DMO requiring further treatment, as determined by the ophthalmologist in the standard of care pathway (sensitivity 95%, 95% CI

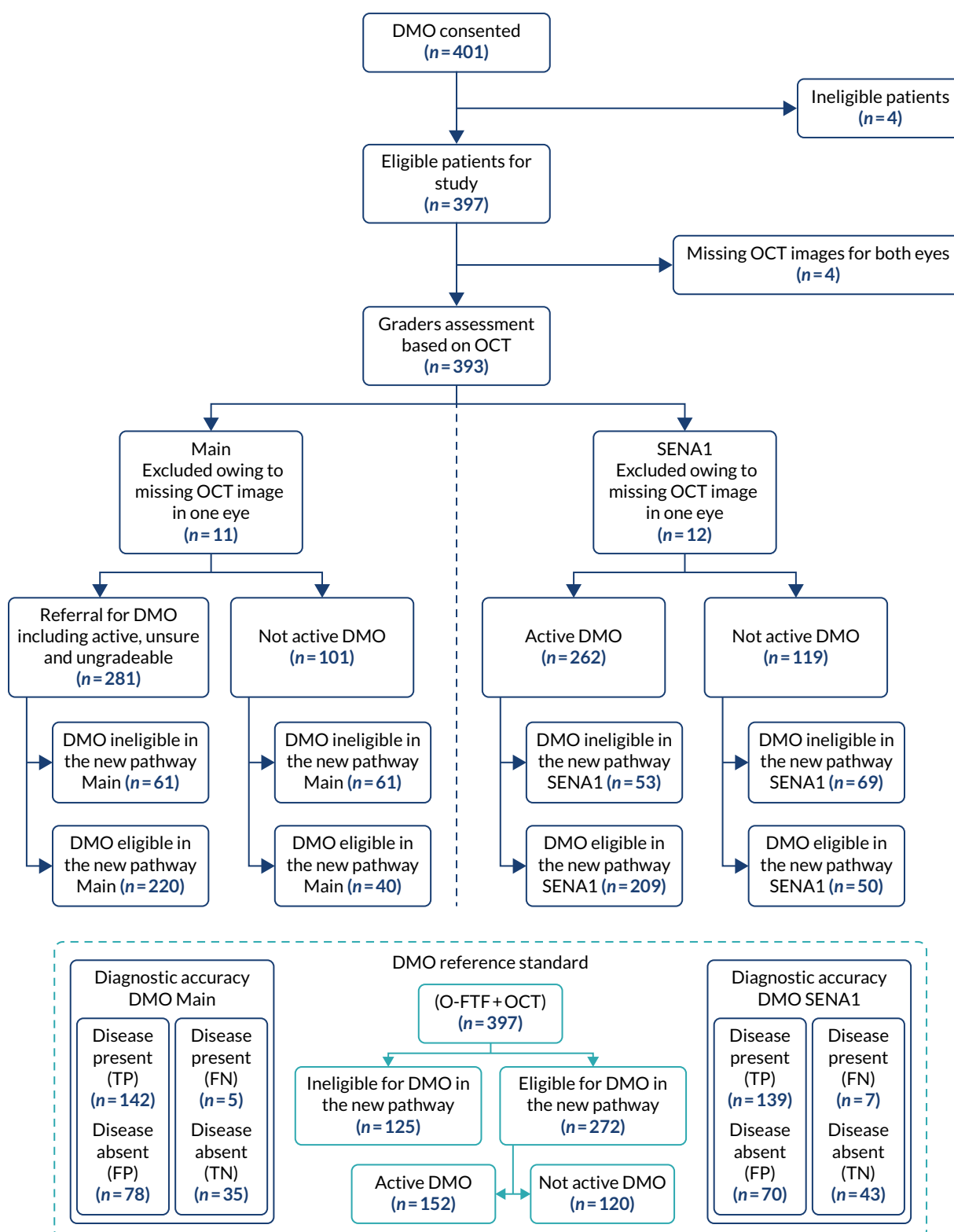


FIGURE 10 Flow diagram: diagnosis as determined by the ophthalmic graders for DMO patients – main and SENA1. Note that this flow diagram presents only the grading results from the graders' assessment based on SD-OCT images. FN, false negative; FP, false positive; TN, true negative; TP, true positive.

TABLE 13 Diagnostic performance for DMO patients

	Test positive	Reference standard	Diagnostic parameter	n/N	Estimate (95% CI)
Main	G-OCT referral ^a for DMO	O-FTF + OCT examination of active DMO in either eye	Sensitivity (%)	142/147	97% (92% to 99%)
			Specificity (%)	35/113	31% (23% to 40%)
			Positive likelihood ratio	–	1.40 (1.23 to 1.59)
			Negative likelihood ratio	–	0.11 (0.04 to 0.27)
SENA1	G-OCT identified active DMO	O-FTF + OCT examination of active DMO in either eye	Sensitivity (%)	139/146	95% (90% to 98%)
			Specificity (%)	43/113	38% (30% to 47%)
			Positive likelihood ratio	–	1.54 (1.32 to 1.78)
			Negative likelihood ratio	–	0.13 (0.06 to 0.27)
SENA2	G-OCT referral for DMO	O-FTF + OCT examination of active DMO in either eye requiring treatment	Sensitivity (%)	81/85	95% (89% to 98%)
			Specificity (%)	36/175	21% (15% to 27%)
			Positive likelihood ratio	–	1.20 (1.10 to 1.31)
			Negative likelihood ratio	–	0.23 (0.08 to 0.62)
SENA3	G-OCT identified central-involving active DMO	O-FTF + OCT examination of central-involving active DMO in either eye	Sensitivity (%)	121/129	94% (88% to 97%)
			Specificity (%)	72/128	56% (48% to 65%)
			Positive likelihood ratio	–	2.14 (1.75 to 2.62)
			Negative likelihood ratio	–	0.11 (0.06 to 0.22)
SENA6	G-OCT referral for DMO in routine clinic	O-FTF + OCT examination of active DMO in either eye of eligible patients in routine clinic	Sensitivity (%)	81/85	95% (89% to 98%)
			Specificity (%)	26/65	40% (29% to 52%)
			Positive likelihood ratio	–	1.59 (1.30 to 1.95)
			Negative likelihood ratio	–	0.12 (0.04 to 0.32)

G-OCT, grader assessment based on SD-OCT images; O-FTF + OCT, reference standard for DMO.

a Grader referral for DMO = 'active' + 'unsure' + 'ungradable'.

89% to 98%; specificity 21%, 95% CI 15% to 27%), and those with central-involving DMO (sensitivity 94%, 95% CI 88% to 97%; specificity 56%, 95% CI 48% to 65%) when compared with the reference standard. Sensitivity and specificity values for the grader's pathway were found to be similar to those presented for the main analysis when referral for active DMO was considered only (excluding 'unsure' and 'ungradable') and when patients were assessed specifically in NHS clinics (vs. in 'research' clinics).

Diagnostic performance for proliferative diabetic retinopathy patients eligible for the new ophthalmic grader pathway

Summary of the ophthalmic graders' grading results

Table 14 shows the UWF OPTOS and seven-field ETDRS-based ophthalmic graders' and ophthalmologists' grading results for PDR patients at the eye level against the reference standard. To mimic the way that the new pathway would function in clinical practice, graders' results of PDR of 'active', 'unsure' and 'ungradable' were considered to require a referral to the clinic to be assessed by an ophthalmologist. Of the 281 PDR patients who were eligible for the new ophthalmic grader pathway, 111 (40%) were identified by the ophthalmologist slit-lamp biomicroscopic examination (O-FTF is the reference standard for PDR) as having active PDR at the time of recruitment in the EMERALD study. A total of 170 (60%) patients were identified by the O-FTF reference standard as having inactive PDR at the time of recruitment. In total, 161 (57%) out of the 281 patients under

TABLE 14 Proliferative diabetic retinopathy classifications: eligible patients for the new pathway (n = 281)

Classification	Ophthalmologist assessment						Grader assessment			
	O-FTF: active PDR ^a	O-OPTOS: active PDR ^b	O-ETDRS: active PDR ^c	ERS ^d	O-FTF: active PDR requiring treatment ^e	O-FTF: active PDR with pre-retinal or vitreous haemorrhage ^f	G-OPTOS: referral ^g for PDR ^h	G-OPTOS: active PDR ⁱ	G-ETDRS: referral for PDR ^j	G-ETDRS: active PDR ^k
Positive	111	96	85	146	95	75	161	109	170	118
Negative	170	166	167	135	186	206	104	155	92	139
Missing/cannot classify	0	19	29	0	0	0	16	17	19	24

G-ETDRS, grader assessment based on seven-field ETDRS fundus images; G-OPTOS, grader assessment based on ultra-wide field fundus images using OPTOS; O-ETDRS, ophthalmologist assessment based on seven-field ETDRS fundus images; O-FTF, reference standard; O-OPTOS, ophthalmologist assessment based on ultra-wide field fundus images using OPTOS.

a Ophthalmologist slit-lamp biomicroscopic examination of active PDR.

b Ophthalmologist's identification of active PDR based on UWF OPTOS fundus images.

c Ophthalmologist's identification of active PDR based on seven-field ETDRS fundus images.

d ERS combined the ophthalmologist slit-lamp biomicroscopic examination, ophthalmologist grading of UWF OPTOS and ophthalmologist grading of seven-field ETDRS imaging. Active PDR was defined as 'Active' detected by any of these three ophthalmologist assessments.

e Ophthalmologist slit-lamp biomicroscopic examination of active PDR in either eye requiring treatment.

f Ophthalmologist slit-lamp biomicroscopic examination disclosed pre-retinal or vitreous haemorrhages in either eye indicative of high-risk PDR.

g Grader referral for PDR = 'active' + 'unsure' + 'ungradable'.

h Grader referral for PDR based on UWF OPTOS fundus images, which includes 'active PDR', 'unsure if PDR active', 'unsure if PDR present' and 'ungradable'. In practice, patients under these four categories would be referred to an ophthalmic clinic for ophthalmological assessment.

i Grader identification of 'active' PDR in either eye by UWF OPTOS fundus images.

j Grader referral for PDR based on seven-field ETDRS fundus images, which includes 'active PDR', 'unsure if PDR active', 'unsure if PDR present' and 'ungradable'. In practice, patients under these four categories would be referred to an ophthalmic clinic.

k Grader identification of 'active' PDR in either eye based on seven-field ETDRS fundus images.

graders' assessment based on OPTOS (G-OPTOS) were referred for further examination. Of these patients, 109 (68%; 109/161) were assessed by the graders based on OPTOS (G-OPTOS) as having PDR that was 'active', whereas 52 (32%; 52/161) were assessed by the graders (G-OPTOS) as 'unsure' or 'ungradable'. In total, 170 (60%) of the 281 patients under graders' assessment based on ETDRS (G-ETDRS) were referred for further examination. Of these, 118 (69%; 118/170) patients were assessed by the graders based on ETDRS (G-ETDRS) as having 'active' PDR, whereas 52 (31%; 52/170) were assessed by the graders (G-OPTOS) as 'unsure' or 'ungradable'. Compared with the reference standard, 146 (52%) of the 281 patients were identified by the enhanced standard as having 'active' PDR.

The flow of study participants who followed the new pathway of graders' assessment of PDR based on OPTOS and ETDRS images is shown in *Figure 11*. Of the 397 EMERALD participants, nine (2%) participants who had missing OPTOS images for both eyes were excluded from the analyses that used OPTOS images. Twelve (3%) participants who had missing ETDRS images for both eyes were excluded from the analyses that used ETDRS images. Of the 388 participants assessed using OPTOS images, 13 (3%) were excluded from the main analysis and 14 (4%) were excluded from SENA1 owing to missing OPTOS images for one eye that led to no overall graders' assessment result at the patient level. In comparison, 15 (4%) and 20 (5%) of the 385 participants assessed using ETDRS images were excluded from the main analysis and SENA1, respectively; this was because of missing ETDRS images for one eye that led to no overall graders' assessment result at the patient level.

In total, 375 participants who had valid OPTOS grading results were available to be included in the main analysis, of whom 207 (55%) were identified by the graders as referring for PDR (i.e. having active PDR or unsure or ungradable). A total of 161 (78%) out of the 207 referrals for PDR were eligible for the new PDR pathway. In total, 168 (45%) out of the 375 participants were identified by the graders as having not active PDR (including no PDR and inactive PDR), of whom 104 (62%) were eligible for the new PDR pathway. A total of 374 participants had valid OPTOS grading results and were available to be included in SENA1. In total, 131 (35%) participants were identified by the graders as having active PDR (i.e. excluding unsure and ungradable), of whom 109 (83%) were eligible for the new PDR pathway. A total of 243 (65%) out of the 374 participants were identified as having not active PDR, of whom 155 (64%) were eligible for the new PDR pathway. Results from the main analysis and SENA1 for OPTOS are also presented in the flow diagram (see *Figure 11*). Owing to missing OPTOS images that led to no overall graders' assessment result, grader's diagnostic performance under main analysis was tested with a slightly smaller number of referrals for PDR (disease present, $n = 105$; TP $n = 87$; FN, $n = 18$) than that for the O-FTF examination of active PDR ($n = 111$), and a slightly smaller number of not active PDR referrals (disease absent, $n = 160$; FP, $n = 74$; TN, $n = 86$) than that for the O-FTF examination of not active PDR ($n = 170$). Graders' diagnostic performance under SENA1 was also tested with a slightly smaller number of active PDR referrals (disease present, $n = 105$; TP, $n = 66$; FN, $n = 39$) than that for the O-FTF examination of active PDR ($n = 111$), and a slightly smaller number of not active PDR referrals (disease absent, $n = 159$; FP, $n = 43$; TN, $n = 116$) than that for the O-FTF examination of not active PDR ($n = 170$).

In total, 370 participants who had valid ETDRS grading results were available to be included in the main analysis, of whom 214 (58%) were identified by the graders as referring for PDR (i.e. having active PDR or unsure or ungradable). A total of 170 (79%) out of the 214 referrals for PDR were eligible for the new PDR pathway. A total of 156 (42%) out of the 370 participants were identified by the graders as having not active PDR (including no PDR and inactive PDR), of whom 92 (59%) were eligible for the new PDR pathway. In total, 365 participants had valid ETDRS grading results and were available to be included in SENA1. A total of 138 (38%) participants were identified by the graders as having active PDR (i.e. excluding unsure and ungradable), of whom 118 (86%) were eligible for the new PDR pathway. A total of 227 (62%) out of the 365 participants were identified as having not active PDR, of whom 139 (61%) were eligible for the new PDR pathway. Results from the main analysis and SENA1 for ETDRS are also presented in the flow diagram (see *Figure 11*). Owing to missing ETDRS images that led to no overall graders assessment result, grader's diagnostic performance under the

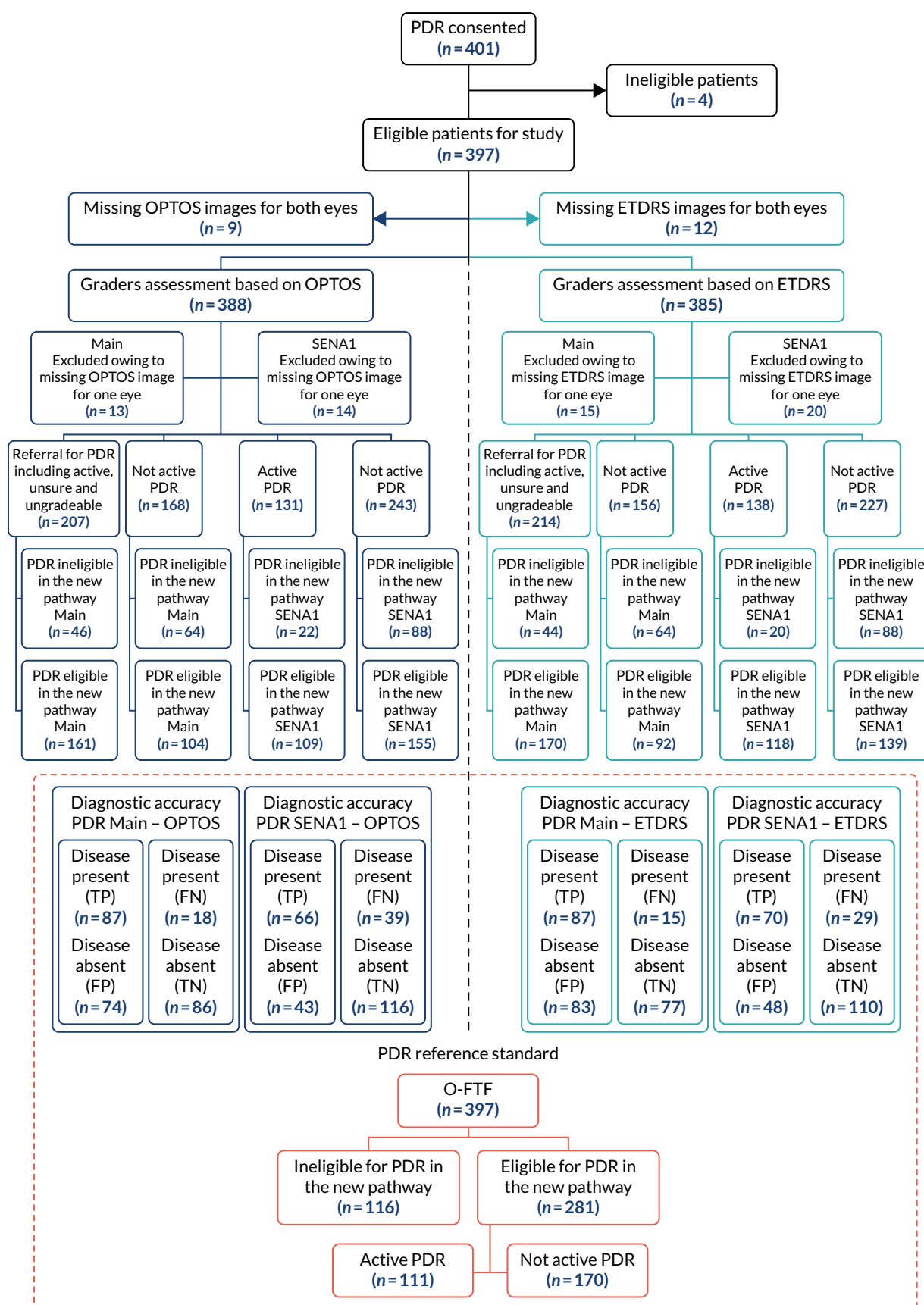


FIGURE 11 Flow diagram: diagnosis as determined by the ophthalmic graders for PDR patients comparing OPTOS and ETDRS – main and SENA1. This flow diagram only presents the grading results from the graders' assessment based on UWF OPTOS images and seven-field ETDRS images.

main analysis was tested with a slightly smaller number of referrals for PDR (disease present, $n = 102$; TP, $n = 87$; FN, $n = 15$) than that for the O-FTF examination of active PDR ($n = 111$), and a slightly smaller number of not active PDR referrals (disease absent, $n = 160$; FP, $n = 83$; TN, $n = 77$) than that for the O-FTF examination of not active PDR ($n = 170$). Graders' diagnostic performance under SENA1 was also tested with a slightly smaller number of active PDR referrals (disease present, $n = 99$; TP, $n = 70$; FN, $n = 29$) than that for the O-FTF examination of active PDR ($n = 111$), and a slightly smaller number of not active PDR (disease absent, $n = 158$; FP, $n = 48$; TN, $n = 110$) than that for the O-FTF examination of not active PDR ($n = 170$).

The diagnostic performance for the main analysis, sensitivity analyses and additional analyses for PDR patients are given in *Table 15*.

TABLE 15 Diagnostic performance for PDR patients

Analysis	Results	Reference standard	Diagnostic parameter	n/N	Estimate (95% CI)
Main	G-OPTOS referral ^a for PDR	O-FTF examination of active PDR in either eye	Sensitivity (%)	87/105	83% (75% to 89%)
			Specificity (%)	86/160	54% (46% to 61%)
			Positive likelihood ratio	–	1.79 (1.48 to 2.16)
			Negative likelihood ratio	–	0.32 (0.20 to 0.50)
	G-ETDRS referral for PDR	O-FTF examination of active PDR in either eye	Sensitivity (%)	87/102	85% (77% to 91%)
			Specificity (%)	77/160	48% (41% to 56%)
			Positive likelihood ratio	–	1.64 (1.39 to 1.95)
			Negative likelihood ratio	–	0.31 (0.19 to 0.50)
SENA1	G-OPTOS identified active PDR	O-FTF examination of active PDR in either eye	Sensitivity (%)	66/105	63% (53% to 71%)
			Specificity (%)	116/159	73% (66% to 79%)
			Positive likelihood ratio	–	2.32 (1.73 to 3.12)
			Negative likelihood ratio	–	0.51 (0.39 to 0.66)
	G-ETDRS identified active PDR	O-FTF examination of active PDR in either eye	Sensitivity (%)	70/99	71% (61% to 79%)
			Specificity (%)	110/158	70% (62% to 76%)
			Positive likelihood ratio	–	2.33 (1.78 to 3.04)
			Negative likelihood ratio	–	0.42 (0.30 to 0.58)
Additional 1	O-OPTOS identified active PDR	O-FTF examination of active PDR in either eye	Sensitivity (%)	74/103	72% (62% to 80%)
			Specificity (%)	137/159	86% (80% to 91%)
			Positive likelihood ratio	–	5.19 (3.46 to 7.80)
			Negative likelihood ratio	–	0.33 (0.24 to 0.45)

TABLE 15 Diagnostic performance for PDR patients (*continued*)

Analysis	Results	Reference standard	Diagnostic parameter	n/N	Estimate (95% CI)
SENA2	O-ETDRS identified active PDR	O-FTF examination of active PDR in either eye	Sensitivity (%)	65/98	66% (57% to 75%)
			Specificity (%)	134/154	87% (81% to 91%)
			Positive likelihood ratio	–	5.11 (3.31 to 7.87)
			Negative likelihood ratio	–	0.39 (0.29 to 0.51)
	G-OPTOS referral for PDR	O-FTF examination of active PDR in either eye requiring treatment	Sensitivity (%)	77/90	86% (77% to 91%)
			Specificity (%)	91/175	52% (45% to 59%)
			Positive likelihood ratio	–	1.78 (1.49 to 2.13)
			Negative likelihood ratio	–	0.28 (0.16 to 0.47)
	G-ETDRS referral for PDR	O-FTF examination of active PDR in either eye requiring treatment	Sensitivity (%)	74/84	88% (79% to 93%)
			Specificity (%)	82/178	46% (39% to 53%)
			Positive likelihood ratio	–	1.63 (1.40 to 1.91)
			Negative likelihood ratio	–	0.26 (0.14 to 0.47)
SENA4	G-OPTOS referral for PDR	O-FTF examination of pre-retinal or vitreous haemorrhage in either eye	Sensitivity (%)	62/71	87% (78% to 93%)
			Specificity (%)	95/193	49% (42% to 56%)
			Positive likelihood ratio	–	1.71 (1.45 to 2.02)
			Negative likelihood ratio	–	0.26 (0.14 to 0.48)
	G-ETDRS referral for PDR	O-FTF examination of pre-retinal or vitreous haemorrhage in either eye	Sensitivity (%)	53/66	80% (69% to 88%)
			Specificity (%)	79/196	40% (34% to 47%)
			Positive likelihood ratio	–	1.35 (1.14 to 1.59)
			Negative likelihood ratio	–	0.49 (0.29 to 0.82)
Additional 2	O-OPTOS identified active PDR	O-FTF examination of pre-retinal or vitreous haemorrhage in either eye	Sensitivity (%)	57/70	81% (71% to 89%)
			Specificity (%)	153/192	80% (73% to 85%)
			Positive likelihood ratio	–	4.01 (2.96 to 5.42)
			Negative likelihood ratio	–	0.23 (0.14 to 0.38)
	O-ETDRS identified active PDR	O-FTF examination of pre-retinal or vitreous Haemorrhage in either eye	Sensitivity (%)	42/64	66% (53% to 76%)
			Specificity (%)	145/188	77% (71% to 83%)
			Positive likelihood ratio	–	2.87 (2.09 to 3.94)
			Negative likelihood ratio	–	0.45 (0.31 to 0.63)

continued

TABLE 15 Diagnostic performance for PDR patients (*continued*)

Analysis	Results	Reference standard	Diagnostic parameter	n/N	Estimate (95% CI)
SENA5	G-OPTOS referral for PDR	ERS	Sensitivity (%)	110/138	80% (72% to 86%)
			Specificity (%)	76/127	60% (51% to 68%)
			Positive likelihood ratio	–	1.98 (1.58 to 2.49)
			Negative likelihood ratio	–	0.34 (0.24 to 0.49)
	G-ETDRS referral for PDR	ERS	Sensitivity (%)	111/135	82% (75% to 88%)
			Specificity (%)	68/127	54% (45% to 62%)
			Positive likelihood ratio	–	1.77 (1.45 to 2.17)
			Negative likelihood ratio	–	0.33 (0.22 to 0.49)
Additional 3	G-OPTOS referral for PDR	O-FTF + O-OPTOS identified active PDR in either eye	Sensitivity (%)	101/125	81% (73% to 87%)
			Specificity (%)	80/140	57% (49% to 65%)
			Positive likelihood ratio	–	1.89 (1.53 to 2.32)
			Negative likelihood ratio	–	0.34 (0.23 to 0.49)
	G-ETDRS referral for PDR	O-FTF + O-ETDRS identified active PDR in either eye	Sensitivity (%)	103/122	84% (77% to 90%)
			Specificity (%)	73/140	52% (44% to 60%)
			Positive likelihood ratio	–	1.76 (1.46 to 2.13)
			Negative likelihood ratio	–	0.30 (0.19 to 0.46)
SENA6	G-OPTOS referral for PDR in routine clinic	O-FTF examination of active PDR in either eye of eligible patients in routine clinic	Sensitivity (%)	63/77	82% (72% to 89%)
			Specificity (%)	47/92	51% (41% to 61%)
			Positive likelihood ratio	–	1.67 (1.32 to 2.11)
			Negative likelihood ratio	–	0.36 (0.21 to 0.60)
	G-ETDRS referral for PDR in routine clinic	O-FTF examination of active PDR in either eye of eligible patients in routine clinic	Sensitivity (%)	60/74	81% (71% to 88%)
			Specificity (%)	41/91	45% (35% to 55%)
			Positive likelihood ratio	–	1.48 (1.19 to 1.83)
			Negative likelihood ratio	–	0.42 (0.25 to 0.71)

G-ETDRS, grader assessment based on seven-field ETDRS fundus images; G-OPTOS, grader assessment based on ultra-wide field fundus images using OPTOS; O-ETDRS, ophthalmologist assessment based on seven-field ETDRS fundus images; O-FTF, reference standard for PDR; O-OPTOS, ophthalmologist assessment based on ultra-wide field fundus images using OPTOS.

a Grader referral for PDR = 'active' + 'unsure' + 'ungradable'.

Results: main analysis for proliferative diabetic retinopathy

The main analysis tested the ophthalmic graders decision of referring patients to ophthalmologists because of the presence of active PDR, they were unsure if PDR was active, or they were unsure if PDR was present or ungradable in either eye based on OPTOS or ETDRS images. The reference standard was the ophthalmologist slit-lamp biomicroscopic examination (O-FTF) identifying active PDR in either eye. Under the main analysis of PDR, graders were found to have similar sensitivity and specificity whether they used UWF fundus images (G-OPTOS: sensitivity 83%, 95% CI 75% to 89%; specificity 54%, 95% CI 46% to 61%) or seven-field ETDRS fundus images (G-ETDRS: sensitivity 85%, 95% CI 77% to 91%; specificity 48%, 95% CI 41% to 56%).

Results: sensitivity analysis for proliferative diabetic retinopathy

Sensitivity analysis 1

Sensitivity analysis 1 (SENA1) assessed the impact of the 'unsure' and 'ungradable' ophthalmic graders' grading results on the diagnostic performance by defining the index test positive as definite assessment of active PDR. Under this analysis, graders were found to have a sensitivity of 63% based on OPTOS (G-OPTOS: sensitivity 63%, 95% CI 53% to 71%; specificity 73%, 95% CI 66% to 79%) and a sensitivity of 71% based on ETDRS (G-ETDRS: sensitivity 71%, 95% CI 61% to 79%; specificity 70%, 95% CI 62% to 76%).

Sensitivity analyses 2, 4, 5 and 6

Values for sensitivity and specificity were very similar to those from the main analysis when graders evaluated patients who required further treatment, as determined by the ophthalmologist in the standard care pathway (G-OPTOS: sensitivity 86%, 95% CI 77% to 91%; specificity 52%, 95% CI 45% to 59%) (G-ETDRS: sensitivity 88%, 95% CI 79% to 93%; specificity 46%, 95% CI 39% to 53%). When determining the sensitivity of the ophthalmic graders' pathway to detect more severe disease (PDR with pre-retinal or vitreous haemorrhage), the sensitivity and specificity of the graders pathway when using ultra-wide-field fundus imaging was slightly higher (G-OPTOS: sensitivity 87%, 95% CI 78% to 93%; specificity 49%, 95% CI 42% to 56%) than that when using seven-field ETDRS images (G-ETDRS: sensitivity 80%, 95% CI 69% to 88%; specificity 40%, 95% CI 34% to 47%) when compared with the reference standard. When determining the sensitivity and specificity of the graders pathway against the ERS, results were similar to the comparison with the reference standard (G-OPTOS sensitivity: 80%, 95% CI 72% to 86%; specificity 60%, 95% CI 51% to 68%) (G-ETDRS: sensitivity 82%, 95% CI 75% to 88%; specificity 54%, 95% CI 45% to 62%). The sensitivity and specificity of graders were found to be similar to those presented in the main analysis when patients were assessed specifically in NHS clinics (vs. in 'research' clinics).

Results: additional analysis for proliferative diabetic retinopathy

Additional analyses 1-3

Three sets of additional ad hoc analyses for PDR patients who were not listed in the statistical analysis plan were also conducted in the light of the results of the pre-planned analysis for PDR to understand the content and potential value of the two imaging modalities better along with the reliability of the O-FTF reference standard for PDR. Additional 1 compared the ophthalmologist identifying active PDR by OPTOS or by ETDRS in either eye with the reference standard of O-FTF assessment of active PDR. The sensitivities of ophthalmologist performance (O-OPTOS: sensitivity 72%, 95% CI 62% to 80%; O-ETDRS: sensitivity 66%, 95% CI 57% to 75%) did not differ greatly from those of graders' performance (SENA1). Ophthalmologist performance was only better than that of the graders when using UWF OPTOS fundus images (difference in sensitivity when comparing the diagnostic performance of ophthalmologists with that of graders based on OPTOS: 9%); whereas graders' performance was

better than that of the ophthalmologists when using seven-field ETDRS fundus images (difference in sensitivity when comparing the diagnostic performance of ophthalmologists with that of graders' based on ETDRS: -5%). However, compared with graders, the ophthalmologists were much better at ruling out not active PDR. The specificity of the ophthalmologists (O-OPTOS specificity 86%, 95% CI 80% to 91%; O-ETDRS specificity 87%, 95% CI 81% to 91%) was higher than that of the graders (G-OPTOS specificity 73%, 95% CI 66% to 79%; G-ETDRS specificity 70%, 95% CI 62% to 76%). Additional 2 compared the ophthalmologist identification of active PDR by OPTOS or by ETDRS in either eye against the reference standard in patients where the O-FTF assessment identified pre-retinal or vitreous haemorrhage (i.e. people with high-risk PDR). Although the specificity of the ophthalmologists (O-OPTOS specificity 80%, 95% CI 73 % to 85%; O-ETDRS specificity 77%, 95% CI 71% to 83%) was much higher than that of the graders (G-OPTOS specificity 49%, 95% CI 42% to 56%; G-ETDRS specificity 40%, 95% CI 34% to 47%), the sensitivity of the ophthalmologists was much lower than that of the graders' pathway when using seven-field ETDRS images (O-ETDRS sensitivity 66%, 95% CI 53% to 76%; G-ETDRS sensitivity 80%, 95% CI 69% to 88%). The sensitivity of the ophthalmologists when using UWF OPTOS fundus imaging was similar to that of the graders (O-OPTOS sensitivity 81%, 95% CI 71% to 89%; G-OPTOS sensitivity 87%, 95% CI 78% to 93%) when compared with the reference standard. Additional 3 compared the ophthalmic graders' referral for either eye by OPTOS or by ETDRS against the O-FTF assessment combined with O-OPTOS or O-FTF assessment combined with O-ETDRS of active PDR in either eye. This additional analysis was aimed at comparing the diagnostic performance of graders using OPTOS or using ETDRS as, in practice, images could be obtained with only one imaging modality. The sensitivity and specificity of the graders were similar to the set reference standard. Some other useful results on the percentage agreement among test positive and reference standard or enhanced standard can be found in *Appendix 1, Table 40*.

Paired comparisons of OPTOS and ETDRS assessment of proliferative diabetic retinopathy

The difference in sensitivity and specificity between OPTOS and ETDRS fundus images assessed by the ophthalmic graders in the main analysis and SENA1 were also compared with corresponding 95% CIs produced using Newcombe's method for paired data³⁰ and McNemar's test.³¹ The results show no evidence of statistically significant difference in sensitivity and specificity between OPTOS and ETDRS fundus images in both analyses (*Tables 16 and 17*). However, a difference of a meaningful amount cannot be ruled out, with 95% CI for difference in sensitivity under SENA1 in particular ranging from -21% to 5%.

TABLE 16 Paired comparisons in sensitivity and specificity between G-OPTOS and G-ETDRS (main analysis)

Main analysis	G-ETDRS		G-ETDRS	
G-OPTOS	Positive	Negative	Positive	Negative
Positive	72	11	45	27
Negative	14	4	35	49
	Sensitivity (n = 101)		Specificity (n = 156)	
OPTOS (95% CI Wilson)	82% (74% to 88%)		54% (46% to 61%)	
ETDRS (95% CI Wilson)	85% (76% to 90%)		49% (41% to 56%)	
Difference (95% CI)	-3% (-14% to 8%)		5% (-5% to 16%)	
McNemar's test <i>p</i> -value	<i>p</i> = 0.55		<i>p</i> = 0.31	
G-ETDRS, grader assessment based on seven-field ETDRS fundus images; G-OPTOS, grader assessment based on ultra-wide field fundus images using OPTOS.				

TABLE 17 Paired comparisons in sensitivity and specificity between grader OPTOS and ETDRS (SENA1)

SENA1	G-ETDRS		G-ETDRS	
G-OPTOS	Positive	Negative	Positive	Negative
Positive	47	14	22	23
Negative	22	15	27	84
	Sensitivity (n = 98)		Specificity (n = 156)	
OPTOS (95% CI Wilson)	62% (52% to 71%)		71% (64% to 78%)	
ETDRS (95% CI Wilson)	70% (61% to 79%)		69% (61% to 75%)	
Difference (95% CI)	-8% (-21% to 5%)		3% (-7% to 12%)	
McNemar's test <i>p</i> -value	<i>p</i> = 0.18		<i>p</i> = 0.57	
G-ETDRS, grader assessment based on seven-field ETDRS fundus images; G-OPTOS, grader assessment based on ultra-wide field fundus images using OPTOS.				

Secondary sensitivity analyses on combined diseases (diabetic macular oedema and proliferative diabetic retinopathy) for all patients

In addition to the principal analyses that were conducted at the person level separately for DMO and PDR, three planned secondary analyses and one additional analysis at person level that focused on the overall referral status of a patient, irrespective of whether it was active (or unsure or ungradable) DMO or PDR that required referral due to disease, were conducted first. These four secondary analyses focused on the entire EMERALD patient population ('All patients') rather than only those patients who were eligible for the new pathway. Table 18 displays the reference standard for the secondary analyses and the ophthalmic graders' grading results for both diseases and for all patients at eye level against the reference standard. The reference standards of O-FTF + OCT assessment for DMO and O-FTF for PDR of those present with active disease were used (either disease constituting active disease at the person level) or active disease which required treatment. Diagnostic performance for the secondary analyses was assessed using the same outcomes and methods as for the principal analyses and the results are given in Table 19.

The results showed an overall high sensitivity for the new ophthalmic grader pathway in detecting DMO and PDR (either or both diseases). The sensitivity of graders referral for combined diseases (SENA2A) was similar among the two combinations of images (G-OCT + G-OPTOS vs. G-OCT + G-ETDRS); however, the specificity was much higher when they used G-OCT + G-OPTOS (sensitivity 96%, 95% CI 92% to 97%; specificity 33%, 95% CI 25% to 42%) than when they used G-OCT + G-ETDRS (sensitivity 95%, 95% CI 92% to 97%; specificity 22%, 95% CI 15% to 30%). Additional 4 was also conducted in the light of the results of the pre-planned analysis, to test graders' assessment of active disease(s) rather than graders' referral of active disease(s) (i.e. including 'unsure' and 'ungradable'). Compared to graders' referral of active disease(s) (including 'unsure' and 'ungradable') (SECA2A), graders' assessment of strictly active disease(s) (Additional 4) presented similar sensitivity (93% compared with 95%), but higher specificity (G-OCT + G-OPTOS, Additional 4: 48% vs SECA2A: 33%; G-OCT + G-ETDRS, Additional 4: 38% vs. SECA2A: 22%). Both SECA2A and Additional 4 demonstrate how the new pathway would function in practice.

Three eye-level secondary analyses were also conducted, and the results are displayed in Appendix 1, Tables 41 and 42.

TABLE 18 Patient-level classifications: all patients (n = 397)

Classification	O-FTF+OCT and O-FTF				Combined test – grader					
	Active DMO ^a	Active PDR ^b	Either eye combined – active DMO or PDR ^c	Either eye combined – active DMO or PDR requiring treatment ^d	G-OCT referral ^e of DMO ^f	G-OPTOS referral of PDR ^g	G-ETDRS referral of PDR ^h	Referral of DMO or PDR based on G-OCT + G-OPTOS ⁱ	Referral of DMO or PDR based on G-OCT + G-ETDRS ^j	Visual acuity ^k
Positive	198	123	272	178	281	207	214	341	350	217
Negative	198	272	125	219	101	168	156	53	40	180
Missing/cannot classify	1	2	0	0	15	22	27	3	7	0

G-ETDRS, grader assessment based on seven-field ETDRS fundus images; G-OPTOS, grader assessment based on ultra-wide field fundus images using OPTOS; G-OCT, grader assessment based on SD-OCT images; O-FTF, reference standard for PDR; O-FTF + OCT, reference standard for DMO.

a 'Active DMO' defined as active DMO (active refers here to 'recurrence' 'de novo' and 'persistence') under ophthalmologist slit-lamp biomicroscopic examination with access to OCT images.

b 'Active PDR' defined as active PDR (active refers here to 'recurrence' 'de novo' and 'persistence') under ophthalmologist slit-lamp biomicroscopic examination.

c Active DMO or PDR (active refers here to 'recurrence' 'de novo' and 'persistence') identified by reference standard in either eye.

d Active DMO or PDR (active refers here to 'recurrence' 'de novo' and 'persistence') identified by reference standard in either eye that requiring treatment.

e Graders' referral for DMO/PDR = 'active' + 'unsure' + 'ungradable'.

f Graders' referral for DMO based on OCT images, which includes 'DMO present', 'Unsure if DMO present' and 'Ungradable'.

g Graders' referral for PDR based on UWF OPTOS fundus images, which includes 'Active PDR', 'Unsure if PDR active', 'Unsure if PDR present' and 'Ungradable'.

h Graders' referral for PDR based on seven-field ETDRS fundus images which includes 'Active PDR', 'Unsure if PDR active', 'Unsure if PDR present' and 'Ungradable'.

i Graders' referral for DMO or PDR based on OCT and UWF OPTOS fundus images, which includes 'active', 'unsure' and 'ungradable'.

j Graders' referral for DMO or PDR based on OCT and seven-field ETDRS fundus images which includes 'active', 'unsure' and 'ungradable'.

k Visual acuity < 6/12 (or equivalent) considered as a referral to ophthalmic clinic.

TABLE 19 Secondary analyses at patient level: all patients

	Results	Reference standard	Diagnostic parameter	n/N	Estimate (95% CI)
SECA2A	G-OPTOS referral ^a for PDR + G-OCT referral for DMO	O-FTF active PDR and O-FTF + OCT active DMO in either eye	Sensitivity (%)	258/270	96% (92% to 97%)
			Specificity (%)	41/124	33% (25% to 42%)
			Positive likelihood ratio	–	1.43 (1.26 to 1.62)
			Negative likelihood ratio	–	0.13 (0.07 to 0.25)
	G-ETDRS referral for PDR + G-OCT referral for DMO	O-FTF active PDR and O-FTF + OCT active DMO in either eye	Sensitivity (%)	253/266	95% (92% to 97%)
			Specificity (%)	27/124	22% (15% to 30%)
			Positive likelihood ratio	–	1.22 (1.10 to 1.34)
			Negative likelihood ratio	–	0.22 (0.12 to 0.42)
Additional 4	G-OPTOS identified active PDR + G-OCT identified active DMO	O-FTF active PDR and O-FTF + OCT active DMO in either eye	Sensitivity (%)	245/263	93% (89% to 96%)
			Specificity (%)	56/117	48% (39% to 57%)
			Positive likelihood ratio	–	1.79 (1.50 to 2.13)
			Negative likelihood ratio	–	0.14 (0.09 to 0.23)
	G-ETDRS identified active PDR + G-OCT identified active DMO	O-FTF active PDR and O-FTF + OCT active DMO in either eye	Sensitivity (%)	242/260	93% (89% to 96%)
			Specificity (%)	45/117	38% (30% to 48%)
			Positive likelihood ratio	–	1.51 (1.31 to 1.75)
			Negative likelihood ratio	–	0.18 (0.11 to 0.30)
SECA2B	G-OPTOS referral for PDR + G-OCT referral for DMO + VA	O-FTF active PDR and O-FTF + OCT active DMO in either eye	Sensitivity (%)	263/271	97% (94% to 99%)
			Specificity (%)	26/124	21% (15% to 29%)
			Positive likelihood ratio	–	1.23 (1.12 to 1.35)
			Negative likelihood ratio	–	0.14 (0.07 to 0.30)
	G-ETDRS referral for PDR + G-OCT referral for DMO + VA	O-FTF active PDR and O-FTF + OCT active DMO in either eye	Sensitivity (%)	262/270	97% (94% to 98%)
			Specificity (%)	20/124	16% (11% to 24%)
			Positive likelihood ratio	–	1.16 (1.07 to 1.25)
			Negative likelihood ratio	–	0.18 (0.08 to 0.41)
					continued

TABLE 19 Secondary analyses at patient level: all patients (*continued*)

	Results	Reference standard	Diagnostic parameter	n/N	Estimate (95% CI)
SECA2C	G-OPTOS referral for PDR + G-OCT referral for DMO + VA	O-FTF active PDR and O-FTF + OCT active DMO in either eye requiring treatment	Sensitivity (%)	174/178	98% (94% to 99%)
			Specificity (%)	30/217	14% (10% to 19%)
			Positive likelihood ratio	–	1.13 (1.07 to 1.20)
			Negative likelihood ratio	–	0.16 (0.06 to 0.45)
	G-ETDRS referral for PDR + G-OCT referral for DMO + VA	O-FTF active PDR and O-FTF + OCT active DMO in either eye requiring treatment	Sensitivity (%)	174/177	98% (95% to 99%)
			Specificity (%)	25/217	12% (8% to 16%)
			Positive likelihood ratio	–	1.11 (1.06 to 1.17)
			Negative likelihood ratio	–	0.15 (0.05 to 0.48)
G-ETDRS, grader assessment based on seven-field ETDRS fundus images; G-OCT, grader assessment based on SD-OCT images; G-OPTOS, grader assessment based on ultra-wide field fundus images using OPTOS; O-FTF, reference standard for PDR; O-FTF + OCT, reference standard for DMO; VA, Visual acuity < 6/12 considered as a referral to ophthalmic clinic.					
a Grader referral for DMO/PDR = ‘active’ + ‘unsure’ + ‘ungradable’.					

Discussion

The diagnostic performance of the EMERALD ophthalmic graders' pathway (proposed new pathway) for the detection of active, recurrent DMO and/or PDR (i.e. in previously successfully treated eyes, reactivation of the disease with fluid present indicative of DMO and active PDR present) was compared for the EMERALD population against the current standard of care, which was considered the reference standard (O-FTF + OCT for DMO and O-FTF for PDR). Participants included in the study had at least one of the two diseases in at least one of their two eyes (except one participant who had an artificial left eye) and had been previously successfully treated (disease became 'inactive' for DMO = no fluid present, for PDR = no signs of active PDR). The EMERALD population included a greater number of participants with active DMO and a smaller number of participants with active PDR. It is important to note that the overall EMERALD cohort should not be taken as a representative reflection of the clinical population. This is due to the balance of groups being intentionally manipulated as part of the study design to increase statistical efficiency in terms of data and to reduce the overall sample size needed. The order in which recruitment to the four target groups ceased does seem to indicate that active PDR is less common among patients being monitored for recurrence of DMO or PDR in hospital eye services (which makes sense clinically as recurrences are infrequent in people with PDR following laser PRP).

The sensitivity of the new pathway to detect active DMO was around 95% (range 94% to 97%) in the main analysis and in all four sensitivity analyses undertaken in EMERALD, suggesting that, for DMO, the new pathway would be safe and could be implemented. The main analysis that was based on graders' decision to refer patients to ophthalmologists had the highest sensitivity across analyses, and a relatively lower specificity. Comparatively, SENA1, which was based on graders' definite detection of active DMO only (excluding the 'unsure' and 'ungradable'), had a small trade-off of higher specificity for a lower sensitivity. The new pathway also performed very well in ruling in and ruling out the presence of central-involving DMO (in contrast to non-central-involving DMO). This is important as central-involving DMO is expected to affect sight and would require identification and treatment earlier than non-central-involving DMO which often does not affect vision and would require time to progress to central-involving DMO. Thus, missing the latter until next evaluation would be expected to have less of a repercussion on the patient than missing central-involving DMO.

One methodological issue to note is that the grader DMO diagnostic accuracy results may be somewhat optimistic in terms of the true performance of a grader's OCT for DMO assessment as the OCT was also part of the reference standard. Nevertheless, this reflects current clinical practice in the NHS and in terms of answering the research question of this study this is not, in our view, a substantive concern. SD-OCT assessments were undertaken by different individuals for the index test (grader) versus its use in the reference standard (determined by the ophthalmologist); this will somewhat reduce the optimism bias associated with the same imaging modality informing the reference standard as well as being the index test. No imputation of any missing data was undertaken in the EMERALD analysis. Given the small number of missing images, this would not have been expected to affect the results presented. The impact of the 'unsure' and 'ungradable' assessment by the graders was explored.

The sensitivity of the new pathway to detect active PDR was generally over 80% (range 63% to 88%) in the main analysis and in five sensitivity analyses undertaken in EMERALD; the exception was SENA1, which evaluated by graders' performance specifically related to the identification of active PDR. The trade-off of lower sensitivity for higher specificity was substantial if graders would only allow referral to the ophthalmologist for those patients who present with active PDR (excluding those in whom they are unsure and/or those that cannot be graded, i.e. ungradable) to ophthalmologist. Comparing the diagnostic performance of the two different types of images that the graders assessed and on which they based their referring decisions, there was no evidence of statistically significant differences in sensitivity and specificity. This suggests that both the ultra-wide field fundus images and the seven-field fundus images could be used, although there may be some minor but still important differences between the two modalities in terms of sensitivity and specificity (circa 10%). However, it should be noted that ultra-wide field fundus imaging had slightly higher sensitivity to detect people with high-risk active PDR (PDR with pre-retinal and/or vitreous haemorrhage), suggesting that if a new pathway involving graders were to be introduced for the surveillance of people with treated PDR, this technology might be preferred to reduce the number of people who could potentially be missed with this more severe disease.

The diagnostic accuracy results for PDR raise a number of issues and potential areas for further research. Graders would seem to find assessment of PDR more difficult than that of DMO, both in terms of differentiating between active and inactive PDR, and also identifying the presence of PDR. The identification of DMO is certainly simpler; on a SD-OCT, fluid is clearly visualised. The lower specificity in DMO for the grader pathway could be explained by the fact that graders may have 'referred' patients even with minimal amount of fluid; these cases may not have been classed by ophthalmologists as 'DMO' (active DMO). Identifying PDR, however, requires the meticulous evaluation of the entire retina; in the case of the graders and ophthalmologists assessing seven-field and ultra-wide fundus images, and the very detailed check of every field of the seven fields in ETDRS and every field of the three fields of the ultra-wide field images for the presence of new vessels. For both diseases there was a small number of eyes 'unclassifiable' due to no view of fundus that led to no assessment under the respective reference standard and these were excluded from the analyses. In practice these may or may not be assessed with imaging prior to referral to an ophthalmologist, if the new pathway were in use.

An assessment of the graders' diagnostic performance was also evaluated against the enhanced reference, in a planned analysis. The findings of the various assessments, and the relative levels of agreement suggest that each modality (including the ophthalmic slit-lamp assessment) may be missing a substantial proportion of people with active PDR. Further work to explore this would be valuable. A number of additional analyses were carried out for PDR to explore the findings of the planned analyses further. These suggest that there is not generally an issue with the grader's assessment for PDR versus an ophthalmologist's assessment for PDR based on both OPTOS and ETDRS imaging (with the exception that the ophthalmologists have a lower number of uncertain assessment and, therefore, rule out active disease more often). This suggests a more inherent issue that the different modalities are not assessing exactly the same thing or at the very least each has its own share (and a roughly similar amount) of disagreeable assessments.

EMERALD used a cross-sectional case-referent design.²⁴ This played a key role in reducing the cost of the study substantially in that it was not longitudinal. It was also efficient in terms of reducing the overall sample size (which also reduced costs). Furthermore, it benefited from being able to use some participants in the main analysis of both DMO and PDR. As the data were readily available and as patients in clinical practice will be assessed for both diseases in both eyes irrespective of history, the full data were collected for all participants. This is reflected in the substantially larger number of people with inactive PDR and active DMO than that set as the target ($n = 104$). It also provided disease-naïve assessments and indication of de novo disease even though this was not the primary reason for an individual to be in the monitoring pathway. The case-referent design, however, has limitations over a cohort design. It does not provide an overall assessment of prevalence of the respective diseases and disease states (active/inactive). An implication of this was the change in the relative proportion of the four groups of interest (DMO inactive and active, and PDR active and inactive) at the time when groups were still being recruited versus the final numbers once recruitment had ceased to all four groups (see *Appendix 1, Table 43*). Furthermore, the number of referrals from this analysis will not translate directly to application in clinical practice due to the use of the case-referral design. The use of a case-referent design might be viewed as potentially increasing the likelihood of a bias sample. However, as consecutive individuals were assessed for eligibility and approached, it is very unlikely that it would have biased the sample in EMERALD. In EMERALD, all patients considered potentially eligible (even those with media opacities, poor co-operation, etc.) were approached to ensure that a consecutive un-biased sample was obtained. Patients not approached and recruited would mainly reflect the incapacity of the EMERALD team to recruit in very busy NHS clinics and, thus, would not be expected to have caused any systematic bias. Assessment of imaging was done independently and separately from the reference standard for both DMO and PDR, avoiding the potential for bias. Similarly, the grading of one imaging technology was done independently of the grading of another, preventing the reading of one imaging modality from affecting the other.

As patients with either disease are routinely assessed in the monitoring for DMO and PDR in both eyes, a secondary analysis explored the performance of a combined overall strategy of grader assessments for DMO and PDR. The overall sensitivity of combining tests was over 96% in all three diagnosis sensitivity analyses, which further supports the claim that the new ophthalmic grader pathway would be safe and could be implemented in clinical practice for the surveillance of people with previously successfully treated DMO and PDR. The various results show the difficulties in using multiple tests and the increasing chance of a false positive result, which would lead to referral to the ophthalmologist. EMERALD was a cross-sectional study; monitoring over time would lead to a substantial number of individuals being unnecessarily referred. However, the poor performance in regard to specificity was more driven by the PDR assessment than by the DMO. When referral related to presence of 'active PDR' only, sensitivity and specificity improved. Given current pressures in the NHS to meet demands, reducing the number of visits by one-third would be a sizeable release of valuable ophthalmologist time (in this regard, see also *Chapter 4*).

Chapter 3 Focus group research

The aim of the focus group work

The aim of the FG work was to explore the acceptability of a new pathway, the ophthalmic grader pathway, for the surveillance and care of current NHS patients with previously treated and stable DMO and PDR.

Focus groups can function in various ways. In this case, they were used to signal key issues that concern patients in their routine encounters with ophthalmology clinics. FGs cannot tell us how those issues are distributed in a wider population, or enlighten us about every issue that might concern patients (other types of research, such as randomised survey research, would be needed to provide data on how opinions are distributed by age, class, ethnicity or gender). However, and as is indicated in *Appendix 2*, there are sound reasons to believe that techniques of the kind used in FGs can capture a considerable amount of information about attitudes, opinions, behaviours, hopes and fears in any given population. Most importantly, they help to sharpen our vision about what is important from 'the patient's point of view', because despite the fact that a FG discussion is guided, it is not determined by a questioner (as is the case in research driven by a structured interview schedule).

Methods

Sampling and recruitment strategy

The original plan was to recruit FG participants via ophthalmology clinics and according to the strategy outlined in *Table 20*. All patients had to have previously successfully treated DMO and/or PDR to be eligible; at the time of enrolment in the EMERALD study their disease could be active or inactive. Some patients had one or the other disease (i.e. only DMO or only PDR), whereas others were under review for both conditions.

Unfortunately, as the recruitment of patients for the main study progressed it became increasingly clear that the targets for FG recruitment were unlikely to be met. Many clinics were seemingly unable to gain consent from more than a handful of patients for a FG meeting. Naturally, some clinics were more successful than others, and towards the end of 2018 it was decided to concentrate on just five localities: Belfast, Fife, Frimley, Gloucester and London. Even then, the prospect of holding FG meetings for different age groups had to be abandoned in four of the five centres. Holding gender-specific meetings was, however, possible in three of the centres.

TABLE 20 Proposed maximum variation sampling strategy for FGs in the UK: number of groups in each category by country and age group

Country	Age group (years)			Total
	18–25	26–60	> 60	
Scotland	1	2	2	5
England	1	2	2	5
Northern Ireland	1	2	2	5
Total number of groups	3	6	6	15
Estimated number of individuals	15–24	30–48	30–48	75–120

Consent is one thing and attending a group is another, and so it proved. Only approximately one-third of those who had consented and agreed to attend a meeting did so. As a result, our final sample was as described in *Table 21*. The average age of all patient attendees was 59.75 years. The lowest average age was 48 years (for a Belfast group) and the highest average age was 70 years (for the group from Fife). In the text that follows, the FG identification numbers do not reflect the order of the groups in *Table 21*.

What happened in the focus groups?

The activities were carried out in five stages:

1. Introduction – following a welcome and first name introductions, the sessions began with a brief outline of the EMERALD project. It was explained to participants that although proceedings would be recorded and subsequently transcribed, whatever was said within the meeting would remain confidential. It was also stated that the doctors who organised their care and treatment were ‘very keen to get their views about the service that patients were offered, and about some changes in the service that were being considered’. (Patient consent had been obtained and information was given to patients by doctors and research nurses at each clinic before the FG meeting.)
2. Phase 1: images – the investigative component of the sessions began with the moderator showing the members of the group a series of photographs. The first photograph was of a doctor and a patient engaged in a slit-lamp examination. Participants were asked what, in particular, they appreciated (or disliked) about ‘this stage’ of routine clinic visits. The second photograph was of a nurse (in uniform) and a patient undertaking an acuity test. Again, respondents were asked what they appreciated about ‘this stage of your visit’. The third photograph was of a room with a chair/bench, in which anti-vascular endothelial growth factor (anti-VEGF) injections (a standard treatment for people with DMO) were given; participants were asked to comment on their experiences. The fourth photograph was of fundus imaging equipment [specifically an ultra-wide field fundus camera (OPTOS) in use and an spectral domain optical coherence tomography (SD-OCT) machine in the background]. The photograph was designed to stimulate discussion about what patients appreciated (or disliked) about their engagement with the photographer. The fifth and sixth photographs were of a waiting room and an empty corridor lined with chairs: typical of almost any corridor queue in a medical clinic. Again, participants were asked to comment on their ‘experiences’ with waiting.
3. Phase 2: vignettes – three vignettes were presented to participants. All of the vignettes concerned the provision of a virtual clinic. The first presented the story of 58-year-old man (Mr Smith) with DR who attended the ophthalmology clinic, had an eye scan and was told that he would get a result by letter. The second was a story of an ophthalmologist (Dr Xi) who wished to reduce waiting times for new patients and, therefore, decided to triage his existing patients so as to concentrate on those whose condition was ‘active’. The third was of a 65-year-old woman (Mrs McWilliams) with DMO who was told that she will need to see the consultant only if the ophthalmic grader detects some changes in her eye condition. All three vignettes served as useful stimuli for discussions on virtual clinics.
4. Phase 3: a ranking exercise – participants were asked, as a group, to rank each of six statements in order of importance to them, and to arrive at an agreed order. The statements are listed in *Table 22*.
5. End of meeting – the meetings ended with an opportunity for participants to raise any issues that they ‘felt had not been dealt with’. They were thanked for their attendance and given £20 in cash per person for travel expenses.

Focus groups for professionals

The FG for consultants and the two FGs for ophthalmic photographers and graders took a different form. Some of the above materials were used, but only in the context of feeding back patient views to the professional group. Participants were encouraged to comment on the views of patients and to suggest ways in which patient concerns might be addressed.

TABLE 21 Number of patients who were invited, gave consent and attended FG meetings at five UK study sites

Study site	Location for FG	Date	Number of patients invited and consented	FG membership	Number of patients who attended
Belfast	Europa Hotel	16 February 2019	3	Males aged > 60 years	2
Belfast	Europa Hotel	16 February 2019	6	All female participants	2
Belfast	Europa Hotel	16 February 2019	24	Males aged 26–60 years	7
Belfast	Europa Hotel	23 March 2019	11	All who have consented	2
QM Fife	Holiday Inn, Dunfermline	9 March 2019	10	All male participants	5
QM Fife	Holiday Inn, Dunfermline	9 March 2019	6	All female participants	1
Frimley	Farnborough Village Club	16 March 2019	21	All who have consented	11
Gloucester	Holiday Inn	6 April 2019	10	All female participants	2
Gloucester	Holiday Inn	6 April 2019	6	All male participants	2
KCL	King's College London, Denmark Hill	11 May 2019	11	All who have consented	2
Total			108		36
KCL, King's College London; QM, Queen Margaret Hospital (Dunfermline, Fife).					

TABLE 22 Ranking of preferences

Statement	Rank group 1	Rank group 2
1. Being told that my eyes are stable or not at the end of clinic visit	3	4
2. Being able to discuss my eye condition with someone at the clinic	4	3
3. More time with the doctor if my eyes get worse	1	2
4. Shorter waiting times for new patients	5	1
5. Spending less time waiting at the clinic	6	5
6. Seeing the eye doctor during every visit	2	6

Analysing the data

The audio files were listened to repeatedly. A number of key issues were subsequently identified by the researcher, and a list of codes or node labels was composed. Node labels were subsequently attached by hand to relevant parts of the data transcripts. Examples of such node labels are present in *Figure 12*. For example, the 'Lampex' node refers to talk surrounding a slit-lamp biomicroscopy phase of a clinical examination. 'Inject' refers to talk concerning anti-VEGF injections and 'Wait' refers to talk about the experience of waiting in clinic for different procedures to be conducted. Most of the nodal issues were introduced to the participants by the researcher. These included issues about waiting times, technology and tools in use, and experiences of interaction with, for example, consultants, photographers and nurses. However, a number of other issues were brought into the discussion by participants independently. One such issue was 'Diabetes'. Thus, participants often identified themselves as 'diabetic', sometimes as type 1 or type 2, or as insulin dependent for X number of years and so forth, and often implied that their diabetes (rather than retinopathy) was their number one health issue. Nevertheless, the most important issue to be brought forward in discussion was related to 'feedback', from the professionals to the patient. We will return to the issue of feedback later. Following the identification of nodal issues or themes, the Microsoft Word® (Microsoft Corporation, Redmond, WA, USA) version of the transcripts was searched

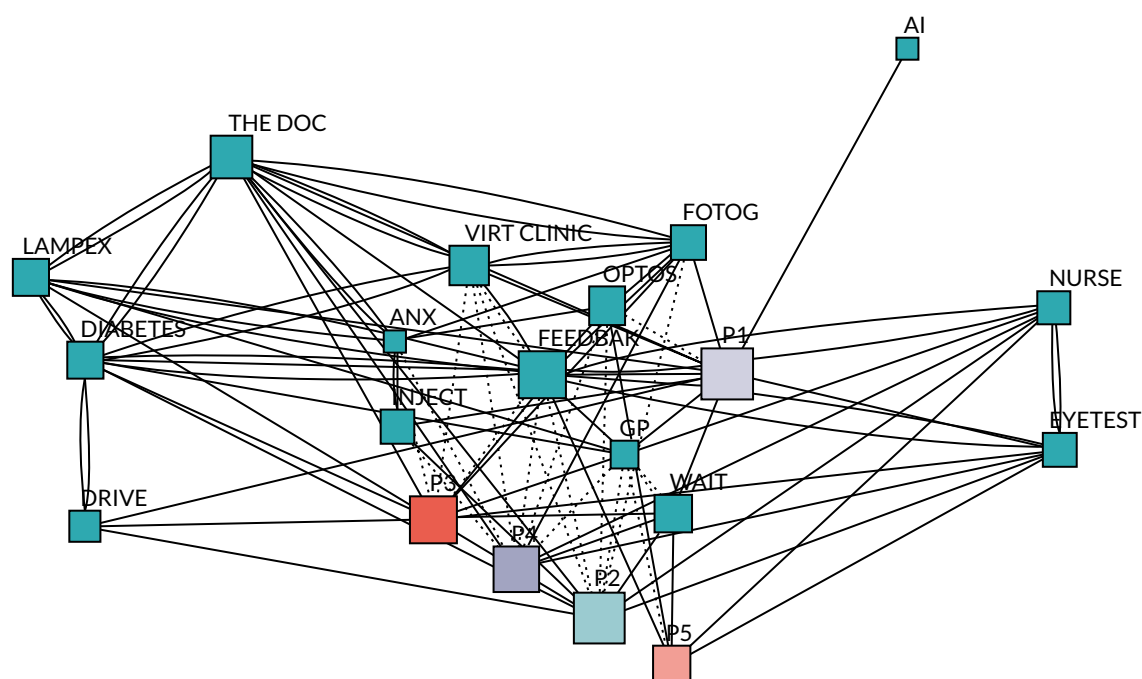


FIGURE 12 Issues discussed by male participants in focus group 5. ANX, anxiety; AI, artificial intelligence for detecting change in retina; DIABETES, my diabetes; DRIVE, car driving; THE DOC, consultant; FEEDBAK, information and results for 'me'; FOTOG, photographer; GP, general practitioner; INJECT, injections; LAMPEX, slit-lamp examination with doctor; OPTOS, Optos ultra-wide field imager; VIRT CLINIC, virtual consultation; WAIT, waiting during routine visits. Participants are labelled 'Pn'.

for key terms (e.g. photographer/grader/scan/image/technician) so as to identify all potential talk about a nodal issue in any one FG meeting.

Figure 12 provides a diagrammatic representation of issues discussed in one FG. The diagram is presented in the form of an 'issue web', outlining the range of issues discussed by participants in FG5. We can see how one issue links to another, which are the most frequently discussed issues and who contributes to the discussion and have a glimpse into whether something is evaluated positively or negatively (broken lines suggest a negative relationship). An explanation of how Figure 12 was constructed is provided in Box 1.

BOX 1 How Figure 12 was constructed

The issue web was constructed using textual codes and numerical counts. Initially, each 'turn' (or phase of talk) in the transcript was linked to an identifiable speaker. Following that, the content of the turn was allocated to a node label or code (sometimes a number of codes). Node labels used included 'my diabetes', 'injections', 'eye test', 'virtual clinic', 'the doctor', 'nurse', 'photographer' and so forth. A simple count of the number of times that a specific speaker could be linked to a code, and the number of times that one code was associated with another in the same turn, was subsequently used as the basis for the construction of a matrix (a 20×20 square matrix in the case of Figure 12). The matrix was then integrated into social network software (using Pajek³²) to generate a graphical representation of the discussion. Within the graph, node size reflects the number of turns that an individual speaker took during the meeting, or the number of times that an issue was referred to. The thickness of the links between nodes (the arcs) reflects the number of times that any one code was associated with another in the responses of participants. Because the diagram was generated using a Fruchterman-Reingold projection, distances between nodes are suggestive of the closeness (or otherwise) of the links between them (unfortunately, overlapping P2, P3 and P4 nodes, had to be separated manually to enhance clarity). Given large variations in the node and arc size, the counts were scaled using a square root transformation.

Issue webs, such as the above, offer a view on the conversational proceedings, albeit one that generates a somewhat truncated image of the entire field. *Figure 12* shows, for example, that some speakers took more turns than others (P5, smaller red box, spoke least; P1, larger yellow box, spoke the most). It also shows that the major concern of these participants was with what they called 'feedback' (large pink box towards the centre of the diagram) or getting information on their eye condition, usually at the time of a clinic visit. Also note that participants introduced topics that were not included in the moderator's script. Examples include references to 'my diabetes', driving and the GP service. More importantly, we can detect differences in attitude to, say, the acceptance of a virtual consultation (dotted lines between nodes signal negative associations) or the use of fundus imaging equipment (specifically the OPTOS ultra-wide field imager). Thus, participants 2, 3 and 4 openly disliked the idea of a virtual clinic: 'For the peace of mind of the patient, they should be seen regularly by the consultant'. Equally, participant 1 'hated [OPTOS] with a passion'. Finally, we can see that participants frequently cross-reference issues (line thickness indicates a strength of association between one topic and another); indeed, they rarely talked in terms of single issues or themes. For the sake of clarity, however, the presentation of findings that follows is presented in terms of single, sequential concerns, rather than one that emphasises cross-referencing and cross-talk.

Findings

Naturally, each FG contained divergent views and raised specific matters of interest. However, given that all groups were led through the same sequence of topics, it is not hard to identify common themes or familiar response patterns emerging from the transcripts. However, the researcher has attempted to be sensitive to diversity of views where such diversity was apparent. In most cases, the origin of a quote is attributed to a specific FG using a number (FGn). In some cases, however, the identification of the FG has been omitted so as to obscure information that could lead to the identification of a participant, hospital or health professional.

Focus group phase 1: the journey through the clinic

As stated above, phase 1 of the FG was built around a set of photographs that were intended to stimulate discussion around a version of current clinic practice in which patients move through various stages of the ophthalmology service (e.g. from the waiting room to the room in which the visual acuity is evaluated, back to the waiting room, to the photographer, SD-OCT, 'Optos' ultra-wide field imager, accepting an anti-VEGF injection and ending with a slit-lamp examination with the consultant). The trajectory and the objects in the pictures were seemingly familiar to all participants.

Waiting

'Waiting' was not generally regarded as a problem by any of the patients in the FGs. Participants accepted that long waits were often to be expected and, therefore, ought to be planned for. Where delays in clinic processes did occur, patients wanted some kind of signal as to the extent of the delay. A number of patients attending for routine appointments aimed to get to their clinics early in the day so as to avoid long waits. Others suggested taking some form of distraction (such as a book or an audio device) to cope with the waiting. In FG1, however, one participant suggested that waiting was not a problem because he was retired, but had he been working then the amount of time spent at a clinic would have been critical. So it may well be that attitudes to waiting were, at least partly, a reflection of the employment status of the FG members. There is some support for that conclusion from the FGs with consultants, ophthalmic photographers and graders, in which some of the participants complained about 'DNAs' (did not attend) and claimed that non-attendance was a significant problem.

'Waiting' also arose in a few other contexts. First, there were issues around waiting for the results of a clinical investigation. Second, there were issues connected to waiting lists (for an appointment). Both sets of issues will be referred to in the sections that follow.

Engagement with the nursing staff

The photograph of a patient undertaking an acuity test was the trigger for discussions of nurse–patient relationships. Patients appreciated interaction with nursing staff and thought that good interpersonal skills of nurses were essential to a good clinic visit. However, it seemed that most interaction was strictly functional: nurses administered, for example, eye drops, injections and acuity tests, and for most people verbal interaction was based on ‘small talk’ or ‘chit chat’. Thus, when asked if they ‘spoke to the nurse’ at all, one participant in FG8 replied ‘Yeah: about the weather’. However, patients did appreciate when nurses provided feedback, such as telling a patient that their acuity test results implied that there was no change in their eye sight or perhaps that it had improved a touch. The male participants in FG5, for example, were very appreciative of their interaction with nurses. For example, the moderator asks about ‘talking to the nurse’ during the acuity phase of the clinic visit (in quotations, Mn refers to Male Participant n):

M1: It's good to get feedback. It's sort of 3/4 months between appointments and it's a case of I'm reading this [line of print], is that the one I read the last time? And they actually tell you, they've got notes there in front of them, you read one smaller than the last time and just give you the feedback, there's a slight deterioration, or not, as the case may be.

M2: When you're actually doing the eye test with the chart, I like the way they ... certainly, my experience is I like the way they try and encourage you to go another one down ...

M3: Yeah.

Similarly, male participants in FG9 stated that they always asked questions of the nurse (with respect to intraocular pressure, acuity test results and so forth) and appreciated ‘knowing what’s what’. However, although nurses might convey bits of information to a patient, they were not regarded as a key source of knowledge about a patient’s underlying eye condition. For that, a consultant was required. ‘I want to ask questions’, said a female participant in FG6, ‘but I don’t until I see the consultant’; ‘I like seeing the main one’.

Engagement with the photographers

Most of the FG participants seemed confused about the status of the photographers. The female participants in FG8 and the male participants in FG3 thought that they were specialist nurses (which was not the case). The male participants in FG9 had no name for the people who took images. A participant in FG4 referred to ‘the guy on the machine’ and a participant in FG1 spoke of ‘the two men and the girl who do the scans’. A participant in FG5 thought that they were dealing with an optometrist. The male participants in FG10 thought that the photographers were doctors (although they also made a reference to a nurse practitioner at a later stage). In FG7, the following exchange took place [in quotations, Fn refers to Female Participant n, and LP refers to Lindsay Prior (who undertook the focus group discussion)]:

LP: When you meet photographers in the clinic, do you discuss anything with them?

F1: What photographers do we meet?

M1: The people who take scans of the back of your eye.

F1: Oh those ...

M2: Oh, yeah, oh yeah.

M3: The lady who runs the scan machine.

LP: The lady who runs the scan machine, OK.

F1: She's quite sweet, and she says 'oh, just go and see the doctor'. Wait for [some] hours.

Some participants referred to photographers as ‘technicians’ (FGs 6 and 7), and only one participant (an active member of a diabetes support group in FG5) mentioned the word ‘grader’. However, whatever word was used to describe the role of the person who took images of the retina, it was clear that the level of information exchange between photographer and patient was minimal:

LP: When you have your photographs taken, who takes the photographs?

F: The technician.

LP: The technician, yeah. Do you speak to the technician at all?

F: Yes. Usually just social conversation.

There was also confusion about the level of understanding and expertise that photographers/technicians/scanners had in determining whether an eye condition was active or stable.

It is worth noting that the technical details of the imaging technology used in clinics appeared to be of little significance to FG participants: patients assessed technology in terms of how it meshed into interactional contexts. Thus, the slit-lamp examination was spoken of favourably because it brought the patient into contact with the consultant, whereas the OPTOS ultra-wide field imager was disliked partly because it introduced a barrier between the patient and the clinic personnel. Generally speaking, ‘technology’ was viewed as a friend and its use was welcomed for the detection of eye problems (see section 3.3.2). Nevertheless, imaging equipment was always evaluated in the context of the social relations that were fostered or hindered by it.

Meeting with the ophthalmologist

It appeared that nearly all of the participants in the FG discussions currently met with an ophthalmologist during their routine clinic visits. Many participants spoke of their relationship with their consultant as a personal one, established perhaps over some time and based on mutual respect. (This was in stark contrast to the ways in which some FG members spoke of their GP, about whom they could be openly negative.)

Meeting the consultant during a clinic visit was above all ‘reassuring’. Thus, one of the male FG participants stated that ‘I think it gives you a lot of reassurance to speak one to one to [consultant X], because there’s questions that you ask that only probably [he] could answer. That’s what I like, to be quite honest’. In addition, a female participant stated that meeting with the consultant offered ‘a great reassurance that you know you’re going to be looked at, and at the end of day, the treatment that I have had, I can’t thank them enough for it, you know’. Some FG participants even mentioned feeling ‘relaxed’ once they had spoken with a consultant.

Asking questions about one’s condition and how problems might develop were regarded as essential by most participants. Thus, one of the male participants in FG3 stated that he appreciated ‘the fact it’s an ophthalmologist there, because it’s not just, you know, the checking of your eyes that’s important at these visits, it’s being able to ask questions’, and a male participant in FG5 claimed that seeing the doctor ‘... is necessary, isn’t it? It’s for the doctor to look into your eye and see and make an opinion. I mean, I don’t see my GP, I haven’t seen my GP for years. I see a nurse specialising in diabetes. I’ve seen [the ophthalmologist] more than I’ve seen my GP’. In FG9, one of the participants spoke of being able to discuss his eye problems ‘in-depth’ whenever he met with his consultant.

As the above extracts suggest, asking questions and getting answers (‘there and then’) constituted both the high point and a suitable end point for routine clinic visits. Indeed, the capacity for getting ‘feedback’ and obtaining ‘information’ was regarded as a key feature of the patient–doctor interaction,

which is something that shows up very clearly in *Figure 12*. It was the potential loss of this capacity (in a virtual clinic) that worried patients most of all. (One male participant in FG3 stated that he would 'go private' rather than miss out on seeing a consultant.)

Administration of the clinic

Although most FG participants were positive about the medical care that they received, there were a number of occasions when they were critical of the 'non-clinical admin side'. Nearly all of the negative responses concerned the process of getting an appointment. In some cases, it seemed, there was a disjunction between the time until next appointment as specified to the patient by a consultant (e.g. 'see you again in 3 months') and the ability of the patient to get an appointment in the specified time frame. Some patients spoke of how they had to 'fight' for an appointment; a member of one FG said:

It's no good being told, when you phone up after 6 months, 'where's my appointment?'. 'Oh, we're not seeing anybody now until May,' or 4 or 5 months down the line, that's really not good enough.

On the other hand, there were some FGs in which both the medical and the administrative arrangements were regarded as praiseworthy, as illustrated by this quotation from a male participant in FG9:

The other thing about X Hospital Eye Department, they keep a very good diary; once they've got you, they don't let you go.

Diabetes and anxiety

Diabetes was not only a condition that had to be managed, but also very clearly a source of identity for many patients. This was particularly evident in the discussions between FG members that occurred before the formal start of proceedings. Individuals would often introduce themselves as a 'type 1' or 'type 2', compare the number of years that they had 'been on insulin' or compare the age at which they were first diagnosed as a person with diabetes. In some cases it was evident that diabetes drove the daily schedule of an individual, in which one course of action had to be 'juggled' with another, and the disorder was uppermost in the thoughts of many FG participants. Some FG members were puzzled as to why the review of their eye condition and the review of their diabetes were undertaken at different times and at different clinics; thus, a male participant in FG10 stated:

Holistic is the only way to deal with diabetes.

Another male participant from FG3 spoke of his new clinic in the following terms:

It was bright, it had a TV screen there that told you a bit about the different eye conditions, there was lots of literature and promotional material for different support groups and things like that, and I found the staff... I have never written a compliment letter in all my time as a diabetic, but I did after my first appointment there, because I thought it was really impressive compared to [the previous clinic].

Anxiety emerged as another cross-cutting theme. Anxiety might be ever present, but it could also be accentuated by clinical or administrative procedures. Not being told about the result of a scan, receiving a clinic re-call letter or the use of an ill-judged phrase could send 'blood pressure through the roof' because 'of all the senses, the sight is probably people's most important one not to lose. I would say that we all worry about our sight being lost' (FG3).

Some further references to anxiety appear in the next section.

Focus group phase 2: the virtual clinic

The acceptability of a 'virtual consultation' or 'virtual clinic' was explored via the use of three vignettes. (However, the word 'virtual' was not used by either the moderator or the participants.) The discussion that ensued from the presentation of the vignettes moved in various directions, as is evident from *Figure 12*.

For most participants, however, the provision of a clinic in which people were assessed for 'stability' and given routine assessments and treatments, but not informed of the result 'on the day', was unacceptable:

Yeah, I mean, anxiety, everybody's the same, but if you are anxious about something like this, it can put your blood sugar up and that screws up your general diabetic health. So I just think it's important to go in, if you find out there's bad news and something needs treated, I mean, they have the stuff there, get it all done [on the day].

M, FG3

Thus, patients spoke of the virtual clinic as 'half a service', a 'backward step' and 'retrograde'. Fears were also expressed about being 'isolated' and 'cast adrift' without an opportunity to 'have a talk [with the consultant] and get any problems in'. One participant in FG8 spoke about the following:

Waiting for the postman to bring a letter and you're worrying, you're anxious. And the thing is, with diabetes, being anxious and stressed doesn't do it any good.

The male participants in FG5 also traced a direct link between waiting and anxiety. Another female participant in FG6 talked of being 'left in limbo' in the absence of any information about her eyesight and the need for 'reassurance' before leaving the clinic:

Walking away not knowing; that's not nice.

In a similar manner, a male participant in FG1 stated that he would not want 'to go away [from the clinic] wondering', and a participant in FG9 stated that he would not want to 'go away thinking, is there something wrong?'. The following exchange occurred in FG2 with respect to the vignette on 'Mr Smith':

F2: *So he doesn't know the results, does he?*

LP: *He doesn't.*

F2: *So he's sat there worried sick.*

Those who regarded the provision of a virtual clinic as acceptable did so on a number of grounds. Some were accepting of the pathway as long as their eyes were stable (although there is something of a 'catch-22' there). A few accepted the proposal on the grounds of 'trust'. Thus, when the moderator asked one of the male participants in FG4 how he would feel in his own case, if an ophthalmic photographer/grader said that he would get a letter about his results, he stated:

That's fine. I trust him, he knows his stuff.

Likewise, a member of FG10 said the following of the nurse practitioner who dealt with the eye scans:

I trust her implicitly. Frankly, I trust her more than I trust some of the younger doctors.

Some participants (e.g. in FG4 and FG10) saw the proposed arrangement as equivalent to and as acceptable as the screening service that they had experienced in other medical settings.

In FG7, divergence emerged in the following exchange, which opens with a participant talking about leaving the clinic without a 'result', and consequently waiting for the letter that will confirm stability or suggest further investigation:

M7: I think that's the problem, is waiting for that letter. I'd rather, on the day, be told rather than have to wait. Like the sword of Damocles hanging over me.

Yes [several answer].

F3: Plus, letters get lost as well.

M2: I must admit, I've been on the reception of that, and it's worked without any problems at all. I got the letter, it told me what I needed to know, and they said 'we'll see you in 3 months,' and I knew exactly there was a plan and that was fine. But I could imagine, as the lady said over there . . .

LP: And you didn't see the consultant at all?

M2: No. And I didn't need to, because again, if it's an objective test and objective results then why do I need to see a doctor?

Exchanges of that kind sometimes led to discussion as to how 'results' might be communicated to patients. As is evident above, patients were not always favourably disposed to traditional mail as a form of communication, and suggested that other media could be effectively used for the transmission of test results, including e-mails, text messages and cell telephone messages; these routes were considered to be quicker than communication by letter. A few participants thought that it would be advantageous to have their scanned images mailed to them in addition to any evaluations.

Linked to fears and anxieties about not seeing the consultant were concerns about the professional skills and expertise of the ophthalmic photographers and graders. As already indicated, the individuals who do the scanning are of uncertain status to the participants; they are seen only through a 'glass darkly', as it were. Consequently, few participants were happy about leaving the decision of whether or not DR was active to a non-medical professional of any kind (the male participant referred to in FG10, above, was unusual in this respect). This is despite the moderator indicating that the individuals who reviewed the images were highly trained. Thus, a female participant in FG2 said of the person who scanned her eyes that he was 'just a photographer'. A participant FG3 suggested that 'someone more trained' should be looking at the images. A participant in FG1 was adamant that decisions of whether or not the retinopathy was 'active' 'should not be left to the photographer/grader'. In FG6, some of the participants stressed the fact that consultants had expert or tacit knowledge that was simply unavailable to others.

In a number of FGs, discussion turned to the issue of whether a 'robot' or an advanced artificial intelligence (AI) system might be able to detect change in the retina better than a human actor (AI appears in *Figure 1*, for example). However, none of those who said that they would feel content with a robotic assessment imagined that the (expert) consultant could be removed from the care pathway; clinical judgements were for clinicians alone.

In nearly all FGs, participants drew a distinction between themselves and future patients. Current patients were used to and comfortable with the existing service and would not want it changed. Future patients, who might experience the virtual service, might not know any better. Changing the care pathway now, and for 'us', would only lead to dissatisfaction. In addition, if change were to occur

then the rationale for such change ought to be fully explained to those affected. It was a position nicely summed up by a member of FG9:

People like [John] and myself, we've been there for some period of time now and we know the procedures and we know what we need to do and we know what they need to do as well, because it becomes a matter of fact, and we look for that all the time. If something changes you're aware of it straightaway. If they want to change the system then they've got to advertise the fact to all people who are under their care so that they're aware, otherwise it falls flat and people get concerned.

Focus group phase 3: the ranking exercise

The ranking exercise was originally intended as a means of stimulating discussion among participants about the key issues, and perhaps reaching a consensus on what was most important to them. In truth, discussion in most groups was minimal, but the final rank order of key statements did serve as something of a confirmatory signal on the 'feeling of the meeting'. For example, in FG7 it seemed to the moderator that two distinct camps emerged on the acceptability of the virtual clinic; therefore, instead of asking everyone present to rank the six statements, the group was divided into two. As can be seen from Table 22, there was indeed a considerable divergence between the two factions on the importance of 'seeing the eye doctor during every visit' (highest rank = 1; lowest rank = 6).

As with most other groups, the issue of waiting was regarded as of low importance. Reducing waiting times for new patients was also of low importance for most groups. The high rank of this item in group 2 was explained to the moderator by saying that 'it was number 1 for others' but not for 'us'. For most of the other FGs, the highest-ranked statement was statement 1.

Responses of consultants

The moderator began the meeting with the six consultants by asking them to outline the potential advantages of the EMERALD care plan for patients, the NHS and themselves. Various responses emerged, but foremost was the idea that the EMERALD study would reduce in-clinic waiting times for patients (as well as referral times for new patients). Given that the patient groups persistently claimed that in-clinic 'waiting' was not a problem, the issue deserves some consideration.

From a consultant stand point, it seems clear that failure to attend clinic appointments (did not attends) are a problem for professionals and patients. Thus, one consultant spoke of patients with very serious conditions who failed to attend the retina clinic because of long in-clinic waiting times, and who had to be chased up in other settings (e.g. the glaucoma clinic). In the light of this, a reduction of routine waiting times could well be an advantage for patients. It may be (as was suggested above) that the FG sample was biased towards people who were retired or not economically active, and for whom in-clinic 'waiting' was a minor inconvenience (however, FG1, FG3, FG5, FG6 and FG10 clearly contained some economically active and self-employed individuals). Either way, there appears to be a strong difference of opinion here on the virtues of reduced in-clinic waiting times. Other advantages of the EMERALD study that were mentioned by the consultants included the potential for reduced (labour) costs for the NHS, and an increased ability to provide personalised treatment to patients with active conditions. This latter advantage is in precise accord with the wishes of the patients (as expressed via the ranking exercise); however, it is equally clear that patients welcome a medical consultation even when their condition is 'stable'.

On patient reactions to the photograph of the slit-lamp examination, consultants adopted a straightforward technical stance. The slit-lamp examination was useful for 'active' eyes only. There was no need to use such methods for patients who were stable, and no real need to meet with such patients. It was further argued that the use of SD-OCT images plus ultra-wide field photography was 'as good as a personal exam' for detecting changes to eye tissue. However, as we have noted above, patients view the clinic as a site for the exchange of information and a source of reassurance, as well as a site for diagnosis and treatment.

Therefore, for patients, the slit-lamp photograph reflected an occasion in which questions could be put and answers obtained from a medical expert. Consequently, one issue that arises out of the provision of virtual clinics is where patients can get advice and reassurance about their condition. Various suggestions were made by consultants. For example, one consultant reported that she currently told the clinic photographers to tell patients that they could ask for her to contact them by telephone if they wished to discuss anything. Others suggested that more 'patient education' was required: informing patients of the merits of SD-OCT and ultra-wide field images over a personal examination.

On the provision of virtual clinics more generally, it seemed as if they were already in operation in some regions. Two consultants reported that although there had been some complaints when virtual clinics were originally organised, there were no complaints any more. Both new and long-term patients (who had experienced the previous care pathway) seemed to accept the reorganised service, and a view was expressed that the new care pathway as proposed by the EMERALD study will work 'eventually' (none of the FG patients was drawn from these clinics).

When asked to discuss how patients might get 'results' of clinic reviews quickly, a number of responses were offered. At the simplest level were proposals to e-mail, telephone or text message results to patients (which is exactly in accord with what patients suggest). More advanced was a proposal that patients could be offered a system whereby they could log-in to a personal account and 'view results for themselves', which would also be in accord with some of the more technically minded patients. Another suggestion was that patients be offered a guarantee that images would be reviewed and assessments returned within 48 hours.

In addition to communication strategies of the above kind, some consultants proposed organisational change. For example, it was suggested that if photographers were expertly trained to grader status, it would be possible for the graders to give 'results' to most patients immediately. One consultant suggested that a grader might sit with the photographer and offer on-the-spot assessments. Whatever technical, administrative and legal difficulties there might be with such proposals, it quickly became clear that they would meet with resource problems of various kinds. For example, one consultant claimed that there were no graders available to her retina clinic as they were all deployed in the screening service. Implicit agreement among the consultants did, however, emerge on some points. First, that it might be possible to get 'results' back to patients in a '1-day virtual clinic' if there were sufficient numbers of trained graders to assess images. Second, that even if results were to be forwarded to patients on a day other than the day of a clinic visit, but in a timely manner, more photographers would need to be trained to grader status.

Finally, issues relating to diabetes emerged during the discussion. The problem of dealing with retinopathy in the context of diabetes was fully recognised, and a number of references were made to a Royal College of Ophthalmologists working-group on the potential provision of a 'one-stop shop' for patients.

Responses of photographers and graders

Two FGs containing five photographers and two graders were held. As with the consultants, the moderator began the meeting by asking participants to outline the potential advantages of the proposed EMERALD care plan for patients, the NHS and themselves. Interestingly, the photographers and graders paid quite a bit of attention to the infrastructure costs of the new pathway, for example purchase of imaging equipment, buying software licenses and training of graders. As with the consultants, however, the primary advantage of the proposed EMERALD pathway for patients was said to be a reduction of in-clinic waiting time; participants voiced surprise at the low ranking of the 'waiting time' item in the patient-group ranking exercise.

Photographers and graders were well aware of patient anxiety regarding getting 'results'. Some said that, currently, they were nearly always asked for indications as to the patients 'stability'. As a consequence, all had developed conversational strategies (or stock replies) for dealing with the problem. For example, a photographer/grader might say to a patient:

Well; we have some very good images now and they just need to be examined by the doctor.

Another response was:

As you know, you're in this clinic because you are low-risk and this is just a routine check-up.

As with consultants, photographers and graders thought that it would be a good idea to explore the use of electronic media (as well as traditional mail) for getting results to patients. An existing time frame for getting results to patients was mentioned as '4 to 6 weeks' in one of the FGs. As to the idea of graders giving results to patients, there was considerable reluctance on a number of grounds. One participant mentioned lack of indemnity insurance as a barrier. Another stated that, even if they were trained to grader status, they would feel too 'vulnerable' to be responsible for making final decisions on images. More fundamentally, both photographers and graders suggested that although it might be OK for suitably trained individuals to give a result to a patient who was evidently 'stable', in cases of active conditions a number of problems arose. Primarily, offering a judgement about new bleeds or changes in the retina would lead the patient to ask further questions about consequences and treatments: questions that the graders felt that they would not be able to answer. It was suggested that intimating to a patient that something had changed, without offering a treatment or advice response there and then, would only intensify anxiety. It was considered far better to notify patients at a later stage, telling them how they could be accommodated in a consultant-led clinic. In addition, nearly all of the photographers and graders pointed out that there were always cases in which they simply did not know whether or not there was something awry; some images required interpretation by a medical expert.

All but one of the photographers and graders seemed to have some experience of virtual clinics, such as maculopathy or glaucoma clinics, and considered them to be both highly efficient and acceptable to most of the patients who they encountered. It was suggested that a virtual retinopathy clinic was comparable to a screening clinic and fulfilled the same public health role, with an emphasis on preventative rather than curative medicine. As with the consultant group, considerable emphasis was placed on the need for 'patient education' if virtual clinics were to be fully acceptable.

Discussion

Patients tend to see the ophthalmology clinic as offering a space in which diagnosis, treatment and the exchange of information concerning DR and other eye conditions takes place. For the participants in this study, all three functions were equally important and a clinic that failed to offer all three was generally regarded as less than good, possibly 'half a service'. Indeed, most members of most of the FGs feared that with a virtual clinic the third function would be missing or delayed. They also voiced fears that a lack of 'engagement' and 'reassurance' would characterise such a clinic. Only the male participants in FG4 and FG10, who spoke of 'trust' and 'confidence' in the clinical service, were unperturbed about the possibility of a virtual clinic.

Although participants did not speak of a gold-standard service, it is not unreasonable to summarise patient opinion in terms of what 'gold'-, 'silver'- and 'bronze'-standard services would offer to patients.

A gold-standard clinic would offer a 'one-stop shop' (FG3). Here, eyes would be investigated (via scans, photographs, etc.), relevant treatment (such as anti-VEGF injections) would be offered on the day,

and information and discussion about eye conditions and diabetes would be available from a medically trained professional: an expert. A patient whose condition was stable would exit the clinic with 'results'. In cases in which new activity was detected, the nature and consequences of such activity would be explained, and treatment options would be offered 'within a reasonable time, say a week' (FG6). It is a system that is summed up by the following male participant in FG6:

I want it over with on the day, not mess around, if and but, maybe there's something, maybe there's not, maybe I've got a problem, on tenterhooks the whole time. Get it sorted on the day. If you are stable, there's no problem, you nip in and see the doctor, 5 minutes, straight out, no problem; it doesn't really take any effort.

A silver- to bronze-standard clinic would be one in which screening and essential treatments were offered and the exchange of information within the clinic was minimal, but in which results of routine check-ups were available within days of a clinic visit, communicated via e-mail, text or telephone. In cases in which new activity was detected, access to an ophthalmologist would be timely: patients would not have to 'struggle to get an appointment'. (According to some consultants, it seems that this kind of service is already in operation in a few centres, although it was not clear how much of a time gap occurred between the patient visit to the clinic and the receipt of a 'result'.)

A zinc-standard clinic would be one in which screening and routine treatments were offered, but in which the exchange of information between the patient and the medical professionals was minimal. 'Results' of routine screening were delayed for more than a few days, and medically trained professionals were rarely met with. In cases of active DR, the zinc-standard clinic would inform the patient by letter that there was a problem, without offering what patients considered a timely appointment for consultation with an ophthalmologist (the longer the delay the greater the anxiety).

One of the key issues that emerged from the FGs was the emphasis that patients placed on getting results from a clinic visit 'on the day'. For a virtual clinic, it would be a critical problem. For example, although rapid results may be well suited to patients whose eye condition is stable, they may not be beneficial for those whose condition is active, especially in the absence of opportunities for further discussion or explanation of the consequences. Indeed, as both the photographers and the graders suggested, any hint or suggestion that things were 'not quite right' would only accentuate anxiety without offering 'solutions'. Thus, it would probably be more acceptable that in the presence of active disease patients were contacted at a later date with a plan for a timely consultation in which 'results' could be fully discussed.

Summary

Patients in the 10 FGs were overwhelmingly positive about the care they currently received and were supportive of the NHS in general. They recognised that the existing ophthalmology service was 'under pressure', but still viewed the service as meeting their needs. They had no wish to see it altered in any fundamental way.

Patients spoke of the ophthalmology clinic as a place in which eye problems are diagnosed and treated, and in which information and advice ('feedback') about their eye condition can be obtained. The provision of information, especially 'results', and advice is regarded as central to a good service.

A good service was also one that communicated the result of investigations on the day of a clinic visit, and provided an opportunity to discuss eye problems with a health professional, preferably a doctor.

Although patients were supportive of the work of all those who worked in the clinics that they attended, they regarded the consultant as the true expert and as the primary source of reliable information about their condition.

The role of photographers and graders in the ophthalmology service was unclear and uncertain as far as patients were concerned. Nearly all of the patients were confused about who was 'behind the scanner' as well as the ability of graders to detect emerging eye problems.

Views on the provision of a 'virtual clinic' or 'virtual consultation' (in which a patient is reviewed via images rather than a face-to-face examination) were mixed. Most participants disliked the idea of such an arrangement and saw it as offering 'half a service'. Fears were expressed about being 'isolated', 'forgotten' and 'cast adrift'.

Those who were unperturbed about the provision of a virtual clinic regarded it as acceptable only as long as 'results', and any necessary treatment, were provided quickly.

The data suggest that in planning 'virtual clinics' for people with previously treated and stable DMO and/or PDR, the following issues ought to be considered:

- Using a range of media (including but not limited to postal communications) to deliver test results and clinical assessments to the patients quickly and efficiently.
- In cases in which 'activity' is identified in the retina by a grader, timely appointments to meet with an ophthalmologist should be offered.
- An organisational capacity to provide 'reassurance' and assuage anxiety, even for patients who are stable or low risk, ought to be recognised as key features of the ophthalmology service.

It appears that the image, role and function of ophthalmic photographers and graders need and deserve sharper definition, no matter what the outcome of the EMERALD trial.

Conclusion and implications

For patients, the ophthalmology clinic is nestled in a set of human relationships. Consequently, the clinical service is not simply one in which pathology can be identified and treated, but also a service in which knowledge and information about eye conditions can be exchanged, reassurances can be given and anxieties assuaged. Consultant-staffed clinics offer all of the above.

Patients fear that a virtual clinic will fall short on offering reassurance, information and advice, and that they are likely to lose from this change if such clinics are rolled out. (However, FG participants recognised that future cohorts of patients who would be unaware of the current system may well accept a virtual pathway without further ado.) The suggestion of the consultants and photographers/graders who have worked with virtual clinics is that patients will come to accept them and to appreciate the shorter waiting times involved.

In the absence of meeting a consultant at a clinic visit, patients want feedback on routine investigations fairly promptly. In fact, probably the best way of 'selling' the virtual clinic to patients would be to guarantee quick 'results'; doing so for patients who are stable is probably feasible. Communicating results to patients for whom some degree of activity is evident in their images, however, is more problematic. The suggestion that something is 'wrong' would undoubtedly generate anxiety in patients, and both photographers and graders suggest that little purpose would be served by patients receiving bad news without having recourse to on-site medical advice.

However, no matter the conclusions of the EMERALD study, it is evident that the identity, role and function of photographers and graders needs to be clarified, both for the benefit of patients and for the good of the imaging specialists. Above all, patients deserve to know the level of expertise of the person who is reviewing their images and who it is that is making decisions about their treatment and care.

Chapter 4 Cost–consequences analysis

Introduction

In this chapter, we examine the costs and consequences of the new ophthalmic grader pathway for the surveillance of people in whom treatment for DMO or PDR has resulted in the condition in their eyes being considered inactive and stable. The purpose of follow-up in these patients is to detect reactivation (i.e. 'recurrence') of DMO or PDR so that further timely treatment can be given, if required.

The current standard of care for the surveillance of patients with DMO is slit-lamp biomicroscopy undertaken by an ophthalmologist who has access to SD-OCT images. In EMERALD, the new pathway for DMO entails the reading of SD-OCT images by trained ophthalmic graders.

The current standard of care for the surveillance of patients with PDR is slit-lamp examination undertaken by an ophthalmologist. The new pathway tested in EMERALD for PDR is the reading of either seven-field ETDRS fundus photographs or ultra-wide field fundus photographs by trained ophthalmic graders.

In addition to the comparison between findings in the grader's pathway and findings in the standard of care pathway for PDR, we examined the costs and consequences of using the ophthalmic grader pathway as compared with the results of the ERS. Under the ERS, a patient was considered to have active PDR if detected on any one of the slit-lamp biomicroscopic examination undertaken by an ophthalmologist, the grading of seven-field fundus photographs by an ophthalmologist or the grading of ultra-wide field fundus images by an ophthalmologist. Data on the comparisons between ophthalmic grader's pathway, reference standard and ERS is contained in *Chapter 2* (see *Table 15*).

Methods

Resource use was captured on trial CRFs at each participant's EMERALD clinic visit to compare the costs of delivering the standard care pathway, the ophthalmic grader pathways and the ERS. The cost analysis took the perspective of the NHS and Personal Social Services and costs are in Great British pounds in 2019/20 prices.

The costs included staff costs, based on the time taken for various procedures, including:

- undertaking best corrected visual acuity
- the time for the photographers/imaging technicians (and their grade) to obtain SD-OCT, seven-field and ultra-wide field fundus images, costed at band-6 salary scale
- for DMO and PDR, the time and grade of the ophthalmologist undertaking slit-lamp biomicroscopy and review of the SD-OCT images (for DMO) as well as the time invested counselling the patient
- the time taken by graders to grade SD-OCT, seven-field and ultra-wide field fundus photographs, costed at band-7 salary scale.
- the time taken for ophthalmologists to grade seven-field and ultra-wide field fundus photographs for the purpose of the ERS.

The time and grade for the above procedures were used to estimate costs of the standard and new care pathway (ophthalmic grader pathway), and the ERS.

The staff costs reflect NHS salary, superannuation, national insurance and overhead costs attributable to different grades of staff (*Table 23*).

Other costs included the equipment required, overheads and consumables.

Costs were applied to each resource item to value total resource use in each pathway using national sources such as the *Unit Costs of Health and Social Care*,³³ published by the Personal Social Services Research Unit (PSSRU) annually, and *NHS Reference Costs 2017/2018*.³⁵

Both the standard care pathway as currently undertaken and the new ophthalmic grader pathway would require a nurse, an optometrist or a visual acuity technician to check visual acuity. Therefore, this cost would not differ between pathways, nor would the cost of taking the OCT images.

The costs of obtaining and grading seven-field ETDRS fundus images were estimated and compared with those for ultra-wide field fundus photography. If there was no difference in diagnostic accuracy between these two imaging modalities, the lower cost one would be favoured.

The costs of the various investigations are not simply the 'direct cost' of time multiplied by salary cost, because many other 'indirect' costs need to be taken into account, including support staff (e.g. clinic reception), medical records, if needed (some clinics have electronic records) or IT support, as well as 'hotel' costs (e.g. buildings, heating and lighting). These are included in the standard reference cost for ophthalmology follow-up visits, which we use as the base-case cost.³⁴

By deriving the direct cost of the ophthalmologist's time for the face-to-face examination and comparing that with the *NHS Reference Costs 2017/2018*,³⁵ we obtained an adjustment factor to apply to the direct costs (salary multiplied by time) of grader or ophthalmologist reading of photographs. This is not a perfect solution because the indirect costs might vary between ophthalmologists and graders, but is a pragmatic approach.

TABLE 23 Staff costs (2019 prices)

Resource item	AfC pay scale band	Unit cost (£)	Measurement unit	Source
Ophthalmic photographer/imaging technician	6	49.00	Per working hour	^a PSSRU 2019, page 143 ³³
Ophthalmic grader	7	59.00	Per working hour	PSSRU 2019, page 143 ³³
Ophthalmologist (consultant medical)	N/A	109.00	Per working hour	PSSRU 2019, page 150 ³³
Associate specialist	N/A	108.00	Per working hour	PSSRU 2019, page 150 ³³
Specialty registrar ^b	N/A	47.00	Per working hour	PSSRU 2019, page 150 ³³
Ophthalmologist (average cost) ^c	N/A	108.50	Per working hour	Average based on unit costs of two grades of staff
Ophthalmologist outpatient follow-up appointment (slit-lamp examination)	N/A	58.00	Per patient contact	NHS 2019/20 <i>National Tariff Payment System</i> ³⁴

AfC, agenda for change; N/A, not applicable.

^a We chose the cost per hourly rate of radiographers as the more relevant unit cost for ophthalmic imaging technicians.

^b Although patients may be seen by specialty registrars, we assumed that unless advanced in training, most will be closely supervised by associate specialists or consultants, hence the average unit cost for an ophthalmologist (used in current analysis) is based on associate specialist and consultant salaries.

^c Average unit cost for an ophthalmologist based on associate specialist and consultant salaries.

In *Table 24* we have included the costs of equipment, such as camera and OCT machines, based on the acquisition and maintenance costs, divided by the lifetime of the equipment and throughputs of 9000 patients per year.³⁶ These estimates are approximate, but, given that they make up a very small proportion of total costs, the approximations are acceptable. Costs are discounted at 3.5% per annum.⁴⁰

The ERS consisted of the evaluation of the patient by the ophthalmologist using slit-lamp biomicroscopy in clinic and, in addition, the ophthalmologist evaluation of the seven-field ETDRS fundus photographs and ultra-wide field fundus images. Obtaining and reading such images is not currently part of standard care. To maintain masking, the ophthalmologist evaluating the patient by slit-lamp biomicroscopy was not the same as the ophthalmologist who graded the fundus images. Furthermore, to avoid the reading of one technology influencing the reading of the other, different ophthalmologists graded the seven-field ETDRS images and the ultra-wide field fundus images. In the ERS, any patient graded as having 'active PDR' based on any one of slit-lamp biomicroscopy, seven-field ETDRS fundus photographs or ultra-wide field fundus images would be considered to have active PDR.

We also provide some exploratory post hoc comparisons not included in the original protocol.¹ We provide these in view of the findings in *Table 15* of the diagnostic accuracy results, which suggest that the ophthalmologist face-to-face examination (O-FTF) (reference standard) may be less sensitive than the ERS.

Outcome measures

Health-related quality of life was assessed using the EQ-5D-5L instrument. The EQ-5D-5L is a generic measure of health status that consists of a descriptive system and the visual analogue scale. The descriptive system includes five questions addressing mobility, self-care, usual activities, pain or discomfort, and anxiety or depression. Each dimension is assessed at five levels ranging from 'no problems' to 'extreme problems'. The EQ-5D-5L utility scores were derived using an EQ-5D-5L index value calculator (version 2.0)⁴¹ that maps the EQ-5D-5L descriptive system data onto the EuroQol-5

TABLE 24 Unit costs of equipment (2019 prices)

Cost variable	Current cost (£)	Lifespan (years)	Annual throughput ^a	Total annual/ discounted costs (£)	Cost per patient (£)
Seven-field ETDRS imaging camera: Topcon TRC-MW-8	14,500: purchase price	8	9000	2431.18	0.27
	500: annual maintenance costs costed from year 3 to year 8 ^b				
Ultra-wide angle imaging equipment: Optos California aqua RG/AF/FA/ICG	88,255: purchase price ^c	10	9000	11,243.17	1.25
	5250: maintenance cost sold as extended warranty (once-off fee)				
Slit lamp: Haag-Streit-BM-900-Table-LED-Slit-Lamp	11,300: purchase price ^d	10	9000	1569.15	0.17
	1750: maintenance cost sold as extended warranty (once-off fee)				

a Annual throughput estimate: University Hospital Southampton NHS Foundation Trust³⁶ and personal communication (Professor Noemi Lois, Wellcome-Wolfson Institute for Experimental Medicine, School of Medicine, Dentistry and Biomedical Sciences, Queen's University Belfast, 2020).

b Equipment price: Topcon Ireland Medical, Dublin, Ireland.³⁷

c Equipment price: Optos Plc, Dunfermline, UK.³⁸

d Equipment price: Veatch Ophthalmic Instruments.³⁹

Dimensions, three-level version (EQ-5D-3L) valuation set.⁴² Detailed description of the mapping methodology is described elsewhere.⁴² We also estimated quality of life using two vision-specific measures: NEI VFQ-25⁴³ and VisQoL questionnaires.⁴⁴ The NEI VFQ-25 is widely validated, and contains 11 subscales (general vision, near activities, distance activities, peripheral vision, colour vision, driving, ocular pain, role difficulties, dependency, social functioning and mental health) and a general health rating item. A composite score was obtained by averaging the scores across the 11 subscales. The VisQoL questionnaire is a six-item-level instrument that focuses on the effects of visual impairment on the risk of injury, coping, friendships, organising assistance, fulfilling roles and everyday activities. A standardised score for the VisQoL was computed using the published algorithm by the Centre for Health Economics, Monash, Australia.⁴⁵ Higher scores represent better function for all three measures.

Our aim was to compare the results obtained from using the much simpler and quicker to administer VisQoL method with NEI VFQ-25 and EQ-5D-5L. Given that this was a cross-sectional study, we simply explore the results of the various outcome measures, including for different levels of visual acuity that are known to be linked with utility levels.^{46,47} We also compare the findings in active DMO and PDR with inactive DMO and PDR using the ophthalmologist's face-to-face (reference standard) classification, with a null hypothesis that there should be no difference.

Results

Costs

The average time taken in minutes to complete specified activities by either the graders or the ophthalmologists is shown in *Table 25*.

Some of the times in the minimum column are very small, but are as recorded in the CRF. They may indicate that problems, such as recurrent PDR, were detected very quickly and a decision was made to refer.

The costs in *Table 23* were obtained from the standard source, PSSRU.³³ It may seem odd that the costs per hour for consultants and associate specialists are almost identical. In addition, no cost is given for staff doctors (sometimes referred to as 'trust doctors').

TABLE 25 Time (in minutes) taken to complete specified activities per patient by ophthalmologists and ophthalmic graders

Time-related activity	Time (minutes)			
	Mean (SD)	Median	Minimum	Maximum ^a
Reference standard: O-FTF	15.3 (7.60)	14	4	50
Ophthalmic photographer/imaging technician taking ETDRS photographs	15.3 (8.37)	13	2	57
Ophthalmic photographer/imaging technician taking ultra-wide field fundus images	10.4 (7.17)	9	1.98	60
Ophthalmic grader reading seven-field ETDRS fundus images	10.6 (6.5)	9	1.98	58
Ophthalmic graders reading ultra-wide field fundus images	9.3 (5.87)	8	1	41
Ophthalmologist reading seven-field ETDRS images	11.9 (4.99)	11	2	31
Ophthalmologists reading ultra-wide field fundus images	10.1 (4.03)	9.5	2	28
Grader reading OCT images	4.2 (3.31)	3.98	0.98	31

^a We excluded entries for which the time to perform task was equal to zero or above 60 minutes, as these values were considered outliers and likely mistakes. Thus, *n* varies slightly across time-related activities.

Patients may be seen by one of four grades of staff:

- consultants – starting salary £80,000 up to £108,000 (after about 20 years; there are yearly increments at years 2–5, then larger increments at 5-yearly intervals)
- associate specialists – starting salary £56,000, potentially up to £96,000 after 20 years
- staff doctors – starting salary £40,000, up to about £76,000 after 20 years
- specialist registrars – under supervision with checking by more senior staff, so costs were not used.

Table 26 presents alternative costing that would give a slightly lower cost for ophthalmologist time. However, in the analysis, we have abided by the standard PSSRU costs. These are based on observed costs, and the absence of a difference in cost between consultants and associate specialists may relate to seniority: if the average seniority of associate specialists is much greater than that of the consultants, they will have accumulated more salary increments.

The NHS reference cost is our default cost. Where no reference cost was available, we used the difference between the NHS reference costs and the time-multiplied salary cost (£26.51) for ophthalmologist time, to derive an indirect cost weight factor of 2.2 to assess the full cost of grader pathways. However, the PSSRU costs include overheads, so the indirect adjustment factor was not necessary. Table 27 shows the costs of the different procedures.

TABLE 26 Alternative costs for staff time

Resource item	AfC pay scale band	Unit cost (£)	Measurement unit	Source
Ophthalmic photographer/imaging technician	6	49	Per hour	^a PSSRU 2019, page 143 ³³
Ophthalmic grader	7	59	Per hour	PSSRU 2019, page 143 ³³
Ophthalmologist (consultant) ^b	N/A	93,764	Per year	Pay and Conditions Circular (M&D) 2/2019 R2, page 13 ^{48,49}
Associate specialist ^c	N/A	74,267	Per year	Pay and Conditions Circular (M&D) 2/2019 R2, page 18 ^{48,49}
Staff grade ^d	N/A	53,287	Per year	Pay and Conditions Circular (M&D) 2/2019 R2, page 23 ^{48,49}
Ophthalmologist (average cost) ^e	N/A	95	Per hour	A calculated average based on unit costs of three grades of staff
Ophthalmologist outpatient follow-up appointment (slit-lamp examination)	N/A	58	Per patient contact	NHS 2019/20 national tariff payment system ³⁴

AfC, agenda for change; M&D, Medical and Dental; N/A, not applicable.

a We chose the cost per hourly rate of radiographers as the most relevant unit cost for ophthalmic imaging technicians.

b We took the mid-point of the consultants' salary scale, i.e. starting salary £80,000 up to £108,000.

c We took the mid-point of the associate specialists' salary scale, i.e. starting salary approximately £56,000 up to £92,000.

d We took the mid-point of the staff doctors' salary scale, i.e. starting salary approximately £37,000 up to £69,500.

e Average cost per hour for an ophthalmologist is based on average salaries of an associate specialist, staff grade doctor and consultant, assuming a 40-hour working week equivalent to 1701 hours. We used an adjustment factor of 2.2 to account for indirect costs.

TABLE 27 The costs of procedures

Item	Cost (SD) per patient (£)	Notes
O-FTF clinic visit for slit-lamp examination (base case)	58.00	NHS 2019/20 national tariff for an ophthalmology outpatient follow-up appointment with a consultant
O-FTF clinic visit for slit-lamp examination (sensitivity analysis)	27.88 (13.75)	Calculated as time cost of conducting assessment × hourly salary rate + equipment cost (per patient)
Ophthalmic photographer/imaging technician obtaining seven-field ETDRS photographs (band 6)	12.78 (6.83)	Time cost × photographer salary + equipment cost
Ophthalmic photographer/imaging technician obtaining ultra-wide field fundus images (band 6)	9.75 (5.85)	Time cost × photographer salary + equipment cost
Reading ETDRS photos by ophthalmic grader (band 7)	10.45 (6.40)	Time cost × grader salary
Reading ultra-wide field fundus images by ophthalmic grader (band 7)	9.10 (5.77)	Time cost × grader salary
Reading ETDRS photos by ophthalmologist	21.58 (9.02)	Time cost × salary
Reading ultra-wide field fundus images by ophthalmologist	18.23 (7.28)	Time cost × salary
Taking OCT	Occurs in both pathways	Cost not included
Reading OCT: grader (band 7)	4.08 (3.25)	Time cost × salary
Equipment cost is a small percentage of total cost. Salary costs include employers costs at 20%. Reading SD-OCT by ophthalmologist is included in the NHS reference cost for ophthalmologist OP visits.		

Diabetic macular oedema: comparison of the ophthalmic grader pathway with current standard of care

The sensitivity of graders as determined in the main analysis of EMERALD was very good, 96.6% in the main analysis (see *Chapter 2, Table 14*), when compared with the standard of care (slit-lamp biomicroscopy undertaken by the ophthalmologist who also had access to SD-OCT images). The specificity of graders, also determined in the main analysis, was not as good, at 31%, although this could still lead to an ophthalmologist seeing approximately one-third fewer patients with inactive DMO in clinics. The main analysis included referrals from graders to ophthalmologists when they detected the presence of DMO (recurrent DMO) but also when graders were not sure if DMO was present (i.e. grading stated 'unsure') or when images could not be graded by the graders (i.e. grader stated 'ungradable'). This would also be the procedure followed if this new pathway were to be used in clinical practice.

However, the savings in ophthalmologist time would apply only if neither eye had active DMO or active PDR, as under these circumstances the patient would still require an ophthalmologist evaluation and treatment. *Table 28* shows the relative costs of follow-up of DMO by ophthalmic graders and ophthalmologists.

We used the results of several sensitivity analyses as reported in *Chapter 2, Table 13*. There were no significant differences in sensitivity, but specificity varied from as low as 21% (i.e. 79% of patients having to be referred to ophthalmologists, reducing the savings) to 56% if graders were asked to identify central-involving active DMO only.

It appears from these analyses that even with low specificity, follow-up by graders could provide useful savings in ophthalmologist time, which could be then redirected to the timely evaluation and treatment of people with sight-threatening disease.

TABLE 28 Relative costs of follow-up of DMO

Pathway	Sensitivity	Specificity	Cost per 100 patients (£) ^a	Cost comparison
Reference standard: O-FTF	Assumed 100%	Assumed 100%	5800.00	
Grader reading (band 7)	97%	31%	$(4.08 \times 100) + (69 \times 58) = 4410.00^b$	Time cost \times salary. Cost compared with standard ophthalmologist outpatient visit
Cost difference			1390.00	Ophthalmologist reference cost for outpatient follow-up visit \times 100 minus (ophthalmologist cost for 69 outpatient visits + grader costs for reading 100 OCTs)

a The costs for grader options take into account the specificity, so that in each 100 patients a proportion will still need to be referred to an ophthalmologist, with the reference standard cost for ophthalmologist follow-up applied. So if specificity of the grader pathway is 31%, then 69% of patients will be seen by ophthalmologists.

b Costs of grader reading images for 100 patients + ophthalmologists' time of reviewing 69 patients classified as possibly active DMO when DMO is inactive, based on specificity.

Proliferative diabetic retinopathy: comparison of grader pathway with current standard of care

The main analysis undertaken in the EMERALD study compares the current standard of care pathway for PDR (i.e. O-FTF with slit-lamp biomicroscopy) with the ophthalmic grader pathway (review of either seven-field ETDRS fundus images or ultra-wide field fundus images by the grader). As in the DMO analysis, in the PDR analysis 'referral' by the grader included those patients identified as having active PDR by the grader and those identified as 'unsure' and 'ungradable' (see the main analysis in Chapter 2, Table 15).

Chapter 2, Table 16, shows a sensitivity of 82% with a specificity of 54% for graders when grading ultra-wide field fundus imaging, and a sensitivity of 85% with a specificity of 48% when grading was based on seven-field ETDRS fundus photographs. The small differences are insignificant (see 95% confidence intervals in Table 16). Ultra-wide field fundus images appear to be more cost-effective given that they provide as good sensitivity and specificity as the ETDRS ones, or in the case of PDR with high-risk characteristics, better sensitivity and specificity, but require less time to take and read, and are therefore less costly.

In SENA1, the diagnostic performance of the graders was assessed based on their specific identification of active PDR ('unsure' and 'ungradable' were not included). Their specificity improved but their sensitivity was much poorer. Therefore, this option would be less costly because the improved specificity means that fewer patients would be referred to ophthalmologists. However, many patients with active disease would be missed. However, this would not be how the grader pathway would be implemented because graders in routine care would be referring patients whether they identified active disease or if they were to be unsure or unable to grade the images.

In SENA2, the ophthalmic grader pathway was assessed by how well it identified eyes requiring treatment [some recurrences of PDR may not need immediate treatment or immediate treatment may not be possible (e.g. in cases of vitreous haemorrhage the view of the back of the eye would be poor and further laser could not be undertaken until the blood clears, which may take weeks to happen); under the latter circumstances, patients will be followed and re-assessed at the next visit]. The graders' sensitivity for detecting these more seriously affected eyes was much better: 86% with ultra-wide field imaging and 88% with seven-field ETDRS photographs. Specificity was 52% and 46%, respectively, which was similar to the main analysis.

Analysis SENA4 shows the sensitivity of ophthalmic graders' assessment of images when the reference standard involves the detection of pre-retinal or vitreous haemorrhages (i.e. people with high-risk PDR). The sensitivity was 87% for ultra-wide field images (95% CI 78% to 93%) and 80% (95% CI 69% to 88%) for seven-field ETDRS images. Specificities were 49% and 40%, respectively.

Table 29 shows costs of care related to PDR as drawn based on the main analysis of EMERALD (comparing the new ophthalmic grader's pathway with the current standard of care: ophthalmologists face-to-face evaluation using slit-lamp biomicroscopy and taking into account referrals by the grader related to 'active PDR', 'unsure' and 'ungradable'). Overall, there could be savings from the graders reading the fundus images compared with the ophthalmologists' face-to-face examinations. Table 30 shows the same comparison but for eyes needing treatment, and Table 31 for eyes with pre-retinal or vitreous haemorrhages.

In both of the above scenarios, the savings are modest with seven-field ETDRS photographs, but higher with ultra-wide field images. However, these scenarios refer only to PDR requiring treatment and PDR with pre-retinal or vitreous haemorrhage; these would be the groups that would be more important not to be missed by the ophthalmic graders.

TABLE 29 Cost comparison: grader pathway and ophthalmologist standard care

Pathway	Sensitivity	Specificity	Cost per 100 patients ^a (£)	Cost saving compared with standard care (£)
O-FTF	Assumed 100%	Assumed 100%	5800.00	–
Grader evaluating seven-field ETDRS fundus images, inclusive referral	85%	48%	5339.00	461.00 ^b
Grader evaluating ultra-wide field fundus images	83%	54%	4611.00	1189.00

a The analysis from which the sensitivity and specificity values are derived in this table are based on either eye being active. If either eye has active PDR, it is referred to the ophthalmologist. The cost will not be increased if both eyes show active PDR.

b Calculated as (time of taking ETDRS photos + costs of grader reading ETDRS photos) × 100 + (52 × O-FTF OP appointment costs). Owing to the specificity, 52% of patients will be referred to the ophthalmologist.

TABLE 30 Cost comparison: grader pathway and detection of eyes needing treatment

Pathway	Sensitivity	Specificity	Cost per 100 patients (£)	Cost saving compared with standard care (£)
O-FTF	Assumed 100%	Assumed 100%	5800.00	–
Grader evaluating seven-field ETDRS fundus images for eyes requiring immediate treatment	88%	46%	5455.00	345.00 ^a
Grader evaluating ultra-wide field fundus images for eyes requiring immediate treatment	86%	52%	4669.00	1131.00

a Calculated as (time of taking ETDRS photos + costs of grader reading ETDRS photos) × 100 + (54 × O-FTF OP appointment costs). Owing to the specificity, 54% of patients will be referred to the ophthalmologist.

TABLE 31 Cost comparison: grader pathway and detection of eyes with pre-retinal of vitreous haemorrhage

Pathway	Sensitivity	Specificity	Cost per 100 patients (£)	Cost saving compared with standard care (£)
O-FTF	Assumed 100%	Assumed 100%	5800.00	–
Grader evaluating seven-field ETDRS fundus images for eyes with haemorrhages	80%	40%	5803.00	No saving ^a
Grader evaluating ultra-wide field fundus images for eyes with haemorrhages	87%	49%	4843.00	957

^a Calculated as (time of taking ETDRS photos + costs of grader reading ETDRS photos) × 100 + (60 × O-FTF outpatient appointment costs). Owing to the specificity, 60% of patients will be referred to the ophthalmologist.

Proliferative diabetic retinopathy: comparison with enhanced reference standard

The cost of the ERS reflects the three elements: O-FTF with slit-lamp biomicroscopy, plus the ophthalmologist reading of both seven-field ETDRS and ultra-wide field fundus images. The ERS in EMERALD was carried out because it was accepted as possible that new vessels could be missed on the ophthalmologist slit-lamp examination but be seen on fundus images. Both sets of photographs, however, could not be carried out routinely in clinic in the NHS. For the cost-consequences analysis, the value of the ERS is that it allows us to compare the costs, sensitivity and specificity of having photographs read by an ophthalmologist or ophthalmic graders. However, for completeness, we give the costs of the ophthalmic grader pathway compared with those of the ERS in Table 32. As before, 'inclusive referral' means that the ophthalmic graders refer to ophthalmologists all eyes in which they diagnose 'active PDR', plus those about which they are 'uncertain' and those with 'ungradable' images.

One interpretation of the above figures is that, assuming that the ERS has 100% sensitivity, the marginal cost of each 1% sensitivity is (£7771.00/21.3) £364.84 for ultra-wide field images. Therefore, using the ERS might seem to be good value if it costs only £364.84 to detect each extra case of PDR. However, the cases of PDR will include a mixture of active PDR requiring immediate treatment and active PDR that is less urgent and might only be brought back at the next regular visit (or earlier).

TABLE 32 Cost and consequences comparison using ERS

Pathway	Sensitivity	Specificity	Cost per 100 patients (£)	Cost saving compared with ERS (£)
Ophthalmologist ERS	Assumed 100%	Assumed 100%	12,034.00 ^a	–
			Note 1	
Grader ETDRS, inclusive referral	82.2%	53.5%	4991.00	7043
Grader OPTOS, inclusive referral	79.7%	58.8%	4263.00	7771

^a The cost of the ERS includes, for each patient, the £58 cost of the O-FTF in the outpatient clinic, plus £12.78 for taking ETDRS images, £9.75 for wide-angle images, £18.23 for reading wide-angle images and £21.58 for reading ETDRS images. The cost would be £2798 lower if only wide-angle images were obtained.

Notes
The ERS comprised the combined findings of the O-FTF, and findings of ophthalmologists reviewing 7-field ETDRS and ultra-wide field images. We have assumed 100% sensitivity and specificity. If the ophthalmologist had only the ultra-wide field images or the seven-field ETDRS images in addition to their O-FTF (as would happen in clinical practice if images were to be added to the current standard of care), the sensitivity might be lower.

The comparison of O-FTF with ERS shows that the ERS identifies more reactivated (recurrent) PDR than was identified using images only. This may not be surprising given that some patients have problems sitting still and focusing on the right point during slit-lamp biomicroscopic examination. However, as shown in *Table 14*, direct O-FTF identifies some eyes with PDR that were not identified from reading the ultra-wide field or seven-field ETDRS fundus images. This suggests that the clinically most effective method of follow-up might be to use a combination of O-FTF and photographs, of which ultra-wide field images appear to be the better choice based on the lower cost of both taking and reading the photographs, and the higher sensitivity and specificity at detecting the more severe form of PDR (i.e. with pre-retinal haemorrhage and/or vitreous haemorrhage).

If the optimal method of follow-up was a combination of O-FTF and reading of ultra-wide field images, one question would be who reads the photographs.

Table 33 shows a simple cost comparison of the reading of photographs by the grader and the ophthalmologist for 100 patients. The data compare ophthalmologist and grader sensitivity for detecting active PDR from ultra-wide field images, using O-FTF as the reference standard. Specificity is higher when ophthalmologists read the photographs, with no overlap of 95% CIs (see *Table 15*). Sensitivity is better by approximately 9% for ophthalmologists reading ultra-wide field images, but CIs overlap. There is little difference in sensitivity between ophthalmologists and graders with ETDRS photographs, with overlapping CIs.

However, in the ERS scenario, the ophthalmologists see all of the patients and would want the images to be available at that time to come to a definite diagnosis and to counsel the patient accordingly. It is likely that the ophthalmologist would want to check areas on the images that had looked suspicious on slit-lamp biomicroscopy. Therefore, even if all of the images came with reports from the ophthalmic graders to the ophthalmologist by the time of the slit-lamp examination, it is likely that the ophthalmologist would still spend time looking at the images and the savings in *Table 33* become an overestimate. The EMERALD study cannot provide data to address this. In any case, logistically, it would be difficult to have a grader grading the images in clinic prior to patients being seen by the ophthalmologists.

Results for EQ-5D-5L, NEI VFQ-25 and VisQoL measures

Tables 34 and *35* show the mean scores for the generic health-related quality-of-life measure EQ-5D-5L and the two vision-specific questionnaires (NEI VFQ-25 and VisQoL) for patients with DMO or PDR. The patients who had active DMO status had slightly higher scores, but these were not statistically significant (see *Table 34*). However, patients who had inactive PDR had slightly higher EQ-5D-5L scores; NEI VFQ-25 and VisQoL scores followed a similar pattern to that observed for DMO. However, the differences were not statistically significant for any of the outcome measures.

TABLE 33 Costs comparison of the reading of images by graders and ophthalmologists for 100 patients

Pathway	Cost per 100 patients (£)
Grader ETDRS	1045.00
Grader OPTOS	910.00
Ophthalmologist ETDRS	2158.00
Ophthalmologist OPTOS	1823.00
Note that in the 'enhanced monitoring' option of O-FTF plus reading of one set of photographs the increased specificity of the ophthalmologist would provide no advantage because all patients would have O-FTF.	

TABLE 34 Mean EQ-5D-5L utility and vision-specific outcome scores by status of DMO (active vs. inactive)

Outcome measure	DMO status						Mean difference	95% CI
	Active			Inactive				
	n	Mean	SD	n	Mean	SD		
EQ-5D-5L utility score	151	0.75	0.28	116	0.74	0.27	0.007	−0.06 to 0.07
EQ-5D-5L visual analogue scale	152	72.05	23.92	120	66.23	25.29	5.83	−0.11 to 11.77
NEI VFQ-25 composite score	152	81.45	19.75	117	78.21	22.58	3.23	−1.95 to 8.42
VisQoL standardised score	146	74.69	21.87	112	70.49	26.75	4.20	−1.93 to 10.33

TABLE 35 Mean EQ-5D-5L utility and vision-specific outcome scores by status of PDR (active vs. inactive)

Outcome measure	PDR status						Mean difference	95% CI
	Active			Inactive				
	n	Mean	SD	n	Mean	SD		
EQ-5D-5L utility score	107	0.72	0.27	165	0.75	0.28	−0.20	−0.09 to 0.04
EQ-5D-5L visual analogue scale score	107	64.03	25.39	165	68.08	25.06	−4.05	−10.12 to 2.03
NEI VFQ-25 composite score	107	80.52	19.91	167	77.32	24.14	3.20	−2.08 to 8.48
VisQoL standardised score	104	71.83	23.47	159	69.62	27.23	2.20	−4.01 to 8.42

Figure 13 shows that there were no significant differences between patients who had active DMO and patients who had inactive DMO in the different NEI VFQ-25 subscales or in the composite score.

Figures 14 and 15 show graphically the scoring of the severity levels for the EQ-5D-5L dimensions. Overall, there was no difference between the active or the inactive eye status for patients with either DMO or PDR.

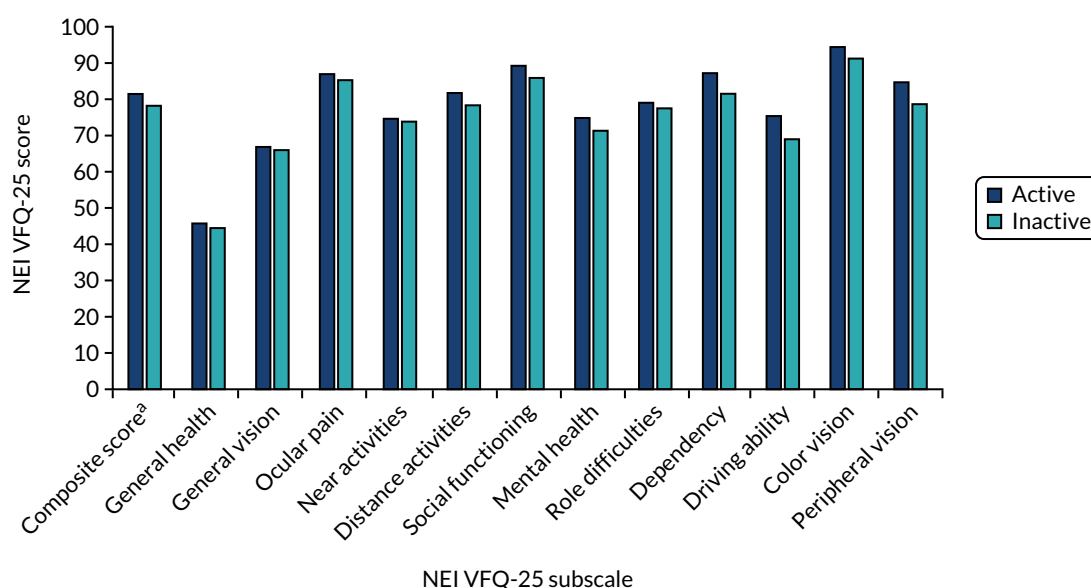


FIGURE 13 Mean NEI VFQ-25 scores among patients eligible for the new ophthalmic grader pathway categorised by DMO status (active vs. inactive). a, Composite score is an average of the vision-targeted sub-scale scores, excluding the general health rating question.

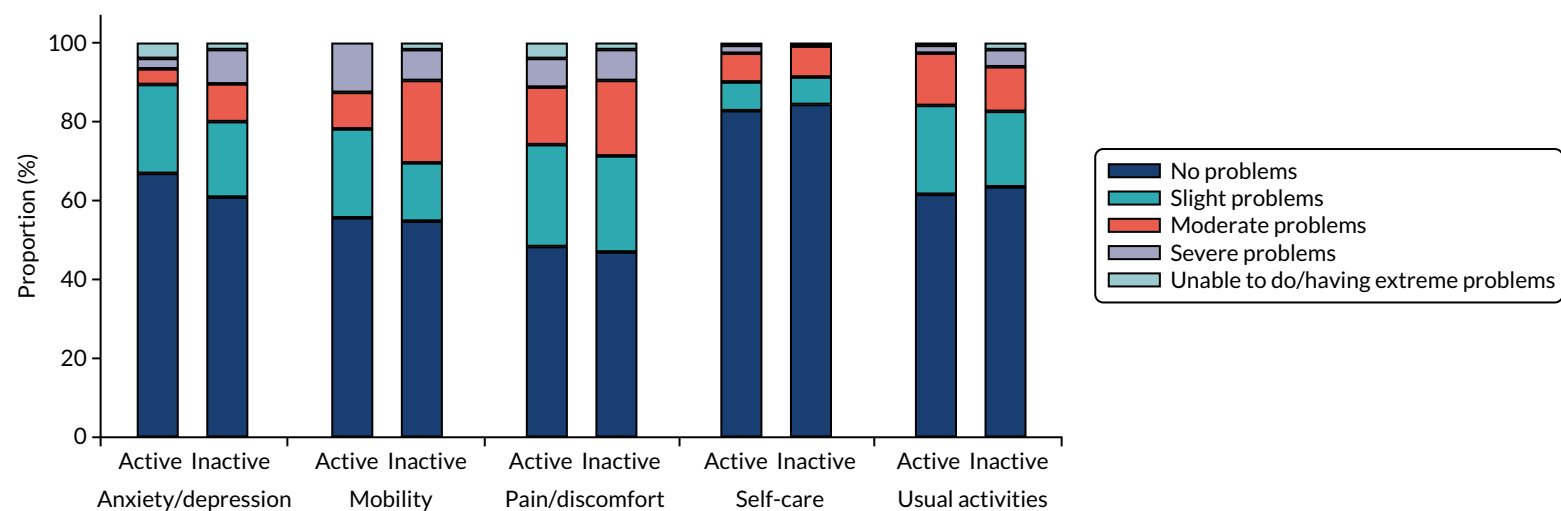


FIGURE 14 Proportion of DMO patients split by EQ-5D-5L dimensions and levels.

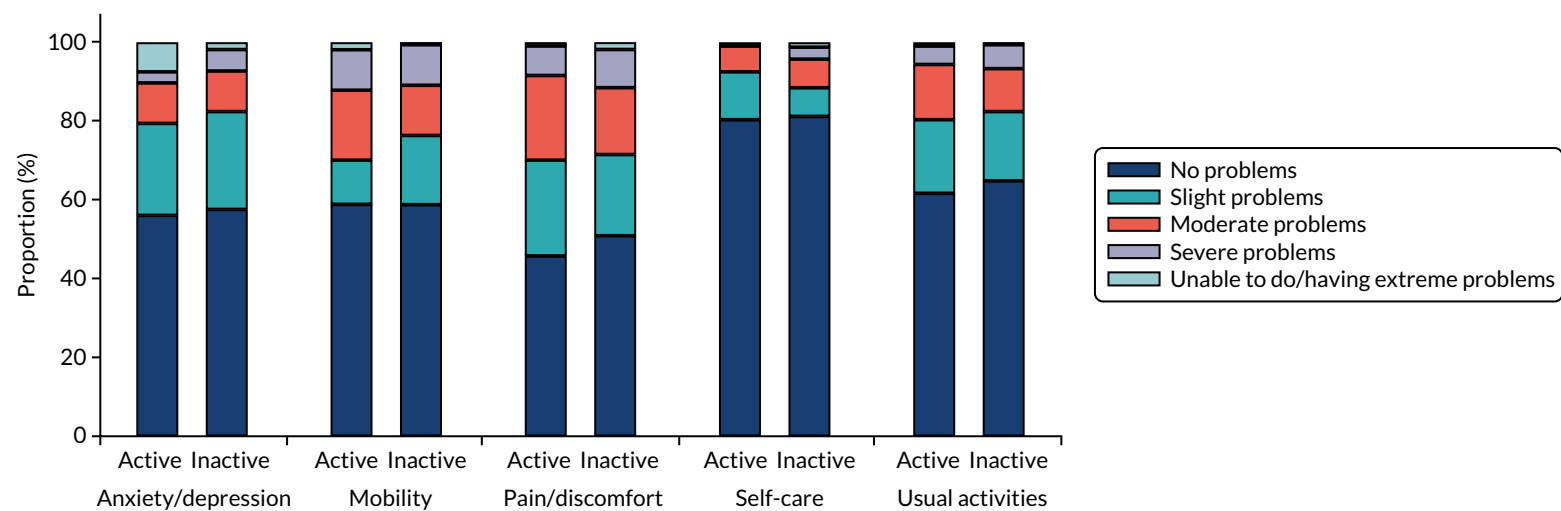


FIGURE 15 Proportion of PDR patients split by EQ-5D-5L dimensions and levels.

Summary/discussion

Diabetic macular oedema

Surveillance of people with DMO by trained ophthalmic graders using SD-OCT was found to have very good sensitivity (97%) but with much poorer specificity, ranging from 31% to 56% depending on the scenario. However, even that level of specificity would still allow useful savings in ophthalmologist time, which could be used for other patients. However, there might be some disutility owing to anxiety and the lack of reassurance gained by seeing an ophthalmologist (see *Chapter 3*). This could be resolved if patients wishing to talk to ophthalmologists could do so by telephone, but there would be a time cost to that.

As noted in *Chapter 2*, the sensitivity and specificity vary according to the comparison. For example, the term 'inclusive referral' meant that graders referred DMO when they consider it to be definitely active, when they were unsure about the activity of the disease and when there were 'ungradable' images. However, in all scenarios in DMO, sensitivity remained over 90%.

Active DMO does not progress rapidly and visual acuity is not expected to drop rapidly either; therefore, significant harm would be unlikely if active DMO were to be missed in 3% of cases at one visit, especially considering that not all eyes with DMO need immediate treatment. The proportion of eyes with DMO requiring treatment is approximately only 58% (85/147) of all active DMO (see *Table 13*).

Proliferative diabetic retinopathy

In PDR, comparison of the grader pathway with the standard care of O-FTF in the main analysis showed a sensitivity that may or may not be considered acceptable in clinical practice. It should be noted, however, that the sensitivity varies according to the grade of PDR, being higher in PDR requiring immediate treatment and reaching nearly 90% if ultra-wide field imaging is used in people with more severe PDR with pre-retinal and/or vitreous haemorrhage.

In PDR, an important finding, albeit from a post hoc analysis, was that the standard care of O-FTF had poorer sensitivity than the ERS. The ERS used both seven-field ETDRS and ultra-wide field images in addition to O-FTF. Obtaining both sets of images would not be feasible in routine care. Further analysis of a post hoc nature would be required to assess the accuracy of O-FTF plus ultra-wide fundus images only, as this is what could be implemented in clinical practice (ultra-wide field fundus images would be the choice when compared with seven-field fundus images considering that the former take less time to take and grade and have higher sensitivity to detect the more severe form of PDR with pre-retinal and vitreous haemorrhage).

One issue is whether or not the grading of images by ophthalmic graders would be seen as taking extra responsibility, meriting an increase in salary band. We have assumed that it would, so estimates presented herein use costs based on graders being on band 7. We have also assumed that the photographers/imaging technicians obtaining the images are on band 6.

Cost would be reduced further if the grader specificity could be improved by feedback and perhaps additional training and by other measures (e.g. patient selection prior to entering the pathway and improved monitors to visualise images).

Our mean time estimates for obtaining ultra-wide field images and seven-field ETDRS photographs were higher than those reported in an earlier study by Silva *et al.*⁴⁹ that, for this purpose, included only 22 patients. That study reported that the mean (\pm SD) time required to acquire non-mydratic ultra-wide field images was 2.8 (\pm 1.3) minutes versus 6.2 (\pm 2.2) minutes for ETDRS photography.⁴⁹ In comparison, our study found that it took 10.4 (\pm 7.2) minutes to take ultra-wide field images and 15.3 (\pm 8.2) minutes to capture seven-field ETDRS photographs. However, the timing by Silva *et al.* appears to start from when the patient was positioned at the camera, whereas in EMERALD the time

was recorded from when the patient entered the room to be imaged; therefore, the times are not comparable. In addition, the study by Silva *et al.* was undertaken in a single centre (rather than being a multicentre study as in EMERALD). Patients included in the Silva *et al.* study had a mean age of 54 years, which is a little lower than EMERALD patients (over half of whom were > 60 years of age). However, both the Silva *et al.* and the EMERALD studies found that taking ultra-wide field images was quicker than obtaining seven-field ETDRS.

Our estimate for the direct cost of the O-FTF is similar to that by Prescott *et al.*,⁵⁰ who estimated a direct cost per slit examination of £27.29 using a similar approach.

The data from the FGs (see *Chapter 3*) showed that patients value being seen by the ophthalmologist. They gain reassurance, have reduction in anxiety and can ask questions. Furthermore, not seeing the ophthalmologist could lead to some disutility, which would persist until the patient was next seen. We do not have data to quantify this. As noted earlier, patients do not understand the professional standing of the photographers/imaging technicians, and have probably never heard of ophthalmic graders. This could perhaps be remedied by the provision of information. In this regard, the EMERALD results demonstrating good sensitivity of graders evaluating DMO images when compared with ophthalmologists (see *Chapter 2, Sensitivity analyses 2, 4, 5 and 6 and Additional analyses 1–3*) should provide reassurance to patients.

Modelling of cost-effectiveness

We had planned that, if there were to be differences among the assessment methods, a Markov model-based cost-utility analysis would be undertaken to examine the cost-effectiveness of the potential surveillance pathways.

The sensitivity of the grader pathway for DMO was very good, and no trade-off between cost savings and vision would be expected if implemented. Specificity was not high, but it is likely that this could be improved (e.g. in the EMERALD study, graders did not have access to any previous clinical information, such as visual acuity, and they did not have access to previous images: both could potentially lead to higher specificity). Poor specificity means that cost savings would be less than hoped for. However, no modelling was required for this part and, thus, it was not conducted.

The overall sensitivity of 80% for the detection of reactivated PDR in the main EMERALD analysis was within the originally accepted level. Indeed, there was the historical precedent of the 80% sensitivity, which was deemed acceptable for DR screening programmes (although it is unclear where this figure came from). Sensitivity was higher for the higher-risk group (high-risk PDR), as it was close to 90% when using ultra-wide field imaging.

Modelling was not possible, however, given that modelling would require knowledge of the following:

- The proportion of the missed reactivated PDR that would progress, and how many patients would have needed treatment if detected.
- How much vision a patient would lose before the next visit, if any.
- What proportion of people would have reactivation detected at the next visit – would some changes be missed again?
- How successful treatment was for those detected at each time point. This could be vision maintained or, for some, improved.
- What would happen if reactivated PDR was missed again at the next visit – what proportion of people would suffer irreversible visual loss and to what extent.
- The quality-of-life consequences for all of the above.
- The costs of treatment of reactivated PDR.
- Differences in costs and quality-adjusted life-years between grader and standard care pathways.

All of the above information would be required for the ophthalmic grader pathway and for the standard face-to-face ophthalmologist pathway (current standard of care). We know from the ERS that the sensitivity of the standard care pathway is not 100%, and that ophthalmologists, for example, may also miss active new vessels (i.e. active PDR) in some patients.

Unfortunately, information regarding most of the above points is not available and this precluded our ability to undertake modelling as initially planned. Furthermore, not everyone with a recurrence of PDR after PRP needs or can have treatment when the recurrence is first noted (e.g. if a vitreous haemorrhage occurs, patients are often managed by observation until the haemorrhage clears). Data on the proportion of patients needing immediate treatment, treatment at a later date or no treatment at all are not available either.

Research is needed into the natural history of treated PDR, including the timelines for recurrence, the forms that recurrence takes, the need for treatment and the effectiveness of treatment.

Chapter 5 General discussion

Statement of principal findings

The EMERALD study evaluated diagnostic accuracy, acceptability to users and cost-effectiveness of a new care pathway, the ophthalmic grader pathway, using multimodal imaging for the surveillance of people with previously treated and stable DMO and/or PDR, when compared with the current standard of care (ophthalmologists evaluating patients in clinic by slit-lamp biomicroscopy, supplemented by SD-OCT in the case of DMO). The study found that for DMO the grader's pathway had excellent and consistent sensitivity of 94% and above based on all pre-planned analyses undertaken, demonstrating that the pathway would be accurate and safe for patients. The specificity for DMO detection was low, ranging from 20% to 56%, depending on the analysis (in the main analysis, 31%). However, even those levels of specificity would still provide a useful release of ophthalmologists' time for other patients. Thus, the pathway could be implemented in the NHS. The EMERALD study found the sensitivity of the ophthalmic grader pathway for PDR, as determined in the main analysis (i.e. grader referring patients to ophthalmologists because of presence of active PDR, if they were unsure or if images were ungradable, which fits with how the pathway would be implemented in practice), to be lower than that for DMO. However, and importantly, the sensitivity of the ophthalmic grader pathway to detect high-risk PDR with pre-retinal and/or vitreous haemorrhage when using ultra-wide field fundus images was higher (87%). The specificity of the new pathway to detect PDR was higher than that for DMO, between 49% and 72% in pre-planned analyses, and was found to save considerable costs and ophthalmologist time. As in the EMERALD study, patients eligible to participate in the ophthalmic grader pathway, if this pathway were to be implemented in clinical practice in the NHS, will have previous treatment with PRP. Hence, the risk of a recurrence and the repercussion of it would not be expected to be as high or as severe as if active disease were to occur in a patient naive to treatment. If a vitreous haemorrhage were to occur, patients would most likely develop floaters and could be instructed to immediately contact hospital eye services to receive timely evaluation and further treatment, if required. In many cases, the course of action of a vitreous haemorrhage is observation until the haemorrhage clears, especially in patients previously treated with PRP, and then fill-in PRP, if required. Few patients previously treated with PRP who develop a vitreous haemorrhage would be expected to require a vitrectomy. TRDs in people previously adequately treated with PRP are infrequent. With this in mind, the ophthalmic grader pathway for PDR may be considered justifiable at the level of sensitivity found in EMERALD, especially in areas of the country where waiting times are preventing patients with active and severe disease from accessing timely care, which has been highlighted to be a real problem currently in the NHS.^{51,52} Finally, and importantly, a secondary planned analysis exploring the performance of the combined ophthalmic grader assessments for both DMO and PDR (rather than of one or other) showed an overall sensitivity of over 96% supporting, further supporting the safety of the new pathway.

Strengths and limitations

EMERALD was carried out in 13 NHS ophthalmology services, with good recruitment and adequate power. Patients were not excluded if they had eye characteristics that would make imaging challenging (e.g. cataracts, poor pupillary dilatation) making the results more applicable to routine care.

The photographers obtaining images were familiar with the standard ETDRS seven-field approach, but were less experienced in taking ultra-wide field images, especially in some centres. Despite this, few images were ungradable.

Images of the iris and anterior chamber angle were not obtained for the evaluation of people with PDR. In very few instances, it would be possible that new vessels could develop in these structures but not be observed in the retina (NVE) or optic disc (NVD). However, this would be unlikely to occur in patients previously treated with laser PRP. If it were to occur, without imaging and evaluation of these structures, it would be missed.

In EMERALD, images were evaluated by ophthalmic graders (and by ophthalmologists undertaking the ERS) without any information about the patient (i.e. masked to any clinical data, including previous images). Although this was a strength in scientific design, it is likely that if clinical information were to be available (e.g. location of previously identified new vessels), including previous images (e.g. images of new vessels following treatment), the sensitivity and specificity of the pathway could be improved. If the pathway were to be implemented in clinical practice, previous clinical information and images would be available to ophthalmic graders.

Implementation of the new care pathway: matters to consider

Before implementation, careful consideration should be given to patients' views, as elicited in our FG discussions. The advice included providing immediate feedback to patients on the status of their disease (i.e. stable or potentially active) using currently available media (e.g. e-mail, telephone, text message) following the reading of the images, and maintaining periodic evaluations by an ophthalmologist, albeit at less frequent intervals. Patients should be informed about the identity, role and function of ophthalmic photographers and graders. Specifically, patients need to know the level of expertise of the grader reviewing their images, and understand that ophthalmologists will still be the ones making decisions about their treatment and care, if an ophthalmic grader pathway were to be implemented. Data obtained in EMERALD, demonstrating that graders perform as well or even better than ophthalmologists at grading images should be reassuring to patients.

There is no clear view on what would be the minimum required sensitivity and specificity that would be acceptable for diagnostic or surveillance pathways. Figures of 80% for sensitivity and 95% for specificity have been quoted in various articles on screening for DR. These seem to have come from a 1997 British Diabetic Association document, based on a consensus conference in 1995, but the document is no longer available (we asked Diabetes UK for access to this document) and so the reasoning behind the figures is not known. As stated above, follow-up of previously treated patients, especially when referring to people with PDR previously treated with PRP, is a rather different scenario to that in DR screening, where naive-to-treatment patients are followed at less frequent intervals. Missing a patient with PDR in a DR screening programme would be expected to have more severe implications than missing an individual with reactivation of PDR in an eye previously treated with PRP. Ten out of the 14 (71%) ophthalmologists in the EMERALD study group, who have extensive expertise in DR and, specifically, PDR, believe that the 80% sensitivity of the ophthalmic grader pathway is considered acceptable in this previously PRP treated population and that, subsequently, the pathway for PDR could, like that of DMO, be implemented.

It may be possible to increase the diagnostic accuracy of the new pathway. To do this, it is essential to understand that the diagnostic accuracy of the new ophthalmic grader pathway goes beyond the ability of the graders to recognise DMO and PDR in fundus images. Thus, it is also dependent on the quality of the images produced, which relies not only on the ophthalmic photographer/imaging technician obtaining them but also on characteristics of the patient and the eye to be imaged. In EMERALD, however, images could be obtained in the great majority of eyes of participants (92–97% of eyes, depending on the imaging technology used), despite the fact that patients were not selected based on whether or not they would be suitable for imaging (i.e. patients with e.g. small pupils, cataract, were not excluded). Furthermore, only a small proportion of images obtained in EMERALD were considered to be ungradable (1–2% of SD-OCTs; 5–7% of seven-field ETDRS images; 4–6% of ultra-wide field images). The diagnostic accuracy of the new

pathway may also be influenced by the platform used to make the images available to the graders, and by the resolution of the screen on which these images were visualised. For logistic purposes (i.e. making all imaging technologies available to each grader in the same location), EMERALD used a commercially available platform (Ophthalsuite, BlueWorks, Coimbra, Portugal) and graders visualised images in available computer screens and, thus, it remains unclear whether the diagnostic performance would have improved if images had been visualised and evaluated using the software available for each of the imaging technologies instead and/or on high resolution screens. As stated above, patients were not excluded if they had eye characteristics that would make imaging challenging (e.g. cataracts, poor pupillary dilatation). This would be expected to affect fundus images (seven-field ETDRS and ultra-wide field images) more so than SD-OCT scans (the number of ungradable images in one or other modality seems to support this), which may have contributed to the lower sensitivity observed for PDR. Currently, patients are imaged routinely with SD-OCT and they are used to this imaging modality; they do not routinely receive ultra-wide field imaging: seven-field ETDRS fundus images are obtained with a standard, conventional fundus camera and, thus, patients would be accustomed to this technology. If ultra-wide field fundus imaging were to be done routinely, it is very likely that improved collaboration could lead to improved image quality. Similarly, many ophthalmic photographers/imaging technicians are not as used to ultra-wide field imaging as they are to standard fundus cameras; further experience on this technology may also lead to better quality ultra-wide field images. Further training of graders and regular feedback following grading would be expected to improve their ability to differentiate between active and inactive disease in previously treated patients, who constitute potentially a more challenging group than those naive to treatment. All of the above may improve further sensitivity and specificity of the new ophthalmic grader pathway.

Potential barriers for implementation of the ophthalmic grader pathway include (1) the identification of individuals that would wish to act as graders (i.e. grader recruitment); (2) the delivery of appropriate training for the graders and the ability of graders to take time to undergo training; (3) the retention of ophthalmic graders once properly trained, once in their role; and (4) efficient systems to report results to patients, and to do so on the same day as images are read. However, the identification and training of graders, as explained in *Chapter 2, Methods*, was possible in EMERALD, suggesting that it is doable. If remunerated appropriately, retention may also be achievable.

Cost-consequences analysis

The cost-consequences analysis identified scenarios where savings could be made. In practice, however, it is unlikely that any savings would be realised. The benefits would be the release of ophthalmologist time to see or treat other patients. The benefits could also include that, for example, patients with treated and stable PDR could be seen earlier (say every 3–4 months) if followed up by ophthalmic graders than if evaluated by ophthalmologists (currently aimed to be done every 6–8 months) although this would have implications on costs of the new ophthalmic grader pathway. We are aware that waiting times for patients to see an ophthalmologist are very long in some places, well beyond 12 months. In such circumstances, being followed by an ophthalmic grader at shorter intervals may be much better, even at the sensitivity observed in EMERALD.

The potential for future automated grading of images

With the increased popularity of artificial intelligence (AI) in the field of ophthalmology and, specifically, in DR, a question arises about whether or not AI could be used instead of graders to determine the presence of active DMO and/or PDR on fundus images and SD-OCT scans. This concept is not new and many groups have worked in this area. Sharp *et al.* proposed and undertook work, funded by the Health Technology Assessment (HTA) programme of the National Institute for Health Research (NIHR), using digital imaging and automated reading for the purpose of grading DR for DR screening programmes.⁵³ Further work conducted by this group, also supported by the NIHR HTA programme, demonstrated

excellent performance of automated detection of DMO and the value of adding OCT for this purpose.⁵⁴ More recent studies undertaken by Abràmoff *et al.*^{55,56} have now demonstrated excellent sensitivity and specificity of AI methods to determine presence of referable DR [defined as presence of moderate and higher stages of non-proliferative DR (NPDR), PDR or DMO] in fundus images when compared with retinal specialists. This AI system (IDx-DR; IDx Technologies Inc., Coralville, IA, USA), with 87% sensitivity, 90% specificity and 96% imageability, has been approved recently by the Food and Drug Administration (FDA) for the automated diagnosis of DR. The studies on which this programme was developed, however, included mostly patients with DMO and/or PDR who had not been previously treated (i.e. treatment naive). Thus, although the programme has been shown to perform well to detect referable DR in an untreated population, it remains to be elucidated whether high diagnostic accuracy could be achieved through this or other AI programmes in the more complex group of previously treated patients.

Imaging modalities

EMERALD compared two different imaging technologies for the evaluation of PDR, seven-field ETDRS and ultra-wide field fundus imaging (for details on each, see *Chapter 1*). For the latter, the Optos system (Optos Inc., Dunfermline, Fife, UK) was chosen as it is the only ultra-wide field imaging technology that allows a 200-degree field of view mode. Other wide-field imaging technologies do not reach the far peripheral retina. The diagnostic accuracy of the new ophthalmic grader pathway to detect PDR was very similar in both sets of images (seven-field ETDRS and ultra-wide field images) under the main analysis (graders referring patients to ophthalmologists if they identified active disease, if they were unsure, or if images were ungradable compared with the standard of care of ophthalmologists evaluating patients on slit-lamp biomicroscopy). Seven-field ETDRS images led to slightly higher sensitivity than ultra-wide field images (71%, 95% CI 61% to 79% for seven-field ETDRS, vs. 63%, 95% CI 53% to 71% for ultra-wide field images) when referral specifically for active PDR was considered. This, however, would not be the way in which the pathway would be running in clinical practice as graders would also be referring patients when uncertain about the activity of the disease (as set in the main analysis). Ultra-wide field imaging tended to have slightly higher specificity. Importantly, there was higher sensitivity and specificity to detect people with more severe PDR (PDR with pre-retinal and vitreous haemorrhage) when ultra-wide field imaging was used. The costs of both taking and reading ultra-wide field images were less than the respective costs related to seven-field ETDRS photographs.

In EMERALD, images were all obtained following pupillary dilatation. It remains to be elucidated whether comparable sensitivity and specificity of the ophthalmic grader pathway could be achieved through multimodal imaging in the undilated eye. This could increase patient satisfaction given that patients would be able to drive safely following the examination as their pupils would not be dilated, and would reduce their time in clinic further, as they would not need to wait for their pupils to dilate prior to imaging. It would also lead to some savings (dilating drops will not be required).

In EMERALD, three-field ultra-wide field imaging was obtained, with fields of the far peripheral superior and inferior retina, in addition to the central field, which already encompasses macula, midperipheral and most of the peripheral retina. It is uncertain whether using these three fields is required; if results were similar using only the central field image, this could reduce time to obtain images and to grade them and would subsequently reduce costs.

One other study compared ultra-wide field images, obtained with the same system used in EMERALD (Optos, Dunfermline, Scotland) with seven-field ETDRS for the detection of new vessels in PDR.⁵⁷ This study evaluated how often new retinal vessels were detected outside areas covered by the seven fields of the ETDRS imaging in a consecutive series of 1562 treatment-naïve eyes from the NHS DR screening service (DRSS) referred to secondary care. After evaluation in secondary care, new vessels were found in 102 of the 1562 eyes (6.5%) of 781 patients. In 72 eyes (71%), new vessels had been detected in the DRSS; in 30 eyes (29%), they were found after examination in secondary care.

In 12 of these 30 eyes, new vessels were detected by wide-angle imaging in areas outside those covered by seven-field ETDRS. The population included in this study (treatment naive individuals), however, was different from the group included in EMERALD (people with PDR previously treated with laser PRP).

Virtual clinics

What have been widely called 'virtual clinics' (evaluation of patients by looking at their images rather than through a face-to-face consultation in clinic with an ophthalmologist) are not new to the NHS. Published studies presented the experience of several groups throughout the UK using this form of evaluation for people with age-related macular degeneration⁵⁸ and other medical retinal diseases,⁵⁹ including DR^{60–62} and glaucoma.⁶³ These studies showed that implementation of virtual clinics was feasible and reduced patients' time in clinic, improving patients' journey, and seemed to increase the efficiency of the service. Most studies evaluating virtual clinics for retinal diseases were based on the assessment of images by ophthalmologists rather than allied health-care staff (e.g. graders),^{58,59,63} and included newly referred patients rather than previously treated ones, and had a selected population, rather than evaluating all comers. In one study evaluating outcomes (proportion of people discharged, referred for further assessment in a face to face clinic undertaken by an ophthalmologist, or continuing care in the virtual clinic) of new referrals to virtual medical retinal clinics, images were reviewed by ophthalmologists (consultants and medical retina fellows), an optometrist and a grader. Diagnostic accuracy of these different members of staff, however, was not investigated. To our knowledge, there have been no studies conducted to date, other than EMERALD, specifically addressing the diagnostic accuracy of a new pathway involving staff other than ophthalmologists for the evaluation of people with previously treated DMO and/or PDR using multimodal imaging.

Despite the fact that some reports on virtual clinics are available, very scarce data exist with regard to the acceptability of virtual clinics by patients⁶¹ and health-care professionals.⁶⁴ Questionnaires were used in these studies to assess acceptability, with the inherent limitations of this method as discussed in *Chapter 3*, including the fact that through questionnaires users are responding to questions previously elaborated, which may or may not address matters that are of importance to them. Furthermore, low ascertainment rates (46–61%) limited the generalisability of the results presented in these studies. One study referred to the use of virtual clinics for the evaluation of patients with DR.⁶¹ A survey was sent to 813 patients attending hospital eye services for review of their DR in Singleton Hospital, Swansea; 498 questionnaires (61%) were returned, fully or only partly filled. Patients were asked to grade their experience with regard to delays in their appointments and waiting times while in clinic, and whether or not they would be willing to attend virtual clinics if appointments were given timely and their time in clinic reduced. A high proportion of patients (86%) responded that they would be happy to be seen in virtual clinics under these circumstances. When asked how happy they would be to attend these virtual clinics (0 being not happy at all and 10 being very happy), the average response was 8.79. A small group of 14% of patients stated they would not be happy to attend virtual clinics at all. It is unclear what information was given to patients with regard to what a virtual clinic involved.

The FG work showed how much patients value seeing the ophthalmologist, and if an option was chosen that did not include the O-FTF, it is likely that patients would incur some disutility due to anxiety. Visual loss is the most feared complication of diabetes. It is not possible to quantify this disutility with our current FG data. The grader FG work also suggested that if photographs showed no reactivation, graders would be happy to reassure patients; however, they would be less happy commenting on changes suggestive of reactivation and answering other potential questions from patients that may follow. These matters need to be considered if the new ophthalmic grader pathway is implemented.

Conclusion

EMERALD supports the implementation of the ophthalmic grader pathway for the evaluation of people with DR with previously treated and stable DMO, especially those with early NPDR (i.e. those with mild or moderate NPDR) who would be unlikely to progress to PDR in a short period of time (*Figure 16*).

The diagnostic accuracy of the ophthalmic grader pathway for PDR was not as good as that for DMO. However, considering that patients going into this pathway would previously have received treatment with PRP, this pathway may be considered safe to be implemented. A pilot phase implementation in selected centres (e.g. those previously participating in EMERALD) prior to its widespread implementation in the NHS could be undertaken. Measures to improve the sensitivity of this pathway (e.g. providing continuous feed-back to graders, selecting patients entering the pathway, enhancing resolution of screens used to view images, etc.) could be introduced and evaluated.

Recommendations for future research

Further research evaluating the performance of AI for the automated detection of DMO and active PDR in previously treated patients would be of great value, using current available commercialised AI methods or through the development of new ones.

It would be useful to determine, either through evidence synthesis or, if insufficient data are available, primary research, the risk of visual loss related to recurrence of disease in people with previously treated PDR, and hence to understand the real risk for patients if the identification of their disease were to be missed until the next evaluation. Although the latter may not be done prospectively, due to ethical considerations, it may be possible to obtain data retrospectively that may inform about this risk (e.g. evaluating outcomes of people with previously treated PDR that develop reactivation of their disease).

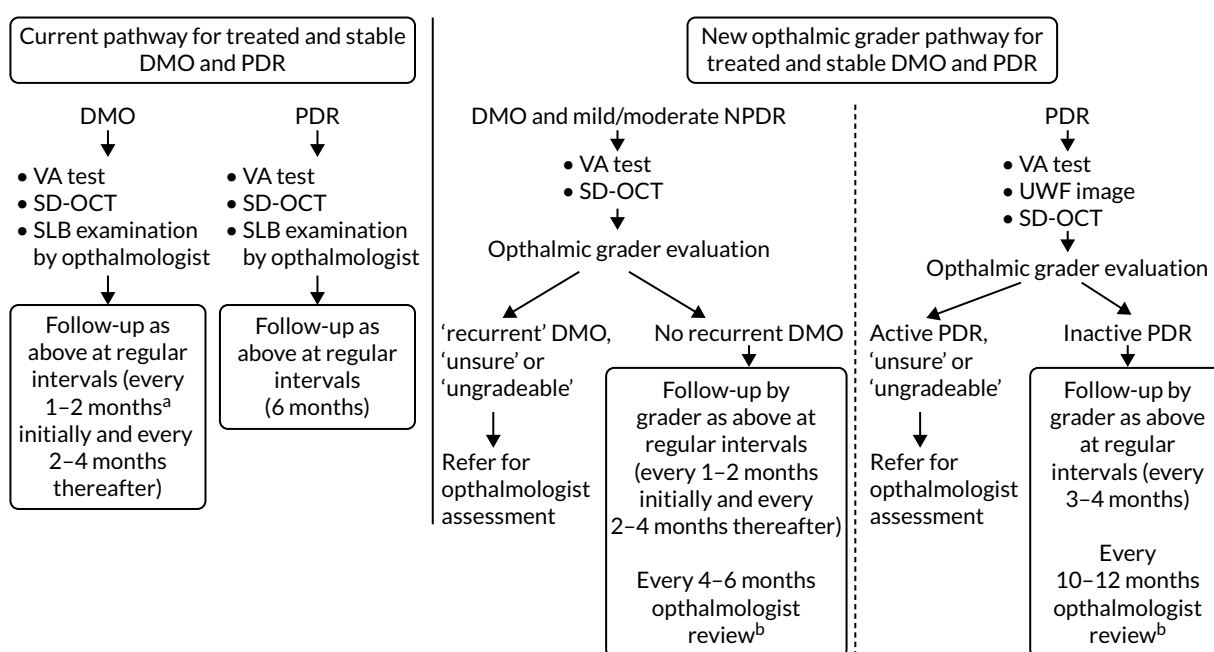


FIGURE 16 Current standard of care and new proposed pathway for people with previously treated and stable DMO and PDR. VA, visual acuity; SD-OCT, spectral domain optical coherence tomography; SLB, slit-lamp biomicroscopy; F/U, follow-up; UWF, ultra-wide field. a, With current anti-VEGFs used reviews are often carried out monthly initially; newer anti-VEGFs may only require reviews every 2 months. b, Telephone calls to patients to discuss findings by ophthalmologist or other members of staff could be arranged in addition, for selected patients, if required.

Studies evaluating means of further improving the sensitivity of the ophthalmic grader pathway for the detection of PDR and the specificity of the ophthalmic grader pathway for the detection of DMO and PDR would also be valuable.

As reported in *Chapter 2*, some reactivated PDR was identified by the ophthalmologists using slit-lamp biomicroscopy but not detected by the reading of the images, and vice versa. Further work exploring the reasons for these differences would be helpful.

If one central ultra-wide field image rather than three, as undertaken in EMERALD, would provide comparable sensitivity and specificity, this would reduce the time taken to obtain and grade images and, subsequently, costs. A study exploring this possibility would be helpful. In addition, a study determining the diagnostic accuracy of the ophthalmic grader's pathway using images obtained in undilated eyes, rather than following pupillary dilatation, as done in EMERALD, would be of interest.

Acknowledgements

The authors would like to thank all patients participating in EMERALD, members of the EMERALD PPI group, especially Tom Rush, Martin Adams, Robert Legge and Anne Ramsay, and Diabetes UK, Northern Ireland, for their support of this diagnostic study. The EMERALD PPI group provided input to EMERALD throughout all its stages, including confirming the importance of the research questions, providing support and input to the elaboration of patient-related materials, such as patient information leaflets and consent forms, as well as elaboration of layman sections of EMERALD-related documents. The PPI group will continue to play a role in the dissemination of EMERALD results.

The authors would like to thank co-investigators, ophthalmic photographers/imaging technicians and graders, and research nurses and research co-ordinators at all participating sites:

- Belfast Health and Social Care Trust, Belfast – Graham Young and Vittorio Silvestri (graders/photographers) and Rebecca Denham and Nuala Jane Lavery (research nurses).
- Bradford Royal Infirmary Bradford Teaching Hospitals NHS Trust, Bradford – Zeid Madanat (co-investigator), Sarah Moss (research nurse) and Nicola Hawes (research nurse).
- Bristol Eye Hospital, University Hospitals Bristol NHS Foundation Trust, Bristol – Danielle Lee (co-ordinator).
- Frimley Park Hospital NHS Foundation Trust, Frimley – Manju Chandran, Sely Mathews and Mohammed Galal (co-investigators), and Teena Kunnath (research nurse).
- Gloucestershire Royal Hospital, Gloucestershire Hospitals NHS Trust – Laura Lodge (co-ordinator), Scott Vallance and Angela Dale (graders/photographers), Jenny Mackenzie and Jade Hearn (admin/recruiter), Gwen George and Paul Dimmock (ophthalmic imagers).
- Kings College Hospital NHS Foundation Trust, London – Noimot Timson (research nurse).
- Manchester Royal Eye Hospital, Central Manchester University Hospitals NHS Foundation, Manchester – Zaria Ali (co-investigator), Goncalo Bento and Natalie Fox (graders/photographers), Charlene Hyde (co-ordinator) and Nattapon Boonarpa (grader/photographer).
- Moorfields Eye Hospital NHS Foundation Trust, Bedford – Piyali Sen (co-investigator), Roxanne Crosby Nwaobi (grader/photographer) and Aljazy Jaber (co-ordinator).
- John Radcliffe Hospital, Oxford University Hospitals NHS Foundation Trust, Oxford – Janette Savage (research nurse).
- Queen's Margaret Hospital, Fife – Susan Fowler and Amanda McGregor (research nurses), and Eric Robinson and Kim Allan (photographers).
- Sheffield Teaching Hospitals NHS Foundation Trust – Stephen Connell and Thomas Evans (graders/photographers), and Helen Pokora (co-ordinator).
- Sunderland Eye Infirmary, City Hospitals Sunderland NHS Foundation Trust – Clair Barbour and Leontia Bell (co-investigators), Steve Dodds and Lauren Gardner (research nurses), and Jill O'Brien, Nicolle Teasdale, Jade Ward and Michelle Young (graders/photographers).
- James Cook University Hospital, South Tees Hospitals NHS Foundation Trust, Middlesbrough – Daniela Vaideanu-Collins (co-investigator), Catherine Williamson and Adele Craggs (co-ordinators), and Kay Henderson, Donna Wilson and Michaela Dickinson (graders/photographers).

The authors would like to thank Optos for providing some of the Optos cameras required for the study, and Anne-Marie Cairns, Kereen Johnston and Robert Moxham for their support. Optos PLC had no role in the design, analysis or reporting of EMERALD. Optos had no role in the study design; collection, management, analysis and interpretation of data, or the writing of this report.

The authors would like to thank Mike Clarke, Lynn Murphy, Mark Wilson, Catherine Campbell and Nuala Hannaway from the NICTU, and Janice Bailie and Alison Murphy for their support, help and assistance with this diagnostic study. The authors would also like to thank the UK Ophthalmology Clinical Research Network, and especially the Northern Ireland Ophthalmology Clinical Research

Network, including Maurice O’Kane, Julie Silvestri, Jonathan Jackson, Paul Biagioni and Louise Scullion, for their support of EMERALD.

The authors thank Dr Pamela Royle for literature searches and Jill Colquitt and Emma Loveman, Effective Evidence (Waterlooville, UK) for their help formatting this monograph.

The authors would like to give special thanks to Professor John Norrie, Dr David Owens, Mrs Florence Findlay-White, Dr Winfried Amoaku and Dr Yemisi Takwoingi, members of the TSC, for their expert advice and support of EMERALD, as well as to Professor Luke Vale, Newcastle University, for commenting on the draft of the health economics analysis plan.

The EMERALD study group

Ahmed Saad, James Cook University Hospital, South Tees Hospitals NHS Foundation Trust; Augusto Azuara-Blanco, Centre for Public Health, Queen’s University, and The Belfast Health and Social Care Trust, Belfast; Caroline Styles, Queen’s Margaret Hospital, Fife; Christine McNally, Andrew Jackson and Rachael Rice, NICTU; Clare Bailey, Bristol Eye Hospital, University Hospitals Bristol NHS Foundation Trust; Danny McAuley, Queen’s University and Royal Victoria Hospital, Belfast Health and Social Care Trust; David H Steel, Sunderland Eye Infirmary, City Hospitals Sunderland NHS Foundation Trust; Faruque D Ghanchi, Bradford Teaching Hospitals NHS Trust; Geeta Menon, Frimley Park Hospital NHS Foundation Trust; Haralabos Eleftheriadis and Stefanos Efraimidis, Kings College Hospital NHS Foundation Trust; Jonathan Cook, Ariel Wang and William Sones, Centre for Statistics in Medicine, University of Oxford; Lindsay Prior, Centre for Public Health, Queens University, Belfast; Nachiketa Acharya, Sheffield Teaching Hospitals NHS Foundation Trust; Noemi Lois, The Wellcome-Wolfson Institute for Experimental Medicine, and the Belfast Health and Social Care Trust, Belfast; Norman Waugh, Hema Mistry and Mandy Maredza, Warwick University; Samia Fatum, John Radcliffe Hospital, Oxford University Hospitals NHS Foundation Trust; Sobha Sivaprasad, Moorfields Eye Hospital NHS Foundation Trust; Stephen Aldington, Peter H Scanlon and Katerina Ivanova, Gloucestershire Hospitals NHS Foundation Trust; Tariq M Aslam, Manchester Royal Eye Hospital, Central Manchester University Hospitals NHS Foundation Trust; and Victor Chong, Royal Free Hospital NHS Foundation Trust, London.

Clinical sites participating in recruitment

The Belfast Health and Social Care Trust, Belfast, Northern Ireland; Bradford Royal Infirmary, Bradford Teaching Hospitals NHS Foundation Trust, Bradford; Bristol Eye Hospital, University Hospitals Bristol NHS Foundation Trust, Bristol; Frimley Park Hospital NHS Foundation Trust, Frimley; Gloucestershire Royal Hospital, Gloucestershire Hospitals NHS Foundation Trust, Gloucester; James Cook University Hospital, South Tees Hospitals NHS Foundation Trust, Middlesbrough; Kings College Hospital NHS Foundation Trust, London; Manchester Royal Eye Hospital, Central Manchester University Hospitals NHS Foundation Trust, Manchester; Moorfields Eye Hospital NHS Foundation Trust, London; John Radcliffe Hospital, Oxford University Hospitals NHS Foundation Trust, Oxford; Queen Margaret Hospital, Dunfermline; Sheffield Teaching Hospitals NHS Foundation Trust, Sheffield; and Sunderland Eye Infirmary, City Hospitals Sunderland NHS Foundation Trust, Sunderland.

Trial Management Group

Professor Noemi Lois (chief investigator), Professor Augusto Azuara-Blanco, Mr Steve Aldington, Dr Danny McAuley, Dr Peter Scanlon, Dr Lindsay Prior, Mrs Clare Newall, Mrs Michelle McGaughey, Mrs Christine McNally, Miss Rachael Rice, Mr Andrew Jackson, Professor Jonathan Cook, Mr William Sones, Professor Norman Waugh, Dr Hema Mistry, Mr Mark Wilson, Mrs Nuala Hannaway and Mrs Catherine Campbell.

Trial Steering Committee

Professor John Norrie (chairperson), Dr David Owens, Mrs Florence Findlay-White, Dr Winfried Amoaku and Dr Yemisi Takwoingi.

Contributions of others

William Sones (Medical Statistician) determined the statistical analysis plan.

Christine McNally (EMERALD Trial Manager), Rachael Rice (Trial Co-ordinator) and Andrew Jackson (Trial Data Manager) provided management co-ordination and data management to the study.

Ahmed Saad, Geeta Menon, Haralabos Eleftheriadis, Nachiketa Acharya and Samia Fatum (Consultant Ophthalmologists) recruited and evaluated patients for the study.

Contributions of authors

Noemi Lois (<https://orcid.org/0000-0003-2666-2937>) (University Professor of Ophthalmology and Honorary Consultant Ophthalmic Surgeon) conceived the study, participated in the study design, determined the sample size for the study, determined the statistical analysis plan, recruited and evaluated patients for the study, drafted this manuscript and revised it to its final format.

Jonathan Cook (<https://orcid.org/0000-0002-4156-6989>) (Associated Professor) participated in the study design, determined the sample size for the study, determined the statistical analysis plan and undertook the statistical analysis of the diagnostic accuracy data.

Ariel Wang (<https://orcid.org/0000-0001-5360-9993>) (Medical Statistician) undertook the statistical analysis of the diagnostic accuracy data.

Stephen Aldington (<https://orcid.org/0000-0003-1808-3580>) (Retinopathy Research and Professional Development Manager) participated in the study design and recruited and evaluated patients for the study.

Hema Mistry (<https://orcid.org/0000-0002-5023-1160>) (University Associate Professor in Health Economics) participated in the study design, planned the health economic evaluation and undertook the health economic analysis.

Mandy Maredza (<https://orcid.org/0000-0002-2030-3338>) (Research Fellow in Health Economics) undertook the health economic analysis.

Danny McAuley (<https://orcid.org/0000-0002-3283-1947>) (University Professor, Former Director of the NICTU) participated in the study design.

Tariq Aslam (<https://orcid.org/0000-0002-9739-7280>) (Professor of Ophthalmology and Interface Technologies) participated in the study design and recruited and evaluated patients for the study.

Clare Bailey (<https://orcid.org/0000-0003-2141-4512>) (Consultant Ophthalmic Surgeon) participated in the study design and recruited and evaluated patients for the study.

Victor Chong (<https://orcid.org/0000-0002-7693-522X>) (Honorary Consultant Ophthalmic Surgeon) participated in the study design.

Faruque Ghanchi (<https://orcid.org/0000-0002-4448-8162>) (Consultant Ophthalmic Surgeon and Honorary Professor of Ophthalmology) participated in the study design and recruited and evaluated patients for the study.

Peter Scanlon (<https://orcid.org/0000-0001-8513-710X>) (Consultant Ophthalmologist and Associate and Visitor Professor) participated in the study design and recruited and evaluated patients for the study.

Sobha Sivaprasad (<https://orcid.org/0000-0001-8952-0659>) (Consultant Ophthalmic Surgeon and Honorary Professor of Ophthalmology) participated in the study design and recruited and evaluated patients for the study.

David Steel (<https://orcid.org/0000-0001-8734-3089>) (Consultant Ophthalmic Surgeon and Honorary Professor of Ophthalmology) participated in the study design and recruited and evaluated patients for the study.

Caroline Styles (<https://orcid.org/0000-0002-6515-1032>) (Consultant Ophthalmic Surgeon) participated in the study design.

Augusto Azuara-Blanco (<https://orcid.org/0000-0002-4805-9322>) (University Professor of Ophthalmology and Honorary Consultant Ophthalmic Surgeon) participated in the study design.

Lindsay Prior (<https://orcid.org/0000-0001-9903-5841>) (Emeritus University Professor) participated in the study design and planned and carried out the focus group discussions.

Norman Waugh (<https://orcid.org/0000-0003-0934-4961>) (Professor of HTA) participated in the study design, planned the health economic evaluation and undertook the health economic analysis.

All authors evaluated and interpreted results, reviewed this manuscript, provided input to it and approved the final submitted version.

Publications

Lois N, Cook J, Aldington S, Waugh N, Mistry H, Sones W, *et al.* Effectiveness of Multimodal imaging for the Evaluation of Retinal oedema And new vessels in Diabetic retinopathy (EMERALD). *BMJ Open* 2019;**9**:e027795.

Lois N, Cook JA, Wang A, Aldington S, Mistry H, Maredza M, *et al.* Evaluation of a new model of care for people with complications of diabetic retinopathy: the EMERALD study. *Ophthalmology* 2021;**128**:561–73.

Data-sharing statement

All data requests should be submitted to the corresponding author for consideration. Access to available anonymised data may be granted following review.

Patient data

This work uses data provided by patients and collected by the NHS as part of their care and support. Using patient data is vital to improve health and care for everyone. There is huge potential to make better use of information from people's patient records, to understand more about disease,

develop new treatments, monitor safety, and plan NHS services. Patient data should be kept safe and secure, to protect everyone's privacy, and it's important that there are safeguards to make sure that it is stored and used responsibly. Everyone should be able to find out about how patient data are used. #datasaveslives You can find out more about the background to this citation here: <https://understandingpatientdata.org.uk/data-citation>.

References

- Lois N, Cook J, Aldington S, Waugh N, Mistry H, Sones W, *et al.* Effectiveness of Multimodal imaging for the Evaluation of Retinal oedema And new vessels in Diabetic retinopathy (EMERALD). *BMJ Open* 2019;**9**:e027795. <https://doi.org/10.1136/bmjopen-2018-027795>
- Lois N, Cook JA, Wang A, Aldington S, Mistry H, Maredza W, *et al.* Evaluation of a new model of care for people with complications of diabetic retinopathy: the EMERALD study. *Ophthalmology* 2021;**128**:561–73. <https://doi.org/10.1016/j.ophtha.2020.10.030>
- Liew G, Michaelides M, Bunce C. A comparison of the causes of blindness certifications in England and Wales in working age adults (16–64 years), 1999–2000 with 2009–2010. *BMJ Open* 2014;**4**:e004015. <https://doi.org/10.1136/bmjopen-2013-004015>
- Minassian DC, Owens DR, Reidy A. Prevalence of diabetic macular oedema and related health and social care resource use in England. *Br J Ophthalmol* 2012;**96**:345–9. <https://doi.org/10.1136/bjo.2011.204040>
- Yau JW, Rogers SL, Kawasaki R, Lamoureux EL, Kowalski JW, Bek T, *et al.* Global prevalence and major risk factors of diabetic retinopathy. *Diabetes Care* 2012;**35**:556–64. <https://doi.org/10.2337/dc11-1909>
- Diabetes UK. *Diabetes UK*. 2020. URL: www.diabetes.org.uk/ (accessed 27 February 2020).
- National Institute for Health and Care Excellence (NICE). *Ranibizumab for Treating Diabetic Macular Oedema*. London: NICE; 2013.
- National Institute for Health and Care Excellence (NICE). *Aflibercept for Treating Diabetic Macular Oedema*. London: NICE; 2015.
- National Institute for Health and Care Excellence (NICE). *Fluocinolone Acetonide Intravitreal Implant for Treating Chronic Diabetic Macular Oedema after an Inadequate Response to Prior Therapy*. London: NICE; 2013.
- The Royal College of Ophthalmologists. *Diabetic Retinopathy Guidelines*. London: The Royal College of Ophthalmologists; 2012.
- Elman MJ, Aiello LP, Beck RW, Bressler NM, Bressler SB, Edwards AR, *et al.* Randomized trial evaluating ranibizumab plus prompt or deferred laser or triamcinolone plus prompt laser for diabetic macular edema. *Ophthalmology* 2010;**117**:1064–77.e35. <https://doi.org/10.1016/j.ophtha.2010.02.031>
- 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 edema. *Ophthalmology* 2011;**118**:615–25. <https://doi.org/10.1016/j.ophtha.2011.01.031>
- Nguyen QD, Brown DM, Marcus DM, Boyer DS, Patel S, Feiner L, *et al.* Ranibizumab for diabetic macular edema: results from 2 phase III randomized trials: RISE and RIDE. *Ophthalmology* 2012;**119**:789–801. <https://doi.org/10.1016/j.ophtha.2011.12.039>
- Korobelnik JF, Do DV, Schmidt-Erfurth U, Boyer DS, Holz FG, Heier JS, *et al.* Intravitreal aflibercept for diabetic macular edema. *Ophthalmology* 2014;**121**:2247–54. <https://doi.org/10.1016/j.ophtha.2014.05.006>
- Diabetic Retinopathy Clinical Research Network, Wells JA, Glassman AR, Ayala AR, Jampol LM, Aiello LP, *et al.* Aflibercept, bevacizumab, or ranibizumab for diabetic macular edema. *N Eng J Med* 2015;**372**:1193–203. <https://doi.org/10.1056/NEJMoa1414264>

16. Virgili G, Menchini F, Casazza G, Hogg R, Das RR, Wang X, Michelessi M. Optical coherence tomography (OCT) for detection of macular oedema in patients with diabetic retinopathy. *Cochrane Database Syst Rev* 2015;1:CD008081. <https://doi.org/10.1002/14651858.CD008081.pub3>
17. Early Treatment Diabetic Retinopathy Study Research Group. Early photocoagulation for diabetic retinopathy. ETDRS report number 9. *Ophthalmology* 1991;98:766–85. [https://doi.org/10.1016/S0161-6420\(13\)38011-7](https://doi.org/10.1016/S0161-6420(13)38011-7)
18. Royle P, Mistry H, Auguste P, Shyangdan D, Freeman K, Lois N, Waugh N. Pan-retinal photocoagulation and other forms of laser treatment and drug therapies for non-proliferative diabetic retinopathy: systematic review and economic evaluation. *Health Technol Assess* 2015;19(51). <https://doi.org/10.3310/hta19510>
19. Moutray T, Evans JR, Lois N, Armstrong DJ, Peto T, Azuara-Blanco A. Different lasers and techniques for proliferative diabetic retinopathy. *Cochrane Database Syst Rev* 2018;3:CD012314. <https://doi.org/10.1002/14651858.CD012314.pub2>
20. Bressler SB, Beaulieu WT, Glassman AR, Gross JG, Jampol LM, Melia M, et al. Factors associated with worsening proliferative diabetic retinopathy in eyes treated with panretinal photocoagulation or ranibizumab. *Ophthalmology* 2017;124:431–9. <https://doi.org/10.1016/j.ophtha.2016.12.005>
21. Keenan TD, Johnston RL, Donachie PH, Sparrow JM, Stratton IM, Scanlon P. United Kingdom National Ophthalmology Database Study: diabetic retinopathy; report 1: prevalence of centre-involving diabetic macular oedema and other grades of maculopathy and retinopathy in hospital eye services. *Eye* 2013;27:1397–404. <https://doi.org/10.1038/eye.2013.196>
22. Silva PS, Cavallerano JD, Haddad NM, Kwak H, Dyer KH, Omar AF, et al. Peripheral lesions identified on ultrawide field imaging predict increased risk of diabetic retinopathy progression over 4 years. *Ophthalmology* 2015;122:949–56. <https://doi.org/10.1016/j.ophtha.2015.01.008>
23. Royal College of Ophthalmologists. *Eye Health Experts Call for Urgent Action on Hospital Eyecare to Reduce Delays and Improve Patient Safety*. London: Royal College of Ophthalmologists; 2020. URL: www.rcophth.ac.uk/2020/02/eye-health-experts-call-for-urgent-action-on-hospital-eyecare-to-reduce-delays-and-improve-patient-safety/ (accessed 18 January 2020).
24. Knottnerus JA, Muris JW. Assessment of the accuracy of diagnostic tests: the cross-sectional study. *J Clin Epidemiol* 2003;56:1118–28. [https://doi.org/10.1016/S0895-4356\(03\)00206-3](https://doi.org/10.1016/S0895-4356(03)00206-3)
25. Whiting PF, Rutjes AW, Westwood ME, Mallett S, Deeks JJ, Reitsma JB, et al. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Ann Intern Med* 2011;155:529–36. <https://doi.org/10.7326/0003-4819-155-8-201110180-00009>
26. Dixon JR Jr. The International Conference on Harmonization Good Clinical Practice guideline. *Qual Assur* 1998;6:65–74. <https://doi.org/10.1080/105294199277860>
27. Obuchowski NA. Sample size calculations in studies of test accuracy. *Stat Methods Med Res* 1998;7:371–92. <https://doi.org/10.1177/096228029800700405>
28. Silva PS, Cavallerano JD, Tolson AM, Rodriguez J, Rodriguez S, Ajlan R, et al. Real-time ultrawide field image evaluation of retinopathy in a diabetes telemedicine program. *Diabetes Care* 2015;38:1643–9. <https://doi.org/10.2337/dc15-0161>
29. Piegorsch WW. Sample sizes for improved binomial confidence intervals. *Comput Stat Data Anal* 2004;46:309–16. <https://doi.org/10.1016/j.csda.2003.10.002>
30. Newcombe RG. Improved confidence intervals for the difference between binomial proportions based on paired data. *Stat Med* 1998;17:2635–50. [https://doi.org/10.1002/\(SICI\)1097-0258\(19981130\)17:22<2635::AID-SIM954>3.0.CO;2-C](https://doi.org/10.1002/(SICI)1097-0258(19981130)17:22<2635::AID-SIM954>3.0.CO;2-C)

31. McNemar Q. Note on the sampling error of the difference between correlated proportions or percentages. *Psychometrika* 1947;**12**:153–7. <https://doi.org/10.1007/BF02295996>
32. de Nooy W, Mrvar A, Batagelj V. *Exploratory Social Network Analysis with Pajek*. Cambridge: Cambridge University Press; 2005. <https://doi.org/10.1017/CBO9780511806452>
33. Curtis L, Burns A. *Unit Costs of Health and Social Care 2019*. Canterbury: Personal Social Services Research Unit, University of Kent; 2019.
34. NHS England and NHS Improvement joint pricing team. *2019/20 National Tariff Payment System: National Prices and Prices for Emergency Care Services*. 2019. URL: <https://improvement.nhs.uk/resources/national-tariff/#h2-201920-national-tariff-payment-system> (accessed 12 February 2020).
35. NHS Improvement. *Reference Costs 2017/2018: Highlights, Analysis and Introduction to the Data*. 2018. URL: <https://improvement.nhs.uk/resources/reference-costs/> (accessed 16 January 2021).
36. University Hospital Southampton NHS Foundation Trust. *Eye Unit*. 2019. URL: www.uhs.nhs.uk/ourservices/eyes/eyeunit.aspx (accessed 30 January 2019).
37. Topcon Ireland Medical. *Equipment Price*. URL: www.topcon-medical.ie/ie/categories/27-diagnostic/ (accessed 20 February 2020).
38. Optos Plc. *Equipment Price*. URL: www.optos.com/en-gb/products/ (accessed 20 February 2020).
39. Veatch Ophthalmic Instruments. *Haag-Streit BM 900 Table LED Slit Lamp Equipment Price*. URL: www.veatchinstruments.com/Haag-Streit-BM-900-Table-LED-Slit-Lamp (accessed 8 January 2020).
40. National Institute for Health and Care Excellence (NICE). *Guide to the Methods of Technology Appraisal*. London: NICE; 2013.
41. EuroQol Group. EQ-5D-5L. URL: <https://euroqol.org/eq-5d-instruments/eq-5d-5l-about/valuation-standard-value-sets/crosswalk-index-value-calculator/> (accessed 20 March 2020).
42. van Hout B, Janssen MF, Feng YS, Kohlmann T, Busschbach J, Golicki D, et al. Interim scoring for the EQ-5D-5L: mapping the EQ-5D-5L to EQ-5D-3L value sets. *Value Health* 2012;**15**:708–15. <https://doi.org/10.1016/j.jval.2012.02.008>
43. Mangione CM, Lee PP, Gutierrez PR, Spritzer K, Berry S, Hays RD, National Eye Institute Visual Function Questionnaire Field Test Investigators. Development of the 25-item National Eye Institute Visual Function Questionnaire. *Arch Ophthalmol* 2001;**119**:1050–8. <https://doi.org/10.1001/archophth.119.7.1050>
44. Peacock S, Misajon R, Iezzi A, Richardson J, Hawthorne G, Keefe J. Vision and quality of life: development of methods for the VisQoL vision-related utility instrument. *Ophthalmic Epidemiol* 2008;**15**:218–23. <https://doi.org/10.1080/09286580801979417>
45. Centre for Economics. *Scoring - Psychometric (unweighted) or Utility (weighted)?* Australia: Monash University; 2020. URL: www.aqol.com.au/index.php/scoring-algorithms (accessed 13 February 2020).
46. Brown GC. Vision and quality of life. *Trans Am Ophthalmol Soc* 1999;**97**:473–511.
47. Brown MM, Brown GC, Sharma S, Landy J, Bakal J. Quality of life with visual acuity loss from diabetic retinopathy and age-related macular degeneration. *Arch Ophthalmol* 2002;**120**:481–4. <https://doi.org/10.1001/archophth.120.4.481>
48. Wallace P. *Pay and Conditions Circular (M&D) 2/2019 R2*. NHS Employers; 2020. URL: www.nhsemployers.org/-/media/Employers/Documents/Pay-and-reward/Pay-and-Conditions-Circular-MD-22019.pdf (accessed 16 January 2021).
49. British Medical Association. *Pay Scales for Consultants in England*. 2019. URL: www.bma.org.uk/advice/employment/pay/consultants-pay-england (accessed 2 February 2020).

50. Prescott G, Sharp P, Goatman K, Scotland G, Fleming A, Philip S, *et al.* Improving the cost-effectiveness of photographic screening for diabetic macular oedema: a prospective, multi-centre, UK study. *Br J Ophthalmol* 2014;**98**:1042–9. <https://doi.org/10.1136/bjophthalmol-2013-304338>
51. Royal National Institute of Blind People. *Report Warns of Crisis in Eye Clinic Capacity as Cancelled Appointments Risks Patients' Sight*. 2018. URL: www.rnib.org.uk/who-we-are/media-centre/latest-media-releases/report-warns-crisis-eye-clinic-capacity-cancelled (accessed 4 March 2020).
52. Impairment A-PPGoEHaV. *See the Light: Improving Capacity in NHS Eye Care in England*. 2018. URL: www.rnib.org.uk/sites/default/files/See%20the%20light_Improving%20NHS%20eye%20care%20capacity%20in%20England.pdf (accessed 4 March 2020).
53. Sharp PF, Olson J, Strachan F, Hipwell J, Ludbrook A, O'Donnell M, *et al.* The value of digital imaging in diabetic retinopathy. *Health Technol Assess* 2003;**7**(30). <https://doi.org/10.3310/hta7300>
54. 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). <https://doi.org/10.3310/hta17510>
55. Abràmoff MD, Lou Y, Erginay A, Clarida W, Amelon R, Folk JC, Niemeijer M. Improved automated detection of diabetic retinopathy on a publicly available dataset through integration of deep learning. *Invest Ophthalmol Vis Sci* 2016;**57**:5200–6. <https://doi.org/10.1167/iovs.16-19964>
56. Abràmoff MD, Lavin PT, Birch M, Shah N, Folk JC. Pivotal trial of an autonomous AI-based diagnostic system for detection of diabetic retinopathy in primary care offices. *NPJ Digit Med* 2018;**1**:39. <https://doi.org/10.1038/s41746-018-0040-6>
57. Talks SJ, Manjunath V, Steel DH, Peto T, Taylor R. New vessels detected on wide-field imaging compared to two-field and seven-field imaging: implications for diabetic retinopathy screening image analysis. *Br J Ophthalmol* 2015;**99**:1606–9. <https://doi.org/10.1136/bjophthalmol-2015-306719>
58. Tsaousis KT, Empeslidis T, Konidaris VE, Kapoor B, Deane J. The concept of virtual clinics in monitoring patients with age-related macular degeneration. *Acta Ophthalmol* 2016;**94**:e353–5. <https://doi.org/10.1111/aos.12832>
59. Lee JX, Manjunath V, Talks SJ. Expanding the role of medical retina virtual clinics using multimodal ultra-widefield and optical coherence tomography imaging. *Clin Ophthalmol* 2018;**12**:2337–45. <https://doi.org/10.2147/OPTH.S181108>
60. Kortuem K, Fasler K, Charnley A, Khambati H, Fasolo S, Katz M, *et al.* Implementation of medical retina virtual clinics in a tertiary eye care referral centre. *Br J Ophthalmol* 2018;**102**:1391–5. <https://doi.org/10.1136/bjophthalmol-2017-311494>
61. Ahnood D, Souriti A, Williams GS. Assessing patient acceptance of virtual clinics for diabetic retinopathy: a large scale postal survey. *Can J Ophthalmol* 2018;**53**:207–9. <https://doi.org/10.1016/j.jcjo.2017.10.035>
62. Kern C, Kortuem K, Hamilton R, Fasolo S, Cai Y, Balaskas K, *et al.* Clinical outcomes of a hospital-based teleophthalmology service: what happens to patients in a virtual clinic? *Ophthalmol Retina* 2019;**3**:422–8. <https://doi.org/10.1016/j.oret.2019.01.011>
63. Kotecha A, Baldwin A, Brookes J, Foster PJ. Experiences with developing and implementing a virtual clinic for glaucoma care in an NHS setting. *Clin Ophthalmol* 2015;**9**:1915–23. <https://doi.org/10.2147/OPTH.S92409>
64. Gunn PJG, Marks JR, Au L, Waterman H, Spry PGD, Harper RA. Acceptability and use of glaucoma virtual clinics in the UK: a national survey of clinical leads. *BMJ Open Ophthalmol* 2018;**3**:e000127. <https://doi.org/10.1136/bmjophth-2017-000127>

65. Prior L. In praise of small N, and of N = 1 in particular. *Critical Public Health* 2016;**26**:115–17. <https://doi.org/10.1080/09581596.2015.1130250>
66. Prior L, Scott D, Hunter R, Donnelly M, Tully MA, Cupples ME, Kee F. Exploring lay views on physical activity and their implications for public health policy. A case study from East Belfast. *Soc Sci Med* 2014;**114**:73–80. <https://doi.org/10.1016/j.socscimed.2014.05.015>
67. Evans MR, Prout H, Prior L, Tapper-Jones LM, Butler CC. A qualitative study of lay beliefs about influenza immunisation in older people. *Br J Gen Pract* 2007;**57**:352–8.
68. Morris PG, Prior L, Deb S, Lewis G, Mayle W, Burrow CE, Bryant E. Patients' views on outcome following head injury: a qualitative study. *BMC Fam Pract* 2005;**6**:30. <https://doi.org/10.1186/1471-2296-6-30>

Appendix 1 Additional tables from Chapter 2 diagnostic accuracy

TABLE 36 Demographics of participants: active diseases based on ophthalmologist slit-lamp examination

Demographic characteristic	Disease status, n (%)			
	Active DMO (N = 198)	Active PDR (N = 123)	Either DMO or PDR active (N = 272)	Both DMO and PDR active (N = 49)
Sex				
Male	131 (66.16)	83 (67.48)	182 (66.91)	32 (65.31)
Female	67 (33.84)	40 (32.52)	90 (33.09)	17 (34.69)
Age (years)				
18–59	83 (41.92)	81 (65.85)	134 (49.26)	30 (61.22)
≥ 60	115 (58.08)	42 (34.15)	138 (50.74)	19 (38.78)
Ethnic origin				
White	171 (86.36)	106 (86.18)	235 (86.40)	42 (85.71)
Black	14 (7.07)	8 (6.50)	18 (6.62)	4 (8.16)
Asian	7 (3.54)	7 (5.69)	13 (4.78)	1 (2.04)
Middle Eastern	3 (1.52)	2 (1.63)	3 (1.10)	2 (4.08)
Other	3 (1.52)	–	3 (1.10)	–

TABLE 37 Findings identified by ophthalmologists (when undertaking slit-lamp biomicroscopy = reference standard) and recorded as potential factors that could lead to inadequate image quality

Factors that could lead to inadequate image quality	Right eye (n = 397)	Left eye (n = 396) (excluded artificial eye)
Pupil size < 3 mm after dilatation		
Yes	102	105
No	293	288
Missing	2	3
Lens status of the eye		
Phakic	298	305
Aphakic	1	0
Pseudophakic	98	90
Missing	–	1
Cataract		
Yes	21	29
No	376	347
Vitreous haemorrhage		
Yes	22	34
No	375	361
Missing	–	1

continued

TABLE 37 Findings identified by ophthalmologists (when undertaking slit-lamp biomicroscopy = reference standard) and recorded as potential factors that could lead to inadequate image quality (*continued*)

Factors that could lead to inadequate image quality	Right eye (n = 397)	Left eye (n = 396) (excluded artificial eye)
Corneal haze/opacity		
Yes	5	6
No	392	389
Missing	–	1
Other	13	13
Poor patient co-operation	3	1
Ptosis	2	1
Dry eyes	1	2
Anterior synechiae	0	1
Hyphema	0	1
Anterior capsular opacification	2	2
Posterior capsule opacification	3	2
Asteroid hyalosis	2	2
Corneal scar	0	1

TABLE 38 Other findings/comorbidities identified by the ophthalmologist when undertaking slit-lamp biomicroscopy

Other findings/comorbidities	Right eye (n = 397)	Left eye (n = 396) (excluded artificial eye)
Epiretinal membrane	10	16
Glaucoma	9	13
Amblyopia	3	3
Ocular hypertension	1	4
Macular hole	1	2
Retinal vein occlusion	3	0
Age-related macular degeneration	1	1
Choroidal naevus	1	1
High myopia	1	1
Vitreomacular traction	0	2
Macular scar	0	1
Neovascularisation of the iris	1	0
Optic atrophy	0	1
Retinal artery occlusion	1	0
Comorbidities with conditions not listed ^a	1	8
Total	33	53

^a Comorbidities with conditions that are not listed referred to those ophthalmologists who said 'Yes' there is eye comorbidities but have not stated what it is.

TABLE 39 Reason for missing eye-level assessments

Reasons for missing data	O-FTF DMO		O-FTF PDR		G-OCT		G-OPTOS		G-ETDRS	
	Right eye	Left eye	Right eye	Left eye	Right eye	Left eye	Right eye	Left eye	Right eye	Left eye
Data/images are available (no data/images)	395 (1)	389 (4)	395 (2)	391 (2)	385 (11)	380 (13)	379 (17)	374 (19)	376 (20)	362 (31)
Missing: patient needs referral	1	4	2	2	6	6	4	6	6	13
No view of fundus: cataract	1		1							
No view of fundus: haemorrhage		3	1			1		1		2
No view of fundus: no reason specified		1		1						
No view of fundus: corneal graft				1				1		1
Poor pupillary dilatation					2		1		2	2
Patient unable to co-operate with imaging							1	1	1	2
Patient unable to sit at the camera							1	2	1	2
Media opacity: type unspecified					4	5	1	1	2	4
Missing: other reason missing	–	–	–	–	5	7	13	13	14	18
Patient did not attend					1	1	4	4	8	8
Only one eye image captured					1	1	2	3	2	3
System error					1	3	3	3	2	4
Other unknown reason					2	2	4	3	2	3
Total	396	393	396	393	396	393	396	393	396	393

TABLE 40 PDR test results using OPTOS or using ETDRS vs. reference standard or enhanced standard

Agreement regarding active PDR status	n/N	Estimate
G-OPTOS referral ^a vs. O-FTF active PDR	87/105	82.86%
G-ETDRS referral vs. O-FTF active PDR	87/102	85.29%
O-OPTOS active PDR vs. O-FTF active PDR that require treatment	62/88	70.45%
O-ETDRS active PDR vs. O-FTF active PDR that require treatment	54/79	68.35%
O-OPTOS active PDR vs. the enhanced standard	96/138	69.57%
O-ETDRS active PDR vs. the enhanced standard	85/131	64.89%
O-FTF vs. the enhanced standard	111/146	76.03%
O-FTF and O-OPTOS vs. the enhanced standard	133/146	91.10%
O-FTF and O-ETDRS vs. the enhanced standard	131/146	89.73%

G-ETDRS, grader assessment based on seven-field ETDRS fundus images; G-OPTOS, grader assessment based on ultra-wide field fundus images using OPTOS; O-ETDRS, ophthalmologist assessment based on seven-field ETDRS fundus images; O-FTF, reference standard for PDR; O-OPTOS, ophthalmologist assessment based on ultra-wide field fundus images using OPTOS.

a Grader referral for PDR = 'active' + 'unsure' + 'ungradable'.

TABLE 41 Secondary analyses on randomly selected eyes: all patients – SECA1A

Secondary analysis	Reference standard	Diagnostic parameter	n/N	Estimate (95% CI)
G-OCT identified active DMO in randomly selected eye	O-FTF + OCT active DMO in randomly selected eye	Sensitivity	116/133	87% (80% to 92%)
		Specificity	174/253	69% (63% to 74%)
		Positive likelihood ratio	–	2.79 (2.30 to 3.39)
		Negative likelihood ratio	–	0.19 (0.12 to 0.29)
G-OPTOS identified active PDR in randomly selected eye	O-FTF active PDR in randomly selected eye	Sensitivity	45/81	56% (45% to 66%)
		Specificity	262/300	87% (83% to 91%)
		Positive likelihood ratio	–	4.39 (3.07 to 6.26)
		Negative likelihood ratio	–	0.51 (0.40 to 0.65)
G-ETDRS identified active PDR in randomly selected eye	O-FTF active PDR in randomly selected eye	Sensitivity	45/75	60% (49% to 70%)
		Specificity	251/295	85% (81% to 89%)
		Positive likelihood ratio	–	4.02 (2.89 to 5.59)
		Negative likelihood ratio	–	0.47 (0.35 to 0.62)
G-ETDRS, grader assessment based on seven-field ETDRS fundus images; G-OCT, grader assessment based on SD-OCT images; G-OPTOS, grader assessment based on ultra-wide field fundus images using OPTOS.				

TABLE 42 Secondary analyses at eye level: all patients

Secondary analysis	Results	Reference standard	Diagnostic parameter	Eye	n/N	Estimate (95% CI)
SECA1B	G-OCT central-involving active DMO at eye level	O-FTF + OCT central-involving active DMO at eye level	Sensitivity	Right	89/106	84% (76% to 90%)
				Left	79/86	92% (84% to 96%)
			Specificity	Right	227/279	81% (76% to 85%)
				Left	229/291	79% (74% to 83%)
			Positive likelihood ratio	Right	–	4.50 (3.48 to 5.84)
				Left	–	4.31 (3.43 to 5.42)
SECA1C	G-OPTOS referral ^a for PDR at eye level	ERS at eye level	Sensitivity	Right	68/116	59% (50% to 67%)
				Left	79/117	68% (59% to 75%)
			Specificity	Right	199/262	76% (70% to 81%)
				Left	205/256	80% (75% to 85%)
			Positive likelihood ratio	Right	–	2.44 (1.87 to 3.17)
				Left	–	3.39 (2.57 to 4.47)
			Negative likelihood ratio	Right	–	0.54 (0.43 to 0.68)
				Left	–	0.41 (0.31 to 0.53)

TABLE 42 Secondary analyses at eye level: all patients (continued)

Secondary analysis	Results	Reference standard	Diagnostic parameter	Eye	n/N	Estimate (95% CI)
	G-ETDRS referral for PDR at eye level	ERS at eye level	Sensitivity	Right	77/115	67% (58% to 75%)
				Left	78/111	70% (61% to 78%)
			Specificity	Right	199/261	76% (71% to 81%)
				Left	188/251	75% (69% to 80%)
			Positive likelihood ratio	Right	–	2.82 (2.19 to 3.63)
				Left	–	2.80 (2.19 to 3.58)
			Negative likelihood ratio	Right	–	0.43 (0.33 to 0.57)
				Left	–	0.40 (0.30 to 0.53)
G-ETDRS, grader assessment based on seven-field ETDRS fundus images; G-OPTOS, grader assessment based on ultra-wide field fundus images using OPTOS; O-FTF, reference standard for PDR; O-FTF + OCT, reference standard for DMO.						
a Grader referral for PDR = ‘active’ + ‘unsure’ + ‘ungradable’.						

TABLE 43 Recruitment proportion by different time points

Time point	Disease							
	DMO				PDR			
	Active, n (%)	Inactive, n (%)	Ineligible, n (%)	Total, n	Active, n (%)	Inactive, n (%)	Ineligible, n (%)	Total, n
First 150 patients	58 (39)	52 (35)	40 (27)	150	18 (12)	87 (58)	45 (30)	150
Until 1 February 2019 ^a	140 (41)	106 (31)	99 (29)	345	64 (19)	167 (48)	114 (33)	345
Total eligible patients	152 (38)	120 (30)	125 (31)	397	111 (28)	170 (43)	116 (29)	397
a 1 February 2019 is the time at which we stopped recruiting patients who contribute only to the inactive DMO and/or inactive PDR groups.								

Appendix 2 Dealing with a small number of participants

Having $n = 36$ (for patients) is, on the face of things, disappointing. However, it would be wrong to dismiss the results of the FG research on the basis of small numbers in a world in which even $n = 1$ has been known to form the basis for discoveries of lasting significance.⁶⁵ More importantly, research in a range of disciplines has repeatedly suggested that the amount of information or data gained from the recruitment of additional cases usually declines to zero as n increases (e.g. the number of new species discovered in sequential river water samples). Accumulation curves can be constructed to illustrate the process, as in *Figures 17–20* (using unpublished data from studies with which the report author was previously involved).

Figure 17 traces the number of perceived barriers to increasing physical exercise according to participants of 14 FGs. The context of the research related to the provision of a 16 km linear park that was redesigned and refurbished for an area of urban deprivation. Barriers to increasing physical exercise mentioned by the FG participants included lack of safety, the weather, sectarianism, pollution and so forth. *Figure 18* shows the number of activities that were considered suitable for the redesigned park; these included walking, cycling and running. The sample comprised different community-based groups (e.g. older people, young mothers and people with disabilities) who agreed to participate in FG discussions at the research location (14 FGs were held with 113 individuals).⁶⁶ *Figure 17* shows that 82% (or 19/23 barriers) were identified by the first three groups. (Groups tended to offer an average of 10 suggestions each.) *Figure 18* indicates that it took just five FGs to identify 82% (or 28/34) of suitable activities. Therefore, the data imply that a relatively small number of FGs would have sufficed to capture the bulk of the information that was ultimately collected.

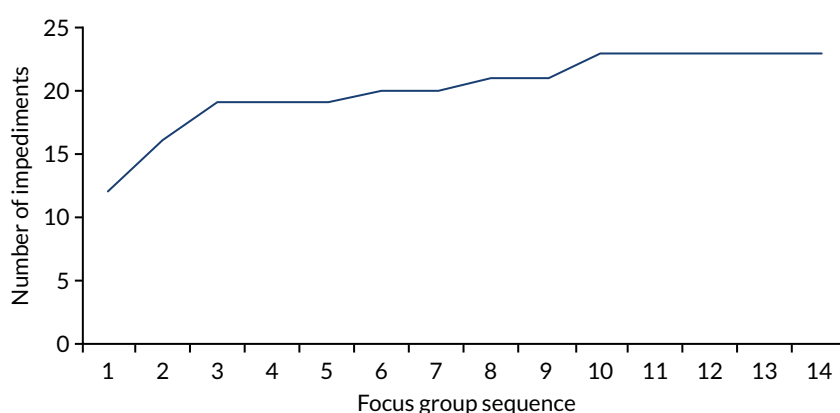


FIGURE 17 Barriers to increasing physical activity identified in 14 FGs.

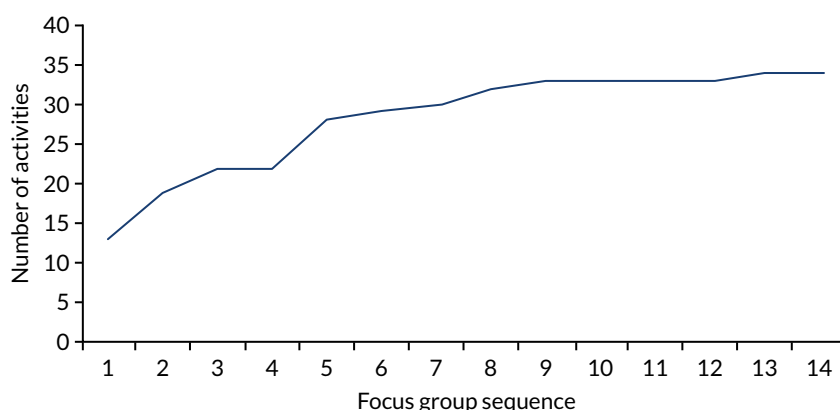


FIGURE 18 Activities for a regenerated park.

Similar accumulation patterns are evident from other studies in which other qualitative methods, and other forms of sampling, have been used. For example, *Figure 19* tracks the number of references to different 'causes' of influenza, such as coughs, sneezes and cold weather, as identified in face-to-face interviews by a sample of older GP patients in south Wales ($n = 54$). The sample was randomly selected from GP lists and stratified according to flu immunisation status (accepted, refused, lapsed or not offered) at time of interview.⁶⁷ It shows that 70% of the presumed causes of flu were mentioned in the first 15 conversational interviews; thereafter, there was a slow accumulation of additional information up to interview 41, after which no new causes were identified.

Figure 20 shows the consequences (e.g. memory loss, depression and loss of job or speech) of traumatic brain injury, as identified by 32 traumatic brain injury patients in south Wales. The sample, drawn from an outpatient clinic list, was a purposive one stratified according to the severity of injury and the time in months since the critical incident.⁶⁸ Of the 55 'consequences' mentioned, 50 (90%) were cited in the first 12 interviews. After interview 24, no new information was obtained.

One implication of these graphs is that although findings from a small n might not be good for estimating how observations are distributed according to gender, age, ethnicity or income group, the findings can nevertheless carry a considerable information load. It is also worth keeping in mind that in the current study, although the participants were self-selecting, they were drawn from five geographically dispersed locations in the UK.

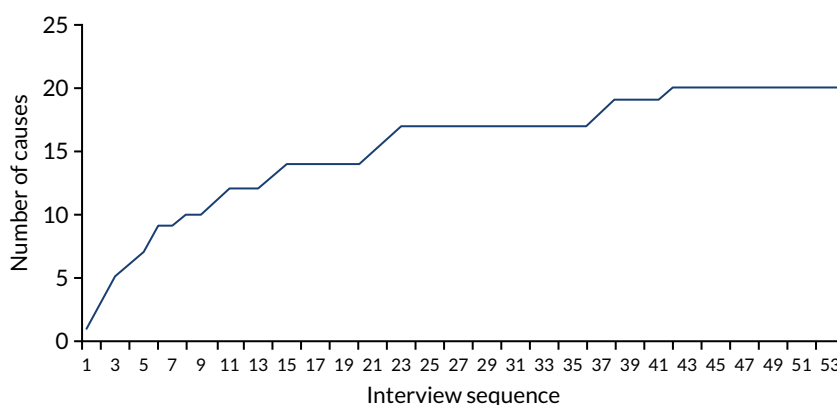


FIGURE 19 Number of new (additional) causes of flu referenced in sequential interviews ($n = 54$).

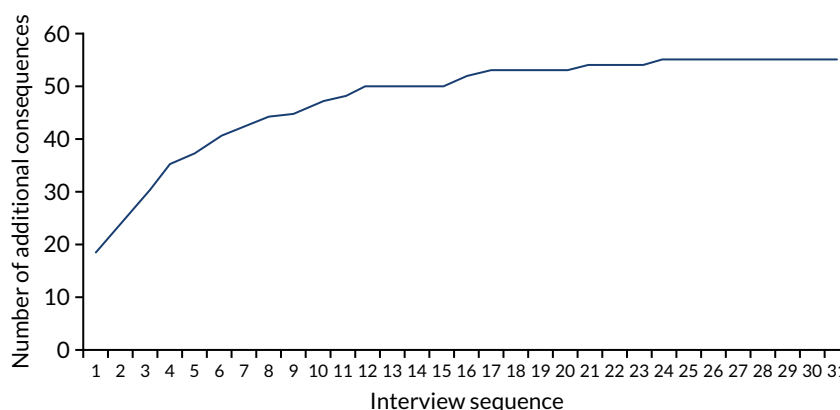


FIGURE 20 Consequences of traumatic brain injury as identified in 32 interviews with traumatic brain injury outpatients.

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 and Social Care*

Published by the NIHR Journals Library