



Health Technology Assessment

Volume 28 • Issue 32 • July 2024

ISSN 2046-4924

Home-monitoring for neovascular age-related macular degeneration in older adults within the UK: the MONARCH diagnostic accuracy study

Ruth E Hogg, Robin Wickens, Sean O'Connor, Eleanor Gidman, Elizabeth Ward, Charlene Treanor, Tunde Peto, Ben Burton, Paul Knox, Andrew J Lotery, Sobha Sivaprasad, Michael Donnelly, Chris A Rogers and Barnaby C Reeves



Home-monitoring for neovascular age-related macular degeneration in older adults within the UK: the MONARCH diagnostic accuracy study

Ruth E Hogg^{1*}, Robin Wickens², Sean O'Connor¹, Eleanor Gidman², Elizabeth Ward², Charlene Treanor¹, Tunde Peto¹, Ben Burton³, Paul Knox⁴, Andrew J Lotery⁵, Sobha Sivaprasad⁶, Michael Donnelly¹, Chris A Rogers² and Barnaby C Reeves^{2*}

¹Centre for Public Health, Queen's University Belfast, Belfast, UK

²Bristol Trials Centre, University of Bristol, Bristol, UK

³James Paget University Hospitals NHS Trust, Great Yarmouth, UK

⁴University of Liverpool, Liverpool, UK

⁵Department of Clinical and Experimental Sciences, Faculty of Medicine, University of Southampton, Southampton, UK

⁶NIHR Moorfields Biomedical Research Centre, Moorfields Eye Hospital NHS Foundation Trust, London, UK

*Corresponding authors

Published July 2024

DOI: 10.3310/CYRA9912

This report should be referenced as follows:

Hogg RE, Wickens R, O'Connor S, Gidman E, Ward E, Treanor C, *et al.* Home-monitoring for neovascular age-related macular degeneration in older adults within the UK: the MONARCH diagnostic accuracy study. *Health Technol Assess* 2024;**28**(32). <https://doi.org/10.3310/CYRA9912>

Health Technology Assessment

ISSN 2046-4924 (Online)

Impact factor: 3.6

A list of Journals Library editors can be found on the [NIHR Journals Library website](#)

Launched in 1997, *Health Technology Assessment* (HTA) has an impact factor of 3.6 and is ranked 32nd (out of 105 titles) in the 'Health Care Sciences & Services' category of the Clarivate 2022 Journal Citation Reports (Science Edition). It is also indexed by MEDLINE, CINAHL (EBSCO Information Services, Ipswich, MA, USA), EMBASE (Elsevier, Amsterdam, the Netherlands), NCBI Bookshelf, DOAJ, Europe PMC, the Cochrane Library (John Wiley & Sons, Inc., Hoboken, NJ, USA), INAHTA, the British Nursing Index (ProQuest LLC, Ann Arbor, MI, USA), Ulrichsweb™ (ProQuest LLC, Ann Arbor, MI, USA) and the Science Citation Index Expanded™ (Clarivate™, Philadelphia, PA, USA).

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.

Criteria for inclusion in the *Health Technology Assessment* journal

Manuscripts 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 article

The research reported in this issue of the journal was funded by the HTA programme as award number 15/97/02. The contractual start date was in October 2017. The draft manuscript began editorial review in April 2022 and was accepted for publication in December 2022. 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' manuscript 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 article.

This article presents independent research funded by the National Institute for Health and Care 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, 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, the HTA programme or the Department of Health and Social Care.

This article was published based on current knowledge at the time and date of publication. NIHR is committed to being inclusive and will continually monitor best practice and guidance in relation to terminology and language to ensure that we remain relevant to our stakeholders.

Copyright © 2024 Hogg *et al.* This work was produced by Hogg *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This is an Open Access publication distributed under the terms of the Creative Commons Attribution CC BY 4.0 licence, which permits unrestricted use, distribution, reproduction and adaptation in any medium and for any purpose provided that it is properly attributed. See: <https://creativecommons.org/licenses/by/4.0/>. For attribution the title, original author(s), the publication source – NIHR Journals Library, and the DOI of the publication must be cited.

Published by the NIHR Journals Library (www.journalslibrary.nihr.ac.uk), produced by Newgen Digitalworks Pvt Ltd, Chennai, India (www.newgen.co).

Abstract

Home-monitoring for neovascular age-related macular degeneration in older adults within the UK: the MONARCH diagnostic accuracy study

Ruth E Hogg^{1*}, Robin Wickens², Sean O'Connor¹, Eleanor Gidman², Elizabeth Ward², Charlene Treanor¹, Tunde Peto¹, Ben Burton³, Paul Knox⁴, Andrew J Lotery⁵, Sobha Sivaprasad⁶, Michael Donnelly¹, Chris A Rogers² and Barnaby C Reeves^{2*}

¹Centre for Public Health, Queen's University Belfast, Belfast, UK

²Bristol Trials Centre, University of Bristol, Bristol, UK

³James Paget University Hospitals NHS Trust, Great Yarmouth, UK

⁴University of Liverpool, Liverpool, UK

⁵Department of Clinical and Experimental Sciences, Faculty of Medicine, University of Southampton, Southampton, UK

⁶NIHR Moorfields Biomedical Research Centre, Moorfields Eye Hospital NHS Foundation Trust, London, UK

*Corresponding authors r.e.hogg@qub.ac.uk and Barney.Reeves@bristol.ac.uk

Background: Most neovascular age-related macular degeneration treatments involve long-term follow-up of disease activity. Home monitoring would reduce the burden on patients and those they depend on for transport, and release clinic appointments for other patients. The study aimed to evaluate three home-monitoring tests for patients to use to detect active neovascular age-related macular degeneration compared with diagnosing active neovascular age-related macular degeneration by hospital follow-up.

Objectives: There were five objectives:

- Estimate the accuracy of three home-monitoring tests to detect active neovascular age-related macular degeneration.
- Determine the acceptability of home monitoring to patients and carers and adherence to home monitoring.
- Explore whether inequalities exist in recruitment, participants' ability to self-test and their adherence to weekly testing during follow-up.
- Provide pilot data about the accuracy of home monitoring to detect conversion to neovascular age-related macular degeneration in fellow eyes of patients with unilateral neovascular age-related macular degeneration.
- Describe challenges experienced when implementing home-monitoring tests.

Design: Diagnostic test accuracy cohort study, stratified by time since starting treatment.

Setting: Six United Kingdom Hospital Eye Service macular clinics (Belfast, Liverpool, Moorfields, James Paget, Southampton, Gloucester).

Participants: Patients with at least one study eye being monitored by hospital follow-up.

Reference standard: Detection of active neovascular age-related macular degeneration by an ophthalmologist at hospital follow-up.

Index tests:

1. KeepSight Journal: paper-based near-vision tests presented as word puzzles.
2. MyVisionTrack®: electronic test, viewed on a tablet device.
3. MultiBit: electronic test, viewed on a tablet device.

Participants provided test scores weekly. Raw scores between hospital follow-ups were summarised as averages.

Results: Two hundred and ninety-seven patients (mean age 74.9 years) took part. At least one hospital follow-up was available for 317 study eyes, including 9 second eyes that became eligible during follow-up, in 261 participants (1549 complete visits). Median testing frequency was three times/month. Estimated areas under receiver operating curves were < 0.6 for all index tests, and only KeepSight Journal summary score was significantly associated with the lesion activity (odds ratio = 3.48, 95% confidence interval 1.09 to 11.13, $p = 0.036$). Older age and worse deprivation for home address were associated with lower participation ($\chi^2 = 50.5$ and 24.3 , respectively, $p < 0.001$) but not ability or adherence to self-testing. Areas under receiver operating curves appeared higher for conversion of fellow eyes to neovascular age-related macular degeneration (0.85 for KeepSight Journal) but were estimated with less precision. Almost half of participants called a study helpline, most often due to inability to test electronically.

Limitations: Pre-specified sample size not met; participants' difficulties using the devices; electronic tests not always available.

Conclusions: No index test provided adequate test accuracy to identify lesion diagnosed as active in follow-up clinics. If used to detect conversion, patients would still need to be monitored at hospital. Associations of older age and worse deprivation with study participation highlight the potential for inequities with such interventions. Provision of reliable electronic testing was challenging.

Future work: Future studies evaluating similar technologies should consider:

1. Independent monitoring with clear stopping rules based on test performance.
2. Deployment of apps on patients' own devices since providing devices did not reduce inequalities in participation and complicated home testing.
3. Alternative methods to summarise multiple scores over the period preceding a follow-up.

Trial registration: This trial is registered as ISRCTN79058224.

Funding: This award was funded by the National Institute of Health and Care Research (NIHR) Health Technology Assessment programme (NIHR award ref: 15/97/02) and is published in full in *Health Technology Assessment*; Vol. 28, No. 32. See the NIHR Funding and Awards website for further award information.

Contents

List of tables	xi
List of figures	xv
List of abbreviations	xvii
Plain language summary	xix
Scientific summary	xxi
Chapter 1 Introduction	1
Background and rationale	1
<i>Test accuracy of tests for self-monitoring neovascular age-related macular degeneration activity</i>	2
<i>The KeepSight Journal</i>	3
<i>MyVisionTrack</i>	4
<i>MultiBit test</i>	4
<i>Reference standard</i>	4
Potential inequalities in uptake	5
Aims and objectives	6
Chapter 2 Methods	9
Study design	9
<i>Changes to trial design after commencement of the trial</i>	9
Participant eligibility	11
<i>Classification of eyes and eligibility of patients with bilateral neovascular age-related macular degeneration</i>	11
<i>Inclusion criteria</i>	12
<i>Exclusion criteria</i>	12
<i>Stratification of study population</i>	12
Settings	12
Index tests	13
<i>The KeepSight Journal</i>	13
<i>MyVisionTrack</i>	13
<i>MultiBit</i>	13
<i>Reference standard</i>	14
Sample size	14
<i>Objective A and C</i>	14
<i>Objective D</i>	15
Outcomes	15
<i>Primary outcome</i>	15
<i>Secondary outcomes</i>	15
Data collection	15
<i>Patient identification</i>	15
<i>Training and equipment</i>	16
<i>Follow-up</i>	16
<i>Technical support</i>	16
<i>Retinal images</i>	16
Features to minimise bias	17

<i>Blinding</i>	17
<i>Bias due to selection of participants</i>	17
<i>Bias in the assessment of the index tests</i>	18
<i>Bias in the assessment of the reference standard</i>	18
<i>Bias due to exclusion of participants or inappropriate intervals between the times of index testing and the reference standard</i>	18
Qualitative methods (Objective B)	18
<i>Participant recruitment and data collection</i>	18
<i>Data management and analysis</i>	19
Methods for Objective E	19
Statistical methods	20
<i>Objective A: analytic approach and methods</i>	20
<i>Objective B</i>	21
<i>Objective C</i>	21
<i>Objective D</i>	22
<i>Objective D</i>	22
Patient and public involvement	22
Chapter 3 Results: study cohort	25
Recruitment	25
Monitoring visits	25
Baseline characteristics	30
Participant withdrawals	34
Protocol deviations	36
<i>Adherence to testing</i>	36
Retinal images collected	38
Effect of coronavirus disease 2019 pandemic on the study	38
Chapter 4 Results: primary outcome (Objective A)	41
Sensitivity analysis 2 – primary outcome based on Central Administrative Research	
Facility grading	43
<i>Secondary outcome – change in lesion activity</i>	48
Chapter 5 Results: Objective C	51
Impacts of inequalities on willingness in principle to participate	51
Impact of inequalities on recruitment	51
Impact of inequalities on ability and adherence to weekly testing	51
Chapter 6 Results: Objective D	59
Analytic approach and methods	59
Primary analysis	60
Sensitivity analysis 1 (data for preceding 4 weeks only)	60
Sensitivity analysis 2 (lesion activity classified by Central Administrative Research	
Facility grading)	62
Chapter 7 Results: Objective E	67
Participant helpline	67
<i>Incoming calls</i>	67
<i>Outgoing calls to participants</i>	67
Issues affecting the availability of the electronic tests	68
Logistical issues with the electronic equipment provided for the study	70

Chapter 8 Results: Objective B (qualitative)	73
Theme 1: the role of home monitoring	74
<i>Subtheme 1: understanding purpose</i>	74
<i>Subtheme 2: perceived impact on eye care</i>	75
Theme 2: suitability of procedures and instruments	75
Theme 3: experience of home-monitoring procedures	75
<i>Subtheme 1: training for home monitoring</i>	75
<i>Subtheme 2: test preferences</i>	76
<i>Subtheme 3: use of MultiBit vision test feedback and data</i>	76
Theme 4: feasibility of regular home monitoring in usual service delivery	76
<i>Subtheme 1: frequency of home monitoring and habit formation</i>	76
<i>Subtheme 2: use of ongoing support</i>	77
Theme 5: impediments to home monitoring	77
<i>Subtheme 1: practical issues</i>	77
<i>Subtheme 2: personal health and social factors</i>	77
Views of informal 'carers' and healthcare professionals – summary	77
Chapter 9 Discussion	81
Main findings: study conduct	81
Main findings: primary outcome (Objective A)	82
<i>Diagnostic accuracy of index tests</i>	82
Main findings: Objective C	82
<i>Potential inequalities in recruitment, ability to perform tests and adherence to testing</i>	82
Main findings: Objective D	83
<i>Diagnostic accuracy of conversion to neovascular age-related macular degeneration in fellow eyes</i>	83
Main findings: Objective B	83
<i>Qualitative evaluation of acceptance of home monitoring</i>	83
Main findings	83
<i>Challenges</i>	83
Equality, diversity and inclusion	85
<i>Participant representation</i>	85
<i>Reflections on the research team and wider involvement</i>	85
Patient and public involvement	85
Strengths and limitations	85
<i>Strengths</i>	85
Lessons for the future	86
Future research	87
Chapter 10 Conclusion	89
Additional Information	91
References	95
Appendix 1 Additional tables and figures	99
Appendix 2 Summary of protocol substantial amendments	123
Appendix 3 Statistical analysis plan	125

List of tables

TABLE 1 Recruitment summary	27
TABLE 2 Reasons for being unwilling in principle to participate	28
TABLE 3 Recruitment into each stratum	29
TABLE 4 Complete and total monitoring visits by site	30
TABLE 5 Baseline demographic characteristics and exposure to technology of consented participants	30
TABLE 6 Demographic characteristics of patients unwilling in principle vs. willing to participate in the study	32
TABLE 7 Baseline ocular histories of study eyes	34
TABLE 8 Combinations of reasons given by participants ($n = 88$) for decisions to withdraw	35
TABLE 9 Baseline images submitted and received by CARF	39
TABLE 10 Model results for the primary outcome of lesion activity in study eyes for all home-monitoring test models and for a no-test model; primary analysis	42
TABLE 11 Model results for the primary outcome of lesion activity in study eyes for all home-monitoring test models and for a no-test model; sensitivity analysis 1	44
TABLE 12 Model results for the primary outcome of lesion activity in right study eyes for all home-monitoring test models and for a no-test model; sensitivity analysis 2	46
TABLE 13 Model results for the primary outcome of lesion activity in left study eyes for all home-monitoring test models and for a no-test model; sensitivity analysis 2	47
TABLE 14 Model results for the secondary outcome of change from inactive to active lesion in study eyes for all home-monitoring test models and for a no-test model	49
TABLE 15 Impact of inequalities on participation	52
TABLE 16 Impact of inequalities on patient ability to perform tests	53
TABLE 17 Impact of inequalities on patient adherence to tests	55
TABLE 18 Number of participants with fellow eye conversions to nAMD and home-monitoring test data	59
TABLE 19 Home-monitoring test accuracy for conversion of fellow eyes against management decisions	61
TABLE 20 Fixed interval home-monitoring test accuracy for conversion of fellow eyes against management decisions	63

TABLE 21 Home-monitoring test accuracy for conversion of fellow eyes against reference standard of CARF grading	65
TABLE 22 Details of incoming helpline calls	67
TABLE 23 Reasons for incoming helpline calls	68
TABLE 24 Details of outgoing calls to patients for whom data had not been received in > 3 weeks or > 2 weeks since consent	68
TABLE 25 Reasons reported by participants for their data not being received	69
TABLE 26 Details of occasions when participants were unable to complete digital tests	69
TABLE 27 Sociodemographic characteristics of participants in the qualitative component of the MONARCH study compared to the rest of the MONARCH study sample	74
TABLE 28 Perspectives of patients on acceptability of home monitoring	78
TABLE 29 Semistructured guide for patient participant interviews	99
TABLE 30 Final scheme for coding of interview data	100
TABLE 31 Number of patients approached to participate in home-monitoring testing	102
TABLE 32 Reasons for ineligibility at screening and training	103
TABLE 34 Visit rate per patient per year	106
TABLE 35 Differences between intended and actual review schedules	106
TABLE 33 Details of re-training; denominators are all consented patients	106
TABLE 36 Reasons for withdrawals from the study	107
TABLE 37 Completion of the KSJ home-monitoring test	109
TABLE 38 Frequency of mVT and MBT home-monitoring test use	109
TABLE 39 Test completion for the mVT and MBT home-monitoring tests	110
TABLE 40 Associations between different KSJ test scores	111
TABLE 41 Model performance for predicting management decisions of lesion activity at thresholds defined by Youden's index for summary test scores for all index tests; primary analysis	111
TABLE 42 Average test scores above and below thresholds defined by Youden's index for the models; primary analysis	112
TABLE 43 Model performance for predicting management decisions of lesion activity at thresholds defined by Youden's index for summary test scores for all index tests; sensitivity analysis 1	112

TABLE 44 Average test scores above and below thresholds defined by Youden's index for the models; sensitivity analysis 1	112
TABLE 45 Model performance for predicting management decisions of lesion activity at thresholds defined by Youden's index for summary test scores for all index tests; sensitivity analysis 1	113
TABLE 46 Average test scores above and below thresholds defined by Youden's index for the models; sensitivity analysis 1	113
TABLE 47 Model performance for predicting management decisions of a change from inactive to active lesion status for thresholds defined by Youden's index for summary test scores for index tests (except for the mVT)	114
TABLE 48 Average test scores above and below thresholds defined by Youden's index for the models predicting management decisions of a change from inactive to active lesion status	114
TABLE 49 Model performance for predicting management decisions of fellow eye conversion for thresholds defined by Youden's index for summary test scores for index tests, primary analysis	115
TABLE 50 Average test scores above and below thresholds defined by Youden's index for the models predicting management decisions of fellow eye conversion, primary analysis	115
TABLE 51 Model performance for predicting management decisions of fellow eye conversion for thresholds defined by Youden's index for summary test scores for index tests, sensitivity analysis 1	115
TABLE 52 Average test scores above and below thresholds defined by Youden's index for the models predicting management decisions of fellow eye conversion, sensitivity analysis 1	116
TABLE 53 Model performance for predicting management decisions of fellow eye conversion for thresholds defined by Youden's index for summary test scores for index tests, sensitivity analysis 2	116
TABLE 54 Average test scores above and below thresholds defined by Youden's index for the models predicting management decisions of fellow eye conversion, sensitivity analysis 2	116
TABLE 55 Relationship between themes and coding of interview data	117
TABLE 56 Perspectives of HCPs	118
TABLE 57 Perspectives of informal 'carers'	119
TABLE 58 Common issues reported by patients during home monitoring	120
TABLE 59 Statistical analysis plan revision history	136

List of figures

FIGURE 1 Example week in the KSJ	3
FIGURE 2 Diagram illustrating the steps when self-monitoring with the mVT app	4
FIGURE 3 Diagram illustrating steps when self-monitoring with the MBT app	5
FIGURE 4 Study schema: Objectives A, C and D	10
FIGURE 5 Central subfield thickness from OCT scan	17
FIGURE 6 Patient flow chart	26
FIGURE 7 Cumulative number of participants recruited for Objectives A, C and D	28
FIGURE 8 Number of monitoring visits recorded on database	29
FIGURE 9 Kaplan–Meier graph of time to withdrawal	36
FIGURE 10 Time to stopping home monitoring with the mVT	37
FIGURE 11 Time to stopping home monitoring with the MBT	37
FIGURE 12 Time to stopping home monitoring with the KSJ	38
FIGURE 13 Receiver operating characteristic curves and AUROCs (95% CIs) for summary test scores for all index tests; primary analysis	43
FIGURE 14 Receiver operating characteristic curves and AUROCs (95% CIs) for summary test scores for all index tests; sensitivity analysis 1	45
FIGURE 15 Receiver operating characteristic curves and AUROCs (95% CIs) for summary test scores for all index tests; sensitivity analysis 2	48
FIGURE 16 Receiver operating characteristic curves and AUROCs (95% CIs) for summary test scores for index tests (except for the mVT) for a change from inactive to active lesion status	50
FIGURE 17 Receiver operating characteristic curves and AUROCs (95% CIs) for summary test scores for index tests for conversion of a fellow to an active lesion; primary analysis	62
FIGURE 18 Receiver operating characteristic curves and AUROCs (95% CIs) for summary test scores for index tests for conversion of a fellow to an active lesion; sensitivity analysis 1	64
FIGURE 19 Receiver operating characteristic curves and AUROCs (95% CIs) for summary test scores for index tests for conversion of a fellow to an active lesion; sensitivity analysis 2	66

LIST OF FIGURES

FIGURE 20	Participant flow – qualitative component of the MONARCH study	73
FIGURE 21	Time from screening to decision on willingness for eligible patients	104
FIGURE 22	Time from willing in principle to training	105
FIGURE 23	Kaplan–Meier graphs of time to withdrawal by site	108

List of abbreviations

AMD	age-related macular degeneration	MDMS	multiple device management software
app	software application	MIFI	Mobile Wireless-Fidelity
AUROC	area under the receiver operating characteristic	MONARCH	monitoring for neovascular age-related macular degeneration reactivation at home
BTC	Bristol Trials Centre	mVT [®]	MyVisionTrack [®] home monitoring test
CARF	Central Administrative Research Facility	nAMD	neovascular age-related macular degeneration
CF	colour fundus	NIHR	National Institute for Health and Care Research
CI	confidence interval	OCT	optical coherence tomography
CONSORT	Consolidated Standards of Reporting Trials	OR	odds ratio
COVID	coronavirus disease	PIL	participant information leaflet
CRF	case report form	PPI	patient and public involvement
CTEU	Clinical Trials and Evaluation Unit	REC	Research Ethics Committee
DTA	Diagnostic Test Accuracy	RNIB	Royal National Institute for the Blind
EDTRS	Early Treatment Diabetic Retinopathy Study	ROC	receiver operating characteristic
EDNA Study	Early Detection of age-related macular degeneration Study	SD	standard deviation
GCP	good clinical practice	SOP	standard operating procedure
HCP	healthcare professional	SSC	Study Steering Committee
HES	Hospital Eye Service	TAM	technology acceptance model
HM	Hand Movements	TFA	theoretical framework of acceptability
HTA	Health Technology Assessment	UTAUT	unified theory of acceptance and use of technology
IMD	Index of Multiple Deprivation	VA	visual acuity
IQR	interquartile range	VEGF	vascular endothelial growth factor
KSJ	KeepSight Journal	VMS grid	visual- and memory-stimulating grid
LogMAR	Logarithm of the minimum angle of resolution		
MBT	Multi-Bit Home-monitoring test		

Plain language summary

What was the question?

Treatment for neovascular age-related macular degeneration, the most common cause of sight loss in those over 50 years, involves regular eye injections and frequent follow-up appointments. This is inconvenient for patients and causes capacity issues in the hospital eye service. Finding tests that could be undertaken at home that could detect if a further injection and hospital appointment was required or not would increase capacity to see those at highest risk of sight loss and also reduce the burden on patients and their carers.

What did we do?

We investigated three different visual function tests, one paper-based and two applications on an iPod Touch_{TM} tablet (Apple, Cupertino, CA, USA). We wanted to see if they could detect an increase in disease activity that would require treatment, compared to the decision by a retinal specialist at a traditional hospital eye outpatient visit based on clinical examination and retinal imaging. To encourage those without a smartphone or home internet to participate, we provided both an iPod Touch and Mobile Wireless-Fidelity device with a mobile contract.

What did we find?

None of the tests performed well enough to safely monitor patients at home. Those who were willing to participate tended to be younger, had previous experience of using smartphones, sending e-mail and internet access and were more well-off than those who chose not to participate. Some participants also experienced difficulties with the devices provided and successfully uploading the data which were not related to the extent of previous information technology experience. There were also significant technical challenges for the research team. The study helpline was used heavily, considerably more than we anticipated.

What does it mean?

These tests are not ready to be used in this context. Future studies involving mobile health technology need to carefully consider how to reach those unlikely to participate and provide sufficient technical support to support long-term follow-up.

Scientific summary

Background

Neovascular age-related macular degeneration (nAMD) is the leading cause of blindness in older adults. Advanced nAMD causes substantial retinal damage, loss of central vision and reduced quality of life.

Several intravitreal drugs [anti-vascular endothelial growth factor (VEGF) that inhibits neovascularisation, i.e. anti-VEGF agents] are used to treat nAMD. Some eyes become fluid free after initial treatment over 3–6 months but relapse is common, and most patients require retreatment at some stage, with the disease typically becoming inactive for a period and then becoming active again. Hence, patients in the maintenance phase with inactive disease still need to be monitored regularly in hospital outpatient clinics for disease reactivation, when treatment is restarted. Monitoring places a substantial burden on hospital resources, patients and their family or carers. This burden would be substantially reduced if patients with inactive disease could self-monitor at home and attend hospital only when the disease reactivates.

Aim

To evaluate three non-invasive test strategies for use by patients at home to detect active nAMD compared to diagnosis of active nAMD during usual monitoring of patients in the Hospital Eye Service (HES).

Objectives

1. Estimate the accuracy of three home-monitoring tests to detect active nAMD.
2. Determine the acceptability of home monitoring to patients and carers and barriers to adhering to regular testing.
3. Describe inequalities in recruitment, participants' ability to self-test and adherence to testing during follow-up.
4. Estimate the accuracy of home monitoring to detect conversion to nAMD in fellow eyes of patients with unilateral nAMD.
5. Describe the challenges experienced when implementing the tests.

Study design

Diagnostic test accuracy cohort study.

Setting

Participants recruited from six UK HES macular clinics (Belfast, Liverpool, Moorfields, James Paget, Southampton, Gloucester).

Methods

Participants

We invited patients to take part who had at least one study eye being monitored by HES for nAMD, were first treated > 6 and < 42 months earlier. We tried to recruit equal numbers by time since starting treatment in the first-treated study eye (6–17 months; 18–29 months; 30–41 months) to ensure test

performance was estimated across this range of duration of nAMD. Patients were followed for at least 6 months.

Reference standard

The reference standard was the reviewing ophthalmologist's decision about the activity status of a study eye at a HES monitoring visit, recorded as active, inactive or uncertain. There were no additional hospital visits for the study. Such decisions are usually based on clinical examination and the results of hospital-based retinal imaging investigations, for example, colour fundus photographs and ocular coherence tomograms (OCTs). The reference standard grouped uncertain with inactive judgements for analyses.

Index tests

Three home-monitoring ('index') tests were evaluated, spanning low to moderate cost and complexity. These were:

1. KeepSight Journal (KSJ): a paper-based booklet of near-vision tests presented as word puzzles developed in the United States and adapted by the study team for use in the study.
2. MyVisionTrack® (mVT®): electronic vision test, intended to be viewed on a tablet device.
3. MultiBit (MBT): electronic vision test, intended to be viewed on a tablet device.

Specific thresholds indicating a significant clinical change were not provided for any of index tests in advance by their developers.

Outcomes

The primary outcome was classification of a study eye at a monitoring visit as having active or inactive disease (active, inactive, uncertain), that is, the reviewing ophthalmologist's decision.

A secondary outcome (new reference standard) for Objective A was a change from inactive to active status from one management visit to the next. This was considered better to represent how home monitoring might be implemented.

For Objective C, outcomes investigated were willingness in principle to participate; ability to carry out index tests; adherence to weekly testing.

For Objective D, the outcome was conversion of a fellow eye to active nAMD as judged by an ophthalmologist, that is, same reference standard.

For Objective E, the following outcomes were described as measures of the technical and logistical challenges identified during the study:

Frequency and reason for incoming calls made to the helpline and outgoing calls made to participants. Frequency and duration of events leading to the digital tests being unavailable for testing.

Other technical and logistical challenges.

Objective B study recruitment and data collection

Recruitment to the qualitative component began 3 months after the monitoring for neovascular age-related macular degeneration reactivation at home (MONARCH) study began recruiting. During the consent process for participation in the study, patients could consent to further contact from the qualitative research team to discuss participation in the qualitative study. Maximum variation sampling ensured a range of perspectives were captured in relation to: age (young-old 50–69 years and older-old > 70 years), gender, one or both eyes with nAMD, time since first treatment (defined as above) and adherence to home monitoring (test data from the two electronic tests were used to categorise

participants into 'regular' testers and 'irregular' testers). Patients who declined to participate in MONARCH but provided consent to be contacted about the qualitative study, informal 'carers', supporters or significant others in the lives of patients and healthcare professionals who interacted with participants at study sites visits were also approached to gather their perspectives about the acceptability of home monitoring.

Statistical analysis

Objective A: The test accuracy of index tests was estimated by fitting a logistic regression model to predict the reference standard from summary test scores for the interval between monitoring visits, adjusting for participants' baseline data. Accuracy was estimated for the primary outcome using all index test data, data only for the 4 weeks preceding the monitoring visit, the reference standard based on reading centre decisions made from OCT images and for the secondary outcome.

Test scores were summarised as: means (MBT and mVT); median (KSJ reported near visual acuity (VA), ordinal six-point scale); proportions (KSJ reported VA, Amsler grid and household object appearance reported worse than baseline vs. same or better). All four scores were fitted in the KSJ model and a single area under the receiving operator curve (AUROC) was estimated. Separate models were fitted for each test for the primary outcome, the two sensitivity analyses and the secondary outcome. Model performance was quantified by the odds ratio (OR) for the index test summary score(s) and the estimate of the AUROC and their respective confidence intervals (CIs). AUROCs were based on predicted probabilities calculated using only the fixed effects in the models. Sensitivity, specificity, positive and negative predictive values and 95% CIs were calculated using cut-off thresholds corresponding to Youden's index for each model, which minimises overall misclassifications. Average test scores above and below the thresholds were also calculated. Analyses took account of the structure within the data, that is, the nesting of visits and eyes within patients.

Objective B: All interviews were audio-recorded and transcribed. A directed content analysis approach based on deductive and inductive coding was used. NVivo version 12 was used to manage data and facilitate the analysis process, which, in summary, included the following stages: (1) independent transcription, (2) data familiarisation, (3) independent coding, (4) development of an analytical framework, (5) indexing, (6) charting and (7) interpreting data.

Objective C: Willingness in principle to participate was defined as an approached eligible patient agreeing to attend a research visit for training. Ability to perform an index test was defined as the proportion of monitoring visits for which some valid index test data were available. Adherence was defined as the proportion of weeks between monitoring visits for which some valid data for an index test were available. The ability and adherence models were performed for each test separately at the level of the patient.

Regression models estimated associations of age, sex, Index of Multiple Deprivation (IMD), stratum of time since first diagnosis and baseline visual acuity at diagnosis on the outcomes of willingness to participate, ability to perform tests and adherence to weekly testing; models for the latter two outcomes were fitted for each index test. Associations were reported with 95% CIs. Analyses of adherence and ability took account of nesting of visits within participants.

Objective D: The test accuracy of the index tests for the reference standard of an ophthalmologist's classification of a fellow eye as having active disease at a monitoring visit, that is, conversion to active nAMD, was estimated by the same methods as described for Objective A. Two sensitivity analyses were carried out: (1) the same reference standard but using test data only for the 4 weeks preceding the management visit and (2) the alternative reference standard of classification of a fellow eye having active disease based on reading centre grading of OCTs carried out during the monitoring visits.

Objective E: This objective used descriptive summary descriptive statistics only.

Results

The study recruited 297 patients (consented participants) between 21 August 2018 and 31 March 2020. Half of recruited participants were first treated for nAMD 6–17 months before consenting, 28% 18–29 months before consenting and 22% 30–41 months before consenting. At the end of the study, data for at least one monitoring visit after starting to use the index monitoring tests were available for 357 study eyes in 297 patients. Data were available for at least one complete monitoring visit after starting to use the index monitoring tests for 317 study eyes of 261 patients. More participants were women (58.6%). Participants' mean age was 74.9 (6.6) years [standard deviation (SD)]. The mean visual acuity of study eyes (better seeing eyes if participants had two study eyes) was 0.2 (0.2) Logarithm of the minimum angle of resolution (LogMAR) (SD).

Objective A: Median testing frequency was three times per month. In the primary analysis, estimated AUROCs were < 0.6 for all index tests, and only KSJ summary score was significantly associated with the lesion activity (OR = 3.48, 95% CI 1.09 to 11.13; $p = 0.036$). Estimated AUROCs were < 0.6 for all tests in both sensitivity analyses and for the secondary outcome of change from inactive to active status between adjacent management visits.

Objective B: Two overarching meta-themes emerged from the qualitative studies related to acceptance or non-acceptance of home monitoring. Meta-theme 1 encompassed four main themes: (1) the role of home monitoring; (2) suitability of procedures and instruments; (3) experience of home monitoring; and (4) feasibility of home monitoring in usual practice. Meta-theme 2 consisted of one main theme covering key inhibitors to acceptability. The main factors influencing acceptability included a participant's understanding about the purpose of home monitoring and their experience of using it. While home monitoring was generally seen as a relatively straightforward exercise to undertake and non-burdensome, training and ongoing support were regarded as essential to its success. Home monitoring was acceptable to patients and its potential to reduce clinic visits during non-active treatment phases was recognised.

Objective C: A minority of patients who were approached were willing in principle to participate. Increasing age and worse deprivation index for home address were associated with being unwilling in principle to participate ($\chi^2 = 50.5$ and 24.3 , respectively, both $p \leq 0.001$). Recruiting site was also associated with willingness in principle to participate, believed to be due to sites adopting different strategies for approaching and recruiting patients. Increasing age and worse deprivation were not consistently associated with either being able to self-monitor with the index tests or adherence to weekly testing (χ^2 for all tests < 5, $p > 0.08$ for ability and adherence, except for worse IMD being associated with better adherence for the KSJ, $\chi^2 = 12.15$, $p = 0.016$). Recruiting site was also associated with being able to test and adhering to weekly testing.

Objective D: There were 132 fellow eyes with data from 544 monitoring visits, 17 of which (12.9%) had nAMD recorded at one or more management visits over about 100 participant-years. This rate of conversion was higher than expected based on epidemiological studies of conversion rates in unaffected fellow eyes, potentially due to study eyes having had nAMD longer ago. Some predictors could not be fitted in models and estimates of associations were imprecise. The no-test model predicted conversion better than for Objective A (AUROC = 0.73) and electronic tests did not increase this (AUROCs = 0.73 and 0.76 for MBT and mVT, respectively). The estimated AUROC for the KSJ was 0.85, due to a strong positive association of the household object summary score with conversion (OR 15.3, $p = 0.036$).

Objective E: Despite two-thirds of the population having previously used a smartphone, there were still a variety of challenges experienced with the electronic devices while testing at home that contributed to both reduced adherence and ultimately withdrawals from the study.

Strengths and limitations

The study had several strengths. Estimates of the diagnostic test accuracy of index tests were at low risk of bias: the study population was appropriate for the intended use of the tests, and summary test scores were not available to ophthalmologists providing the reference standard, which was judged after the index test data were collected.

Limitations

A smaller-than-planned sample size (less than half the target number of monitoring visits); nonetheless, 95% CIs for AUROCs were narrow (± 0.04) and estimates were able to rule out tests providing adequate accuracy for diagnosing active nAMD to enable patients to be monitored without hospital review.

Tests were sometimes not available for technical reasons that were beyond the control of the research team.

The study had no control over monitoring visits and participants are likely to have reported their subjective visual experience to their consultants, which might have influenced the reference standard.

We could not define test thresholds a priori, and instead estimated AUROCs. We did not compare AUROCs for tests due to their poor accuracy.

The ways in which patients were approached and screened varied across sites, generating a site effect in analyses of potential inequalities; variations may have reflected the pre-conceptions of research staff regarding the capabilities of patients to use the electronic tests.

Conclusions

Based on the detection of lesion activity assessed by clinicians in the clinic, we have shown that none of the index tests provides acceptable test accuracy for home monitoring in this context. Associations of increasing age and deprivation index for home address with unwillingness in principle to participate despite provision of hardware highlight the potential for inequality with interventions of the kind evaluated. While a proportion of nAMD patients are willing and interested in the potential for home monitoring, substantial practical and technological issues are encountered in the implementation of such, requiring a significant support infrastructure, including a study helpline.

Future work

Future research should focus on the methodological challenge of efficiently evaluating mobile health technologies which deal with constantly emerging new technology. The clear evidence of inequalities in participation and retention should prompt future research on ways to encourage participant and adoption of mobile health technologies by underserved populations. Focus should also be placed on methods to improve adherence and retention in longitudinal studies involving electronic testing, particularly around the nature of feedback to participants.

Trial registration

This trial is registered as ISRCTN79058224.

Funding

This award was funded by the National Institute of Health and Care Research (NIHR) Health Technology Assessment programme (NIHR award ref: 15/97/02) and is published in full in *Health Technology Assessment*; Vol. 28, No. 32. See the NIHR Funding and Awards website for further award information.

Chapter 1 Introduction

Some text in this chapter has been reproduced from our study protocol, Ward E, Wickens RA, O'Connell A, Culliford LA, Rogers CA, Gidman EA, *et al.* Monitoring for neovascular age-related macular degeneration (AMD) reactivation at home: the MONARCH study. *Eye (Lond)* 2021;35(2):592-600. <https://doi.org/10.1038/s41433-020-0910-4>, published under the Creative Commons Attribution License (CC BY 4.0), which permits unrestricted use, distribution and reproduction in any medium, provided the original work is properly cited (<https://creativecommons.org/licenses/by/4.0/>).

Background and rationale

Despite significant therapeutic advances, neovascular or wet age-related macular degeneration (AMD) remains the leading cause of blindness in older adults.¹ While the early stages of AMD have subtle visual symptoms, the advanced stages [neovascular age-related macular degeneration (nAMD) and geographic atrophy (GA)] result in substantial retinal damage with accompanying loss of central vision and reduced quality of life.²

A range of intravitreal drugs that inhibit vascular endothelial growth factor (VEGF; anti-VEGF antibodies) for nAMD are currently available, with treatment generally starting with a loading phase of three injections over 3 consecutive months. A proportion of eyes become fluid free in the subsequent maintenance phase, but relapse is common and most patients require retreatment in affected eyes at some stage, with the disease typically becoming inactive for a period and then becoming active again.³ Patients in the maintenance phase with inactive disease still need to be monitored in hospital by measurement of best corrected visual acuity and optical coherence tomography (OCT). When disease reactivation is detected, treatment is restarted.

A large study has shown that, while many patients have many months of treatment-free periods, a significant burden falls on hospitals (as well as patients) with respect to the need for regular and repeated review.³ Thus, methods that might allow the patient to self-monitor at home would reduce the burden on hospitals.

When diagnosis of active nAMD is confirmed, treatment with anti-VEGF therapy is almost always initiated. In most cases, patients receive three injections every 4–6 weeks initially (loading phase) and then patients are reassessed at each subsequent visit in the treatment cycle to determine lesion activity and decide whether retreatment is necessary (maintenance phase). Monitoring visits use a combination of visual acuity, clinical biomicroscopic examination and OCT to determine if the neovascular lesion is active (wet) or inactive (dry). It is these monitoring appointments which are causing a significant strain on NHS outpatient clinics in eye hospitals.

Pre-coronavirus disease (COVID), during the maintenance phase, patients were monitored for relapse at regular monitoring outpatients' visits at eye hospitals. The frequency of monitoring depended on the drug being used (ranibizumab or aflibercept) and the preferred treatment regimen [treat (active disease)-as-necessary or treat-and-extend]. Ranibizumab is licensed for monthly treatment as required, and aflibercept every 2 months, during the maintenance phase. For treat-as-necessary regimens, a lesion found to be active is treated and a further monitoring visit is arranged; treatment is withheld if the nAMD lesion is inactive, and a further monitoring visit is arranged. For treat-and-extend regimens, 'prophylactic' treatment is administered to an eye with an inactive lesion, extending the interval between monitoring visits providing the disease remains inactive; if the nAMD lesion is found to be active, the interval between monitoring visits returns to the standard interval (1 month for ranibizumab or 2 months for aflibercept until the lesion becomes inactive, and the interval is then extended again). A disadvantage of the treat-and-extend treatment regimen is that it can lead to unnecessary overtreatment.

The seminal clinical trials of anti-VEGF therapy for nAMD used change in best corrected distance visual acuity as their primary end point.⁴⁻⁶ This was measured using strict protocols to standardise refraction methods, visual acuity recording, distance measurements and illumination between trial sites. Visual function was used as a surrogate biomarker for the therapeutic response to treatment and other functional tests together with retinal imaging parameters were used as secondary outcomes. The clinical trials mainly reported results at 2 years, though most continued to follow patients to provide longer-term data and information of time to reactivation was obtained from the IVAN study (personal communication). A database observational study of treat-and-extend patients in Australia reported that the risk of reactivation rose from 2.2% for a 6-week interval to 15.6% for 20 weeks.³

In this diagnostic test accuracy study, monitoring of patients continued as usual in eye hospitals. Ophthalmologists in the six centres continued their preferred drug and treatment regimen to monitor and treat nAMD in their patients. The study added weekly home monitoring, using three different tests (time 20–40 minutes), to the usual care pathway. Where possible, data from both eyes were collected.

The reference standard was the usual care clinical decision about the activity of nAMD in the study eye at the hospital outpatient appointment. Each participant remained in the study throughout the follow-up period. Therefore, there will be some monitoring visits when the study eye is judged by the participant's ophthalmologist to have active disease and some visits when the study eye is judged to have inactive disease. The primary quantitative analysis compared the results of the home-monitoring tests during the interval preceding the monitoring visit with the reference standard assessed at the monitoring visit.

Because of (1) the clinic workload in treating and monitoring nAMD patients and (2) the high cost of establishing a robust reference standard for people at high risk of nAMD but not currently being monitored by the NHS, we decided that the most urgent priority was to identify a home-monitoring test that could detect reactivation in the patients currently being managed in the NHS. We envisaged that, ideally, after diagnosis, patients would have their initial loading-dose injections in a hospital clinic and would then be discharged with the home-monitoring test; if the test subsequently indicated a deterioration in their vision, they would arrange an urgent appointment. The focus of NHS hospital nAMD clinics would then shift to providing urgent appointments to administer treatment, rather than regular monitoring.

Since the start of the study, several other studies have evaluated home monitoring in AMD; this has also been facilitated by the coronavirus disease 2019 (COVID-19) pandemic which necessitated a focus on home monitoring and forced an increase in digital literacy in the older age groups. A recent study which recruited patients with diabetic macular disease prior to the lockdown in October 2019 and after the lockdown in February 2020 provided patients with the Alleeye (Oculocare, Switzerland) app on their own smartphone or provided them with it preloaded on an apple iPod Touch™ (6th generation, Apple, Cupertino, CA, USA) reported that the smartphone home tests were able to indicate worsening pathology and the need for treatment.⁷

Test accuracy of tests for self-monitoring neovascular age-related macular degeneration activity

The advent of tablet computers and mobile/wireless technology has led to the development of devices for self-monitoring of visual function in nAMD.⁸ The disadvantages of the standard Amsler chart have long been recognised; its sensitivity to detect the onset of nAMD has been estimated to be only 50–70%.⁹ Perceptual completion¹⁰ and the inability of patients to understand the test or reliably report the results are thought to contribute to poor performance.

Reactivation of nAMD is more difficult to detect because some patients have distortion due to scarring and photoreceptor disorganisation in the absence of disease activity; therefore, a test has to enable patients to perceive an increase in distortion rather than solely its presence. Newer technologies such as visual- and memory-stimulating grids (VMS grid),¹¹ preferential hyperacuity perimetry home devices,^{12,13}

and shape discrimination tests¹⁴⁻¹⁶ have been reported to quantify distortion more accurately than either the Amsler grid or visual acuity in clinical settings.¹⁷

During the application and commissioning stage of this study, the literature was searched and all tests with longitudinal data (including usability data) in patients with nAMD were considered for inclusion. This study investigated the test accuracy of home-monitoring 'index' tests to detect reactivation. The tests chosen spanned a range of complexity and cost, one used paper and pencil and two used modern information technology, implemented as software applications (apps) on an iPod Touch device.

These were:

1. KeepSight Journal (KSJ) originally developed by Mark Rosner and collaborators at KeepSight adapted for UK use for this study (a paper-based booklet of near vision tests).
2. MyVisionTrack® (mVT®) electronic vision test, owned by Genentech Inc., a member of the Roche Group.
3. MultiBit (MBT) electronic vision test, developed by Visumetrics, licensed by Novartis.

The KeepSight Journal

The KSJ encouraged weekly monitoring using a paper journal. It included three different monitoring strategies, viewed one eye at a time. Firstly, near-visual acuity was assessed using a puzzle (crossword or word search) employing a variety of font sizes (an example is shown in [Figure 1](#)). Secondly, patients were encouraged to view objects with straight lines in the home to check for distortion (wall panelling, floor tiles, venetian blinds, etc.). Finally, they used a modified Amsler chart (VMS grid) to record areas of distortion or scotoma in their vision. The KSJ has been used before; 198 patients with intermediate AMD (at high risk of progression to late stage) were randomised to use the KSJ to self-monitor or usual care, with follow-up at 6 and 12 months to assess adherence.¹¹ The results showed significantly better adherence in the journal group with the findings supporting the efficacy of the journal for increasing vision self-monitoring adherence and confidence while promoting persistence in weekly monitoring.


1. As you play word scramble, check the circle next to each line you can read.

In the Kitchen

Left Eye		Right Eye
1 <input type="radio"/>	<u>milk</u> LIMK	WERB 1 <input type="radio"/>
2 <input type="radio"/>	GEG	AFN 2 <input type="radio"/>
3 <input type="radio"/>	CKOO	BFEE 3 <input type="radio"/>
4 <input type="radio"/>	VOEN	LOIB 4 <input type="radio"/>
5 <input type="radio"/>	LEPE	KFRO 5 <input type="radio"/>
6 <input type="radio"/>	LRIG	SOVE 6 <input type="radio"/>

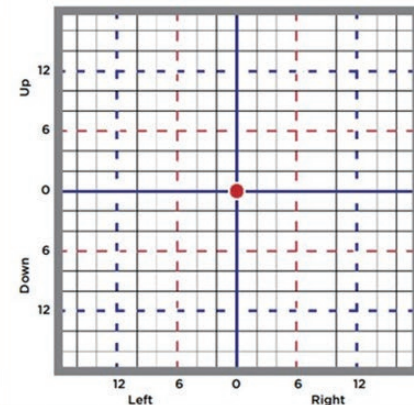
Compare to your acuity baseline test.
 • Left eye Better Same Worse
 • Right eye Better Same Worse

2. Look at objects in your home with straight lines.



Compare to your home object baseline test.
 • Left eye Better Same Worse
 • Right eye Better Same Worse

3. Focus on the center dot.
Look for wavy, blurry or spotty lines.



Compare to your grid baseline test.
 • Left eye Better Same Worse
 • Right eye Better Same Worse

FIGURE 1 Example week in the KSJ.

MyVisionTrack

MyVisionTrack is a software application (app) on an iPod Touch device. It is a shape discrimination test which measures hyperacuity, by displaying four circles, one of which is radially deformed ('bumpy' rather than perfectly circular). Viewing the display monocularly, the patient has to identify the odd one out (Figure 2). Studies have shown that the task implemented on an iPod Touch can distinguish between intermediate and advanced nAMD and a survey reported that 98% of patients found the test easy to use.¹⁸ Studies at the Liverpool site led by Paul Knox (co-investigator) successfully used the test in macular clinics and patients have found the test straightforward to complete.^{19,20}

MultiBit test

MultiBit test (MBT) is also an app on an iPod Touch. It is a near-acuity threshold test of neuroretinal damage. Traditional tests fail to detect such damage because they are suprathreshold. The MBT displays receptive field-sized dots or 'rarebits', which provide a miniscule amount of information to the visual system compared to conventional targets (Figure 3). Patients are presented with pairs of numbers; they state the numbers that they see out loud and the numbers are then represented at high contrast together with a recording of the patient's responses. MBT is the only test with published data describing its performance to alongside changes in nAMD activation.²¹ It was used to track 29 patients during treatment and monitoring in NHS outpatient clinics (average 39 weeks follow-up), with patients monitoring themselves at home with an iPod Touch.²¹ MBT performance improved gradually after treatment, stabilised during periods of disease inactivity and deteriorated gradually preceding reactivation. MBT performance also agreed well with retinal imaging clinical assessments but not with visual acuity (known to be an insensitive test of reactivation).

Reference standard

The reference standard (sometimes called the 'gold standard') is a test that classifies an observation in a way that is considered 'definitive'. In the case of monitoring for neovascular age-related macular degeneration reactivation at home (MONARCH), this was the nAMD status of a study eye being monitored at a clinical visit. The reference standard is sometimes imperfect but represents how diagnostic decisions are currently being made.

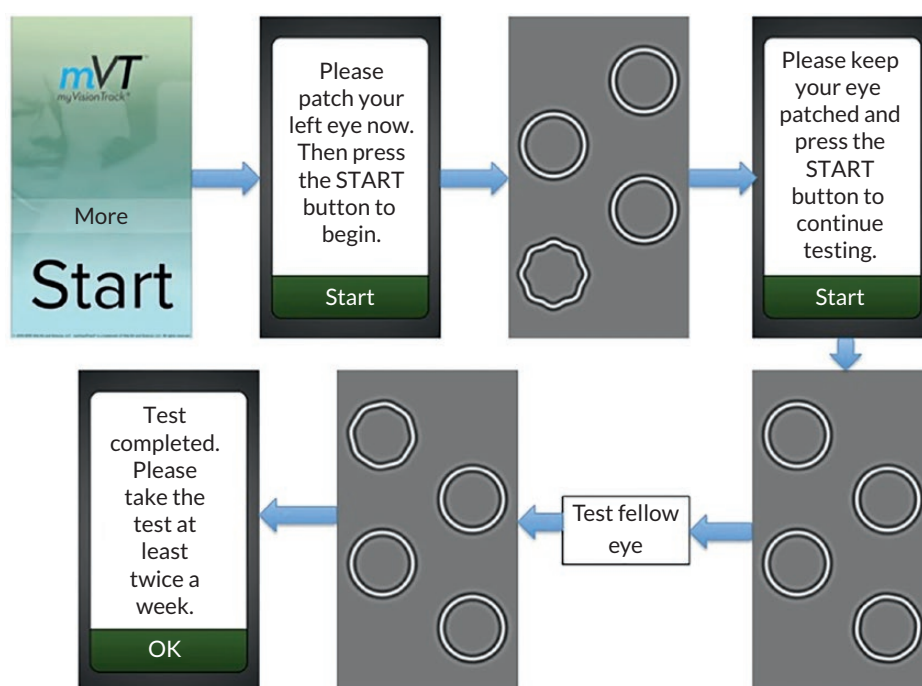


FIGURE 2 Diagram illustrating the steps when self-monitoring with the mVT app.

The reference standard test for the study was defined as the reviewing ophthalmologist's decision at a monitoring visit about the activity status of a study eye. This decision was made on the basis of clinical examination and the results of hospital-based retinal imaging investigations such as colour fundus (CF) photographs and OCT images. We recognise that it is possible that the reviewing ophthalmologist sometimes misjudges the status of a study eye at a monitoring visit, the judgements required are complex and can be difficult; even experts can disagree when judging the activity status of a nAMD lesion,²² but the decisions made by ophthalmologists currently represent the best reference standard.

Potential inequalities in uptake

The study also aimed to address the question: How do demographic, socioeconomic and visual function factors influence the uptake of home-monitoring tests for detecting active nAMD?

A survey by Age UK in 2013 found that internet use among people aged 65 years or over varied across the UK, with a 'north-south' divide; more than 50% in the south (Surrey, Bedfordshire, Buckinghamshire,

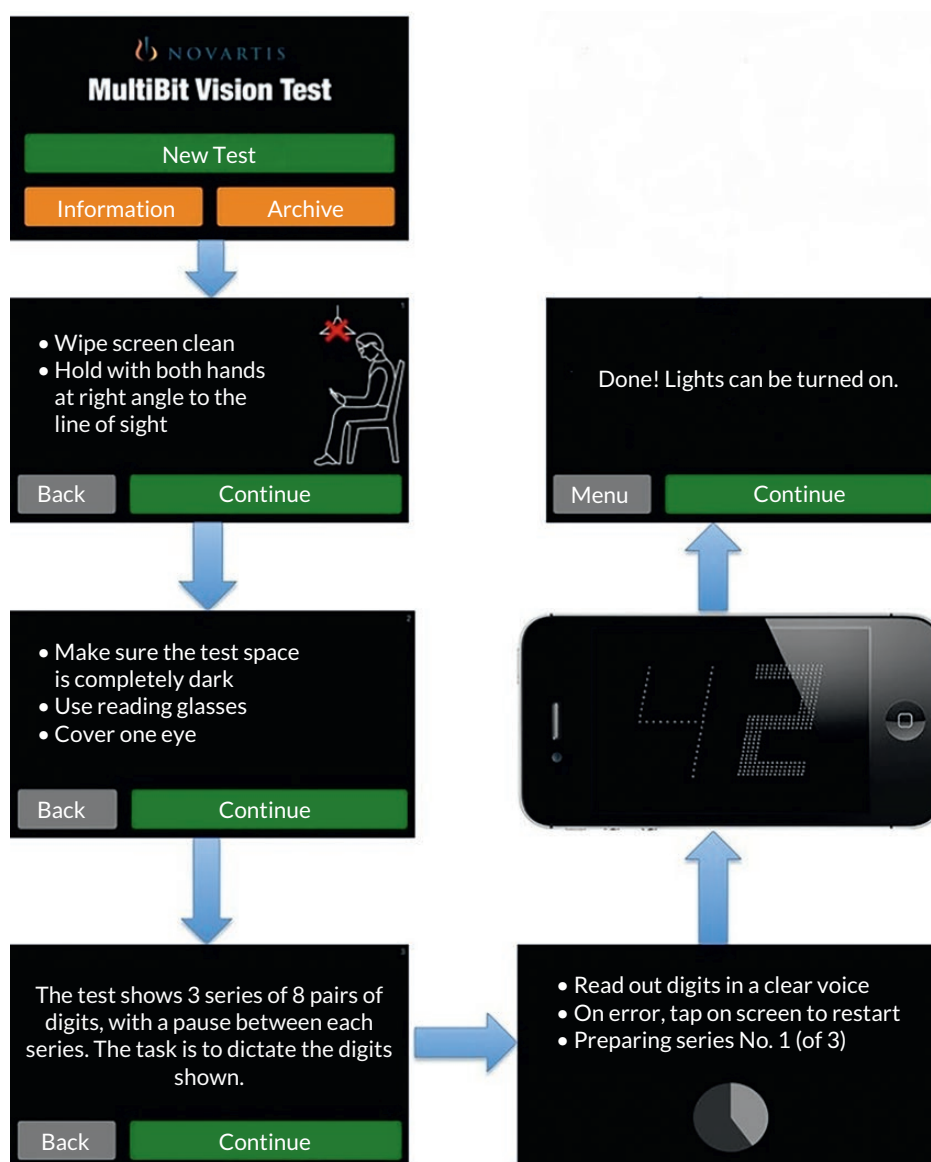


FIGURE 3 Diagram illustrating steps when self-monitoring with the MBT app.

Suffolk and Oxfordshire) used the internet, but less than a third in the north (Cumbria, Yorkshire, Hull, Tyne and Wear).²³ With respect to smartphone use, only 20% of 65- to 74-year-olds used such a device to access the internet in 2013;²⁴ perhaps more importantly, this percentage had increased from only 12% in 2012, suggesting that the situation is changing rapidly over time. The potential importance of failure to access the internet has been highlighted by a study of men and women in the English Longitudinal Study of Ageing from 2004 to 2011 internet use was found to be significantly 'protective against health literacy decline'.²⁵

During the study design phase, the small percentage of regular internet and smartphone users was considered a potential threat to the study; we were especially concerned that potential participants might feel alienated by the technology and would not be prepared to try out the solutions we proposed. Therefore, we sought to determine the extent to which the technology was a barrier to consent and participation in order to project wider adoption of home monitoring in the future if it is found to have satisfactory performance. Moreover, among participants, it is possible that some tests will be easier to do for participants with limited experience of smart devices and the internet. This was an important factor to weigh against test performance if differences in test performance were found to be small.

We designed the study to include the following features to try to minimise the extent to which technology could be a barrier to home monitoring:

1. We included a simple paper-based home-monitoring test, which we hoped would feel familiar to participants. This test involves a series of puzzles which require participants to use their near-vision correction.
2. We also provided a mobile broadband device so that participation was not limited by the lack of home Wi-Fi. The device had a simple on/off switch; the only things that a participant needed to remember to do were to keep the device charged (a main micro-USB charger was provided) and to switch on the device before performing the home-monitoring tests that use the iPod. The iPod interacted with the mobile broadband device automatically to transmit data.

We explained the use of the devices during an initial training and information session with each potential participant and also provided a help line for participants to call in the event of the experiencing difficulty.

Aims and objectives

The aim of the MONARCH study is to quantify the performance of three non-invasive test strategies for use by patients at home to detect active nAMD compared to diagnosis of active nAMD during usual monitoring of patients in the Hospital Eye Service (HES).

The study had five objectives:

- A. Estimate the test accuracy of three tests to self-monitor reactivation of nAMD compared to the reference standard of detection of reactivation during hospital follow-up with OCT imaging, clinical examination and Early Treatment Diabetic Retinopathy Study (EDTRS) visual acuity.
- B. Determine the acceptability of the tests to patients and carers and their adherence to home-monitoring testing regimens.
- C. Explore whether inequalities (by age, sex, socioeconomic status and visual acuity) exist in recruitment to the study and impact the ability of participants to do the tests during follow-up and the adherence of participants to weekly testing.
- D. Provide pilot data for the use of home monitoring to detect conversion to nAMD in the fellow eyes of patients with unilateral disease, compared to the reference standard of detection of conversion during hospital follow-up with EDTRS visual acuity and OCT imaging.

- E. Describe the challenges of implementing and evaluating home-monitoring technologies (see [Changes to trial design after commencement of the trial](#)).

The study population recruited for Objective B (the integrated qualitative study) differed from that required for Objectives A, C and D. Only selected sites, Belfast, Moorfields and James Paget, participated in procedures and data collection for the patient and carer aspects of Objective B. All participating NHS centres took part in the healthcare professional (HCP) aspect of Objective B.

Chapter 2 Methods

Parts of this chapter are reproduced with permission from Ward *et al.*²⁶

Study design

The MONARCH study was a multicentre diagnostic test accuracy cohort study to estimate the sensitivity and specificity of home-monitoring tests to detect active nAMD in patients previously diagnosed with nAMD and quiescent after treatment.

MONARCH was designed to compare the results of the home-monitoring tests being evaluated ('index tests', see [Index tests](#)) with the results of a reference standard (see [Clarification of reference outcome](#)). These comparisons allowed the accuracy of the index tests to be quantified with respect to the reference standard.

Participants were followed for at least 6 months, accruing on average six clinic attendances at which home monitoring and reference test results were compared.

The nature of active nAMD may change over time since diagnosis, if the disease progresses despite monitoring and treatment. Therefore, the study population was stratified by time since first treatment of nAMD in the first-treated study eye (see [Stratification of study population](#)). This design was to avoid the prolonged duration of follow-up which would be required if, instead, we were to follow participants from diagnosis to an equivalent time in the natural history of their condition. These design features are shown in [Figure 4](#).

All study data collections were planned before the first participant was recruited, although additions were made to the case report forms (CRFs) in the first month to clarify some data items, for example, eye classification. Case report forms are available at: <https://fundingawards.nihr.ac.uk/award/15/97/02>.

Changes to trial design after commencement of the trial

Visual acuity eligibility

When recruitment began, the exclusion criteria have been amended to specify that study eyes with vision equal to or worse than Snellen score 6/24, LogMar0.64 or 53 letters are ineligible for the study due to concern that participants with poorer visual acuity would 'perform' less well than other participants. This change was agreed with the funder and Study Steering Committee (SSC) with the proviso that the Study Management Group carefully monitor performance of the patients at the lower end of the visual acuity cut-off threshold and extend recruitment criteria accordingly if these patients do appear to cope with the home eye tests, in order to recruit as broad a patient population as possible. In February 2019, data reviewed by the SSC of the first 59 participants showed that participants with lower visual acuity performed no less well than participants with higher visual acuity. The SSC agreed with the recommendation to lower the visual acuity eligibility threshold to 6/60. This change was approved on 11 May 2019.

New objective

Due to the technical difficulties faced during this study, the SSC supported the addition of a new Objective E when producing the final report, namely, to describe the challenges of implementing and evaluating home-monitoring technologies. This additional objective has not been added to the protocol. The objective used data that were available from processes instituted during the study, for example, a

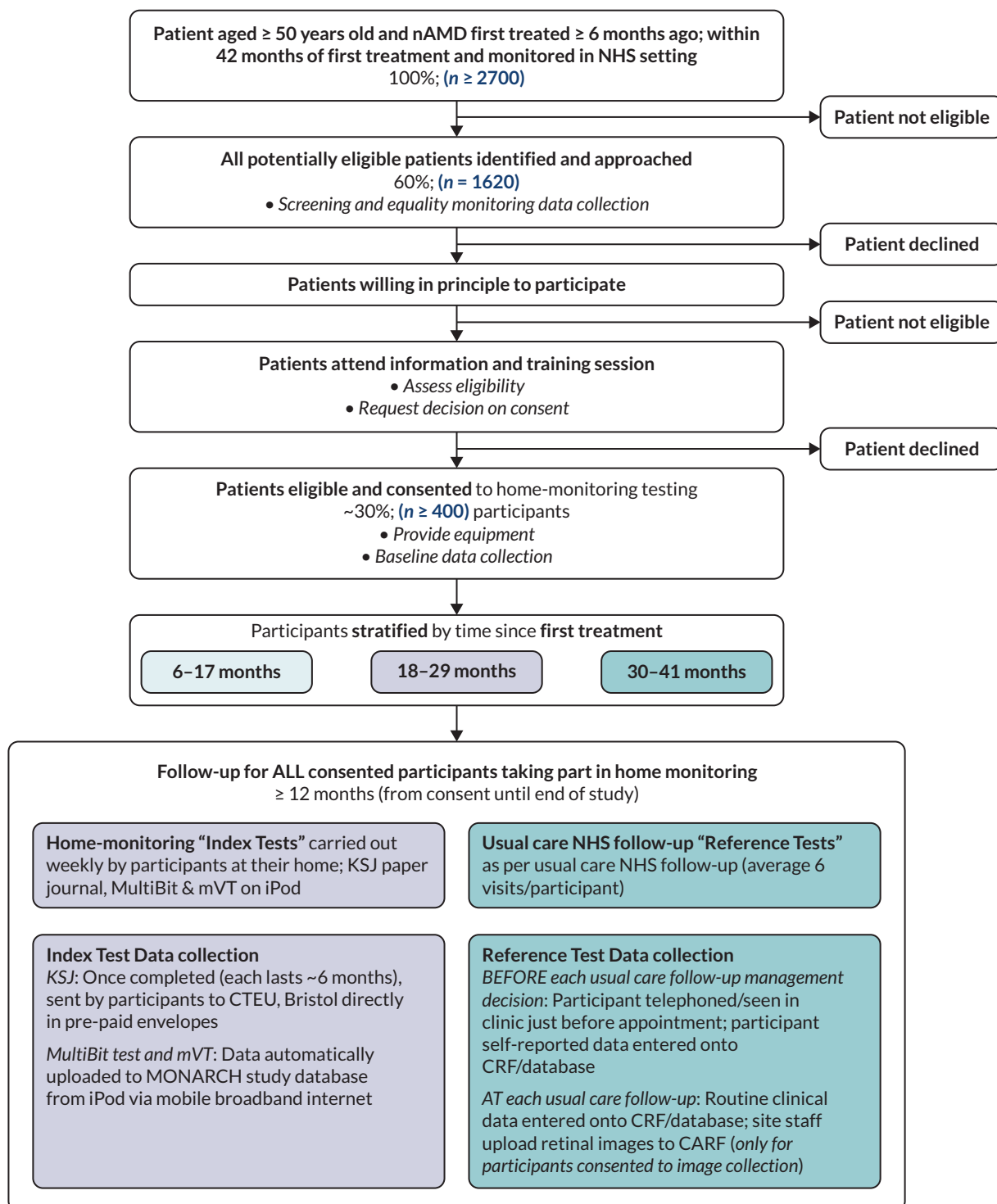


FIGURE 4 Study schema: Objectives A, C and D.

telephone helpline, and accrued other data during the conduct of the study. This objective describes both anticipated and unanticipated challenges of implementing and using the digital apps and the methods used to capture these challenges, mindful that they may be applicable in a variety of contexts in which home monitoring is being attempted beyond ophthalmology.

Clarification of reference outcome

As stated in the grant application, the reference standard is classification of eyes being monitored by the HES as active/inactive nAMD at a management visit. However, consideration of how home monitoring

might be implemented (e.g. patient with an inactive lesion discharged with test equipment, with an appointment subsequently triggered by test data indicating reactivation) led us to define a secondary outcome, change from inactive to active lesion status between adjacent management visits. Analyses for this outcome had low power, mainly due to lesions in study eyes being classified as active in a much higher proportion of management visits than expected.

Recruitment and follow-up periods

Following discussions with the SSC in June and September of 2019 regarding the slower-than-anticipated rate of recruitment, it was agreed that the recruitment period would be extended by 6 months (to March 2020), while retaining the original end of follow-up date (September 2020), effectively shortening the minimum follow-up period to 6 months. The purpose of this was to maximise recruitment and the number of monitoring visits without a funded extension. Two hundred and ninety-seven patients of the target 400 were recruited.

Frequency of active lesion status

Monitoring during the study found that the percentage of study eyes classified as active at clinics was much higher than anticipated (75% of study eyes at baseline and 60% of subsequent management visits, compared to the expected 30%).²⁶ This prompted concern about whether ophthalmologists' classification decisions at monitoring visits reflected usual care or were being modified due to the study. As a result, some of the underspend from research costs for sites due to lower than expected required was allocated to Central Administrative Research Facility (CARF) to grade OCTs with respect to activity status. The difference between observed and expected percentages was suspected to be due to clinicians treating eyes in such a way that a small amount of fluid in the retina was maintained because this is believed to protect against the development of macular atrophy; such eyes might be classified as having an active lesion. Grading of activity by CARF permitted sensitivity analyses for Objectives A and D.

Stratum recruitment

The protocol stated that study would aim to recruit equally to the three strata (6–17, 18–29 and 30–41 months since first treatment). This approach was relaxed once it became clear that after the pool of existing patients had been approached for the study, only new patients were being recruited as they became eligible. This led to a larger proportion of patients in the 6–17-month strata as recruitment to the 30–41- and 18–29-month strata became more difficult.

Participant eligibility

The study population for the quantitative part of the study were patients with at least one study eye being monitored by HES for nAMD, stratified by time since starting treatment in the first-treated study eye (6–17 months; 18–29 months; 30–41 months) to ensure test performance was estimated across this range of duration of nAMD.

Classification of eyes and eligibility of patients with bilateral neovascular age-related macular degeneration

A patient must have had at least one study eye to enter MONARCH.

A study eye was an eye with a nAMD diagnosis meeting the study eye eligibility criteria (below). Participants with bilateral nAMD may have had two study eyes if both eyes met the criteria for eligibility.

A fellow eye was an eye without a nAMD diagnosis, but which met all other eligibility criteria.

An excluded eye is an eye which is not eligible to be a study eye or a fellow eye.

It was possible at the time of consent for a participant to have had one study eye and one excluded eye and for the excluded eye to become eligible later during follow-up. An excluded eye of a participant

could become eligible if either: (1) the time since last treatment was < 6 months at screening but enough time had passed while the participant was in the study so that the time since last treatment became ≥ 6 months and/or (2) the participant had surgery in the excluded eye in the 6 months prior to screening but enough time had passed while the participant was in the study that time since surgery became ≥ 6 months. In such instances, the excluded eye of a study participant was rescreened and became a study eye if all eligibility criteria were met.

If a participant had two study eyes, stratification by time since first treatment was based on the date of first treatment in the eye first treated for nAMD (see [Stratification of study population](#)).

Participants were asked to complete the home-monitoring tests for study and fellow eyes.

Inclusion criteria

Participants could enter the study if they were ≥ 50 years old and had at least one study eye, that is, meeting the inclusion criteria.

All of the following had to apply for an eye to be considered for inclusion:

1. first treated for active nAMD ≥ 6 months ago
2. first treated for active nAMD not more than 42 months ago
3. currently being monitored for nAMD disease by the NHS.

Exclusion criteria

A patient could not enter the study if the patient did not have a study eye.

A potential study eye was excluded if ANY of the following applied:

1. vision in the potential study eye was worse than Snellen score 6/60, LogMar 1.04 or 33 letters – comment here on the change in visual acuity (VA) eligibility over time
2. vision in the potential study eye was limited by a condition other than nAMD
3. surgery in the potential study eye in the previous 6 months
4. refractive error in the potential study eye $> -6D$
5. retinal or choroidal neovascularisation in the potential study eye not due to nAMD
6. in addition, a participant will be excluded if ANY of the following apply
7. inability to do one or more of the proposed tests as assessed during ‘further information and training’ session (see [Training and equipment](#))
8. unable to understand English
9. home or personal circumstances are unsuitable for home testing.

Stratification of study population

Home monitoring to detect active nAMD is relevant at any stage of the condition after diagnosis, apart from an initial phase of treatment (usually 3 months). To recruit a study population that evaluates home monitoring across the time spectrum of monitoring, recruitment was stratified into three strata according to time since first treatment for nAMD in the study eye (time since the first-treated study eye for participants with two study eyes): (1) 6–17 months; (2) 18–29 months; (3) 30–41 months.

Settings

The study was run in the homes of patients being monitored by HES for nAMD at participating NHS hospitals and in the participating NHS hospitals.

Participants were recruited in secondary care (HES clinics). During the study, the reference standard for an eye being monitored was determined at HES clinic visits. During intervals between clinical visits, participants used home-monitoring tests to test their vision themselves. Participants were asked to complete the home-monitoring tests themselves weekly at home.

Index tests

There were three home-monitoring ('index') tests spanning low to moderate cost and complexity. These were:

1. KSJ test: a paper-based booklet of near-vision tests presented as word puzzles developed in the USA by KeepSight and adapted by the study team for use in the study (UK wording and defining test scores for estimating diagnostic test accuracy).
2. mVT: electronic vision test, owned by Genentech Inc., a member of the Roche Group, intended to be viewed on a tablet device.
3. MultiBit test (MBT): electronic vision test, developed by Visumetrics, licensed by Novartis, intended to be viewed on a tablet device.

Each electronic test took approximately 5 minutes (2–3 minutes per eye) to complete, though some participants took longer. The KSJ took between 10 and 20 minutes (5–10 minutes per eye) depending on the difficulty of the puzzle. Therefore, weekly monitoring was expected to take 20–40 minutes (10–20 minutes per eye). Participants were asked to try to complete all three of the index tests for both eyes but could opt to stop up to two of the three tests and continue with the study.

Each KSJ booklet lasted approximately 6–7 months. Every 6 months, participants were sent a new booklet and instructed to return the used booklet in a pre-paid envelope to the study management team at Clinical Trials and Evaluation Unit (CTEU) Bristol.

Data were collected automatically from the electronic tests.

The KeepSight Journal

The KSJ encourages weekly monitoring using a paper booklet covering 6 months. It includes three different monitoring strategies, viewed one eye at a time. Firstly, near-visual acuity is assessed formally or in a puzzle (crossword, anagrams or word search) employing a variety of font sizes. Secondly, patients are encouraged to view objects with straight lines in the home to check for distortion (wall panelling, floor tiles, venetian blinds, etc.). Finally, they use a modified Amsler chart (VMS grid) to record areas of distortion or scotoma in their vision (see [Figure 1](#)). Different booklets were prepared for each 6-month follow-up period (available on request). New booklets were posted to participants towards the end of a period, with a reply-paid envelope to return the completed one.

MyVisionTrack

MyVisionTrack is a software application (app) viewed on Apple devices. It is a shape discrimination test which measures hyperacuity, by displaying four circles, one of which is radially deformed ('bumpy' rather than perfectly circular). Viewing the display monocularly, the patient has to identify the odd one out (see [Figure 2](#)).

MultiBit

The MBT is also an app viewed on an iPod Touch. It is a near-acuity threshold test of neuroretinal damage. Traditional tests fail to detect such damage because they are suprathreshold. The MBT displays receptive field-sized dots or 'rarebits', which provide a miniscule amount of information to the visual system compared to conventional targets (see [Figure 3](#)). Patients are presented with pairs of numbers;

they state the numbers that they see out loud and the numbers are then represented at high contrast together with a recording of the patient's responses. The patient's stated responses are recorded by the app and, after testing, must be played back and scored by the patient, that is, a patient scores their own performance.

Reference standard

The reference standard represents the classification that the index tests are intended to 'diagnose'. The reference standard is sometimes imperfect but represents how diagnostic decisions are currently being made. The rationale of home-monitoring tests is that a test can 'diagnose' the status classification that will be made at the next usual care monitoring clinic review. Therefore, in MONARCH, this reference standard was the reviewing ophthalmologist's decision at a monitoring visit about the activity status of a study eye; response options were inactive, active or uncertain. These decisions were made on the basis of clinical examination and the results of hospital-based retinal imaging investigations such as CF photographs and OCT images. It is possible that the reviewing ophthalmologist sometimes misjudged the status of a study eye at a monitoring visit [the judgements required are complex and can be difficult; even experts can disagree when judging the activity status of a nAMD lesion (20)], but the decisions made by ophthalmologists currently represent the best reference standard. The reference standard grouped uncertain judgements with inactive ones for all analyses.

Sample size

Objective A and C

In MONARCH, each monitoring visit (and the immediately preceding period of home monitoring) represented the unit of observation. We wanted to recruit enough participants to accrue sufficient visits to allow the study to have 90% (80%) power to detect a difference of 0.06 (0.05) in the area under the receiver operating characteristic (AUROC) curves for two tests if the AUROC is 0.75.²⁷

We had to make several assumptions to estimate the sample size required:

1. The reference standard would be 'active' for 30% of monitoring visits.
2. Correlations between the index test and reference standard classifications (unknown) would be 0.6 for both active and inactive lesions.²⁷
3. Participants would on average have index test and reference standard data for six monitoring visits, based on durations of follow-up between 12 and 30 months and information about treatment regimens at participating sites.
4. Each monitoring visit (and the immediately preceding period of home monitoring) represents the unit of observation; due to nesting of monitoring visits within participants and the consequent lack of independence of observation, we assumed that the effective average number of visits per participant with data would be 4 rather than 6.
5. The measurement errors of index tests were unknown, and we simply inflated the calculated sample size by 30% to make some allowance for dilution of the power to discriminate differences in performance between tests.
6. The ratio of the standard deviation (SD) of the negative group to the positive group for a home-monitoring test would be 1.
7. Attrition of participants during follow-up (e.g. due to mortality, changes in participants' circumstances or participants' unwillingness to continue participating) $\leq 5\%$.

To achieve the desired power (above), we estimated that we should recruit ≥ 400 participants. With, on average, 6 visits per participant (about 2400 visits with data for at least one index test and the reference standard, allowing for 5% attrition), we estimated that these parameters would yield an effective sample size of about 1200 visits after allowing for (4) and (5) above.

Objective D

Estimates of the rate of conversion to nAMD among fellow eyes vary, ranging from 4% to 16%.^{9,28,29} Assuming the risk in unselected patients is 5–6% per year, up to 50 of the 400 target patients were expected to have nAMD in both eyes at the time of recruitment. Among the remaining 350 patients, it was expected to identify conversion of fellow eyes to nAMD in about 25–30 patients. We recognised that test accuracy estimates based on such a small number of true positives would inevitably be imprecise but nevertheless useful as pilot data for future research.

Outcomes

Primary outcome

The primary outcome was classification of a study eye at a monitoring visit as having active or inactive disease, that is, the reviewing ophthalmologist's decision at a monitoring visit about the activity status of the study eye (active, inactive, uncertain). Index tests generated a scaled or categorical score on each occasion when they were used for testing, often many times between monitoring visits. Summary scores were derived for each test for each interval preceding a monitoring visit (see [Objective A: analytic approach and methods](#)).

Secondary outcomes

A secondary outcome for Objective A was defined, namely change in classification of lesion activity from inactive at the previous management visit to active as assessed at a management visit.

For Objective C, the following outcomes were investigated as measures of uptake of home-monitoring tests:

- impacts of inequalities on willingness in principle to participate;
- impact of inequalities on ability and adherence to weekly testing.

For Objective D, the test accuracy of the three tests to self-monitor conversion of fellow eyes to nAMD compared to reference standard was investigated.

For Objective E, the following outcomes were described as measures of the technical and logistical challenges identified during the study:

- frequency and reason for incoming calls made to the helpline and outgoing calls made to participants;
- frequency and duration of events leading to the digital tests being unavailable for testing;
- other technical and logistical challenges.

Data collection

Patient identification

Potential study participants were identified by local teams from established clinical databases of patients and by reviewing lists for outpatient clinics. Potential participants were screened for eligibility by the healthcare team through review of their medical notes and any existing imaging.

Potential participants were sent by post or given an invitation letter and patient information leaflet (PIL) describing the study. An appropriately trained and qualified member of the local research team (e.g. study clinician/research nurse/optometrist) discussed the study with them by telephone or in person. The potential participant will have had time to read the PIL and to discuss their participation with others outside the research team (e.g. relatives or friends). All potential participants who were provided a PIL were given a unique study number against which details, including reason(s) for non-participation (e.g.

reason for being ineligible or patient refusal) along with equality monitoring data [age, sex, ethnicity, Index of Multiple Deprivation (IMD), most recent visual acuities for both eyes] were collected.

Training and equipment

Verbal consent to attend the further information and training session was taken by a member of the local research team and recorded in the patient's hospital record. The information and training session was led by an appropriately qualified member of the local research team with an experience of working with patients. At the information and training session, the potential participant was shown the equipment and how it should be used for the study. The local research team member answered any further questions, checked and confirmed the participant's eligibility and took written informed consent if the potential participant was eligible and agreed to participate.

Participants were also asked for consent to their address and telephone number being held at the study coordination centre for receipt of new KSJs and pre-paid envelopes and participant newsletters, to receive optional SMS text reminders during the study, and an optional copy of the final study results at the end of the study. Following consent, the participant was provided with the following to take home: an iPod Touch, a lens cloth, an eye patch, stylus pen, the KSJ and a mobile broadband router. Eye classification (study eye, fellow eye, excluded eye) was confirmed, visual acuities were updated if necessary and information about the participant's use, 'at least weekly', of widespread technologies (e.g. electronic devices, e-mail, internet at home) was collected. The local study team sent a letter to the participant's general practitioner (GP) to inform them of study participation.

Follow-up

Patients were followed up for at least 6 months. There was no specific follow-up schedule required for the study. Participants continued to have usual care (i.e. review of disease activity and treatment if required) in NHS monitoring clinics. Retinal imaging was also be carried out as required to inform usual care management decisions. Local site teams collected data for study and fellow eyes at each usual care follow-up visit.

Imaging, review and treatment may have all be completed on the same day or depending on the usual care arrangements at a site, may have been carried out over a number of days. Retinal imaging could be reviewed (i.e. a management decision made) in the absence of a patient, with the patient being advised that they only needed to return if the nAMD was active and treatment was required. A management decision was a decision about the status of a nAMD lesion and/or the treatment plan.

Participants were contacted before the management decision for each follow-up visit was made. Participants were telephoned before (maximum of 5 working days before retinal imaging) or seen in clinic before having an appointment. A member of the local research team asked the participant questions on how they felt their vision has been since their last visit, whether the participant had been carrying out home monitoring, whether the participant had experienced any problem with home monitoring, to confirm the participant's willingness to continue and whether there was need for retraining.

Technical support

The study management team operated a technical support phone line during standard office hours Monday to Friday (except holidays) to resolve issues with use of the applications, iPod touches, mobile broadband devices, automatic data upload or other technical queries. Calls received to the support line are documented and reasons reviewed. The study management team also made periodic calls to patients for whom data had not been received in > 21 days (or > 14 since consent). Reasons for the absence of test data are documented.

Retinal images

Patients were asked to provide consent to retinal image collection (optional). For participants who agreed to retinal image collection, all retinal images (e.g. CF, OCT) taken during follow-up were submitted to CARF to be stored for use in future ethically approved research.

Given the higher-than-expected proportion of active lesions, we decided that the stored retinal images should be assessed for presence of activity by senior staff from the NetWOCR UK reading centre (www.networcuk.com/) according to a standardised protocol. All available images were used to make the assessment. Presence of the following features was assessed in both eyes:

- Presence of subretinal fluid.
- If present, was it more or less than previous visit?
- Central subfield thickness ([Figure 5](#)) is recorded from the OCT machine automatic algorithm.

Features to minimise bias

Risk of bias was considered with respect to bias domains described in an appraisal tool for diagnostic accuracy studies.³⁰ Information here should be read in conjunction with information about the proposed methods of analyses (see [Methods for Objective E](#)).

Blinding

All personnel carrying out usual care NHS monitoring were masked to data from index tests. However, participants may have gained an impression that their vision had changed from home-monitoring tests and may have communicated their impressions to their ophthalmologists.

Bias due to selection of participants

Bias in this domain was avoided by using a cohort study design and recruiting a representative sample of eligible patients. We cannot guarantee that consecutive eligible patients were recruited but factors such as absence of research staff (e.g. annual leave) or other logistical issues were very unlikely to be associated with the characteristics of patients. Therefore, we anticipated that, when staff were available, research teams would invite consecutive eligible patients to take part and hence recruit a representative sample of patients. The exclusion criteria were appropriate, that is, they would prevent a person self-monitoring using one or more of the tests even if the test(s) were implemented as part of usual care (if shown to detect nAMD reactivation accurately).

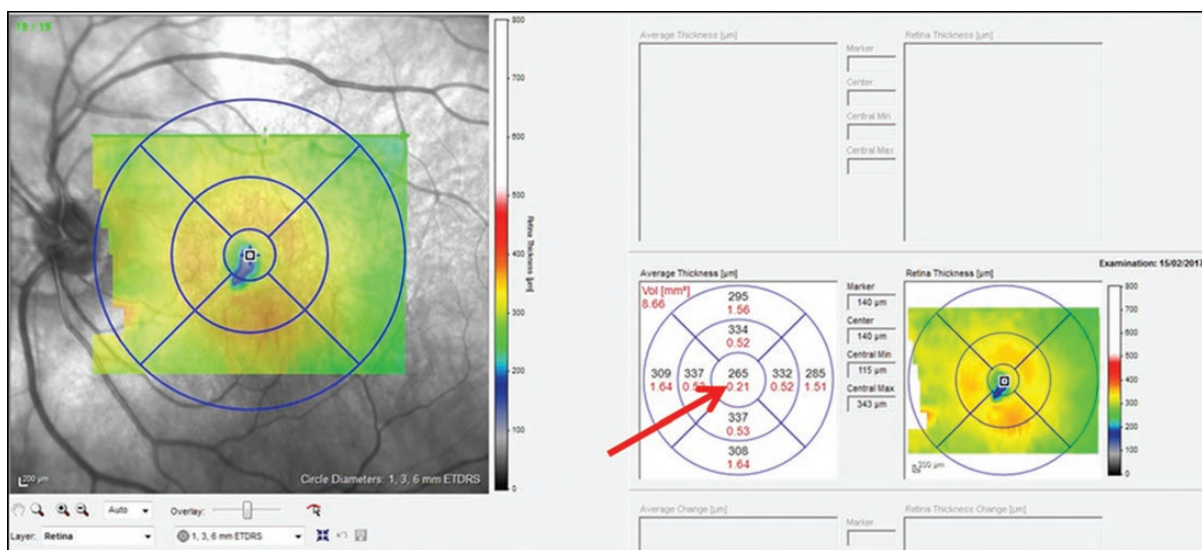


FIGURE 5 Red arrow shows central subfield thickness from OCT scan.

Bias in the assessment of the index tests

Bias in this domain was avoided by ensuring that the index tests were to be 'scored' without knowledge of the results of the reference standard. (Index tests were performed before monitoring visits; so, they could not have been influenced by the reference standard classification.) We chose methods to summarise test scores before carrying out any analyses and estimated the AUROCs based on these, across the range of summary test scores. We did not pre-specify threshold scores for classifying test summary results as active/inactive and report sensitivity and specificity based on Youden's index, acknowledging the latter takes no account of the relative clinical and other implications of false-positive and false-negative misclassifications. We could not specify test thresholds at the outset because there were no available data to inform these definitions.

Bias in the assessment of the reference standard

Bias in this domain was avoided by ensuring that the reference standard was assessed without knowledge of the results of the index tests. The reference standard represents a usual care decision about the reactivation of nAMD and, although these decisions will not always have been accurate,²² we assumed that these decisions could reasonably be considered likely to classify participants correctly with respect to reactivation of nAMD.

Bias due to exclusion of participants or inappropriate intervals between the times of index testing and the reference standard

Bias in this domain was avoided by ensuring the analysis included all follow-up visits for which the reference standard was assessed and by carefully describing the time intervals between index tests and the reference standard. We also accounted for all patients recruited into the study, for example, using a flow diagram and tables as appropriate. We acknowledge potential differences between participating centres and present information that may characterise this, for example, variation in methods used to obtain the reference standard and the centre-specific rate of reactivation of nAMD.

Qualitative methods (Objective B)

Semistructured interviews were conducted face to face and via telephone. The interview schedule (see [Appendix 1, Table 29](#)) was based on the experience of the research team and was informed by models and theories of technology acceptance. The members of the team that collected and analysed the data have extensive experience in the application of qualitative methods in healthcare research. The study followed the consolidated criteria for reporting qualitative research (COREQ) criteria.³¹

Participant recruitment and data collection

Recruitment to the qualitative component began 3 months after the MONARCH study began recruiting. During the consent process for participation in the diagnostic accuracy study, individuals who consented to further contact to discuss participation in this qualitative study were approached via a telephone call from a qualitative researcher. Researchers confirmed if participants were happy to take part and arranged an interview date. Informed consent was obtained prior to interviews and following an explanation of study procedures. Maximum variation sampling was used to ensure that a range of perspectives were captured in relation to age category (young-old 50–69 years and older-old 70+ years), gender, laterality of nAMD (unilateral and bilateral) and time since first treatment (6–17, 18–29 and 30–41 months). In addition, test usage data were used to sample participants with different levels of adherence to home monitoring. Usage data from the two digital index tests were examined by the qualitative researchers to categorise participants into two groups: (1) 'Regular' testers who completed weekly home monitoring without two or more gaps in testing of greater than 3 weeks, and (2) 'Irregular' testers who stopped and started testing on more than two occasions or stopped testing completely. Patients who declined to participate in MONARCH but provided consent to be contacted about the qualitative study were also approached. We approached informal 'carers', supporters or significant others in the lives of patients, and HCPs

who interacted with participants at study sites visits in order to gather their perspectives about the acceptability of home monitoring.

Data management and analysis

All interviews were audio-recorded and transcribed. A directed content analysis approach based on deductive and inductive coding was used.³² An initial coding scheme was developed based on a synthesis of relevant theories, including the unified theory of acceptance and use of technology (UTAUT),³³ the technology acceptance model (TAM)³⁴ and the theoretical framework of acceptability (TFA).³⁵ The coding framework underwent iterative development as individual transcripts were reviewed and rereviewed during data familiarisation. Following line-by-line coding of each transcript, findings were summarised based on the coding scheme and these summaries were used to revise initial codes if necessary and develop new codes based on emerging data (see [Appendix 1, Table 30](#) for final coding scheme). A third researcher coded a random sample of 10% of the transcripts and subsequently discussed and compared coding with CT (Charlene Treanor) and SOC (Sean O'Connor) in order to ensure adequate rigour and reflexivity. Related codes were then clustered and grouped into initial themes. Narrative summaries were written for each theme, then reviewed and discussed to refine main themes and subthemes to ensure coherence. NVivo version 12 was used to manage data and facilitate the analysis process, which, in summary, included the following stages: (1) independent transcription, (2) data familiarisation, (3) independent coding, (4) development of an analytical framework, (5) indexing, (6) charting and (7) interpreting data.

Methods for Objective E

Due to the age range of the patient population, the study team anticipated that participants with little or no experience in using electronic devices may have difficulty operating and interacting with the study equipment [iPod and Mobile Wireless-Fidelity (MiFi) router]. As described, participants' exposure to various types of electronic technology was recorded at consent. Reasons for not consenting were also recorded when participants did not want to engage with the technologies required for the study.

Throughout the study, at or before each follow-up appointment, the local research staff asked participants if they were experiencing any issues with the equipment and additional training sessions were offered. This information was captured in the CRFs and was used as a means of troubleshooting participants technical issues and maintaining participant testing.

Participants were encouraged to contact the study team on the technical support to resolve technical issues (incoming calls). Details of all calls received were logged, including call length, issues discussed, actions taken and resolution status. The helpline acted not only to capture the details and frequency of various technical issues, but also to resolve them and maintain participant testing.

The study team realised early on during the study that data were not being received from some participants, often immediately following consent. To try to identify and resolve any technical issues that participants may be facing, periodic calls were made from the helpline to participants for whom data had not been received in more than 3 weeks or more than 2 weeks since consent (outgoing calls, only to participants who had consented to being contacted by telephone). Logging these calls documented the challenges participants were experiencing, and also meant that members of the research team could try to resolve any technical issues, or prompt participants to test. The same details were recorded as for incoming calls.

In addition to periodic calls to participants, a SMS text message notified a participant if test data had not been received in more than 3 weeks since last test or more than 2 weeks since consent, or to thank them if their data were being regularly received (when consent to send these messages had been given). These messages were intended to act as a prompt to participants to test or to contact the study team if they were experiencing technical issues. All notifications sent to participants were logged.

The study team experienced a range of different logistical and technical issues relating directly to the study equipment and applications used. All issues with the study equipment and the apps were logged with as much detail as possible and are described in the results, including time frames, impact on participants and resolutions, where possible.

Statistical methods

A pre-specified analysis plan was written before analyses were carried out.

Objective A: analytic approach and methods

The test accuracy of all three home-monitoring tests was examined using three outcomes:

- **Primary analysis:** the reference standard of the classification of a study eye at a complete management visit as having either active or inactive disease, using all available test data.
- **Sensitivity analysis 1:** the reference standard of the classification of a study eye at a complete management visit as having either active or inactive disease, using test data for the 4 weeks preceding the management visit.
- **Sensitivity analysis 2:** the alternative reference standard of the classification of a study eye as having either active or inactive disease based on the OCT image taken during a complete management visit, using all available test data.
- **Secondary outcome:** the reference standard of a change in classification of a study eye from inactive at the previous complete management visit (including baseline assessment) to active at the subsequent complete management visit, using all available test data.

Two test summary scores (predictors) for each test were calculated per management visit per patient. The first was an average (mean, proportion or median) of the test raw scores across the entire preceding intervisit follow-up period. The second was calculated in the same way, but only using raw scores during the 4-week period immediately preceding the management decision, since the periods between visits were not of fixed length (affecting the precision of the summary score and containing older test information that might be out of date).

Models used all available data. The reference standard was always available for a complete management visit, but summary tests scores were not because participants did not always do all index tests.

Means of MBT and mVT test scores were fitted in all models. The KSJ generated four test raw scores: ordinal near VA (on a scale of 1–6; added before carrying out analyses to make use of the scaled nature of these data); VA worse than baseline; Amsler grid worse than baseline; household object appearance worse than baseline. The latter three binary scores were created by combining 'same' or 'better' responses (score = 0). KSJ summary test scores for a complete visit were created as a median for near VA and percentages of the three other scores that were worse across the respective preceding interval. All four scores were fitted in the KSJ model and a single AUROC was estimated. Correlations between the four KSJ summary tests scores were estimated.

Separate models were fitted for each test/summary test score for the primary outcome, the two sensitivity analyses and the secondary outcomes. All models included the fixed effects of sex, stratum for time since first treatment for nAMD at baseline, age, visual acuity stratum at baseline and days since baseline management visit. Each model was fitted, where possible, with a random intercept and random slope on calendar quarter since baseline visit at the participant level, and a random intercept at the eye level.

Higher MBT scores represent better performance (requiring less information to identify numbers correctly). Higher mVT scores (on the logMAR scale) represent poorer performance (more pronounced

deviations required to identify the bumpy circle correctly). Test scores from the mVT application were multiplied by a factor of 100 before fitting the models, so that odds ratio (OR) estimates for this test correspond to changes in ORs for every 0.01 change in mVT score. For the KSJ, a higher median represented better near VA; higher percentages of worse raw scores for the other outcomes represented poorer performance. Unless otherwise stated, the fixed effect of days since baseline was estimated after applying a natural log transformation due to non-linearity and improvements in model specification. No interactions were tested.

Model performance in all cases was examined by inspecting the OR for the index test summary score(s) and the estimate of the AUROC and their respective confidence intervals (CIs). AUROCs were based on predicted probabilities calculated using only the fixed effects in the models. Finally, home-monitoring score thresholds associated with predicting active lesion status were calculated by identifying the point on the AUROC curve that maximised sensitivity and specificity based on Youden's index. The predicted probability associated with the cut-off point was then applied to the data to estimate average test scores above and below the threshold.

The overall performance of the tests was quantified by the AUROC. We intended to compare AUROCs for the tests to determine if one or more tests is superior to one or more of the others but did not do this due to the emerging results. Sensitivity, specificity, positive and negative predictive values of each test were also reported with 95% CIs. Analyses took account of the structure within the data, that is, the nesting of visits (and eyes) within patients.

Objective B

Qualitative research methods were used to explore individual responses, views and experiences around Hand Movements (HM) acceptability, as well as to examine variations in personal contexts.³² Semistructured interviews were conducted face to face and via telephone. Participants were not known to the researchers who conducted the interviews. The interview schedule (see [Appendix 1, Table 29](#)) was based on the experience of the research team and was informed by models and theories of technology acceptance. Several theories have been developed to improve understanding about technology acceptance, including the original and extended versions of the UTAUT,³³ the TAM,³⁴ the TFA³⁵ and the senior technology acceptance model (STAM).³⁶ The study followed the COREQ criteria.³¹

Recruitment to the qualitative component began 3 months after the MONARCH study. During the consent process for participation in the diagnostic accuracy study, individuals who consented to further contact to discuss participation in this qualitative study were approached via a telephone call from a qualitative researcher. Researchers confirmed whether participants were happy to take part and arranged an interview date. Informed consent was obtained prior to interviews and following an explanation of study procedures. Maximum variation sampling was used to ensure that a range of perspectives were captured in relation to age category (young-old 50–69 years and older-old 70+ years), gender, laterality of nAMD (unilateral and bilateral) and time since first treatment (6–17, 18–29 and 30–41 months). In addition, test usage data were used to sample participants with different levels of adherence to HM. Usage data from the two digital index tests were examined by the qualitative researchers to categorise participants into two groups: (1) 'Regular' testers who completed weekly HM without two or more gaps in testing of greater than 3 weeks, and (2) 'Irregular' testers who stopped and started testing on more than two occasions or stopped testing completely. Patients who declined to participate in MONARCH but provided consent to be contacted about the qualitative study were also approached. We approached informal 'carers', supporters or significant others in the lives of patients, and HCPs who interacted with participants at study sites visits in order to gather their perspectives about the acceptability of HM.

Objective C

Three outcomes were considered for this objective: (1) willingness in principle to participate in principle among approached patients, (2) consent to participate when eligible and willing in principle and

(3) ability to perform tests and adherence to weekly testing. Ability to perform an index test was defined as the proportion of qualifying management visits for which some valid index test data are available. Adherence was defined as the proportion of weeks between qualifying management visits for which some valid data for an index test were available. The ability and adherence models were performed for each test separately at the level of the patient.

Regression models explored the influences of age, sex, IMD, stratum of time since first diagnosis and baseline visual acuity at diagnosis on the outcomes of: willingness to participate in screening (when first approached), consent to take part (among all patients approached); ability of a participant to complete a test, analysed separately for each index test (among all participants); and adherence to the study protocol (among all participants). The influence of these factors is reported as ORs with 95% CIs. Analyses of adherence and ability took account of nesting of visits within participants.

The IMD was used as an indicator of participant socioeconomic status. However, IMD ranks cannot be directly compared between UK countries. To allow comparison of Belfast Northern Ireland with English IMD ranks, adjusted IMD ranks were used, by normalising 2010 Northern Ireland IMD data to the 2015 English IMD. The approach required back-converting available (English) IMD ranks on the MONARCH database to Lower Layer Super Output Areas geographies, allowing linkage to an adjusted IMD data source (<https://data.bris.ac.uk/data/dataset/1ef3q32gybk001v77c1ifmty7x>). During this process, seven patients were identified as having erroneous IMD ranks: four due to residing in the Isle of Man and three due to having out of range IMD ranks. These patients were excluded from analyses using IMD.

The exposure to technology indicator was based on whether the participant stated that they used at least weekly any of smartphone, tablet, laptop/home computer, internet, e-mail or social media. As these questions were only asked after consent, the indicator could not be examined in the analysis of inequalities on willingness in principle to participate.

Objective D

The test accuracy of the index tests was examined for the reference standard of an ophthalmologist's classification of a fellow eye having active disease at a management visit, that is, conversion to active nAMD. Two sensitivity analyses were carried out:

- **Sensitivity analysis 1:** the same reference standard outcome but using test data for the 4 weeks preceding the management visit only.
- **Sensitivity analysis 2:** the alternative reference standard of classification of a fellow eye having active disease at a management visit based on grading by CARF of OCT images taken during the management visits.

In other respects, the analyses used the same methods as for Objective A.

Objective D

This objective used descriptive summary descriptive statistics only.

Patient and public involvement

From the beginning we recognised that effective patient and public involvement (PPI) was central to the delivery of high-quality research. Discussions with patients, clinical and academic colleagues informed the initial proposal and, in particular, the frustration articulated by patients about the burden of monthly monitoring hospital visits, both for them and their carers. This was the stimulus for our focus on home monitoring for patients already being treated for nAMD. During the design phase, both the Royal National Institute for the Blind (RNIB) and the Macular Society were contacted about the proposal and provided helpful feedback. One of the tests was also demonstrated to, and discussed with, a group of

current AMD patients. They welcomed the development of technologies for home monitoring and said that they would be interested in using the test at home to monitor their vision. Feedback from a larger group of patients who had used the test as part of a formal study was also sought, and they confirmed that the test was easy to use, even for older patients with little or no experience of using 'smart' devices like the iPod Touch. Throughout the study, we sought to use the INVOLVE principles and resources to guide our active involvement with patients and the public. A representative from both the RNIB and the Macular Society agreed to sit on the SSC as these are both key stakeholders for patients with AMD and severe sight loss.

A group consisting of both four patients and two carers was also established [managed by Dr Hogg and a post doctoral research associate (PDRA) based in Belfast] provided crucial feedback on appropriateness of the study materials, with particular focus on the design of the training materials for using the apps and the study equipment. They also provided feedback on content of visits and frequency of procedures. Input from the PPI group was very helpful for the content of the KSJ, firstly in adapting the original one for a UK audience and then for the development of the subsequent versions.

Specific feedback that proved very useful included:

- Removal of the motivational quotes from the KSJ as the PPI group member did not like them and felt they were not appropriate for a UK audience.
- Requested removal of italicised text in the equipment instruction sheet.
- Members did not like the term 'carer' as they felt very few patients would describe their spouse or friend/family member who helps them as a carer and requested the patient information be changed accordingly.
- They felt there was too much reading in the PIL and so it was explained that a lot of it was necessary for legal purposes.
- They requested removal of coloured text from the study newsletter and replaced by either black on white or black on yellow text.
- They expressed no strong preference as to how they were contacted throughout the study but felt that updates via a study website would not be effective, so it was decided to not produce a dedicated study website.

Chapter 3 Results: study cohort

Recruitment

Details of patients approached and screened are described in [Appendix 1](#) (see [Table 31](#)). The flow of patients approached and screened, and the numbers of patients who were willing in principle, trained to use the index tests, eligible and who consented to take part are shown in [Figure 6](#) and summarised in [Table 1](#). Recruitment over the lifetime of the study (21 August 2018 to 31 March 2020) is shown in [Figure 7](#). The last patients were recruited in March 2020, before the first UK lockdown due to COVID-19 (see [Effect of COVID-19 pandemic on the study](#)). A total of 297 patients (participants) consented to take part.

Overall, eligibility was high at screening (943/974, 97%), about two-fifths of screened patients were willing in principle to participate (312/814, 38%), and of those willing, the majority attended training and consented (297/312, 95%).

Reasons for ineligibility at screening and training are shown in [Appendix 1](#) (see [Table 32](#)). The most common reasons for ineligibility at screening were due to patients not being previously treated or monitored for active nAMD (20/31, 65%) or to vision being limited by another condition (18/31, 58%). The most common reason for ineligibility at training was due to patients being unable to perform the MultiBit (MBT) electronic test (8/9, 89%).

Reasons for being unwilling in principle to participate are shown in [Table 2](#). Being 'put off by technology' was the most common reason (19%, 95/502); other commonly cited reasons included 'personal reasons' (17%, 86/502), not interested (17%, 83/502) and 'other' (17%, 87/502).

The time taken from screening to a receipt of a willingness decision, and from expressing willingness to training, are shown as a survival analysis by site and overall in [Appendix 1](#) (see [Figures 21](#) and [22](#)). Most centres received a decision from around 80% of their patients within 200 days. Exceptions were James Paget, with received decisions from 80% of patients after < 100 days, and Moorfields with a maximum of 60% of decisions received in total. About 90% patients had received their training within 50 days of notifying their willingness; patients from Moorfields and James Paget received their training on the day when they expressed willingness.

Recruitment by stratum of time since first treatment is shown in [Table 3](#). Half of recruited participants were first treated for nAMD 6–17 months before consenting, 28% 18–29 months prior to consent and 22% 30–41 months prior to consent. Eighteen of 297 participants (6%) were retrained (see [Appendix 1](#), [Table 33](#)), most of whom were being managed at James Paget (11/18, 61%).

Monitoring visits

At the end of the study, data for at least one monitoring visit after starting to use the index monitoring tests were available for 357 study eyes in 297 patients. Data were available for at least one complete monitoring visit after starting to use the index monitoring tests for 317 study eyes in 261 patients. The cumulative numbers of monitoring visits, and complete visits (including an OCT), over time are shown in [Figure 8](#). The number of complete visits decreased after the first UK COVID-19 lockdown in March 2020 because many appointments after this date were virtual consultations or treatment was administered without formal review. This situation slowly improved in the following months. Despite the challenges of the COVID-19 pandemic, 91% of the recorded monitoring visits were classified as complete for study eyes. Further reporting of the results focuses on the complete visits, except where indicated. Further information about total visits is described in [Appendix 1](#) (see [Table 34](#)).

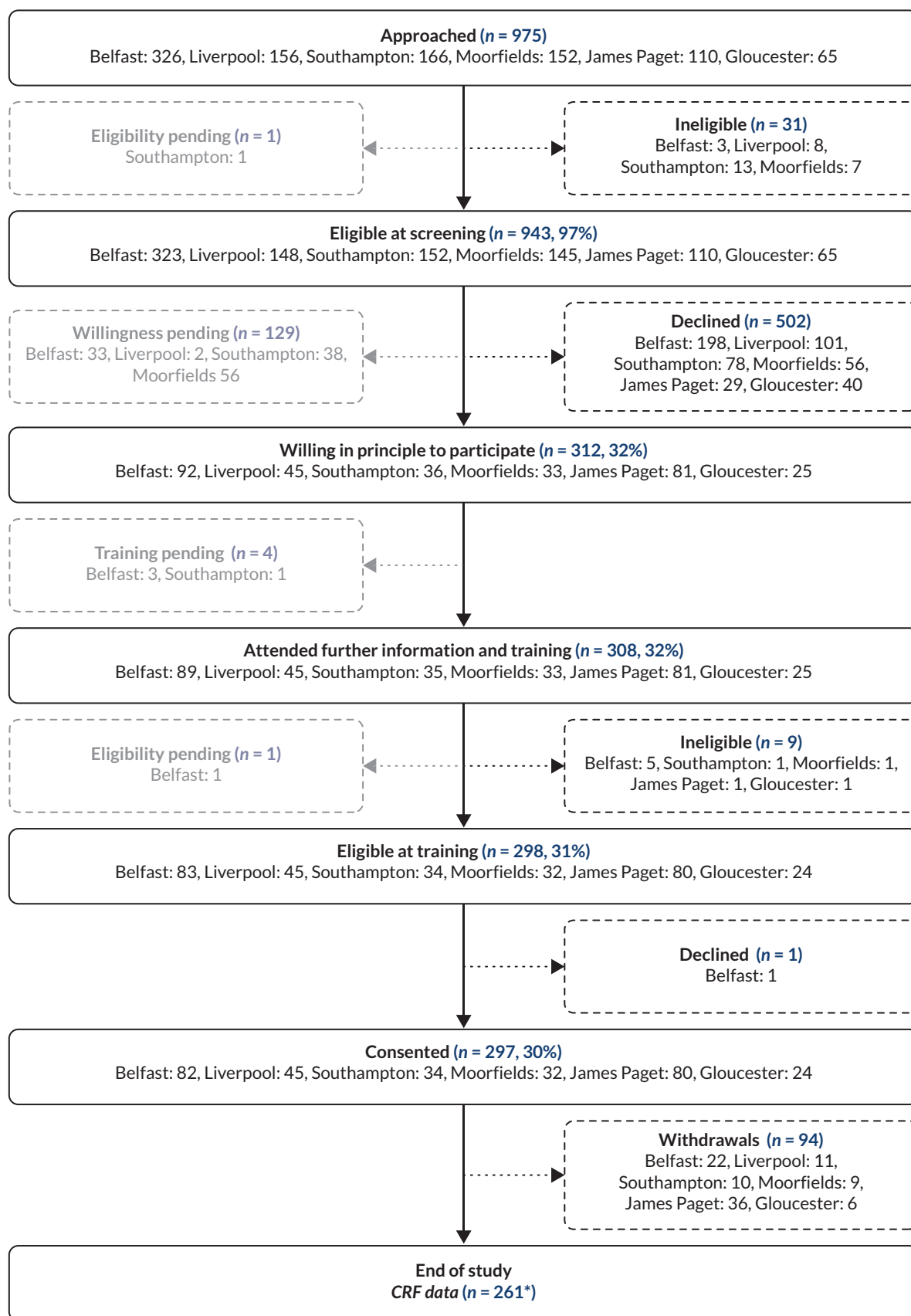


FIGURE 6 Patient flow chart. (1) Figure counts all participants with at least one management decision visit, irrespective of whether the management visit was complete (includes imaging and lesion status information) or whether they withdrew before the end of the study. (2) Percentages above are all calculated with respect to the number approached.

TABLE 1 Recruitment summary

Stage in recruitment pathway	Belfast n (%)	Liverpool n (%)	Southampton n (%)	Moorfields n (%)	James Paget n (%)	Gloucester n (%)	Overall n (%)
Approached	326	156	166	152	110	65	975
<i>Eligibility pending at screening</i>	0/326 (0)	0/156 (0)	1/166 (1)	0/152 (0)	0/110 (0)	0/65 (0)	1/975 (0)
Ineligible at screening ^a	3/326 (1)	8/156 (5)	13/165 (8)	7/152 (5)	0/110 (0)	0/65 (0)	31/974 (3)
Eligible at screening^a	323/326 (99)	148/156 (95)	152/165 (92)	145/152 (95)	110/110 (100)	65/65 (100)	943/974 (97)
<i>Willingness pending</i>	33/323 (10)	2/148 (1)	38/152 (25)	56/145 (39)	0/110 (0)	0/65 (0)	129/943 (14)
Declined (at screening) ^b	198/290 (68)	101/146 (69)	78/114 (68)	56/89 (63)	29/110 (26)	40/65 (62)	502/814 (62)
Willing in principle to participate^b	92/290 (32)	45/146 (31)	36/114 (32)	33/89 (37)	81/110 (74)	25/65 (38)	312/814 (38)
Attended further information and training	89/92 (97)	45/45 (100)	35/36 (97)	33/33 (100)	81/81 (100)	25/25 (100)	308/312 (99)
<i>Eligibility pending after training</i>	1/89 (1)	0/45 (0)	0/35 (0)	0/33 (0)	0/81 (0)	0/25 (0)	1/308 (0)
Ineligible after training ^c	5/88 (6)	0/45 (0)	1/35 (3)	1/33 (3)	1/81 (1)	1/25 (4)	9/307 (3)
Eligible^c	83/88 (94)	45/45 (100)	34/35 (97)	32/33 (97)	80/81 (99)	24/25 (96)	298/307 (97)
<i>Consent pending</i>	0/83 (0)	0/45 (0)	0/34 (0)	0/32 (0)	0/80 (0)	0/24 (0)	0/298 (0)
Declined consent ^d	1/83 (1)	0/45 (0)	0/34 (0)	0/32 (0)	0/80 (0)	0/24 (0)	1/298 (0)
Consented^d	82/83 (99)	45/45 (100)	34/34 (100)	32/32 (100)	80/80 (100)	24/24 (100)	297/298 (100)

a Denominators are approached minus eligibility pending patients.

b Denominators are eligible minus willingness-pending patients.

c Denominators are attended further information and training minus eligibility after training-pending patients.

d Denominators are eligible minus consent-pending patients.

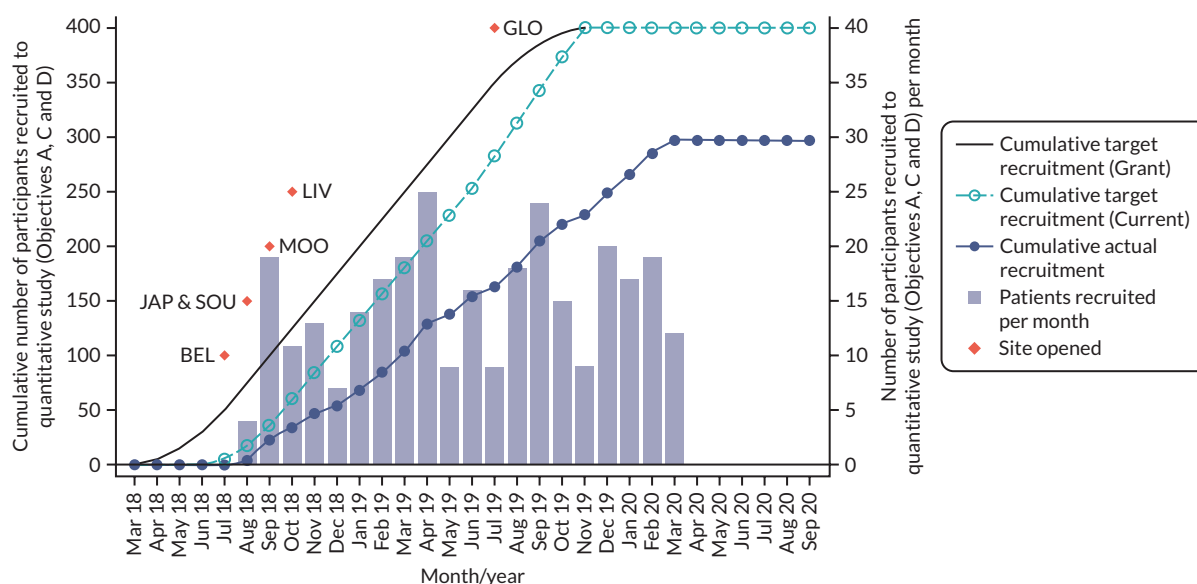


FIGURE 7 Cumulative number of participants recruited for Objectives A, C and D.

TABLE 2 Reasons for being unwilling in principle to participate

Reason for not willing	Belfast n (%)	Liverpool n (%)	Southampton n (%)	Moorfields n (%)	James Paget n (%)	Gloucester n (%)	Overall n (%)
Ineligible	- (-)	3 (3)	5 (6)	- (-)	2 (7)	- (-)	10 (2)
No reason given	32 (16)	18 (18)	12 (15)	- (-)	4 (14)	10 (25)	76 (15)
Not interested	18 (9)	11 (11)	16 (21)	19 (34)	6 (21)	13 (33)	83 (17)
Put off by technology	44 (22)	11 (11)	8 (10)	20 (36)	10 (34)	2 (5)	95 (19)
No benefit in taking part	6 (3)	- (-)	2 (3)	- (-)	1 (3)	- (-)	9 (2)
Personal reasons	21 (11)	8 (8)	32 (41)	10 (18)	6 (21)	9 (23)	86 (17)
Not enough time to consider	- (-)	1 (1)	- (-)	- (-)	- (-)	- (-)	1 (0)
Insurance invalidated	- (-)	- (-)	- (-)	- (-)	- (-)	- (-)	- (-)
Unable to agree to consent questions	- (-)	- (-)	- (-)	1 (2)	- (-)	- (-)	1 (0)
Too much of a commitment	37 (19)	11 (11)	1 (1)	5 (9)	- (-)	- (-)	54 (11)
Other ^a	40 (20)	38 (38)	2 (3)	1 (2)	- (-)	6 (15)	87 (17)
Overall (within site)	198 (100)	101 (100)	78 (100)	56 (100)	29 (100)	40 (100)	502 (100)

a Other reasons include: BEL - 17 Medical reason, 1 Poor memory, 5 Too far to travel, 8 Feels they are too old, 1 Poor mobility, 1 Does not want internet, 1 Does not feel capable, 2 Lives alone, 1 Initially agreed, but changed mind, 2 Not interested, 1 Too complicated; LIV - 11 Medical reason, 2 Too far to travel, 3 Does not feel capable, 7 Not interested, 1 Too complicated, 13 COVID-19 recruitment pause, 1 Not eligible; SOU - 1 Too far to travel, 1 Discharged from follow-up; MOO - 1 Did not attend training; GLO - 1 Poor mobile signal at home, 4 COVID-19 recruitment pause, 1 Patient died.

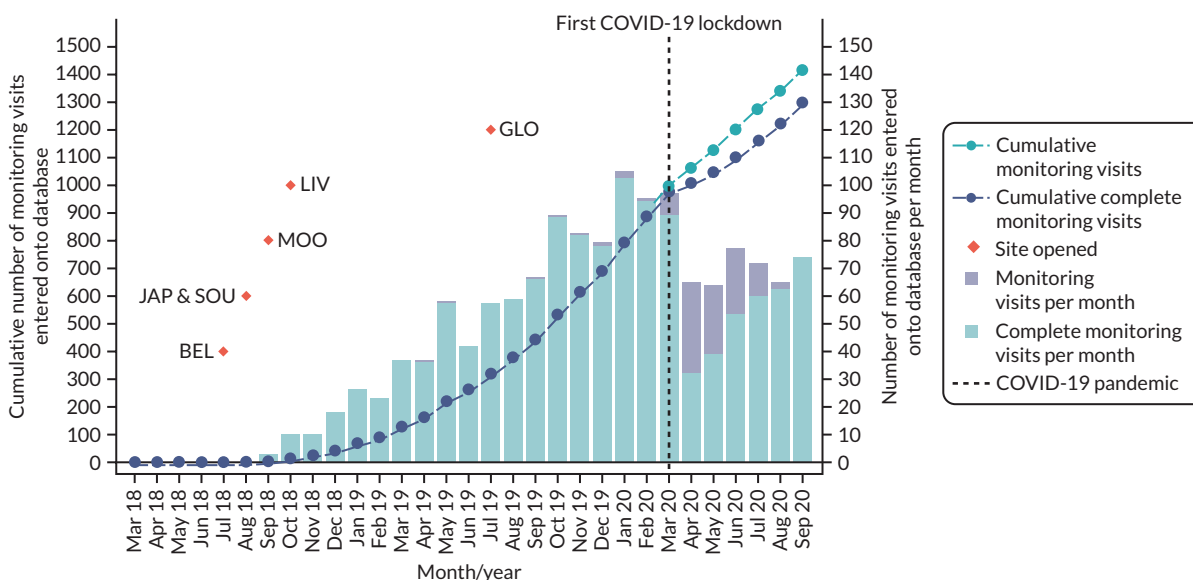
The percentages of monitoring visits classified as complete are shown overall and by site in [Table 4](#). Moorfields had the lowest proportion of complete monitoring visits for study eyes (259/331, 78%) and Southampton (153/156, 98%) and James Paget (563/563, 100%) had very high proportions. For fellow

TABLE 3 Recruitment into each stratum

Site	Months in study	Expected overall recruitment	Total recruitment n (%) ^a	Stratum of time since first treatment		
				6–17 months n (%) ^b	18–29 months n (%) ^b	30–41 months n (%) ^b
BEL	27	80.0	82 (103)	40 (49)	21 (26)	21 (26)
LIV	24	80.0	45 (56)	23 (51)	10 (22)	12 (27)
SOU	26	80.0	34 (43)	14 (41)	15 (44)	5 (15)
MOO	25	80.0	32 (40)	17 (53)	11 (34)	4 (13)
JAP	26	50.0	80 (160)	42 (53)	20 (25)	18 (23)
GLO	15	30.0	24 (80)	13 (54)	7 (29)	4 (17)
Overall	143	400.0	297 (74)	149 (50)	84 (28)	64 (22)

a Percentage is of expected recruitment for site.

b Percentage is of total actual recruitment within site.



Complete monitoring visits are defined as having an OCT taken and clinical decision given for all study eyes.

FIGURE 8 Number of monitoring visits recorded on database.

eyes, the overall proportion of complete visits was 92%, with a similar pattern across sites. The rates of total monitoring visits per participant per year (not influenced by site-specific decisions about how to provide care during the first lockdown) was 4.01 consistent with the assumption set out in the sample size estimate in the protocol. However, this rate varied markedly by site (2.03–6.39 management visits per patient per year; see [Appendix 1, Table 34](#)). Sites also varied with respect to agreement between intended and actual review schedules (see [Appendix 1, Table 35](#)). Overall, 53% of monitoring visits occurred on or before the intended review date.

The reference standard was available for all complete management visits for both study and fellow eyes. For study eyes, the classifications were: active = 932 (60.1%), inactive = 576 (37.2%), uncertain = 40 (2.6%) and no lesion = 1 (0.1%). For fellow eyes, the classifications were: no lesion = 561 (90.6%), active = 34 (5.5%), inactive = 19 (3.1%) and uncertain = 5 (0.8%).

TABLE 4 Complete and total monitoring visits by site

Site	Patients, n	Study eyes ^a		Fellow eyes	
		Eyes, n	Visits, complete ^b /total ^c (%)	Eyes(n)	Visits, complete ^b /total (%)
Belfast	82	101	248/265 (93.6)	43	99/107 (92.5)
Liverpool	45	53	192/248 (77.4)	26	94/118 (79.7)
Southampton	34	38	153/156 (98.1)	21	86/88 (97.7)
Moorfields	32	42	259/331 (78.2)	20	93/110 (84.5)
James Paget	80	91	563/563 (100.0)	37	170/170 (100.0)
Gloucester	24	32	134/140 (95.7)	15	77/79 (97.5)
Overall	297	357^c	1549/1703 (91.0)	162	619/672 (92.1)

a Three hundred and seventeen study eyes in 261 patients had at least one complete visit.

b 'Complete' visits are those with a lesion activity decision and an OCT image.

c The overall total (1549/1703) is the number of study eyes with at least one complete/incomplete management visit, so visits were counted twice when a patient had two study eyes.

Baseline characteristics

Overall baseline demographic characteristics and exposure to technology of consented participants are described in [Table 5](#). More participants were women (58.6%). The mean age was 74.9 (6.6) years (SD) and the mean visual acuity of study eyes (better seeing eyes if participants had two study eyes) was 0.2 (0.2) Logarithm of the minimum angle of resolution (LogMAR) (SD). Many participants had known risk factors for nAMD; for example, 10.1% were current smokers and 53.2% had hypertension requiring treatment. Participants reported a range of comorbidities, including 20.1% malignancy, 2.4% type 1 diabetes, 10.4% type 2 diabetes and 5.1% other conditions that may have affected their ability to adhere to home monitoring. Most participants had substantial exposure to technology before entering the study; 66.6% had a smartphone, 66.2% had a tablet and 85.1% had internet at home.

Objective C investigated inequalities that may have arisen with home monitoring (see [Statistical analyses and report content](#)). The demographic characteristics of patients who were willing in principle versus those who were unwilling in principle are reported in [Table 6](#) by site.

TABLE 5 Baseline demographic characteristics and exposure to technology of consented participants

Baseline characteristic	Overall (n = 297)	
	n	%/SD
Sex		
	Male	123/297 41.4
	Female	174/297 58.6
Age	Mean (SD) years	74.9 6.6
Number of participants with	Two study eyes ^a	50 16.8
	One study and one fellow eye ^b	161 54.2
	One study and one excluded eye ^c	86 29.0

TABLE 5 Baseline demographic characteristics and exposure to technology of consented participants (*continued*)

		Overall (n = 297)	
		n	%/SD
Visual acuity in study eye ^d	Mean (SD) ETDRS	72.9	10.7
Smoking history	Current smoker	30/297	10.1
	Ex-smoker (> 1 month)	137/297	46.1
	Never smoked	130/297	43.8
Medical history			
Congestive cardiac failure		11/297	3.7
Myocardial infarction		19/297	6.4
Peripheral vascular disease		7/297	2.4
Cerebrovascular disease		21/297	7.1
Hypertension requiring treatment		158/297	53.2
Chronic pulmonary disease		28/297	9.4
Rheumatological disease		53/297	17.8
Renal disease		25/297	8.4
Liver disease		7/297	2.4
Neurological dysfunction		12/297	4.0
Malignancy		60/297	20.2
Diabetes – type 1		7/297	2.4
Diabetes – type 2		31/297	10.4
Other conditions that may affect ability to perform testing		15/297	5.1
Exposure to technology			
Television		294/296	99.3
Simple mobile phone		130/296	43.9
Smartphone		197/296	66.6
Tablet		196/296	66.2
Laptop/home computer		184/296	62.2
Internet at home		252/296	85.1
E-mail		213/296	72.0
Social media		97/296	32.8
TV streaming/on-demand services		146/296	49.3

a Four of the 50 participants had no follow-up hospital management visits and their study eyes were excluded from diagnostic accuracy analyses.

b Twenty-three participants with a fellow eye had no follow-up hospital management visits and their study eyes were excluded from diagnostic accuracy analyses; fellow eyes of 17/138 participants converted to nAMD during follow-up and all are included in analyses for Objective D.

c Nine participants with one excluded eye had no follow-up hospital management visits and their study eyes were excluded from diagnostic accuracy analyses; five patients had eyes excluded at baseline (due to diagnosis of nAMD within the last 6 months) but which became eligible during follow-up and were included as study eyes thereafter.

d For participants with two study eyes, better seeing eye is used.

TABLE 6 Demographic characteristics of patients unwilling in principle vs. willing to participate in the study

	BEL (n = 290)		LIV (n = 146)		SOU (n = 114)		MOO (n = 90)		JAP (n = 110)		GLO (n = 65)		Overall (n = 815)	
	Not willing n (%)	Willing n (%)	Not willing n (%)	Willing n (%)	Not willing n (%)	Willing n (%)	Not willing n (%)	Willing n (%)	Not willing n (%)	Willing n (%)	Not willing n (%)	Willing n (%)	Not willing n (%)	Willing n (%)
n	198	92	101	45	78	36	56	33	29	81	40	25	503	312
Gender														
Male	52 (26)	37 (40)	40 (40)	19 (42)	28 (36)	13 (36)	24 (43)	15 (45)	12 (41)	34 (42)	9 (23)	12 (48)	165 (33)	130 (42)
Female	146 (74)	55 (60)	61 (60)	26 (58)	50 (64)	23 (64)	32 (57)	18 (55)	17 (59)	47 (58)	31 (78)	13 (52)	338 (67)	182 (58)
Age														
Mean (SD)	79.3 (6.7)	75.5 (6.6)	77.2 (8.1)	73.6 (7.5)	80.8 (7.5)	78.3 (7.7)	77.8 (6.5)	72.4 (5.2)	77.8 (5.2)	76.4 (6.5)	72.9 (4.7)	72.0 (4.2)	78.3 (7.2)	75.2 (6.8)
Minimum age	58	57	51	56	63	60	63	60	68	54	57	65	51	54
Maximum age	94	91	94	88	97	93	88	83	87	89	80	79	97	93
Ethnicity														
White/Caucasian	- (-.-)	2 (2)	100 (99)	45 (100)	78 (100)	36 (100)	32 (57)	21 (64)	29 (100)	81 (100)	40 (100)	25 (100)	280 (56)	210 (67)
Asian/Asian British	- (-.-)	- (-.-)	- (-.-)	- (-.-)	- (-.-)	- (-.-)	1 (2)	- (-.-)	- (-.-)	- (-.-)	- (-.-)	- (-.-)	1 (0)	- (-.-)
Mixed/multiple ethnic groups	- (-.-)	- (-.-)	1 (1)	- (-.-)	- (-.-)	- (-.-)	5 (9)	3 (9)	- (-.-)	- (-.-)	- (-.-)	- (-.-)	6 (1)	3 (1)
Data not available	198 (100)	90 (98)	- (-.-)	- (-.-)	- (-.-)	- (-.-)	18 (32)	9 (27)	- (-.-)	- (-.-)	- (-.-)	- (-.-)	216 (43)	99 (32)

	BEL (n = 290)		LIV (n = 146)		SOU (n = 114)		MOO (n = 90)		JAP (n = 110)		GLO (n = 65)		Overall (n = 815)	
	Not willing n (%)	Willing n (%)	Not willing n (%)	Willing n (%)	Not willing n (%)	Willing n (%)	Not willing n (%)	Willing n (%)	Not willing n (%)	Willing n (%)	Not willing n (%)	Willing n (%)	Not willing n (%)	Willing n (%)
IMD rank^a														
25th percentile	224.0	381.5	1968.0	5810.0	18,534.0	18,317.0	9140.0	13,554.0	11,025.0	8661.0	17,767.0	17,873.0		
Median	501.0	693.5	12,909.0	13,612.0	25,653.0	26,473.5	16,901.0	18,089.0	13,002.0	13,357.0	21,806.0	24,395.0		
75th percentile	712.0	792.0	20,374.0	25,063.0	29,501.0	31,044.5	22,740.5	25,290.0	17,852.0	20,028.0	27,759.0	26,869.0		
Distance visual acuity^b														
Snellen ≥ 6/18	167 (84)	80 (87)	82 (81)	36 (80)	68 (87)	34 (94)	48 (86)	29 (88)	25 (86)	77 (95)	27 (68)	21 (84)	417 (83)	277 (89)
Snellen < 6/18 and > 6/24	16 (8)	5 (5)	10 (10)	5 (11)	7 (9)	2 (6)	5 (9)	3 (9)	3 (10)	3 (4)	5 (13)	2 (8)	47 (9)	20 (6)
Snellen ≤ 6/24	15 (8)	7 (8)	9 (9)	4 (9)	3 (4)		3 (5)	1 (3)	1 (3)	1 (1)	8 (20)	2 (8)	39 (8)	15 (5)

a Site percentages given are across all patients assessed as eligible at screening and should be referred to the denominators at the top of each column.

b IMD rankings are not comparable between Northern Ireland and England. This precludes reporting an overall median.

Note

The eye with the best visual acuity was used.

Details of baseline ocular histories of study eyes are shown in [Table 7](#). (Note that the total of 347 study eyes includes two study eyes for 86 participants.) At baseline, 75% of study eyes were recorded as having an active nAMD lesion, 71% of study eyes were being treated with or were most recently treated with Aflibercept and 10% of study eyes had a cataract.

Participant withdrawals

In almost all instances, decisions by participants were the reason for withdrawal (88/94, 94%; see [Appendix 1, Table 36](#)). The reason for withdrawal was very rarely discharge from NHS review (2/94, 2% of all withdrawals). One participant withdrew due to inadequate mobile phone network coverage and one died. The reason for withdrawal was missing for the other two.

The most common reason given by participants was personal reasons (50/88, 57% of participants' decisions to withdraw and 53% of all withdrawals), followed by testing being too time-consuming (25 participants, 28%) or other reason (25 participants, 28%; [Table 8](#)).

TABLE 7 Baseline ocular histories of study eyes

		Overall (n = 347)	
		n	%
Lesion status	Present	345/347	99.4
	Active ^a	257/345	74.5
	Inactive ^a	84/345	24.3
	Uncertain ^a	4/345	1.2
Current/most recent treatment	Aflibercept (Eylea)	246/347	70.9
	Ranibizumab (Lucentis)	67/347	19.3
	Not applicable	2/347	0.6
	Bevacizumab (Avastin)	32/347	9.2
	UNK	0/347	0.0
Current cataract		34/347	9.8
Complication of cataract surgery		2/347	0.6
Glaucoma		0/347	0.0
Diabetic eye disease	Present	1/347	0.3
	Laser treatment ^b	0/1	0.0
Retinal detachment		1/347	0.3
Vascular occlusion		2/347	0.6
Corneal problems		2/347	0.6
Lazy eye (amblyopia)		7/347	2.0
Other		10/347	2.9
Information in ophthalmological review of any other condition that limits vision	Anterior segment	6/347	1.7
	Posterior segment	2/347	0.6

UNK, Unknown.

a Percentage given is only for those study eyes with a lesion present.

b Percentage given is only for those study eyes with diabetic eye disease present.

Note

Overall number is given for all study eyes across patients.

TABLE 8 Combinations of reasons given by participants (n = 88) for decisions to withdraw

Reasons for withdrawal	n	%
Personal reasons only	25	28.41
Home eye tests too time-consuming	9	10.23
Other ^a	7	7.95
Home eye tests too time-consuming, personal reasons	5	5.68
Personal reasons, other ^a	5	5.68
Unwilling to give reason	5	5.68
Difficulties operating equipment	4	4.55
Difficulties operating equipment, personal reasons	3	3.41
Difficulties operating equipment, home eye tests too time-consuming	2	2.27
Difficulties operating equipment, personal reasons, other ^a	2	2.27
Difficulties operating equipment, home eye tests too time-consuming, personal reasons	1	1.14
Difficulties operating equipment, home eye tests too time-consuming, unhappy with mVT, personal reasons	1	1.14
Difficulties operating equipment, home eye tests too time-consuming, unhappy with mVT, unhappy with MBT, other ^a	1	1.14
Difficulties operating equipment, home eye tests too time-consuming, unhappy with mVT, unhappy with MBT, unhappy with KSJ	1	1.14
Difficulties operating equipment, home eye tests too time-consuming, unhappy with mVT, unhappy with MBT, unhappy with KSJ, other ^a	1	1.14
Difficulties operating equipment, unhappy with MBT, other ^a	1	1.14
Difficulties operating equipment, unhappy with mVT, unhappy with MBT	1	1.14
Difficulties operating equipment, unhappy with mVT, unhappy with MBT, personal reasons	1	1.14
Difficulties operating equipment, unhappy with mVT, unhappy with MBT, personal reasons, other ^a	1	1.14
Difficulties operating equipment, unhappy with mVT, unhappy with MBT, unhappy with KSJ	1	1.14
Home eye tests too time-consuming, personal reasons, other ^a	1	1.14
Home eye tests too time-consuming, unhappy with MBT, other ^a	1	1.14
Home eye tests too time-consuming, unhappy with mVT, other ^a	1	1.14
Home eye tests too time-consuming, unhappy with mVT, unhappy with MBT, personal reasons	1	1.14
Personal reasons, unwilling to give reason	1	1.14
Unhappy with MBT, other ^a	1	1.14
Unhappy with MBT, personal reasons	1	1.14
Unhappy with mVT, unhappy with MBT, personal reasons, other ^a	1	1.14
Unhappy with mVT, unhappy with MBT, unhappy with KSJ	1	1.14
Unhappy with mVT, unhappy with MBT, unhappy with KSJ, other ^a	1	1.14
Unhappy with mVT, unhappy with MBT, unhappy with KSJ, personal reasons, other ^a	1	1.14

^a Other reasons include: BEL – 1 Tests too difficult, 2 Tests too difficult to see in light, 3 Medical reason, 2 Thought study only lasted for a year, 1 Made patient more aware of vision difficulties, 2 Tests too frustrating; LIV – 2 Patient unwell, 1 Tests too difficult to see in light, 1 Thought study only lasted for a year, 1 COVID-19 reasons; SOU – 1 Vision too poor; MOO – 1 Tests too frustrating; JAP – 1 Tests too difficult, 1 Tests too fiddly, 3 Tests too time consuming, 1 Equipment difficulties, 1 Vision too poor.

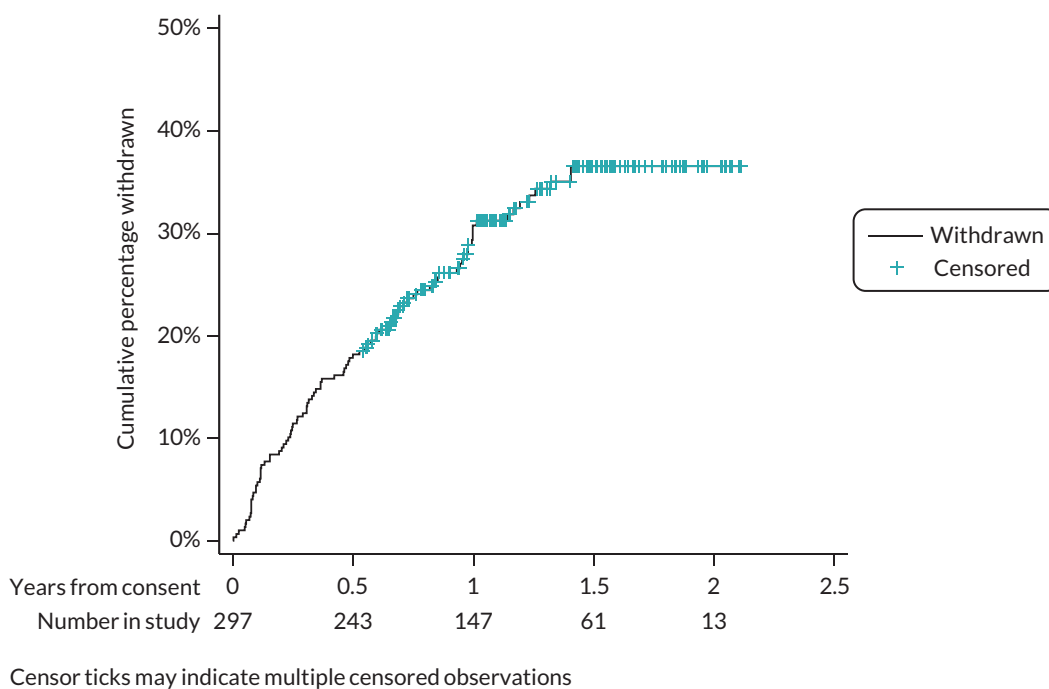


FIGURE 9 Kaplan–Meier graph of time to withdrawal.

Time to withdrawal overall is shown in [Figure 9](#) (and by site in [Appendix 2, Figure 23](#)). Overall, 32% of participants withdrew.

Protocol deviations

Protocol deviations were few apart from monitoring visits without OCTs during the last months of follow-up, due to the COVID-19 pandemic. Other protocol deviations comprised:

1. patients who were recruited when the mVT app was not working. These patients trained to use the mVT at a later date;
2. retinal images that were not sent to CARF; percentages of monitoring visits for which images were expected are shown by site in [Table 9](#);
3. patients' self-reported data that were not collected before patients attended monitoring visits.

Adherence to testing

Frequencies of testing and test completion for the paper KSJ home-monitoring test are shown in [Appendix 1](#) (see [Table 37](#)). It was intended that participants should report KSJ results once each week. Overall, participants completed the booklet fully for 87% of weeks, with < 1% of weeks with all KSJ data missing. The median time between adjacent completion dates was 7 days.

Frequency of testing and weeks when tests were completed for the electronic home-monitoring tests (mVT and MBT) are also shown in [Appendix 1](#) (see [Tables 38](#) and [39](#), respectively). The median testing frequencies for both tests for the three strata of time since starting treatment were almost all 3 or 4 per month. For study eyes, 56.2% and 59.3% of expected weekly tests were completed for MBT and mVT, respectively. There was evidence that testing frequencies/text completion differed by site: 33.7% (MBT) and 39.8% (mVT) of expected tests were performed by participants at one site compared to over 60% for both tests for all other study sites. Times to stopping testing with each home-monitoring test are shown in [Figures 10, 11](#) and [12](#); tests were being completed for about 60% of study eyes by 18–24 months after starting to test.

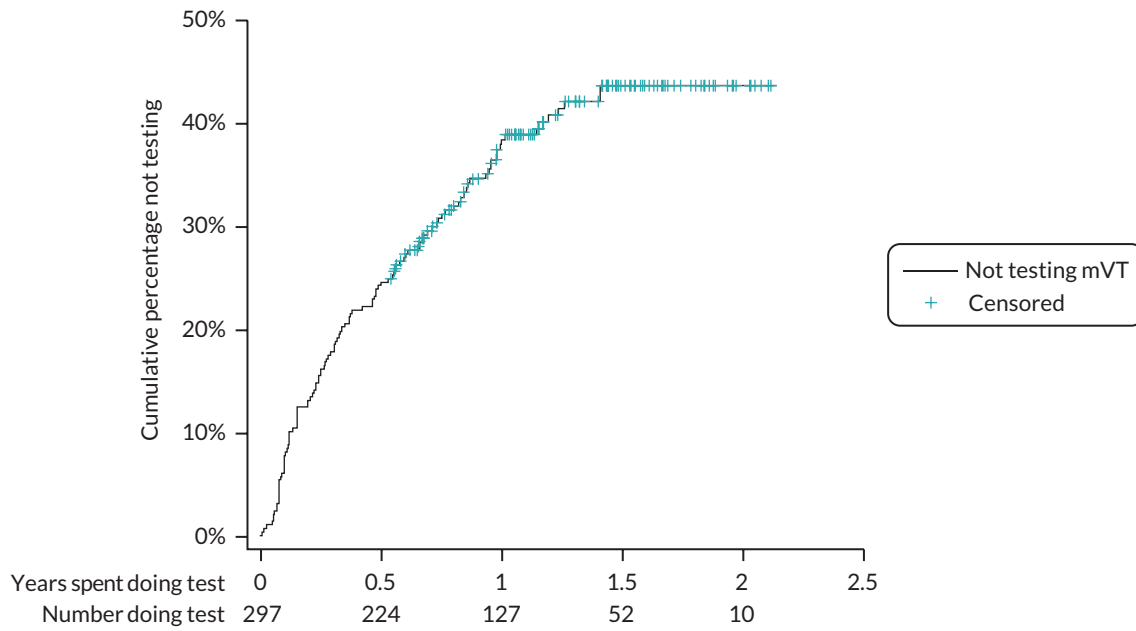


FIGURE 10 Time to stopping home monitoring with the mVT.

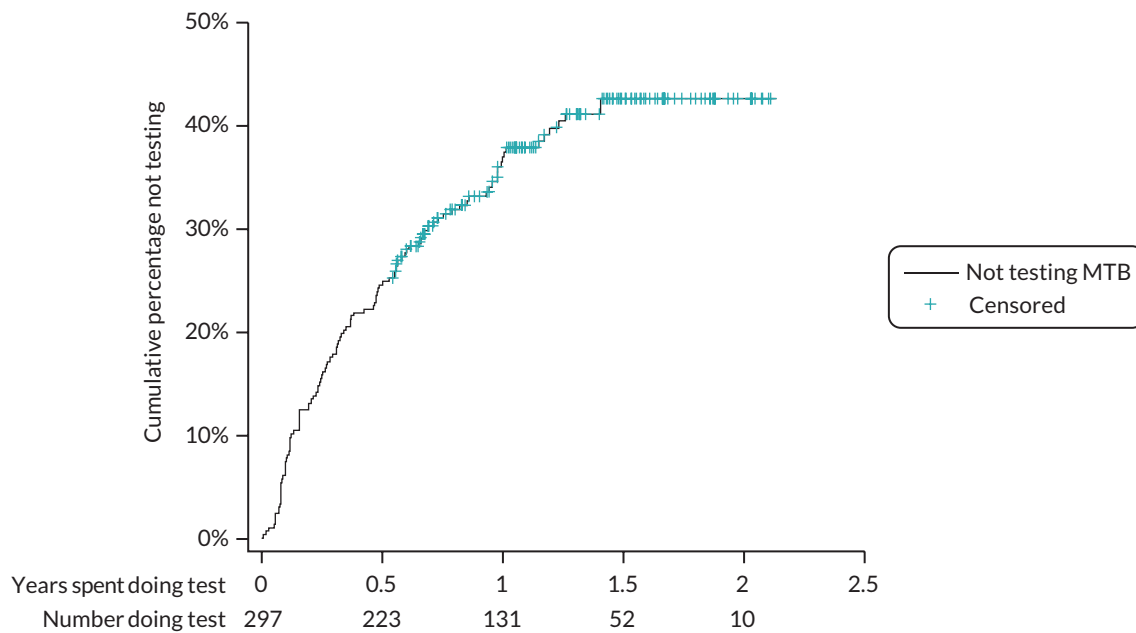


FIGURE 11 Time to stopping home monitoring with the MBT.

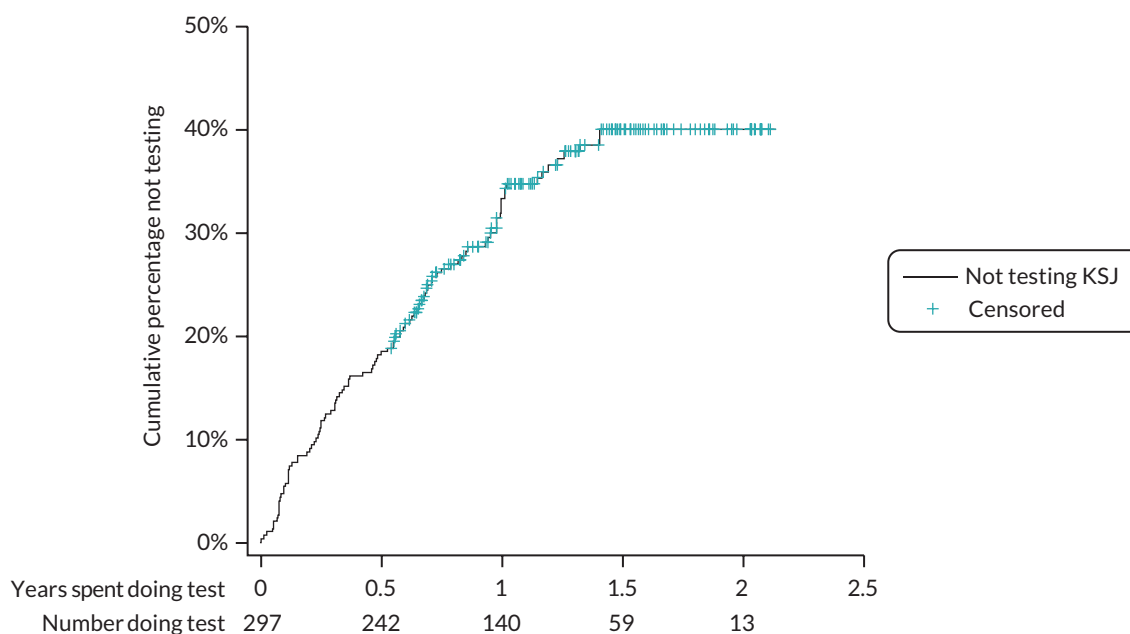


FIGURE 12 Time to stopping home monitoring with the KSJ.

Retinal images collected

The numbers of baseline and follow-up images submitted to and received by CARF are shown in [Table 9](#).

Effect of coronavirus disease 2019 pandemic on the study

The study was not greatly affected by the COVID-19 pandemic. Recruitment was scheduled to end on 31 March 2020. A small number of patients were not recruited who might otherwise have been. However, even if they had been recruited, they might not have had a complete management visit due to the lockdown (so they could not have contributed to the primary analysis of Objective A).

The first lockdown had a more noticeable effect on the number of complete management visits. We estimate that 100–150 additional complete visits would have been documented if the lockdown had not come into effect (see [Figure 8](#)). Outgoing calls to participants to prompt testing or to find out why a participant was unable to test, and SMS text notifications, were suspended from March to June 2020 because the research team could not access the NHS-hosted MONARCH database when working from home.

Details of site-specific strategies for managing patients during the first 6 months of the pandemic were not collected (after which data collection stopped) and were not evident from the data collected (apart from the absence of an OCT). Anecdotally, strategies seemed to vary across sites. Some continued to provide face-to-face outpatient clinics; some reviewed patients virtually (or in person) but without performing an OCT to minimise patients' time in the hospital; some treated patients (a quick procedure in an aseptic room) without assessing activity status, again to minimise potential exposure in hospital.

TABLE 9 Baseline images submitted and received by CARF

Site (participants)	Patients consented to image collection (n)	Number of monitoring visits attended (n)	Eyes assessed ^a		Images received by CARF	
			Study eyes (n)	Fellow eyes (n)	Study eyes, n (%)	Fellow eyes, n (%)
Baseline						
Belfast (n = 82)	82	82	100	43	31 (31.0)	12 (27.9)
Liverpool (n = 45)	45	45	51	26	47 (92.2)	22 (84.6)
Southampton (n = 34)	34	34	38	21	28 (73.7)	15 (71.4)
Moorfields (n = 32)	32	32	39	20	0 (0.0)	0 (0.0)
James Paget (n = 80)	79	79	88	35	85 (96.6)	35 (100.0)
Gloucester (n = 24)	0	0	0	0	0 (0)	0 (0)
Overall (n = 297)	272	272	316	145	191 (60.4)	84 (57.9)
Follow-up						
Belfast (n = 82)	82	217	265	107	247 (93.2)	100 (93.5)
Liverpool (n = 45)	45	220	248	118	194 (78.2)	95 (80.5)
Southampton (n = 34)	34	141	156	88	151 (96.8)	86 (97.7)
Moorfields (n = 32)	32	233	331	110	0 (0.0)	0 (0.0)
James Paget (n = 80)	79	474	551	158	551 (100.0)	158 (100.0)
Gloucester (n = 24)	0	0	0	0	0 (0)	0 (0)
Overall (n = 297)	272	1285	1551	581	1143 (73.7)	439 (75.6)

a Includes complete and incomplete management decisions.

Chapter 4 Results: primary outcome (Objective A)

Data were available for 261 patients and 317 study eyes for at least one complete management visit after starting to use the index tests. These study eyes contributed 1549 complete visits to the primary analysis population. These data form the analysis population for the primary analysis to address Objective A. All available dates were included, that is, visit records did not depend on having complete data. Uncertain and no-lesion reference standard classifications were recoded as inactive for all analyses. The number of visits included in models fitted for each index test varied because participants did not always complete all tests at all visits ([Table 10](#)). Correlations between the four KSJ summary tests scores are shown in [Appendix 1](#) (see [Table 40](#)).

Results from the models for the primary outcome of lesion activity in study eyes as assessed during a complete management visit are given in [Table 10](#) for all home-monitoring test models and for a model that excluded the home-monitoring tests (no-test model).

None of participants' sex, age, stratum of time since first treatment at baseline or visual acuity strata at baseline were seen to affect lesion activity, irrespective of whether a home-monitoring test score was included or excluded. Days since baseline were significantly associated with lesion activity irrespective of the model; in the no-test model, a doubling in days since baseline was associated with a 36% reduction in the odds of a lesion being recorded as active (equivalent to OR 0.52, 95% CI 0.36 to 0.74; $p < 0.001$ for 1 natural log unit). The strength of this association was consistent between models.

Among the index tests, only the KSJ VA worse score had a statistically significant association with lesion activity. Patients who reported their vision to be 'worse than baseline' at all time points throughout an inter-follow-up period were 3.5 times more likely to have a lesion assessed as active compared to patients who reported the 'same' or 'better' vision across all time points (OR 3.48, 95% CI 1.09 to 11.13; $p = 0.036$).

Receiver operating characteristic (ROC) curves and AUROCs for summary test scores for all tests are shown in [Figure 13](#); further details of each model, including model performance at thresholds identified by Youden's index, are shown in [Appendix 1](#) (see [Table 41](#)).

The AUROCs from the models that included the home-monitoring test summary scores ranged from 0.57 to 0.59, only marginally higher than the model that excluded the tests (0.55); 95% CIs overlapped across all four models. Sensitivities at thresholds identified by Youden's index were low for all models, ranging from 0.36 for the no-test and KSJ models to a maximum of 0.47 in the model for the MBT test; specificity was 0.73 in the no-test model, lower for the MBT and mVT and 0.77 for the KSJ.

Average test scores above and below the predicted probability thresholds are shown in [Appendix 1](#) (see [Table 42](#)). Mean MBT test scores were lower below the identified threshold (prediction of inactive lesion) than above (prediction of active lesion), contrary to expectation (higher MBT scores should be indicative of better vision). Conversely, higher mVT score was associated with inactive lesions, when higher mVT scores are indicative of poorer vision. For the KSJ, median near VA was marginally higher above the threshold (again contrary to expectation), but percentages of visits when VA, Amsler and object recognition were reported to be worse were higher above the threshold, as expected.

Sensitivity analysis 1, predictor 2 (test data for preceding 4 weeks only).

The same analysis population was used as for the primary analysis. However, the numbers of visits included in models fitted for each index test varied because participants did not always complete all tests at all visits and data were not always available for the 4 weeks immediately preceding the management visit ([Table 11](#)). Results for the primary outcome of lesion activity in study eyes as assessed during a management visit are given in [Table 11](#) for all home-monitoring test models and the no-test model.

TABLE 10 Model results for the primary outcome of lesion activity in study eyes for all home-monitoring test models and for a no-test model; primary analysis

Numbers in models	No test		MBT		mVT		KSJ	
Visits	1413		1213		1249		1238	
Participants	252		233		237		224	
Eyes	302		279		284		270	
Predictor in model	OR (95% CI)	<i>p</i> -value	OR (95% CI)	<i>p</i> -value	OR (95% CI)	<i>p</i> -value	OR (95% CI)	<i>p</i> -value
Mean MBT score			1.02 (1.00 to 1.04)	0.083				
Mean mVT score ^a					0.99 (0.98 to 1.01)	0.432		
Median KSJ VA score							1.33 (0.97 to 1.84)	0.080
% KSJ VA worse ^b							3.48 (1.09 to 11.13)	0.036*
% KSJ Amsler grid worse ^b							3.39 (0.70 to 16.48)	0.130
% KSJ household object worse ^b							1.14 (0.20 to 6.54)	0.880
Sex								
Male	1		1		1		1	
Female	0.72 (0.29 to 1.78)	0.481	0.73 (0.29 to 1.81)	0.496	0.82 (0.33 to 2.02)	0.67	0.51 (0.20 to 1.29)	0.153
Time since first treatment at baseline strata								
6–17 months	1		1		1		1	
18–29 months	0.67 (0.25 to 1.81)	0.429	0.96 (0.35 to 2.64)	0.943	0.87 (0.32 to 2.36)	0.781	0.60 (0.22 to 1.68)	0.331
30–41 months	0.45 (0.15 to 1.39)	0.165	0.50 (0.16 to 1.55)	0.228	0.47 (0.15 to 1.47)	0.194	0.54 (0.17 to 1.70)	0.290
Age	0.99 (0.93 to 1.06)	0.744	1.00 (0.93 to 1.07)	0.946	0.98 (0.91 to 1.05)	0.537	1.00 (0.94 to 1.07)	0.906
Baseline VA strata								
Better than or equal to 6/18	1		1		1		1	
Worse than 6/18 and better than 6/24	0.72 (0.15 to 3.52)	0.684	1.86 (0.34 to 10.12)	0.472	1.14 (0.22 to 6.05)	0.875	1.00 (0.2 to 5.03)	0.997
Worse than or equal to 6/24	0.64 (0.10 to 4.10)	0.638	0.77 (0.11 to 5.19)	0.787	0.91 (0.14 to 5.93)	0.918	0.96 (0.14 to 6.59)	0.970
Days since baseline ^c	0.52 (0.36 to 0.74)	< 0.001	0.49 (0.33 to 0.72)	< 0.001	0.54 (0.37 to 0.78)	0.001	0.52 (0.36 to 0.75)	< 0.001

* $p < 0.05$.^a Transformed by multiplying by 100.^b Proportion of scores across the respective interval that reported vision as being 'worse'.^c Natural log transformed.

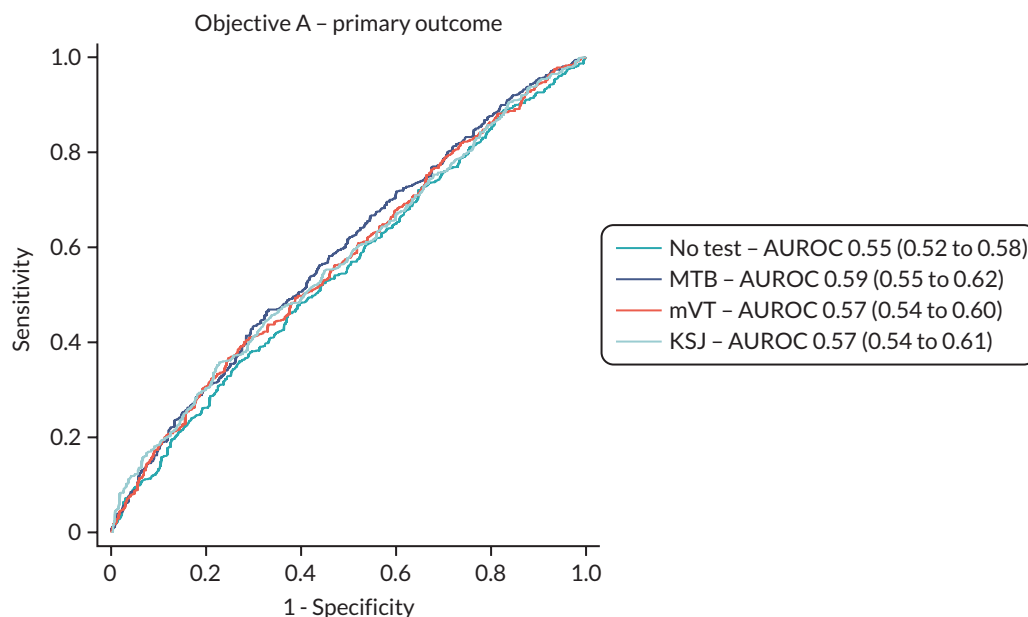


FIGURE 13 Receiver operating characteristic curves and AUROCs (95% CIs) for summary test scores for all index tests; primary analysis.

The baseline model results were the same as in the primary analysis since all predictors and the outcome were identical. No test summary score was seen to have a significant association with lesion activity. Results from the models are shown in [Table 11](#) for all the home-monitoring tests. The KSJ median VA score and VA worse score approached statistical significance, consistent with the primary analysis.

Receiver operating characteristic curves and AUROCs for summary test scores for all tests are shown in [Figure 14](#); further details of each model, including model performance at thresholds identified by Youden's index, are shown in [Appendix 1](#) (see [Table 43](#)). As in the primary analysis, the AUROCs from the models that included the home-monitoring tests were only marginally improved from the model that excluded the tests, with 95% CIs that overlapped across all models. Sensitivities at thresholds identified by Youden's index were low for all models, ranging from 0.36 for the base model to 0.46 or 0.47 for models that included summary test scores; specificity was 0.76 in the no-test model, and lower (0.65 or 0.67) for models including the summary test scores.

Average test scores above and below the predicted probability thresholds are given in [Appendix 1](#) (see [Table 44](#)). As in the primary analysis, mean MBT test scores were lower below the identified threshold than above, and mVT scores were higher (less negative) below the identified threshold than above. For the KSJ, median near VA was marginally higher above the threshold (again contrary to expectation) but percentages of visits when VA, Amsler and object recognition were reported to be worse were higher above the threshold, as expected.

Sensitivity analysis 2 – primary outcome based on Central Administrative Research Facility grading

The analysis population was again the same as for the primary analysis, but with lesion activity at the management visit based on OCT images taken during the visits rather than ophthalmologists' decisions. The same number of complete visits was available, but models including data for all study eyes could not be fitted due to the failure to converge. Instead of sampling one eye randomly (see [Appendix 3](#)), separate models were fitted for left and right study eyes; we decided to do this to make best use of the data collected. The numbers of visits included in models fitted for each index test varied because some participants had a right study eye, some a left study eye and some had two study eyes ([Tables 12](#) and [13](#)). Test summary scores for the whole preceding interval since the previous management visit were modelled.

TABLE 11 Model results for the primary outcome of lesion activity in study eyes for all home-monitoring test models and for a no-test model; sensitivity analysis 1

Numbers in models	No test		MBT		mVT		KSJ	
Visits	1413		1213		1134		1173	
Participants	252		233		210		218	
Eyes	302		279		253		260	
Predictor in model	OR (95% CI)	<i>p</i> -value	OR (95% CI)	<i>p</i> -value	OR (95% CI)	<i>p</i> -value	OR (95% CI)	<i>p</i> -value
Mean MBT score			1 (0.98 to 1.03)	0.679				
Mean mVT score ^a					0.99 (0.98 to 1.01)	0.379		
Median KSJ VA score							1.30 (0.94 to 1.79)	0.112
% KSJ VA worse ^b							2.57 (0.99 to 6.71)	0.053
% KSJ Amsler grid worse ^b							2.10 (0.59 to 7.47)	0.252
% KSJ household object worse ^b							2.38 (0.54 to 10.56)	0.255
Sex								
Male	1		1		1		1	
Female	0.72 (0.29 to 1.78)	0.481	0.8 (0.31 to 2.07)	0.642	0.71 (0.29 to 1.77)	0.466	0.57 (0.22 to 1.47)	0.242
Time since first treatment at baseline strata								
6–17 months	1		1		1		1	
18–29 months	0.67 (0.25 to 1.81)	0.429	0.96 (0.33 to 2.77)	0.938	0.98 (0.35 to 2.69)	0.961	0.54 (0.19 to 1.54)	0.251
30–41 months	0.45 (0.15 to 1.39)	0.165	0.44 (0.13 to 1.48)	0.185	0.47 (0.15 to 1.50)	0.201	0.57 (0.17 to 1.88)	0.352
Age	0.99 (0.93 to 1.06)	0.744	0.99 (0.92 to 1.06)	0.776	0.98 (0.91 to 1.05)	0.566	1.01 (0.94 to 1.09)	0.712
Baseline VA strata								
Better than or equal to 6/18	1		1		1		1	
Worse than 6/18 and better than 6/24	0.72 (0.15 to 3.52)	0.684	1.10 (0.20 to 6.06)	0.912	1.25 (0.24 to 6.63)	0.791	0.71 (0.14 to 3.63)	0.683
Worse than or equal to 6/24	0.64 (0.10 to 4.10)	0.638	0.54 (0.08 to 3.68)	0.531	0.82 (0.13 to 5.31)	0.838	0.81 (0.12 to 5.51)	0.827
Days since baseline ^c	0.52 (0.36 to 0.74)	< 0.001	0.48 (0.32 to 0.73)	< 0.001	0.54 (0.37 to 0.79)	0.001	0.51 (0.34 to 0.74)	0.001

a Transformed by multiplying by 100.

b Proportion of scores across the respective interval that reported vision as being 'worse'.

c Natural log transformed.

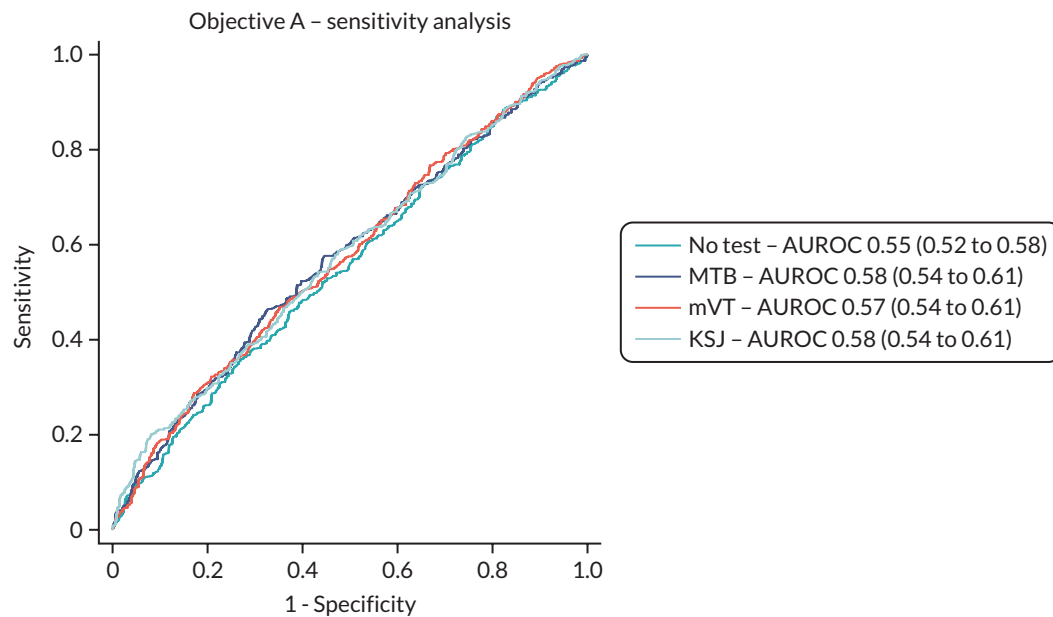


FIGURE 14 Receiver operating characteristic curves and AUROCs (95% CIs) for summary test scores for all index tests; sensitivity analysis 1.

The baseline VA strata of 'Worse than or equal to 6/24' was seen to predict CARF-graded lesion activity perfectly in left eyes, resulting in these observations (about 5%) being excluded from the model.

Results for the no-test model and for all home-monitoring test models are shown in [Tables 12](#) and [13](#) for right and left eyes, respectively. In the baseline model, age was a significant predictor of activity for right eyes but not left eyes. No test summary score was seen to have a significant association with lesion activity. Days since baseline treatment was not significant for either model, contrary to the primary analysis.

Area under the receiver operating characteristics and model performance at thresholds identified by Youden's index for each model are shown in [Figure 15](#); further details of each model are shown in [Appendix 1](#) (see [Table 45](#)). As in the primary analysis, the AUROCs from the models that included the home-monitoring tests were only marginally improved from the model that excluded the tests, with 95% CIs that overlapped across all models irrespective of eye. AUROCs and sensitivity were marginally better in models based on right eyes.

A significant association between the MBT summary score and CARF-graded lesion activity was seen in right eyes (OR 1.05, 95% CI 1.00 to 1.1; p : 0.038), but not in left eyes (OR 0.98, 95% CI 0.94 to 1.02; p = 0.366). For the mVT summary score, a significant association was seen in left eyes (OR 1.04, 95% CI 1.01 to 1.08; p : 0.015), but not in right eyes (OR 0.98, 95% CI 0.95 to 1.00; p : 0.097). An increase in mVT score of 0.01 was associated with a 1.04 increase in odds of lesion activity in the left eye (i.e. in the expected direction). No significant associations were seen for the KSJ home-monitoring test.

As in the primary analysis, the direction of the association observed for the MBT for right eyes is contrary to expectation; the average test score above the cut-off threshold was higher than below the threshold (see [Appendix 1, Table 46](#)). The direction of this association was inverted for right and left eyes. For the mVT, the direction of the association observed for left eyes is contrary to expectation; the average test score above the cut-off threshold was lower than below the threshold (see [Appendix 1, Table 46](#)). The direction of this association was also inverted for right and left eyes.

TABLE 12 Model results for the primary outcome of lesion activity in right study eyes for all home-monitoring test models and for a no-test model; sensitivity analysis 2

Numbers in models	No test		MBT		mVT		KSJ	
Predictor in model	OR (95% CI)	<i>p</i> -value	OR (95% CI)	<i>p</i> -value	OR (95% CI)	<i>p</i> -value	OR (95% CI)	<i>p</i> -value
Visits	643		561		583		552	
Participants	147		135		140		132	
Mean MBT score			1.05 (1 to 1.1)	0.038*				
Mean mVT score ^a					0.98 (0.95 to 1.00)	0.097		
Median KSJ VA score							1.49 (0.78 to 2.83)	0.224
% KSJ VA worse ^b							0.62 (0.09 to 4.37)	0.634
% KSJ Amsler grid worse ^b							0.41 (0.03 to 5.79)	0.507
% KSJ household object worse ^b							3.41 (0.20 to 59.2)	0.399
Sex								
Male	1		1		1		1	
Female	0.58 (0.13 to 2.65)	0.479	0.44 (0.09 to 2.18)	0.312	0.56 (0.13 to 2.42)	0.436	0.66 (0.13 to 3.46)	0.622
Time since first treatment at baseline strata								
6–17 months	1		1		1		1	
18–29 months	2.91 (0.52 to 16.1)	0.222	2.01 (0.34 to 11.90)	0.441	2.95 (0.56 to 15.6)	0.203	3.76 (0.57 to 24.5)	0.167
30–41 months	1.41 (0.22 to 9.22)	0.719	2.40 (0.33 to 17.49)	0.386	2.08 (0.33 to 13.0)	0.434	1.89 (0.25 to 14.3)	0.538
Age	0.86 (0.76 to 0.97)	0.015*	0.93 (0.82 to 1.04)	0.206	0.89 (0.79 to 1.00)	0.053	0.89 (0.79 to 1.01)	0.070
Baseline VA strata								
Better than or equal to 6/18	1		1		1		1	
Worse than 6/18 and better than 6/24	12.40 (0.35 to 436)	0.166	25.9 (0.63 to 1059)	0.086	16.3 (0.55 to 484)	0.107	88.2 (1.28 to 6077)	0.038
Worse than or equal to 6/24	0.81 (0.05 to 12.5)	0.880	0.9 (0.05 to 15.00)	0.940	1.66 (0.10 to 27.6)	0.722	2.37 (0.09 to 61.7)	0.603
Days since baseline ^c	1.10 (0.60 to 2.02)	0.761	1.01 (0.54 to 1.89)	0.970	1.22 (0.67 to 2.23)	0.516	1.11 (0.55 to 2.24)	0.769

* $p < 0.05$.^a Transformed by multiplying by 100.^b Proportion of scores across the respective interval that reported vision as being 'worse'.^c Natural log transformed.

TABLE 13 Model results for the primary outcome of lesion activity in left study eyes for all home-monitoring test models and for a no-test model; sensitivity analysis 2

Numbers in models	No test		MBT		mVT		KSJ	
Predictor in model	OR (95% CI)	<i>p</i> -value	OR (95% CI)	<i>p</i> -value	OR (95% CI)	<i>p</i> -value	OR (95% CI)	<i>p</i> -value
Visits	716		607		620		632	
Participants	148		136		137		130	
Mean MBT score			0.98 (0.94 to 1.02)	0.366				
Mean mVT score ^a					1.04 (1.01 to 1.08)	0.015		
Median KSJ VA score							0.87 (0.44 to 1.73)	0.689
% KSJ VA worse ^b							1.56 (0.14 to 17.3)	0.719
% KSJ Amsler grid worse ^b							2.52 (0.08 to 77.2)	0.596
% KSJ household object worse ^b							3.93 (0.06 to 240)	0.515
Sex								
Male	1		1		1		1	
Female	1.44 (0.26 to 7.87)	0.677	1.19 (0.22 to 6.28)	0.841	1.14 (0.21 to 6.05)	0.882	1.83 (0.31 to 10.9)	0.507
Time since first treatment at baseline strata								
6–17 months	1		1		1		1	
18–29 months	0.52 (0.08 to 3.34)	0.490	1.26 (0.20 to 7.81)	0.803	0.50 (0.08 to 3.18)	0.464	0.48 (0.07 to 3.28)	0.452
30–41 months	0.44 (0.05 to 4.20)	0.473	0.60 (0.07 to 5.17)	0.645	0.28 (0.03 to 2.53)	0.258	0.51 (0.05 to 5.52)	0.582
Age	1.08 (0.95 to 1.23)	0.244	1.05 (0.92 to 1.20)	0.441	1.04 (0.91 to 1.18)	0.566	1.04 (0.91 to 1.18)	0.598
Baseline VA strata								
Better than or equal to 6/18	1		1		1		1	
Worse than 6/18 and better than 6/24	3.52 (0.17 to 74.2)	0.418	8.90 (0.19 to 415)	0.265	0.8 (0.03 to 24.82)	0.896	1.45 (0.06 to 35.01)	0.820
Days since baseline ^c	0.68 (0.35 to 1.31)	0.249	0.66 (0.36 to 1.20)	0.174	0.78 (0.42 to 1.43)	0.418	0.81 (0.38 to 1.72)	0.588

a Transformed by multiplying by 100.

b Proportion of scores across the respective interval that reported vision as being 'worse'.

c Natural log transformed.

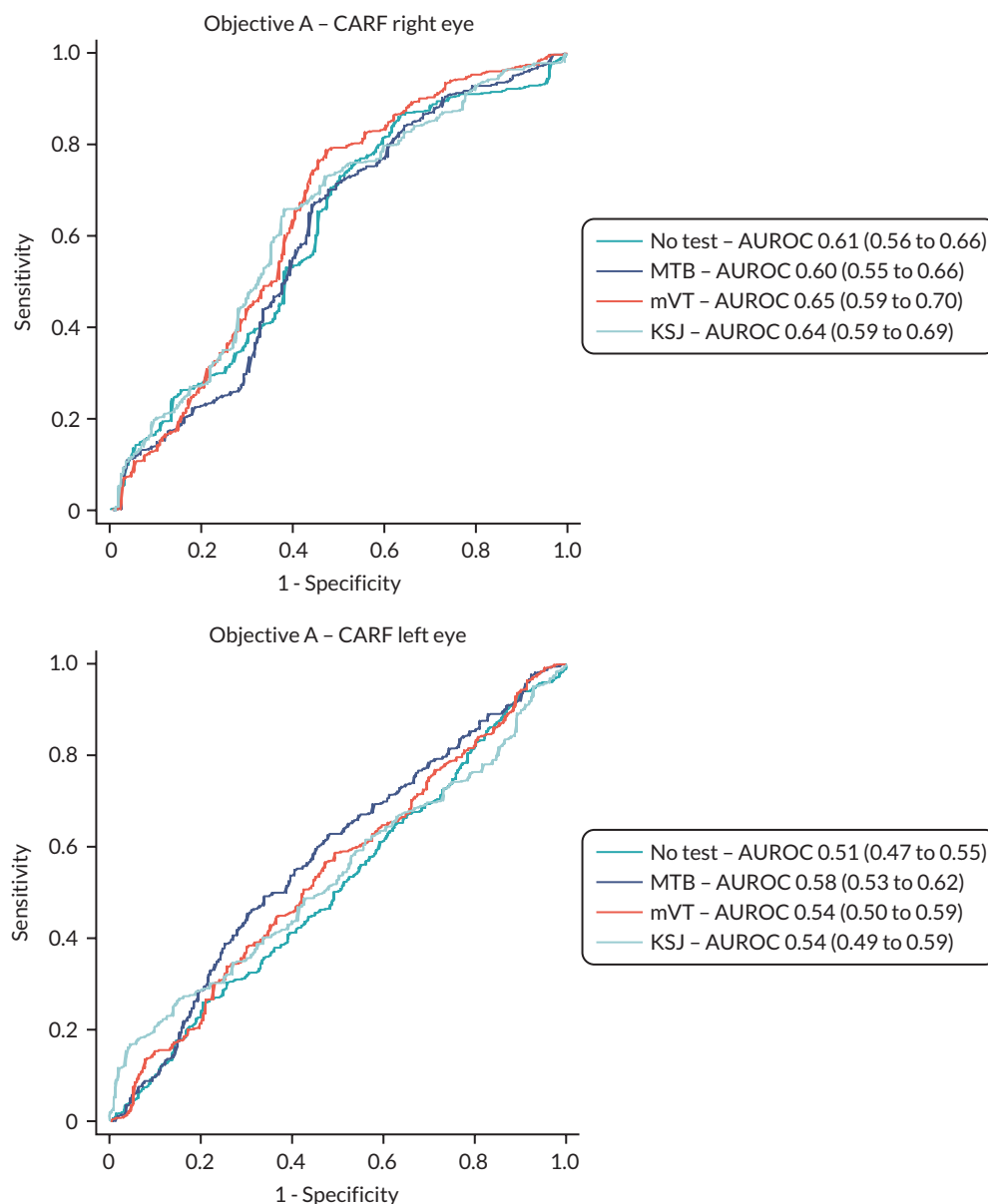


FIGURE 15 Receiver operating characteristic curves and AUROCs (95% CIs) for summary test scores for all index tests; sensitivity analysis 2.

Secondary outcome - change in lesion activity

The ability of test summary scores to predict the secondary outcome, a change in lesion activity from inactive to active as determined by management decisions at pairs of management visits adjacent in time, was quantified. Lesion activity at baseline was used to define the first pair of management visits. Only complete visits, that is, visits on which an OCT image was taken, were included.

Models were constructed using the same approach taken in the primary analysis, except days since baseline was not natural log transformed. A model for the mVT summary scores did not converge. Therefore, there are no results for this test.

The analysis population was the same as for the primary analysis. There were similar numbers of pairs of visits as single visits due to using lesion activity classification at the baseline visit for the first pair in a participant's follow-up. All available dates were included, that is, visit records did not depend on having complete data. The number of visits included in models fitted for each index test varied

because participants did not always complete all tests at all visits (Table 14). Results from the models for this secondary outcome are shown in Table 14. No significant associations were seen for any home-monitoring test or other covariate.

TABLE 14 Model results for the secondary outcome of change from inactive to active lesion in study eyes for all home-monitoring test models and for a no-test model

	No test		MBT		KSJ	
Visit pairs	1413		1213		1238	
Inactive to active transitions	131 (9.3%)		104 (8.6%)		120 (9.7%)	
Participants	252		233		224	
Predictor in model	OR (95% CI)	p-value	OR (95% CI)	p-value	OR (95% CI)	p-value
Mean MBT score			1.00 (0.99 to 1.02)	0.601		
Median KSJ VA score					1.11 (0.89 to 1.38)	0.372
% KSJ VA worse ^b					1.21 (0.43 to 3.40)	0.715
% KSJ Amsler grid worse ^b					0.64 (0.15 to 2.71)	0.543
% KSJ household object worse ^b					1.82 (0.37 to 9.05)	0.464
Sex						
Male	1					
Female	1.05 (0.69 to 1.61)	0.815	1.18 (0.75 to 1.84)	0.475	1.00 (0.64 to 1.54)	0.989
Time since first treatment at baseline strata						
6–17 months						
18–29 months	0.79 (0.49 to 1.28)	0.349	0.76 (0.45 to 1.26)	0.284	0.81 (0.50 to 1.29)	0.37
30–41 months	0.95 (0.54 to 1.68)	0.860	0.96 (0.53 to 1.75)	0.903	0.91 (0.50 to 1.64)	0.747
Age	1.02 (0.99 to 1.05)	0.190	1.02 (0.99 to 1.06)	0.194	1.03 (0.996 to 1.06)	0.091
Baseline VA strata						
Better than or equal to 6/18						
Worse than 6/18 and better than 6/24	0.56 (0.21 to 1.52)	0.255	0.77 (0.28 to 2.17)	0.625	0.66 (0.24 to 1.85)	0.428
Worse than or equal to 6/24	0.93 (0.31 to 2.82)	0.904	0.80 (0.23 to 2.82)	0.727	1.06 (0.33 to 3.42)	0.928
Days since baseline [*]	1.001 (0.999 to 1.002)	0.236	1.001 (1.000 to 1.003)	0.112	1.001 (0.999 to 1.002)	0.452

* $p < 0.05$.

a Transformed by multiplying by 100.

b Proportion of scores across the respective interval that reported vision as being 'worse'.

c Natural log transformed.

Note

mVT test not shown because no model achieved convergence.

Receiver operating characteristic curves and AUROCs for summary test scores for tests (but not for mVT) are shown in [Figure 16](#); further details of each model, including model performance at thresholds identified by Youden's index, are shown in [Appendix 1](#) (see [Table 47](#)). AUROCs from the models that included the home-monitoring tests were only marginally improved compared to the no-test model, with 95% CIs that overlapped across all models. Sensitivities and specificities were broadly similar for the no-test model and the MTB model. The KSJ model showed decreased sensitivity and increased specificity compared to the no-test model. Average test scores above and below the predicted probability thresholds are shown in [Appendix 1, Table 48](#). As in previous analyses, the average MBT score above the threshold (signifying lesion activity transition) was higher than below the threshold, contrary to expectations.

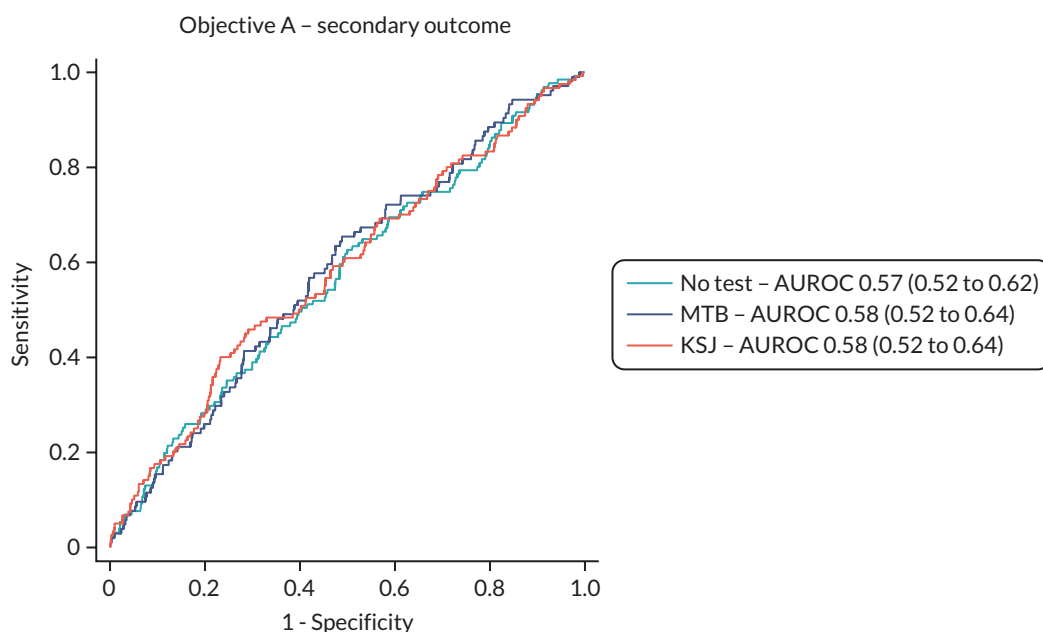


FIGURE 16 Receiver operating characteristic curves and AUROCs (95% CIs) for summary test scores for index tests (except for the mVT) for a change from inactive to active lesion status.

Chapter 5 Results: Objective C

Impacts of inequalities on willingness in principle to participate

During mapping to IMD, six patients were identified as having erroneous IMD ranks: four due to residing in the Isle of Man and two due to having out of range IMD ranks. These patients were excluded from analysis. Due to the small number of participants under 60 years of age (11 participants; no participation = 6, participation = 5) and over 89 years of age (32 participants; no participation = 30, participation = 2), participant age was split into three categories: under 70 years, 70 – 79 years and 80 years and older.

This analysis was performed on 936 patients (not willing, including 129 for whom willingness was pending = 645, willing = 291). Characteristics of these patients are described in [Table 6](#). The results of the analysis are given in [Table 15](#).

There was no significant effect of time since first treatment for nAMD on the odds of willingness to participate in the study. There was a 21% decrease in the odds of women participating in the study, although this was not significant (OR 0.79, 95% CI 0.58 to 1.09). An overall effect of IMD quintile was highly significant ($\chi^2 = 24.3$, $p < 0.001$). Patients from the most deprived areas (IMD quintile 5) had a 47% decrease in odds of being willing compared to those from the middle quintile-deprived areas (OR 0.53, 95% CI 0.32 to 0.88); those in the second quintile had a 1.8-fold increase in odds of participation (OR 1.76, 95% CI 1.12 to 2.76).

Older patients were seen to have significantly decreased odds of being willing compared to patients < 70 years old (overall of the effect of age category, $\chi^2 = 50.5$, $p < 0.001$). Patients between 70 and 79 years of age had 37% decreased odds (OR 0.63, 95% CI 0.41 to 0.97) and patients aged 80 years and over had 79% decreased odds (OR 0.21, 95% CI 0.13 to 0.35).

Decreased odds of participation were seen when potential study eyes had visual acuity worse than 6/18 at baseline compared to better than or equal to 6/18. However, this impact was not significant overall ($\chi^2 = 1.90$, $p = 0.387$) or when either of the two strata with poorer vision were compared with the stratum of potential study eyes with vision better than or equal to 6/18.

Significant effects of study centre were seen on participation when compared with Belfast (overall of the effect of age category, $\chi^2 = 98.0$, $p < 0.001$), most notably with Moorfields showing a 41% decrease in odds of patient willingness (OR, 0.59, 95% CI 0.35 to 0.97) and James Paget showing a 7.83-fold increase (OR 7.83, 95% CI 4.62 to 13.28).

Impact of inequalities on recruitment

[Figure 6](#) shows that, although 312 patients were willing in principle, only one patient declined to take part when formal consent was sought. Of the other 14, 9 were found to be ineligible when attending for information and training and eligibility was pending for 4 (did not attend for information and training). Thus, 297/298 patients consented. No further analysis of the outcome consent to participate was undertaken.

Impact of inequalities on ability and adherence to weekly testing

Ability to perform an index test was defined as the proportion of qualifying management visits for which some valid index test data were available. Adherence was defined as the proportion of weeks in

TABLE 15 Impact of inequalities on participation

Predictor	OR	95% CI	p-value
Time since first treatment at baseline strata			
6–17 months	1.00		
18–29 months	0.95	(0.66 to 1.36)	0.784
30–41 months	1.13	(0.75 to 1.70)	0.553
Sex	0.79	(0.58 to 1.09)	0.149
Male			
Female			
IMD quintile			
1	1.47	(0.86 to 2.50)	0.157
2	1.76	(1.12 to 2.76)	0.015
3	1.00		
4	0.96	(0.59 to 1.57)	0.879
5 – most deprived	0.53	(0.32 to 0.88)	0.015
Age			
< 70 years	1.00		
70–79 years	0.63	(0.41 to 0.97)	0.037
≥ 80 years	0.21	(0.13 to 0.35)	< 0.001
Baseline VA strata			
Better than or equal to 6/18	1.00		
Worse than 6/18 and better than 6/24	0.73	(0.41 to 1.30)	0.278
Worse than or equal to 6/24	0.74	(0.38 to 1.43)	0.368
Study centre			
Belfast	1.00		
Liverpool	1.14	(0.71 to 1.82)	0.584
Southampton	0.69	(0.40 to 1.18)	0.174
Moorfields	0.59	(0.35 to 0.97)	0.039
James Paget	7.83	(4.62 to 13.28)	< 0.001
Gloucester	0.90	(0.48 to 1.66)	0.729

between qualifying management visits for which some valid data for an index test were available. The ability and adherence models were performed for each test separately at the level of the patient. Most participants used a smartphone, tablet, laptop/home computer, internet, e-mail or social media at least weekly [232/259 (89.6%)].

Frequency of testing and weeks when tests were completed are also shown in [Appendix 1](#) (see [Tables 37–39](#)). Analyses of the impact of inequalities on ability to perform test and adherence to weekly testing are based on the population of 297 consented participants. The results of the analysis of participant ability to perform tests are shown in [Table 16](#) and the results of the analysis of participant adherence to perform tests weekly in [Table 17](#).

TABLE 16 Impact of inequalities on patient ability to perform tests

Predictor	MBT			mVT			KSJ		
	OR	95% CI	p-value	OR	95% CI	p-value	OR	95% CI	p-value
Time since first treatment for nAMD									
6–17 months	1.00			1.00			1.00		
18–29 months	0.66	(0.31 to 1.41)	0.284	0.73	(0.35 to 1.53)	0.410	0.97	(0.43 to 2.20)	0.939
30–41 months	1.38	(0.60 to 3.15)	0.449	0.99	(0.42 to 2.33)	0.974	1.04	(0.46 to 2.35)	0.933
Sex									
Male	1.00			1.00			1.00		
Female	0.67	(0.33 to 1.39)	0.285	0.90	(0.44 to 1.81)	0.762	0.93	(0.46 to 1.85)	0.833
IMD quintile									
1 – least deprived	2.05	(0.75 to 5.59)	0.162	2.30	(0.81 to 6.55)	0.120	1.00	(0.36 to 2.81)	0.995
2	1.37	(0.52 to 3.64)	0.522	0.90	(0.36 to 2.27)	0.825	1.18	(0.48 to 2.89)	0.718
3	1.00			1.00			1.00		
4	1.08	(0.45 to 2.61)	0.865	0.86	(0.36 to 2.06)	0.732	0.88	(0.33 to 2.35)	0.797
5 – most deprived	0.66	(0.23 to 1.91)	0.438	0.91	(0.30 to 2.70)	0.860	4.07	(0.95 to 17.5)	0.059
Age									
< 70 years	1.00			1.00			1.00		
70–79 years	0.95	(0.36 to 2.51)	0.922	1.29	(0.51 to 3.24)	0.586	0.99	(0.40 to 2.48)	0.985
80+ years	0.61	(0.20 to 1.82)	0.371	0.86	(0.31 to 2.41)	0.778	2.13	(0.70 to 6.51)	0.185
Baseline VA strata									
Better than or equal to 6/18				1.00					
Worse than 6/18 and better than 6/24				0.54	(0.20 to 1.44)	0.219			
Worse than or equal to 6/24				5.84	(0.70 to 48.7)	0.103			

continued

TABLE 16 Impact of inequalities on patient ability to perform tests (continued)

Predictor	MBT			mVT			KSJ		
	OR	95% CI	<i>p</i> -value	OR	95% CI	<i>p</i> -value	OR	95% CI	<i>p</i> -value
Exposure to technology									
No exposure									
Exposure	2.05	(0.69 to 6.07)	0.194	1.62	(0.60 to 4.36)	0.336	0.47	(0.16 to 1.40)	0.175
Study centre									
Belfast	1.00			1.00			1.00		
Liverpool	1.15	(0.33 to 4.03)	0.831	1.22	(0.44 to 3.42)	0.704	2.62	(0.89 to 7.69)	0.079
Southampton	1.04	(0.32 to 3.34)	0.949	1.09	(0.33 to 3.59)	0.892	0.95	(0.33 to 2.75)	0.929
Moorfields	0.89	(0.27 to 2.88)	0.843	0.72	(0.23 to 2.27)	0.577	0.60	(0.23 to 1.61)	0.311
James Paget	0.21	(0.08 to 0.54)	0.001	0.18	(0.07 to 0.47)	< 0.001	1.04	(0.43 to 2.53)	0.935
Gloucester	0.71	(0.16 to 3.23)	0.658	0.99	(0.22 to 4.45)	0.992	2.35	(0.67 to 8.29)	0.183

TABLE 17 Impact of inequalities on patient adherence to tests

Predictor	MBT			mVT			KSJ		
	OR	95% CI	p-value	OR	95% CI	p-value	OR	95% CI	p-value
Time since first treatment for nAMD									
6–17 months	1.00			1.00			1.00		
18–29 months	1.13	(0.70 to 1.83)	0.615	1.16	(0.73 to 1.86)	0.521	1.33	(0.79 to 2.23)	0.281
30–41 months	1.09	(0.64 to 1.88)	0.746	1.04	(0.60 to 1.80)	0.890	0.87	(0.52 to 1.45)	0.584
Sex									
Male	1.00			1.00			1.00		
Female	0.90	(0.58 to 1.38)	0.620	1.03	(0.68 to 1.58)	0.883	1.09	(0.70 to 1.7)	0.693
IMD quintile									
1 – least deprived	1.18	(0.60 to 2.33)	0.627	1.32	(0.69 to 2.54)	0.403	0.88	(0.44 to 1.76)	0.724
2	0.95	(0.52 to 1.73)	0.857	0.87	(0.47 to 1.60)	0.653	0.90	(0.49 to 1.63)	0.716
3	1.00			1.00			1.00		
4	0.69	(0.37 to 1.29)	0.245	0.68	(0.36 to 1.29)	0.240	0.74	(0.39 to 1.42)	0.369
5 – most deprived	0.98	(0.46 to 2.08)	0.949	1.15	(0.56 to 2.38)	0.699	2.11	(1.07 to 4.16)	0.032*
Age									
< 70 years	1.00			1.00			1.00		
70–79 years	0.97	(0.58 to 1.64)	0.917	1.05	(0.62 to 1.77)	0.850	0.81	(0.46 to 1.42)	0.454
80+ years	0.85	(0.47 to 1.55)	0.593	1.12	(0.62 to 2.03)	0.715	1.45	(0.74 to 2.84)	0.279
Baseline VA strata									
Better than or equal to 6/18	1.00			1.00			1.00		
Worse than 6/18 and better than 6/24	1.12	(0.55 to 2.28)	0.757	1.09	(0.51 to 2.31)	0.824	1.10	(0.52 to 2.31)	0.808
Worse than or equal to 6/24	4.14	(2.25 to 7.62)	< 0.001	3.10	(1.60 to 6.01)	0.001	2.22	(1.15 to 4.30)	0.018

continued

TABLE 17 Impact of inequalities on patient adherence to tests (continued)

Predictor	MBT			mVT			KSJ		
	OR	95% CI	<i>p</i> -value	OR	95% CI	<i>p</i> -value	OR	95% CI	<i>p</i> -value
Exposure to technology									
No exposure	1.00			1.00			1.00		
Exposure	1.66	(0.81 to 3.37)	0.164	1.77	(0.91 to 3.44)	0.093	0.66	(0.30 to 1.45)	0.301
Study centre									
Belfast	1.00			1.00			1.00		
Liverpool	1.11	(0.57 to 2.17)	0.750	1.01	(0.53 to 1.92)	0.967	1.73	(0.89 to 3.37)	0.104
Southampton	0.85	(0.44 to 1.62)	0.617	0.72	(0.37 to 1.38)	0.321	0.96	(0.50 to 1.82)	0.894
Moorfields	0.77	(0.39 to 1.49)	0.433	0.73	(0.37 to 1.46)	0.379	0.57	(0.29 to 1.13)	0.109
James Paget	0.29	(0.16 to 0.52)	< 0.001	0.31	(0.17 to 0.57)	< 0.001	0.77	(0.44 to 1.34)	0.348
Gloucester	1.06	(0.46 to 2.44)	0.892	1.07	(0.47 to 2.43)	0.874	1.55	(0.69 to 3.46)	0.288

* *p* < 0.05.

For ability to test, there were overall associations with VA stratum at baseline for the KSJ and MBT ($\chi^2 = 768$ and 672 , respectively, both $p \leq 0.001$, although the associations were not in a consistent direction), and with centre for the mVT ($\chi^2 = 21.7$, $p < 0.001$). Participants from James Paget showed a significant decrease in odds of ability when compared to participants from Belfast, with a 79% reduction in odds of ability for MBT (OR 0.21, 95% CI 0.08 to 0.54) and an 82% reduction in odds of ability for mVT (OR 0.18, 95% CI 0.07 to 0.47). There were no associations of ability to test with time since first treatment for nAMD, sex, age, IMD or exposure to technology on participant ability to perform home testing was seen for any test.

For adherence to testing, there were overall associations with VA stratum at baseline for the MBT and mVT ($\chi^2 = 20.9$ and 11.3 , $p < 0.001$ and 0.004 , respectively, although the associations were not in a consistent direction), and with centre for the MBT and mVT ($\chi^2 = 21.8$ and 17.8 , respectively, both $p < 0.001$). Participants from James Paget showed a significant decrease in odds of adherence when compared to participants from Belfast, with a 71% reduction in odds of adherence for MBT (OR 0.29, 95% CI 0.16 to 0.52) and a 69% reduction in odds of adherence for mVT (OR 0.31, 95% CI 0.17 to 0.57). IMD had a significant impact on KSJ testing adherence ($\chi^2 = 12.5$, $p = 0.016$), with participants in the most-deprived quintile showing 2.1 times increased odds of adherence compared to those from the middle quintile (OR 2.11, 95% CI 1.07 to 4.16). There were no associations of adherence to testing with time since first treatment for nAMD, sex, age or exposure to technology.

Chapter 6 Results: Objective D

Analytic approach and methods

The test accuracy of the index tests was examined for two outcomes:

- **Primary analysis outcome:** the reference standard of an ophthalmologist's classification of a fellow eye having active disease at a management visit, that is, conversion to active nAMD.
- **Sensitivity analysis 1:** the reference standard of the classification of a study eye at a complete management visit as having either active or inactive disease, using test data for the 4 weeks preceding the management visit.
- **Sensitivity analysis 2:** the alternative reference standard of classification of a fellow eye having active disease at a management visit, that is, conversion to nAMD, based on grading by CARF of OCT images taken during the management visits ([Table 18](#)).

A sensitivity analysis of the primary outcome that using test summary scores calculated for the 4-week period immediately prior to management visit was also carried out.

Out of 132 participants with a fellow eye and some home-monitoring test data, 17 (12.9%) were recorded as having nAMD during at least one management visit. This rate of conversion was higher than expected based on epidemiological studies of conversion rates in unaffected fellow eyes, potentially due to study eyes having had nAMD longer ago. No lesion, uncertain and inactive reference standard classifications were combined for all analyses; only active classifications were considered to represent conversion. (Each participant could only have one fellow eye because each participant had to have a study eye.) All available dates were included, that is, visit records did not depend on having complete data. However, participants did not have home-monitoring data for all tests. The numbers of participants with home-monitoring data and fellow eye conversions to nAMD disease for the three tests are shown in [Table 18](#).

Separate models were generated for each test, summary test score and outcome and for a no-test model. Baseline VA strata could not be included because the better than or equal to 6/18 stratum predicted success perfectly and the worse than or equal to 6/24 stratum exhibited collinearity. Days since baseline was not included as a fixed component because its inclusion suggested model misspecification. For the KSJ models, only the household object and median near VA score could be included due to convergence issues or high SDs with the % Amsler grid worse and % VA worse scores. In addition, postestimation of the KSJ model suggested model misspecification that could not be controlled. All models were constructed with independent covariance, as other covariance structures resulted in failure of convergence. Each model fitted a random effect of participant and allowed random slopes for days since baseline nested within participant. As for Objective A, test summary scores for the mVT were multiplied by 100 before fitting the models, so that OR estimates correspond to changes

TABLE 18 Number of participants with fellow eye conversions to nAMD and home-monitoring test data

Home-monitoring test	Management decisions		CARF grading	
	Participants with data	Participants with conversions (%)	Participants with data	Participants with conversions (%)
MBT	118	17 (14.4)	119	17 (14.3)
mVT	126	17 (13.5)	127	17 (13.4)
KSJ	116	16 (13.8)	117	16 (13.7)

in ORs for every 0.01 change in mVT score. No interactions were tested. For eyes that converted, follow-up data (reference classifications and test summary scores) were not included after conversion to active nAMD was identified.

Model performance in all cases was examined by assessing OR estimates along with their CIs and by examining model AUROCs. AUROCs were calculated conditional on the random effects and on the fixed effects only.

Primary analysis

[Table 19](#) shows the numbers of participants and visits with data for models fitted for each index test.

Odds ratios estimates for index test summary scores and AUROCs for each model are also shown in [Table 19](#). Neither participants' sex nor time since first treatment at baseline was associated with conversion of fellow eyes to nAMD disease in any model. Increasing age at baseline was associated with conversion, significantly so in the model that included the KSJ summary test scores (OR 1.18 per year of age, 95% CI 1.02 to 1.36; $p = 0.023$) and with borderline significance in the MBT and no-test models.

Only the KSJ home object summary score was significantly associated with conversion; patients who reported more often that their vision was 'worse than baseline' were more likely to have a fellow eye conversion (OR 15.33 for a household object percentage worse summary score of 100 vs. 0, 95% CI 1.19 to 196.83; $p = 0.036$). However, this estimate should be interpreted with caution, due to the imprecise CI and potential misspecification of the model.

AUROCs and model performance at thresholds identified by Youden's index for each model are shown in [Figure 17](#), with further details in [Appendix 1](#) (see [Table 49](#)). The AUROCs from the models that included the MBT and mVT home-monitoring tests were only marginally higher than for the no-test model, while the AUROC from the KSJ model was moderately higher. All AUROCs had wide 95% CIs that overlapped across all models. Sensitivity and specificity for all models were reasonable (≥ 0.69).

Average test scores above and below the predicted probability thresholds are shown in [Appendix 1](#) (see [Table 50](#)). Mean MBT test scores were lower above the identified threshold (prediction of fellow eye conversion) than below, with non-overlapping CIs. Mean mVT scores were higher above the identified threshold (prediction of fellow eye conversion) than below. In contrast to several previous MBT and mVT models, these differences in means above and below test thresholds were in the expected direction associated with predictions fellow eye conversion, again with non-overlapping CIs. The percentage KSJ home object summary score above the threshold as 12% (home object test 'worse than baseline') compared to 1% below the threshold. The median self-scored VA assessment was very similar above and below the threshold.

Sensitivity analysis 1 (data for preceding 4 weeks only)

[Table 20](#) shows the numbers of participants and visits with data for models fitted for each index test. ORs estimates for index test summary scores and AUROCs for each model are also shown in [Table 20](#).

As in the primary analysis, baseline VA strata and days since baseline treatment could not be included in the models. The no-test model was the same. For the KSJ models, the Amsler grid percentage worse summary score was not included due to a very high SD and model misspecification that could not be controlled. All models were constructed with independent covariance, as models with other covariance structures did not converge.

TABLE 19 Home-monitoring test accuracy for conversion of fellow eyes against management decisions

Numbers in models	No test		MBT		mVT		KSJ	
Visits	544		468		479		480	
Participants	132		118		126		115	
Predictor in model	OR (95% CI)		p-value		OR (95% CI)		p-value	
Mean MBT score			0.99 (0.96 to 1.03)		0.628			
Mean mVT score ^a					1.02 (0.99 to 1.06)		0.137	
Median KSJ VA score							0.60 (0.25 to 1.44)	
% KSJ household object worse ^b							15.33 (1.19 to 197)	
Sex								
Male	1		1		1		1	
Female	3.9 (0.68 to 22.23)		0.126		3.58 (0.78 to 16.53)		0.102	
2.79 (0.67 to 11.62)							0.158	
2.25 (0.52 to 9.64)							0.275	
Time since first treatment at baseline strata								
6–17 months	1		1		1		1	
18–29 months	0.94 (0.18 to 4.91)		0.939		0.95 (0.22 to 4.16)		0.945	
30–41 months	0.46 (0.03 to 6.35)		0.559		0.47 (0.05 to 4.90)		0.529	
Age	1.13 (1 to 1.28)		0.050		1.11 (1.00 to 1.24)		0.053	
					1.09 (0.98 to 1.22)		0.121	
							1.18 (1.02 to 1.36)	
							0.023	

^a Transformed by multiplying by 100.

^b Proportion of scores across the respective interval that reported vision as being 'worse'.

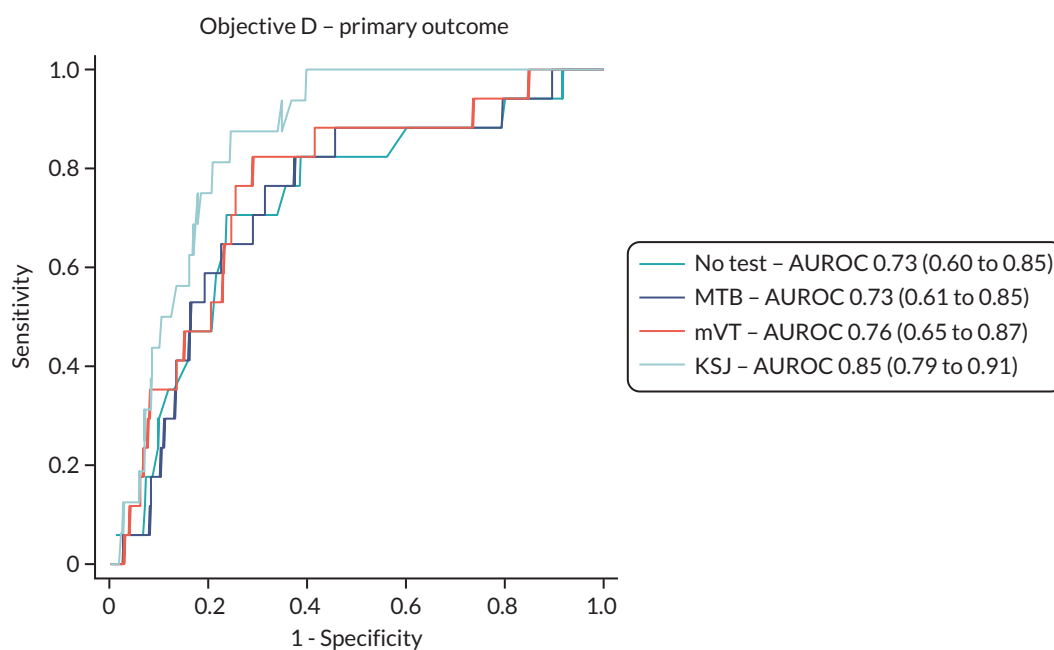


FIGURE 17 Receiver operating characteristic curves and AUROCs (95% CIs) for summary test scores for index tests for conversion of a fellow to an active lesion; primary analysis.

Neither sex nor time since first treatment at baseline was associated with conversion of fellow eyes to nAMD, in any model. Increasing age at baseline was associated with conversion, again significantly so in the model that included the KSJ summary test scores (OR 1.21 per year of age, 95% CI 1.04 to 1.40; $p = 0.014$) and with borderline significance in the no-test model (OR 1.13 per year of age, 95% CI 1.00 to 1.28; $p = 0.050$).

As in the primary analysis, only the KSJ home object test summary score was significantly associated with conversion (OR 15.33 for the percentage worse summary score of 100 vs. 0, 95% CI 1.19 to 196.83; $p = 0.036$). This estimate should be interpreted with caution, due to the imprecise CI and potential misspecification issues.

AUROC and model performance at thresholds identified by Youden's index for each model are shown in [Table 20](#) and [Figure 18](#), with further details in [Appendix 1](#) (see [Table 51](#)). The AUROCs from the models that included the home-monitoring tests were similar to those in the primary analysis and, again, only marginally improved from the model that excluded the tests, with wide 95% CIs that overlapped across all models.

Average test scores above and below the predicted probability thresholds are shown in [Appendix 1](#) (see [Table 52](#)). The pattern was the same as in the primary analysis.

Sensitivity analysis 2 (lesion activity classified by Central Administrative Research Facility grading)

[Table 21](#) shows the numbers of participants and visits with data for models fitted for each index test. OR estimates for index test summary scores and AUROCs for each model are also shown in [Table 21](#).

Model fitting was constrained as for the primary analysis and first sensitivity analysis, except that days since baseline could be included. Predictors that were not to do with the tests showed similar associations; the association of age with conversion was of borderline significance. In this analysis, the mVT and KSJ median VA test summary scores were significantly associated with conversion

TABLE 20 Fixed interval home-monitoring test accuracy for conversion of fellow eyes against management decisions

Numbers in models	No test		MBT		mVT		KSJ	
Visits	544		442		447		456	
Participants	132		109		114		114	
Eyes	544		442		447		456	
Predictor in model	OR (95% CI)	<i>p</i> -value	OR (95% CI)	<i>p</i> -value	OR (95% CI)	<i>p</i> -value	OR (95% CI)	<i>p</i> -value
Mean MBT score			0.97 (0.93 to 1.01)	0.184				
Mean mVT score ^a					1.04 (0.99 to 1.09)	0.106		
Median KSJ VA score							0.69 (0.25 to 1.90)	0.478
% KSJ VA worse ^b							9.66 (0.89 to 105)	0.062
% KSJ household object worse ^b							15.11 (1.94 to 118)	0.010
Male	1		1		1		1	
Female	3.90 (0.68 to 22.2)	0.126	3.82 (0.58 to 25.1)	0.163	3.13 (0.47 to 21.0)	0.240	1.03 (0.22 to 4.76)	0.968
Time since first treatment at baseline strata								
6–17 months	1		1		1		1	
18–29 months	0.94 (0.18 to 4.91)	0.939	0.85 (0.13 to 5.38)	0.863	0.94 (0.13 to 6.69)	0.954	1.68 (0.33 to 8.51)	0.529
30–41 months	0.46 (0.03 to 6.35)	0.559	0.39 (0.03 to 5.88)	0.500	0.49 (0.03 to 9.19)	0.631	0.63 (0.08 to 5.12)	0.664
Age	1.13 (1.00 to 1.28)	0.050	1.11 (0.97 to 1.26)	0.129	1.09 (0.95 to 1.24)	0.228	1.21 (1.04 to 1.4)	0.014

^a Transformed by multiplying by 100.

^b Proportion of scores across the respective interval that reported vision as being 'worse'.

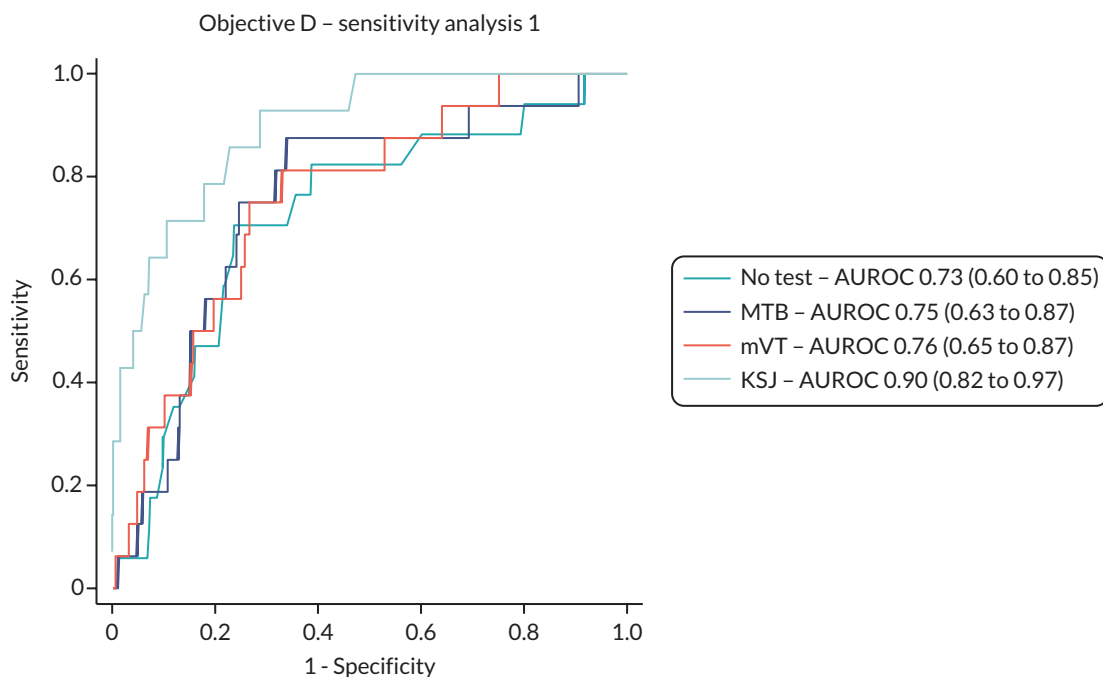


FIGURE 18 Receiver operating characteristic curves and AUROCs (95% CIs) for summary test scores for index tests for conversion of a fellow to an active lesion; sensitivity analysis 1.

(respectively, OR 1.06 per 0.01 logMAR unit, 95% CI 1.01 to 1.12; $p = 0.012$ and OR 0.19 per one VA step increase, 95% CI 0.04 to 0.97; $p = 0.046$). The latter estimate was unexpected, as higher KSJ VA scores relate to better VA.

Area under the receiver operating characteristics and model performance at thresholds identified by Youden's index for each model are shown in [Figure 19](#), with further details in [Appendix 1, Table 51](#). The AUROCs from the models that included the home-monitoring tests were worse than those in the primary analysis, including the no-test model. However, as in other analyses, AUROCs for models including summary test scores were not markedly larger than for the no-test model, with wide 95% CIs that overlapped across all models.

Average test scores above and below the predicted probability thresholds are shown in [Appendix 1](#) (see [Table 52](#)). The pattern was the same as in the primary analysis and the first sensitivity analysis.

TABLE 21 Home-monitoring test accuracy for conversion of fellow eyes against reference standard of CARF grading

Numbers in models	No test		MBT		mVT		KSJ	
Visits	537		461		472		470	
Participants	131		117		125		114	
Eyes	537		461		472		470	
Predictor in model	OR (95% CI)	<i>p</i> -value	OR (95% CI)	<i>p</i> -value	OR (95% CI)	<i>p</i> -value	OR (95% CI)	<i>p</i> -value
Mean MBT score			0.97 (0.92 to 1.03)	0.365				
Mean mVT score ^a					1.06 (1.01 to 1.12)	0.012		
Median KSJ VA score							0.19 (0.04 to 0.97)	0.046
% KSJ VA worse ^b							1.37 (0.02 to 89.7)	0.883
% KSJ household object worse ^b							0.76 (0.01 to 76.5)	0.908
Sex								
Male	1		1		1		1	
Female	1.74 (0.25 to 12.2)	0.576	3.06 (0.43 to 21.9)	0.265	2.26 (0.32 to 15.9)	0.412	1.34 (0.18 to 10.3)	0.778
Time since first treatment at baseline strata								
6–17 months	1		1		1		1	
18–29 months	0.08 (0.01 to 1.03)	0.053	0.14 (0.01 to 1.65)	0.119	0.10 (0.01 to 1.46)	0.093	0.2 (0.02 to 2.31)	0.197
30–41 months	0.14 (0.01 to 2.37)	0.172	0.11 (0.01 to 2.11)	0.142	0.23 (0.02 to 3.42)	0.287	0.1 (0.00 to 2.55)	0.161
Age	1.10 (0.94 to 1.28)	0.237	1.08 (0.92 to 1.25)	0.348	1.05 (0.90 to 1.23)	0.499	1.10 (0.93 to 1.30)	0.270
Days since baseline	1.00 (0.99 to 1)	0.079	1.00 (0.99 to 1.00)	0.066	1.00 (0.99 to 1.00)	0.074	1.00 (0.99 to 1.00)	0.073

^a Transformed by multiplying by 100.

^b Proportion of scores across the respective interval that reported vision as being 'worse'.

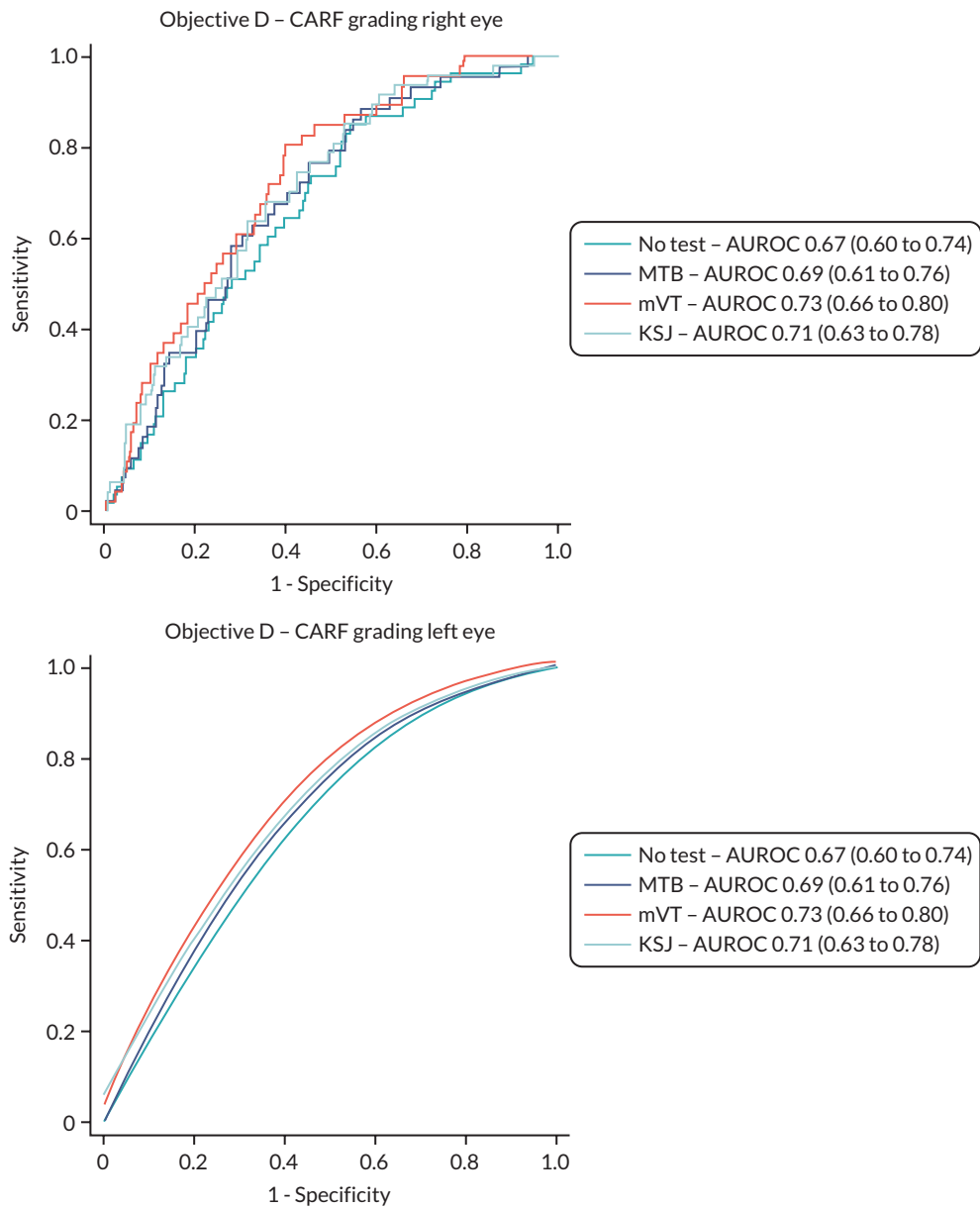


FIGURE 19 Receiver operating characteristic curves and AUROCs (95% CIs) for summary test scores for index tests for conversion of a fellow to an active lesion; sensitivity analysis 2.

Chapter 7 Results: Objective E

It is important to keep in mind the participants' characteristics when considering the challenges experienced with the index tests that were being evaluated. The study population had used widely available technologies quite extensively: 67% used a smartphone, 85% used the internet at home and 72% used e-mail at least weekly (see [Table 5](#)).

The challenges that participants experienced in testing at home with the electronic mVT and MBT apps contributed to withdrawals (see [Table 8](#)) and to the participants' ability to adhere to the requested testing frequency.

Participant helpline

Incoming calls

Details of incoming calls to the participant helpline are given in [Table 22](#), with the reasons for incoming calls detailed in [Table 23](#). For each call to the helpline, a maximum of two reasons were recorded. A total of 353 incoming calls were received, all in relation to electronic testing, although a small proportion (3.7%) were for reassurance about success of data transmission; 46% of reasons concerned one or both electronic apps and 26% concerned the devices or connectivity. A total of 141 (47.5%) participants made at least one call to the helpline and the percentage of participants varied by site (27.5–71.4%).

Outgoing calls to participants

Calls were made to patients in batches if app data had not been received within 2 weeks of consent or, thereafter, if no new test data were received by 3 weeks after the previous test; participants who had been called recently were sometimes not called again immediately if they triggered this criterion. Details of outgoing calls and reasons for them are described in [Tables 24](#) and [25](#). For each outgoing call, a maximum of two reasons for lack of data were recorded. A total of 272 calls were made with 66% being answered. Frequent reasons for lack of data were connectivity issues (23.9%) and personal or home circumstances (22.5%). Outgoing calls were suspended between March and June 2020 during the COVID-19 pandemic. In total, at least one outgoing call was made to 131 participants (44.1%). The number of calls was roughly proportional to the number of participants recruited at each site and participants who spent longer total time in study required on average more outgoing calls to be made.

TABLE 22 Details of incoming helpline calls

Centre	Total number and duration of calls, <i>n</i> (minutes)	Median call duration, minutes (IQR)	Total time of participants in study (months)
Belfast	135 (784.0)	5 (3–8)	1090.2
Liverpool	39 (227.0)	4 (2–7)	446.0
Southampton	61 (426.0)	5 (3–10)	508.5
Moorfields	37 (207.0)	5 (3–8)	411.3
James Paget	39 (154.0)	3 (2–5)	937.6
Gloucester	30 (124.0)	3 (2–5)	206.1
Unknown	12 (59.0)	5 (2–6)	
Overall	353 (1981.0)	5 (2–7)	3599.6

TABLE 23 Reasons for incoming helpline calls

Reason for call ^a	Number of calls, n (%)
Multitbit	121 (27.8)
mVT	67 (15.4)
Connectivity	58 (13.3)
SMS5	56 (12.9)
None	42 (9.7)
iPod	32 (7.4)
MiFi	22 (5.1)
Data check	16 (3.7)
Both applications	12 (2.8)
Portal issue	5 (1.1)
Broken device	4 (0.9)
Overall	435 (100)

^a Reasons were logged in relation to 353 calls; up to two reasons could be logged for one call. Eight calls had no recorded reason.

TABLE 24 Details of outgoing calls to patients for whom data had not been received in > 3 weeks or > 2 weeks since consent

Centre	Total number and duration of calls n (minutes)	Number of calls answered n (%)	Number of calls not answered n (%)	Number of voicemails left n (%)	Total time of participants in study (months)
BEL	79 (358.0)	59 (75)	7 (9)	13 (16)	1090.2
LIV	28 (88.0)	18 (64)	8 (29)	2 (7)	446.0
SOU	32 (142.8)	28 (88)	3 (9)	1 (3)	508.5
MOO	27 (100.5)	13 (48)	7 (26)	7 (26)	411.3
JAP	98 (265.0)	60 (61)	16 (16)	22 (22)	937.6
GLO	8 (14.0)	2 (25)	4 (50)	2 (25)	206.1
Overall	272 (968.3)	180 (66)	45 (17)	47 (17)	3599.6

Issues affecting the availability of the electronic tests

Frequencies of testing and test completion for electronic tests described in [Adherence to testing](#) include occasions when participants were **unable** to complete the electronic tests. Details of downtime for each test are shown in [Table 26](#). These occasions accrued to 15 days of testing for the MBT test and 30 days for the mVT test. To put these numbers in context, total testing time in the study spanned 1318 days, although the number of participants testing varied over the course of the study.

On 25 October 2018, the company which created the mVT app was acquired by the Roche Holding AG (Basel, Switzerland). Upon acquisition, the mVT servers were switched off as they prepared to move to the Roche servers. Consequently, the mVT app could not be activated on new iPods and existing iPods

TABLE 25 Reasons reported by participants for their data not being received

Reason for call ^a	Number of calls, n (%)
Connectivity	52 (23.9)
Not tested due to personal or home circumstances	49 (22.5)
Multibit app not working	26 (11.9)
None	21 (9.6)
Unable to contact the participant	13 (6.0)
Both apps not working	12 (5.5)
Holiday	12 (5.5)
iPod issue	12 (5.5)
MiFi issue	11 (5.0)
mVT app not working	10 (4.6)
Overall	218 (100)

^a Reasons were logged in relation to 272 calls; up to two reasons could be logged for one call. Ninety-two calls had no recorded reason.

TABLE 26 Details of occasions when participants were unable to complete digital tests

App affected	Participant testing impacted	Participant recruitment impacted	Reason	Days until resolved
mVT	Yes	Yes	Server migration	22
mVT	Yes	Yes	Expiry of database website security certificate	7
MBT	Yes	No	Expiry of database website security certificate	1
mVT	Yes	Yes	Database issue	1
MBT	Yes	No	Expiry of domain certificate	14
				Total: 45

with the app activated were not compatible with the latest version of iOS installed on the iPods. This issue meant that prospective participants attending an information and training session were unable to have the mVT app activated on their iPods. In discussion with the SSC, it was decided that recruitment should continue and affected patients should have their mVT app activated at the earliest opportunity when the new servers were switched on. This issue coincided with an iOS update from v11.0 to v12.0. The mVT app was incompatible with v12.0 during the server switchover period, preventing the mVT app from working on any iPod that was updated to v12.0 at that time.

The mVT team assured the research team that both issues would resolve when the server transfer was complete. They had been unable to provide advance notice due to the confidentiality of the acquisition and had not anticipated that there would be any impact on the study. The server switch over was completed on 16 November 2018 when the mVT app was again available. iPods were updated to v12.0, a process that was quite time consuming for both sites and the post-doctoral fellow at Queen's University Belfast (QUB) and multiple device management software (MDMS) was set to install future updates automatically. The study team updated the equipment instructions for patients with what to do if their iPod needed to install an update. For patients who joined the study without an activated mVT app, the study team phoned the patients to talk them through activating the app remotely or the local site team arranged to see the patients when they next attended clinic.

Around this time, sites and participants reported a variety of technical issues such as error messages, inability of the iPod to connect to the internet, the MBT app failing to work on two iPods following an update, iPods of a certain batch being unable to update to v12.0 and three iPods with an mVT app error preventing activation of the app. The team's ability to guide a participant through problem solving remotely over the telephone varied substantially by patient.

On 15 April 2019, a participant reported that the two-port USB charger had exploded in a power socket while charging the devices. The charger casing had split into two pieces and the electricity to the participant's house was tripped. The participant was unharmed and there was no damage to the house or the iPod and MiFi devices. Subsequently, all participants were posted a letter with new instructions on charging the study devices, advising that devices are charged for a maximum of 3 hours per week while at home (i.e. not to leave the charging devices unattended) and to unplug the charger once charging is complete. The Sponsor, site research teams and the study management group were also informed of the incident and kept informed about actions taken. The letter was subsequently submitted as a non-substantial amendment and approved.

Feedback from the charger supplier reported that such an incident had not been reported previously. They specified that the chargers were fitted with a surge protector which was designed to disable the plug to preserve the devices in the event of a power surge. The Sponsor arranged for all chargers (approximately 200) that had not yet been dispatched to sites to have a portable appliance test; all these chargers passed testing except for one charger which split similarly during testing. The research team was subsequently made aware that seven other chargers being used in the study had broken apart: one at a participant's home, one at a site during equipment set up and five when returned to testing and being unpackaged by the Sponsor. The Sponsor decided that all participants should have a tested charger. Tested chargers were posted to all sites to replace their local stock and chargers belonging to their participants. The chargers were either swapped at clinic visits or posted to participants.

For both applications, a security certificate was required for the servers that host the online portals. The certificates expired, causing both test portals to become temporarily unavailable. Renewal of the certificate was the responsibility of the app providers. For the MBT app, the portal was down for less than 1 day (9 August 2019) and the issue was resolved by 3 p.m. The helpline received three phone calls about this issue. For the mVT app, the portal was down for 1 week (from 6 August 2019 to 13 August 2019). The delay was due to administrative processes within Roche. During this time, participants could test as normal, but new devices could not be activated. This issue effected four participant training sessions, causing sessions to be rearranged, participants being sent home without equipment or without the mVT app activated. In addition, patient test data were not received by the study team during this week but were stored locally on the device and became available after the issue was resolved.

Test data for the MBT app were unavailable to be downloaded for 11 days in October/November 2019 due to the portal server being at full capacity. The server capacity was increased, causing the cost of hosting the portal to rise.

Participants could not test using the MBT app and the online portal could not be accessed from 14 August 2020 to 28 August 2020. The issue was caused by the expiry of a domain certificate required for the host database. The delay in resolving the issue arose because Novartis had to approve the update, leading to over 50 participants contacting the study team during the affected period.

Logistical issues with the electronic equipment provided for the study

As we supplied a limited data contract with the MiFi device, we were keen to limit the use of the iPod to study testing only; therefore, we decided to use a MDMS which meant we retained control of all devices during the study. We chose to use the free Apple product as commercial custom-designed

versions appeared expensive in comparison; however, on hindsight, the limited functionality of the basic version was a disadvantage, and a commercial version would have been better. In particular, we were unable to stop Apple updates being pushed to devices which caused a lot of confusion for technology-naïve participants and calls to the helpline. The MDMS also meant all the iPods had to be loaded with the apps in Belfast while connected to the QUB network and then transported to the sites. This was problematic as there are strict regulations governing the transport of the lithium-ion batteries by air and ferry with only specific carriers able to accept them and only two batteries could be present per package. We had budgeted to send the equipment packs to sites in batches of 10 by a standard carrier, whereas the subsequent regulations meant we had to send individual equipment packs (both iPod and MiFi device contained a battery) at a cost of £6.27 each with a specialist carrier, significantly increasing the costs and time required to administer the process.

Mobile phone connectivity using the MiFi device was required for app data to be transmitted. Mobile phone coverage was checked at the outset, but connectivity was still a problem for some participants some of the time, either due to variability in the strength of mobile phone coverage or difficulty in using the MiFi device. We also had considerable problems at one of our sites which had no network coverage within the hospital preventing set up occurring at the training visit. This required negotiation with the hospital IT department to provide special permission to access a sufficiently secure Wi-Fi network. In September 2019, the cost of a MiFi device increased from £33 to £63. The cost of the Vodafone tariff also rose on 16 September 2019 from £6 per month to £9 per month for each activated MiFi device. The study was able to bear these cost increases only because the total number of participants recruited (and hence packs of equipment in operation) was smaller than expected.

Chapter 8 Results: Objective B (qualitative)

Thirty per cent (89/297) of participants who were recruited to the study agreed to be contacted about the qualitative component and received study information. Follow-up telephone calls enquired further about participation in a semistructured interview – 3/89 did not wish to take part, and 8/89 could not be contacted after at least three calls. The remainder (26%; 78/297 of MONARCH study participants) were contacted, provided informed consent and subsequently took part in an interview (Figure 20), including participants categorised as regular home testers ($n = 63$), irregular testers (ITs) ($n = 14$) and non-testers who declined to take part in HM but consented to take part in qualitative component of the study ($n = 1$). Overall index test usage was similar for each index test over the duration of the study. Demographic characteristics of patient participants ($n = 78$) were comparable to those not taking part in the qualitative study (Table 27). In addition to the 78 patient participants, 11 informal 'carers', supporters or significant others in the lives of patients, and 9 HCPs who interacted with participants at study site visits were interviewed (6/11 informal 'carer' interviews took place in the presence of the patient participant). A total of 98 interviews were completed (patients, carers and health professionals). Interviews were conducted, flexibly, in-person at patients' homes ($n = 51$), at clinical sites for health professionals ($n = 4$) or via telephone ($n = 45$) between October 2018 and September 2020 and lasted an average of 36 minutes (range: 25–78 minutes).

Views about the acceptability of HM appeared to be represented by five overarching themes (and nine associated subthemes) that emerged from the inductive and deductive analysis of interview data: 1. *The role of HM*; 2. *Suitability of procedures and instruments*; 3. *Experience of HM*, and 4. *Feasibility of HM in usual practice*; 5. *Impediments to home monitoring*. Aspects of some themes were overlapping and concepts interacted with one another. The relationships between themes and coding of interview data are shown in Appendix 1, Table 55. Each theme is presented in sections *Theme 1: the role of home monitoring* to *Theme 5: impediments to home monitoring*, with selected illustrative quotes from patient participants (Table 28). The description focuses on presenting the patient perspective though the views of HCPs and informal 'carers' are summarised and individual quotes are presented in Appendix 1, Tables 56 and 57, respectively.

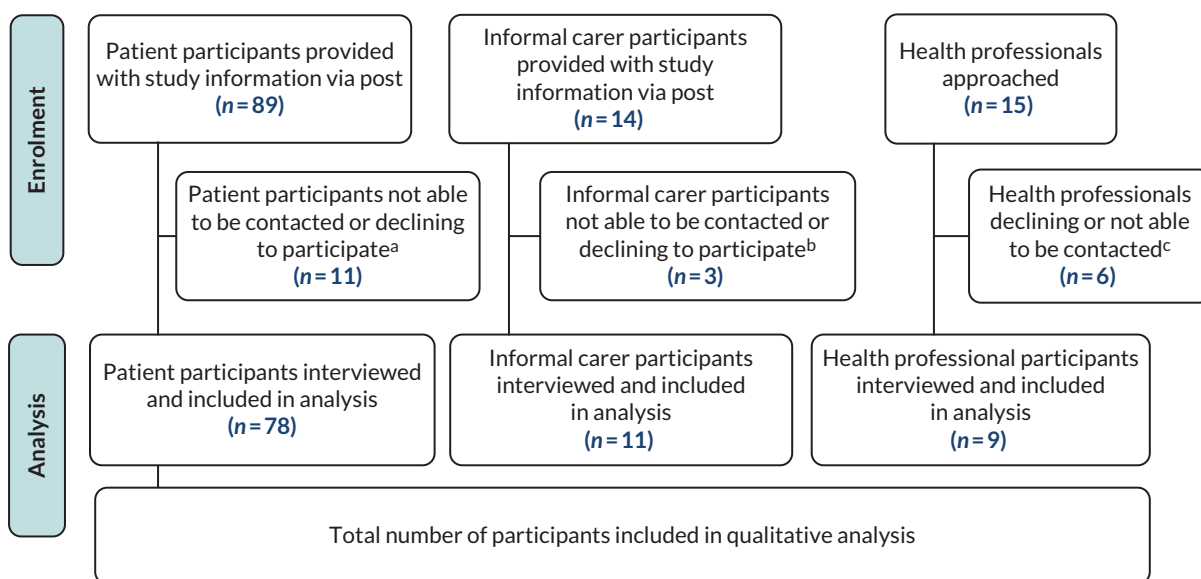


FIGURE 20 Participant flow – qualitative component of the MONARCH study. IT, irregular testers (completed one, two or three tests, or stopped-and-started testing or withdrew from diagnostic accuracy study); NT, non-testers (declined to take part in home monitoring); RT, regular testers (completed weekly testing without significant gaps in testing during the study period). a, Declined to participate ($n = 3$), not contactable by telephone ($n = 8$). b, Not contactable by phone ($n = 3$). c, Moved post ($n = 1$), not contactable ($n = 5$).

TABLE 27 Sociodemographic characteristics of participants in the qualitative component of the MONARCH study compared to the rest of the MONARCH study sample

		Qualitative sample (<i>n</i> = 78) ^a		Remaining MONARCH study participants (<i>n</i> = 221)	
		<i>n</i>	%	<i>n</i>	%
Baseline characteristics					
Sex	Male	30	38.5	93	42.1
	Female	48	61.5	128	57.9
Age	Mean (SD) years	74.3 (6.8)		75.1 (6.6)	
Visual acuity ^b	Mean (SD) LogMAR	0.2 (0.2)		0.2 (0.2)	
Smoking history	Current smoker	7	9.1	23	10.4
	Ex-smoker (> 1 month)	44	57.1	94	42.5
	Never smoked	26	33.8	104	47.1
Exposure to technology					
	Television	75	97.4	220	100.0
	Simple mobile phone	24	31.2	106	48.2
	Smartphone	53	68.8	145	65.9
	Tablet	55	71.4	142	64.5
	Laptop/home computer	53	68.8	132	60.0
	Internet at home	68	88.3	185	84.1
	E-mail	62	80.5	152	69.1
	Social media	30	39.0	68	30.9
	TV streaming/on-demand services	36	46.8	110	50.0

a Calculations are based on *n* = 77 as overall qualitative sample includes *n* = 1 participant who declined to take part in home monitoring but consented to take part in the qualitative part of the study.

b For patients with two involved eyes, better seeing eye is used.

Theme 1: the role of home monitoring

Subtheme 1: understanding purpose

The role or function of HM was clear from the perspective of participants who viewed it as facilitating the 'self-measurement' of changes in visual acuity and who consistently commented that HM was important because it helped identify differences in vision that might be due to degenerative changes and require additional clinical investigation and treatment. HM was viewed also as providing a patient with 'ownership' or 'personal control' over their visual health, in the sense that they 'knew their own eyes better than anyone else'.

Participants described how they 'self-monitored' their vision before the study including noting '... changes when looking out at a street sign visible from my window', or checking for increased difficulties when reading or watching television. This practice, sometimes referred to clinically as an 'environmental Amsler test', was acknowledged by patient participants as being less able to detect small changes in vision, and HM using digital testing was seen as able to potentially provide more consistent and accurate measurements.

Subtheme 2: perceived impact on eye care

Home monitoring was perceived as supporting usual clinical care in terms of providing a further assessment of possible deterioration in vision, in addition to clinic-based diagnostic testing. Participants emphasised that they did not regard HM as a 'replacement' for usual care, but viewed it as part of their eye care, particularly during non-active treatment periods. Participants indicated that HM could also 'fill the gap' between appointments given that, in their view, the time between appointments was long. In addition, it was recognised that HM could potentially reduce the frequency of clinic visits. This was an acceptable and positive outcome, as it was highlighted that this could save time and effort without inconveniencing family and friends who are required to transport patients to clinics.

The potential for HM to reduce health service costs and relieve burden on services were noted also, and it was recognised by participants that these 'savings' might allow additional appointments and treatments to be targeted at patients with high-level clinical need, specifically, patients with a higher risk of deterioration. There was a view that HM by itself might be appropriate for patients who had 'stable' visual acuity though the definitions of 'low risk' and 'stable' in the context of HM were not clear. Concerns were expressed that a reliance on HM alone might delay necessary treatment depending on the frequency or degree to which results were reviewed by clinic staff. Participants highlighted the importance of contact with clinicians identify accurately any deterioration, and underlined that there needed to be a clear pathway to routine clinic appointments, particularly if there were noticeable changes in visual acuity during HM.

Theme 2: suitability of procedures and instruments

The iPod Touch and app-based index tests were viewed as novel innovations in eye care while recognising that they reflected the increasing pervasiveness of technology. Participants' pre-HM use of such devices varied, and experience with more recent devices, including smartphones, was relatively limited. Despite this experience, even those with minimal or no previous exposure to similar technologies reported that they viewed HM as realistic and suitable. Overcoming unfamiliarity with technology or any hesitancy with using it was regarded as 'just something that needed to be done' as part of their eye care.

Participants suggested demonstration of the device and index tests, prior to, and separate from the formal 'training' in HM procedures, provided reassurance that it was 'simple enough to do', and increased belief in their capacity to undertake HM. In addition to this technological apprehension, participants reflected that being older and unfamiliar with technology might also result in hesitation about engaging in HM and generate concern that HM was too complex or a 'burden'.

Theme 3: experience of home-monitoring procedures

Subtheme 1: training for home monitoring

Training to ensure familiarisation and appropriate use of the iPod Touch and the app-based index tests was provided before participants started HM. Participants described the training as essential for 'getting going' with HM but as relatively 'information heavy'. 'Refresher' sessions were suggested by participants to help them overcome difficulties recalling correctly the training information that was provided earlier.

Experiential learning was important as participants reported that HM became easier and more routine with time and practice. Participants described as valuable the occasional, 'informal' support and advice that health professionals provided during clinic-based study monitoring visits, especially when an individual encountered a problem with HM.

Subtheme 2: test preferences

Participants typically viewed the iPod and index tests as easy to use but suggested that a larger device could increase usability. The digital tests were referred to, positively, as 'feeling like a game' that reduced the 'boredom' of repeated testing. A clear preference was expressed for MultiBit (MBT) relative to mVT. This preference was linked closely to the 'feedback' (represented as a percentage score) that was provided by the MBT test (discussed under subtheme 3 below). There were other factors that influenced participants' views of the digital index tests. For the MBT, there was a lack of clarity among participants about the purpose of the later stages of the test which were described sometimes as 'disheartening' as stages became progressively 'too fast'. The mVT test was viewed as too subtle (when discriminating between the 'distorted' circles), and as being difficult to complete, which led participants to 'guess' responses. In addition, the mVT test and the paper-based KSJ were perceived to be less 'interesting' or engaging compared to the MBT test.

Subtheme 3: use of MultiBit vision test feedback and data

Overall, MBT test scores were viewed as providing valuable feedback and as a way of 'self-monitoring' changes in vision. This was apparent despite participants not being encouraged to use the percentage scores or being provided with specific information on their purpose or meaning. Participants described how they noted results, made comparisons overtime and used the feedback to 'try and beat my last score'. This sense of 'self-competition' was seen as helpful for keeping engaged with testing as well as having other indirect or unintended effects.

However, perceived ambiguity about the meaning of scores also produced uncertainty among participants about how to respond to changes and generated doubts regarding how results were used and regarding their role as a patient. 'Retesting' was described as a useful way of confirming a change in scores. Negative changes were attributed sometimes to tiredness or becoming distracted during testing but consistently lower results, even when small in magnitude, were interpreted as a concern about their eye health.

Theme 4: feasibility of regular home monitoring in usual service delivery***Subtheme 1: frequency of home monitoring and habit formation***

Participants highlighted how their views about using HM improved as they became more familiar with testing. Weekly HM was seen as realistic and feasible, and usually lasted between 10 and 15 minutes each time. Establishing a HM 'routine' was considered important by participants who used several methods to help them continue regular HM, including use of written reminders or prompts, such as a diary, or 'a note left on the fridge'. Participants described how family members would visit at a set time each week while tests were being completed and this routine functioned as a reminder when to test, and as a means of getting help with setting up devices.

Participants referred to using other forms of digital 'self-monitoring', including blood pressure and respiratory rate measurements; and said this made it easier to set up a HM routine, and also acted as a reminder when to do the vision tests.

Participants acknowledged that since changes in vision could happen gradually overtime, HM would need to be continued in the longer term and that it would need to be 'easy, and not a burden' to achieve sustainability and high adherence.

Subtheme 2: use of ongoing support

The support that was provided by the study telephone helpline was perceived to be a key component of successful HM. Occasionally, support and advice were provided by health professionals during clinic visits and this additional support was deemed to be valuable. Family members and friends, too, were a source of technical and problem-solving support.

Theme 5: impediments to home monitoring

Subtheme 1: practical issues

Practical or technical issues with HM were reported relatively infrequently and tended to be minor issues such as devices needing to be recharged frequently and sometimes before weekly testing could be completed (see [Appendix 1, Table 58](#)). Occasionally, technical issues stopped participants from testing and returning to HM – restarting even after a brief period of non-testing could be challenging. Participants reported that, in general, these issues could be overcome via 'problem solving', effort and persistence. Regarding the MBT specifically, participants addressed the requirement of dark conditions for completion of the test by changing their physical environment, such as placing a blanket over their head while testing or undertaking testing in a windowless space. Sometimes, these adaptations were described by participants as making the test challenging and 'awkward'.

A more substantial issue with HM was that participants were sometimes unsure if tests had been completed successfully or if data had been transmitted, and it was suggested that an improvement might be to provide instantaneous confirmation of both aspects.

Subtheme 2: personal health and social factors

Other issues that negatively affected HM were related to health concerns or functional limitations such as fine motor tremors, fatigue and concentration problems, each of which made it harder to undertake and complete tests.

Participants who had caregiving responsibilities such as providing childcare or caring for a partner or spouse reported difficulties finding sufficient time to test regularly and consistently.

Views of informal 'carers' and healthcare professionals – summary

In general, there was a high level of concurrence between the perspectives of informal 'carers' and HCPs and views of patients (see [Table 28](#)). HCPs suggested that increasing age may be linked to initial reluctance or hesitancy about HM while acknowledging that this view may also be related to their underlying assumptions about older people and use of technology. HCPs agreed with the patients' views that training was essential to the successful uptake and use of HM, and that training needed to be adapted to the needs of each individual patient and their previous experience of technology. There was a concern that HM could be resource intensive in terms of the technical support needed (see [Appendix 1, Table 28](#) HCP quotes). Informal 'carers' or family members saw their role in HM as a supportive and facilitative one, and they valued the feedback from a test performance as a way of assessing the response to the nAMD treatment that was received by their family member (see [Appendix 1, Table 29](#) carers quotes).

TABLE 28 Perspectives of patients on acceptability of home monitoring

Perspectives of patients	Theme/subtheme	Supporting quote(s) from patients
<ul style="list-style-type: none"> - HM viewed as providing 'ownership' or 'personal control' - HM could reduce the frequency of clinic visits - Clear pathways to routine clinic appointments are needed if there are changes in visual acuity 	<p>Theme 1. The role of home monitoring Subtheme 1: Understanding purpose Subtheme 2: Perceived impact on eye care</p>	<p>'... it is to put you in charge. I could judge if I needed help, if I saw deterioration in my vision when I did the test, or if I noticed a change by myself'. (Female, Regular HM, 62 years, #53) 'I would feel, yes, I'm doing the tests and that's okay. At the minute, I'm only going (to the clinic) four times a year, so even two or three times would be okay. I'd be happy enough now [To home monitor], you know? ... Providing nothing happens'. (Female, Regular HM, 78 years, #37) '... I don't think it would always work because it's near impossible to get an appointment, you know? I mean, I've done that. I've seen a change in shape, not when I was in this study but before. I asked for an appointment but didn't get it, so is the purpose is to try and put people more in charge of saying what they can see, saying if they need help or not?' (Male, Regular HM, 82 years, #24)</p>
<ul style="list-style-type: none"> - Overcoming unfamiliarity with technology regarded as 'something needing to be done' - Unfamiliar with technology might result in hesitation about engaging in HM 	<p>Theme 2. Suitability of procedures and instruments</p>	<p>'... technology is a funny thing to lots of people my age, some have embraced it, now of course it's a necessary evil, so I'm on catch up' (Male, Regular HM, 76 years, #08) '... if this (the test device) was just given to me, I would be a bit lost but I'm always trying to keep an open mind with technology and do what I can, you know'. (Male, Irregular HM, 79 years, #38) '... I mean it's no problem because I'm not too bad. I've got an iPad and an iPod, but I can see lots of people couldn't do it. A lot of them don't even like using the computer do they?' (Female, Regular HM, 81 years, #68) '... Well, mostly it's the elderly people that have got it (AMD) and most of them are not okay with computers and things. I mean I'm not brilliant, but I can do it. As you get older you can't learn these things so easily'. (Female, Regular HM, 79 years, #82)</p>
<ul style="list-style-type: none"> - Refresher training could help overcome difficulties recalling information - mVT and paper-based KSJ tests were perceived as less engaging than the MBT - MBT test feedback seen as helpful for keeping engaged with HM - Lower test scores, even when small, were interpreted as a concern about their eye health 	<p>Theme 3. Experience of home-monitoring procedures Subtheme 1: Training for home monitoring Subtheme 2: Test preferences Subtheme 3: Use of MBT feedback and data</p>	<p>'... and so (the clinic staff) demonstrated it ... I thought that actually looks easy, but a week later when I'm on my own, I just said 'what did they say?' (Female, Regular HM, 71 years, #49) '... well, I found that test (MBT) ... first of all it was very quick. You had to be so alert and I could be pressing away and it was doing nothing because it was too fast for me'. (Female, Regular HM, 76 years, #17) '... but the test with the flashing numbers (MBT), I actually liked that. I couldn't stand the other test (mVT) because you get four shapes and one of them is sort of out of sync. The first three are easy, then it gets more and more tricky. It gets to the stage where I just had me guess. I actually found that annoying because I didn't know how I was doing. The other one you get a percentage, which is good'. (Male, Regular HM, 80 years, #46) '... so you see benefits instantly because you've got a result, not only have I done an exam, I have a result instantly, the minute you finish and put your stuff away, the mental benefits are there'. (Male, Regular HM, 75 years, #87) '... if I get less than 90(%) then I absolutely know that there's something wrong. I'm not happy with 92, it's always been 94 or 96, 98 or 100. So that did worry me, but I will do it again, just to check, and I've got an appointment on the second anyway'. (Male, Irregular HM, 77 years, #50)</p>

Perspectives of patients	Theme/subtheme	Supporting quote(s) from patients
<ul style="list-style-type: none"> Several methods used to help continue regular HM, including use of reminders or prompts Using other forms of digital 'self-monitoring', including blood pressure measurements; made it easier to set up a HM routine HM needs to be 'easy, and not a burden' to achieve sustainability and high adherence Family members were a source of support 	<p>Theme 4. Feasibility of regular home monitoring in usual service delivery</p> <p><i>Subtheme 1: Frequency of home monitoring and habit formation</i></p> <p><i>Subtheme 2: Use of ongoing support</i></p>	<p>'... and (my granddaughter) would get it set up for me and then when that test is finished, switch over on to the next but she doesn't have to stand over me, you know'. (Male, Irregular HM, 79 years, #38)</p> <p>'... I have used that (monitoring device for tracking COPD symptoms) for about 18 months, so this can also helped me know when I'm getting bad, because they were reading it and then they were ringing back and checking with me. That made me feel better, being in touch with people'. (Female, Regular HM, 62 years, #53)</p> <p>'... when I first went back to [eye hospital] they gave me the bag and then when I went to [hospital] they gave me a blood pressure monitor, so what I do is, I have to check my blood pressure regularly you see, so I stick this in with my machine because I'm doing them both weekly at the minute and it all works out well, I don't forget'. (Female, Regular HM, 74 years, #34)</p> <p>'... my son has got me using smart phones and what not. I am ok with an iPad and an iPhone, no problem. I can handle anything in medical terms, I am keeping tabs on my medications on a daily basis. I have a little app that reminds me every hour, every two hours, what I have to do for the day' (Male, Regular HM, 70 years, #136)</p> <p>'...You don't do for enjoyment you're doing it to see how it goes. I don't look at it as a pleasure that I can't wait to do, and think, oh I must go up and do my wobbly circles. I just think it's time I did those, I'll go up and do them now'. (Female, Regular HM, 66 years, #62)</p> <p>'... I had a lot of trouble at one point, but my husband said, 'let me have it', and he diddled about with the buttons, one of which was the light intensity so I had probably turned the light down without realising it. He helped a lot. He said 'you go through it and see what you get stuck on. He didn't just take over, he just said call me when you need me'. (Female, Regular HM, 72 years, #58)</p> <p>'... so I had to ring [the helpline], he was very nice and went through it all. My son lives down the road and is into computers and I said well, I could ask my son again, but it was all sorted before my son appeared'. (Female, Regular HM, 76 years, #16)</p>
<ul style="list-style-type: none"> Some adaptations made HM challenging and 'awkward'. Other health concerns or functional limitations made it harder to undertake HM Caregiving responsibilities made it difficult to find time for regular HM 	<p>Theme 5. Impediments to home monitoring</p> <p><i>Subtheme 1: Practical issues</i></p> <p><i>Subtheme 2: Personal health and social factors</i></p>	<p>'... it was difficult, I just couldn't get it dark enough. I racked my brain and thought I've got a big wool rug. I got under that and did my best but there's also the claustrophobia, it just got me annoyed in the end'. (Female, Irregular HM, 77 years, #33)</p> <p>'...and I have a tremor, when I'm holding it (the iPod), you don't know where the numbers are going to come from on the screen ... so you're sort of anticipating you know? And this means you just don't catch it'. (Female, Regular HM, 71 years, #71)</p> <p>'... I have had problems with my health, my heart scare, lots of things all happening, a lot of times I think this leaves me feeling really really tired ... I'm staring, not knowing if I even hit the buttons'. (Male, Irregular HM, 74 years, #29)</p> <p>'... it's because I have been caring for (a relative) and I don't even remember. It's not high on my list of priorities. I have been doing it, but it's when I get to it, not when it gets to me'. (Female, Regular HM, 72 years, #83)</p>

COPD, Chronic Obstructive Pulmonary Disease, HM, home monitoring.

Chapter 9 Discussion

Main findings: study conduct

To our knowledge, this is the first multisite deployment of home-based visual function monitoring in nAMD patients in the UK adhering to best practice for diagnostic test accuracy studies. All index tests had supporting peer-reviewed evidence of their potential to identify active nAMD lesions in similar populations when the study was conceived. Both app-based tests had formal spin-out companies and at the time of trial initiation, both had been acquired by pharma companies for further application and development. Hence, they were in a relatively well-developed stage when we began the trial. Despite this, we experienced a wide range of practical and logistical challenges in the delivery of the trial:

1. Device-based challenges

- A. During the initial recruitment period, the company that created the mVT app was acquired by a pharma company, upon acquisition, the mVT servers were switched off and for a period the mVT app or servers were not available for new participants nor were compatible with the iOS update which happened around the same time.
- B. Regulations limiting the transport of lithium-ion batteries occurred after the grant submission. Rather than sending a bulk pack of 10 sets of equipment to sites, we had to send them individually. Packaging and arranging postage proved very time consuming for study staff.
- C. Malfunctioning of chargers – iPod Touch devices do not come with their own charger unlike iPhones, and these needed to be purchased and provided separately. Since both the MiFi and iPod Touch needed charging, we opted for a double USB charger rather than an Apple version. Despite checking all the appropriate certifications were in place, we had a report of a device overheating and exploding. This required a temporary stop of the trial while all chargers were replaced, and Portable Appliance Testing (PAT) tested.
- D. iOS management – we used the standard Apple MDMS for the study devices. With hindsight, a more expensive custom version tailored to the trial would have been much better. We were unable to block iOS updates going to the devices and for technology-naïve participants this caused a lot of confusion and concern and calls to the helpline. One of the updates made the apps stop working and so there was a delay while they were updated appropriately. We promised participants they could keep the devices if they completed the required follow-up as an incentive, but this proved very problematic as the MDMS meant they had to be returned to Belfast and connected to QUB network for files to be deleted, and this took considerable time when staff were working from home. For as many participants as possible, we sent new devices which were spare given the sample size was smaller than expected.

2. Different approaches at sites to patient recruitment/inclusion.

There was some variation in how different sites managed patients in terms of frequency of monitoring visits and treatment strategies that may have contributed to the variation in recruitment across sites. During study initiation visits, we also encountered several research staff who were very sceptical about providing older nAMD patients with electronic devices and we felt this may have contributed to how they approached potential participants in clinics.

3. Recruitment timescale.

It quickly became apparent that sites had to wait for patients to attend routine visits before approaching them, and if they were willing in principle to take part, wait until the next routine visit to attend an information and training session. Such a protracted recruitment pathway was not factored into the recruitment prediction. Therefore, recruitment was slower than expected. We also ended up with fewer participants in the two strata of longer time since diagnosis because, once the study was underway, patients were recruited when they became eligible. Sites were able to recruit patients eligible for the 30- to 41-month stratum mainly early on during recruitment; recruitment became increasingly difficult as the study progressed as eligible patients did reach this length of time from diagnosis.

1. Index tests generated many scores between management visits. These needed to be summarised to fit models because models could not fit raw data, as well as respecting the structure of eyes and management visits nested within participants. Ultimately, we calculated very simple summary scores (averaging multiple scores) but considered several issues, for example, controlling for the time between management visits and number of scores being averaged, weighting proximal scores (closer in time to the reference standard visit) more highly than those that were more distal/older (addressed in the first set of sensitivity analyses) alternatives to a simple average such as mode.
5. *Feedback to participants on performance.* The MBT provided a numerical score on performance, whereas mVT did not; this was unable to be altered at study onset and as a study team, we would have preferred that neither provided a score to prevent concern or anxiety for the participants. However, some participants really liked getting feedback from MBT and it seemed to motivate adherence and retention, whereas for others it caused concern.
6. *Pandemic:* Recruitment was due to end on 31 March 2020. Therefore, the COVID-19 pandemic only impacted the final 2–3 weeks of recruitment. The frequency of the monitoring visits was significantly impacted with different sites adopting different strategies to managing care pathways, including one site which stopped retinal imaging altogether.

Main findings: primary outcome (Objective A)

Diagnostic accuracy of index tests

1. There was no effect of time since first starting treatment, though days since baseline was significant. In the second sensitivity analysis, there were inconsistent/contradictory effects of age by right/left eye which we attribute to chance.
2. None of the tests had adequate Diagnostic Test Accuracy (DTA) for the primary outcome to enable patients to be monitored without hospital review. This conclusion was unaltered by the two sensitivity analyses (more recent test data vs. all data for preceding interval between visits; same outcome but using CARF grading).
3. None of the tests had adequate DTA for the secondary outcome, better approximating how we believe tests are likely to be used, although estimates were imprecise.

Main findings: Objective C

Potential inequalities in recruitment, ability to perform tests and adherence to testing

Our findings were largely consistent with prior expectations:

1. A minority of patients who were approached were willing in principle to participate.
2. Increasing age and deprivation index for home address were associated with being unwilling in principle to participate. Recruiting site was also associated with willingness in principle to participate; this was believed to be due to different strategies for approaching patients and in some cases may have reflected the research staff's own preconceptions regarding those suitable for a digital health intervention.
3. IMD quintile and age were not consistently associated with either being able to self-monitor with the index tests, or adherence to weekly testing. There was some indication that participants with the worst vision at baseline were less able and adherent. Recruiting site was also associated with being able to self-monitor and adhering to weekly testing.
4. Recruitment varied across sites which complicated the inequality relationships for which we tested. One of the six sites recruited much more freely than other sites, weakening the associations observed for willingness in principle to take part. Conversely, inequality associations were more apparent at this site than other sites for ability to perform tests and adherence to testing.

The COVID-19 pandemic significantly hastened the adoption of telemedicine and digital health interventions alongside heightened concern about the impact of the digital divide and its potential to increase health inequalities.^{37,38} Mitigating potential inequalities in participation was central to our study design by providing both the hardware for accessing the apps (iPod Touch) as well as internet access (MiFi device and network contract). Despite this, those who chose to participate had a high prevalence of digital literacy and internet access; so, providing these resources alone was insufficient to entice those without such experience.

Main findings: Objective D

Diagnostic accuracy of conversion to neovascular age-related macular degeneration in fellow eyes

We hypothesised that index tests should be better able to identify conversion of fellow eyes to active nAMD because conversion is expected to cause a step change in visual symptoms (e.g. distortion of objects); this should be easier to perceive than a change in the extent of distortion. Weak relationships were evident but estimates of DTA were very imprecise. We judged that, as with the primary objective, tests did not have adequate DTA compared to prediction without test information to warrant application to this context. This conclusion was unaltered by the two sensitivity analyses (more recent test data vs. all data for preceding interval between visits; same outcome but using CARF grading). Moreover, given the findings for Objective A, patients with one fellow eye and one eye with a nAMD lesion would still need to attend HES clinics for the latter eye to be monitored.

About 14% and 20% of fellow eyes converted based on management visit decisions and CARF grading, respectively. This rate of conversion was higher than expected based on epidemiological studies of conversion rates in unaffected fellow eyes^{9,39} and the total time for which participants were followed in the study. This may be due to having recruited participants with study eyes that had longer than average times since first treatment.

Main findings: Objective B

Qualitative evaluation of acceptance of home monitoring

Two overarching meta-themes emerged from the qualitative studies related to acceptance or non-acceptance of home monitoring. Meta-theme 1 encompassed four main themes: (1) the role of home monitoring; (2) suitability of procedures and instruments; (3) experience of home monitoring; and (4) feasibility of home monitoring in usual practice. Meta-theme 2 consisted of one main theme covering key inhibitors to acceptability. The main factors influencing acceptability included a participant's understanding about the purpose of home monitoring and their experience of using it. While home monitoring was generally seen as a relatively straightforward exercise to undertake and non-burdensome, training and ongoing support were regarded as essential to its success. The response to feedback was interesting in that the numerical values provided by the MBT for some were very important for motivation and adherence, whereas for others increased anxiety. In general, the qualitative study demonstrated that home monitoring was acceptable to patients and its potential to reduce clinic visits during non-active treatment phases was recognised.

Main findings

Challenges

Despite two-thirds of the population reporting having previously used a smartphone, there were still a variety of challenges experienced while testing at home with the electronic devices that contributed

to both reduced adherence and ultimately withdrawals from the study. The top two reasons for calls by participants to the study helpline were issues with the two electronic tests followed by connectivity problems. The addition of the MiFi device increased the technical complexity of the home-testing procedures which on hindsight outweighed any benefit gained by improving equity of participation. Calls to participants in response to significant gaps in data transfer highlighted connectivity issues as the primary reason followed by personal or home issues. Various issues related to database and server problems resulted in participants being unable to use the electronic tests for 3.4% of the total testing time.

There were also challenges relating to equipment itself and the logistics of distributing it to sites. The two-port USB chargers were particularly problematic including a rather spectacular failure in which one exploded in a power socket while charging, while there were no serious consequences, we arranged for all chargers not yet dispatched to be tested and then used to replace all those currently in use. This was a very costly and time-consuming episode for the research fellow. Mobile phone connectivity was also a problem for a few participants and was initially problematic at one of our sites as the absence of signal in the hospital meant the iPod could not be properly set-up and apps activated at the training visit until special permissions were granted to access a Wi-Fi network. There were also substantial increases in the cost of MiFi devices and the phone tariff during the study which was only able to be absorbed due to the reduced sample size.

Other studies have also evaluated the utility of tablet- or phone-based self-tests for monitoring macular disease. A tablet-based retinal function test was evaluated in nAMD eyes and at-risk fellow eyes and the relationship between retinal sensitivity and retinal features present at corresponding SD-OCT locations.⁴⁰ Of note, significant relationships between reduced retinal sensitivity and features related to non-nAMD were observed such as atrophy, retinal pigment epithelium disruption and absent ellipsoid zone and not with typical nAMD features such as the presence of subretinal fluid or intraretinal fluid demonstrating the difficulties in detecting changes in function in relation to nAMD activity. While sample size may have contributed to the finding, it may also be because in these eyes, there is already substantial visual disturbance; therefore, the noticeable difference threshold is higher than an eye without pathology or newly developed pathology.

The Alleye app (Oculocare, Zurich, Switzerland) has been evaluated in a series of pragmatic studies in Moorfields Eye Hospital in London. A small-prospective study⁴¹ of 73 eyes from 56 patients receiving antiangiogenic treatment for nAMD or diabetic macular oedema (DMO) reported a false alarm rate of 6.1% and a positive predictive value of 80% which was promising. However, approximately 20% of data points had to be discarded post hoc due to various factors relating to test performance. We did not undertake any attempt at cleaning the data of outliers or anomalous results as no criteria were provided by the developers to enable this to be pre-specified within the analysis plan. Patients showed enthusiasm for the task in that they continued to complete the testing even after the study period had finished, reinforcing our finding that elderly patients with impaired vision have both the ability and motivation to undertake tablet/smartphone-based testing at home. A subsequent study⁷ undertaken in response to the COVID-19 pandemic evaluated 245 patients undergoing antiangiogenic treatment. During the follow-up period, 85 eyes triggered an alarm, of which 66 eyes (78.5) showed either a drop in VA or increase in central macular thickness on clinical examination and 60% received an injection, demonstrating that the Alleye app has the potential to reliably detect worsening nAMD pathology and the need for intravitreal therapy. This study was not a formal diagnostic test accuracy study and much of the on-boarding procedures had to be done remotely using video and telephone given the ongoing restrictions. Results from another app⁴² (OdySight, Tilak, France) in 60 patients with a variety of oedematous maculopathies demonstrated that decreases in visual acuity during follow-up could be detected by the app with 92% sensitivity and a specificity of 99%. To improve adherence and patient satisfaction, this app included both puzzles and eye test modules, performance of the eye tests earns credits to progress to the more interesting and enjoyable puzzle game. Some patients engaged

with the puzzles even outside testing times and those who did so showed better long-term adherence and retention. This highlights the importance of incentivising adherence and participation through gamification of the process and boredom with the monotonous nature of weekly testing was mentioned during interviews by several of our participants.

Equality, diversity and inclusion

Participant representation

This study was planned and initiated prior to the National Institute for Health and Care Research (NIHR) adoption of the INCLUDE and INCLUDE ETHNICITY framework; so, while effort was made to ensure that participation was optimised generally and in particular for those lacking digital skills, there wasn't a specific strategy for targeting any potential gender or racial imbalances. Of those deemed ineligible at screening, none were because of English language. There were more females than males in both the overall study (57.8% vs. 42.1%) and the qualitative sample (61.5% vs. 38.5%), although AMD is more common in females. In future mixed-methods studies, particular attention should be paid to ensuring sufficient males are included in qualitative components.

Reflections on the research team and wider involvement

The research team was well balanced from a gender perspective; at the beginning of the study, the co-chief investigator, the study manager and study post-doctoral research fellow were all female. The study timeline also accommodated a maternity leave for the co-chief investigator.

Patient and public involvement

Age-related macular degeneration patients who were members of our PPI group provided useful feedback regarding the KSJ, the style, format and content of the PIL and study newsletter. PPI feedback was most useful for the KSJ as the US style motivational quotes were strongly rejected by the PPI group. Feedback on the design and operation of the electronic tests was not very useful as the PPI participants had minimal experience of smartphones or computer access and so seemed unwilling to criticise their content. Future studies evaluating mobile devices should ensure PPI groups include participants with a range of exposure and familiarity with the potential technology. The strong rejection of the term 'carer' by the PPI group members was also an important criticism which we acted upon and should be considered for future studies.

Strengths and limitations

Strengths

The study was multicentre, with sites that were representative of ophthalmic services in which macular/retinal HES clinics are provided to monitor nAMD. Estimates of the diagnostic test accuracy of index tests were at low risk of bias: the study population was appropriate for the intended use of the tests, and summary test scores were not available to ophthalmologists providing the reference standard, which was judged after the index test data were collected. All test and reference standard data were included. We attempted to minimise inequality by providing a device for self-monitoring compared to most other studies of home monitoring to date, which required participants to use their own devices. We attempted to avoid access to internet being a barrier to participation by providing an additional MiFi device. All tests evaluated were developed independently of the study investigators and the developers had no role in the conduct of the study or the analysis plan. Peer-reviewed longitudinal data were available for all index tests at the outset which supported their inclusion in the study. The study also contained a strong qualitative component to explore issues of acceptability, barriers etc.

There were several limitations. Recruitment was slower than anticipated and was closed before the target number of participants or management visits was met; about 75% of the target number of participants were recruited but less than half the target number of management visits were documented, despite the rate of management visits overall being as expected. The target sample size was intended to provide adequate power to compare the DTA of different tests. We did not make such comparisons due to the manifest poor DTA of all tests and the AUROCs for individual tests accuracy in diagnosing active lesions were estimated precisely; 95% CIs for AUROCs were narrow (± 0.04) and estimates ruled out tests providing adequate accuracy for diagnosing active nAMD to enable patients to be monitored without hospital review.

We could not define test thresholds a priori, due to having no information to weight the importance of false and negative misclassifications and no recommendations from test developers about cut-off thresholds. Instead, we estimated AUROCs over the entire range of summary test scores. Standard DTA indices were reported based on Youden's index for comparability across tests but should not be viewed as recommended test thresholds.

Study eyes were classified as active about twice as often as expected. This reduced the number of pairs of visits with inactive to active changes in lesion status, that is, the secondary outcome. Consequently, DTA estimates for this outcome were imprecise.

Tests were sometimes not available for technical reasons that were beyond the control of the research team. Interruptions occurred sporadically during the study and, hence, were very unlikely to have introduced bias. The reduction in overall testing time and data was small and did not affect our conclusions about the overall performance of tests.

Sites were trained to provide information and provide training to participants about how to perform home monitoring. However, we were unable to monitor the quality of training provided by sites or decisions to offer retraining.

The study had no control over monitoring visits and participants are likely to have reported their subjective visual experience to their consultants, which might have influenced the reference standard.

The ways in which patients were approached and screened varied across sites. One site recruited more freely than other sites, with a higher subsequent withdrawal rate and poorer average adherence. Variations may have reflected preconceptions of research staff regarding the capabilities of patients to use the electronic tests. We think that both approaches to recruitment were valid and consistent with the protocol, but note that more inclusive recruitment, while more equal with respect to participation, was more wasteful with respect to duration of participation and test data available for analysis.

Lessons for the future

It is clear from the outcome of this study that independent validation of such tests is essential. Robust pilot data about the DTA of potential tests are essential before designing a full-scale diagnostic test accuracy study. Most pilot studies describe associations rather than predictions, with analyses that have not been directed by a prespecified data analysis plan. Therefore, there is a high risk of selection of the reported result from multiple analyses,⁴³ for example, reflecting post hoc manipulations of the data with respect to outliers, uncertain or missing reference classifications or other missing data. In addition, studies are typically designed to optimise test performance or in populations or settings that bias test performance. The topic under which this study was commissioned was advertised twice. The commissioning process was abandoned on the first attempt when the manufacturers of a test of interest declined to collaborate. The topic was readvertised despite this experience and this study was commissioned on the basis of the peer-reviewed evidence cited in our application. In our view, the

development of the tests and the evidence to support them was below the usual level required by the Health Technology Assessment (HTA) programme for a commissioned topic.

Given the additional complexity that provision of the iPod device and MiFi added to the execution of the study, it is likely that future studies will concentrate on providing access to an app that can be loaded direct to patients own smartphone. However, some patients may find that intrusive or unacceptable and those without devices will face exclusion.

It is clear from the outcome of this study that independent validation of such tests is essential. However, evaluations need to be agile to react quicker to poor performance or add/swap included tests as technologies emerge. Adaptive study designs may be considered in future,⁴⁴ in conjunction with an ongoing review of literature (akin to a Cochrane living review) to identify promising tests and efficiently evaluate them.

Future research

Given the clear potential for digital exclusion in these types of studies, further research should concentrate on interventions to improve participation in mobile health research beyond providing access to hardware.

Further research is also warranted to identify pros and cons of different methods to summary test score when the results of multiple tests are likely to contribute to DTA for a subsequent reference standard. Considerations include: the best summary measure, how time over which monitoring has occurred/ number of tests should be considered, whether test scores to application of the reference standard should be weighted more strongly than test scores that are more remote in time.

Further research should also focus on the nature of personal feedback to home-monitoring tasks and how it can be used to help or hinder adherence and retention. It was clear that different patients reacted differently to the feedback on test performance with some really liking it and others finding it stressful.

Given the superiority of retinal imaging biomarkers over functional testing in the Early Detection of age-related macular degeneration (EDNA) study, further research focusing on low-cost portable home OCT devices would be worthwhile. Several devices are in development,^{45,46} but none have been deployed in a large-scale study yet.

The longitudinal imaging data collected in MONARCH can also be further analysed later. Of particular interest is whether central macular thickness on the visit prior to an active clinical decision is predictive of lesion reactivation.

Chapter 10 Conclusion

None of the index tests evaluated was suitable for home monitoring in the intended patient population and setting. The principle of home monitoring was understood and broadly acceptable to patients and the overall adherence and length of follow-up achieved encouraging. There is a proportion of AMD patients able and willing to undertake electronic device-based home monitoring and so pursuing and refining this model using different technologies is a worthwhile aim particularly as this proportion is likely to increase with time due to increasing digital literacy in future cohorts. Home-based retinal imaging devices may be particularly worthwhile to focus on given the striking findings from the EDNA study, which demonstrated the superiority of retinal imaging over any of the visual function or patient-reported tests for detecting conversion to nAMD in the fellow eye.⁹ Several groups have reported early prototypes for home-based OCT devices⁴⁵⁻⁴⁷ and given that fluid on OCT is one of the main parameters by which lesion activity is assessed, this could be very useful, particularly if combined with artificial intelligence-based image interpretation to trigger referral.⁴⁸

Additional Information

Contributions of authors

Ruth E Hogg (<https://orcid.org/0000-0001-9413-2669>) (Senior Lecturer, Co-chief Investigator) co-led the conception and design of the study, contributed to the interpretation of the results, led the writing of the introduction and discussion and contributed to the writing and editing of the report.

Robin Wickens (<https://orcid.org/0000-0003-4374-3747>) (Trial management, Bristol CTU) was responsible for the day-to-day management of the study, contributed to the interpretation of the results, led the writing of the methods and challenges section and contributed to the writing/editing of the report.

Sean O'Connor (<https://orcid.org/0000-0001-6805-8899>) (Post-doctoral fellow, QUB) was responsible for the day-to-day management of the qualitative component, managed the device logistics and MDMS, contributed to the interpretation of the qualitative results, led the writing of the qualitative chapter and contributed to the writing/editing of the report.

Eleanor Gidman (<https://orcid.org/0000-0002-5261-5213>) (Statistician, Bristol CTU) conducted the statistical analysis study, contributed to the interpretation of results, led the statistical methods section and contributed to the writing/editing of the report.

Elizabeth Ward (Study management, Bristol CTU) was responsible for the day-to-day management of the study, contributed to the interpretation of the results, led the writing of the methods and challenges section and contributed to the writing/editing of the report.

Charlene Treanor (<https://orcid.org/0000-0002-3209-8060>) (Post-doctoral fellow, QUB) was responsible for the day-to-day management of the qualitative component, managed the device logistics and MDMS and contributed to the writing/editing of the report.

Tunde Peto (<https://orcid.org/0000-0001-6265-0381>) (Consultant Ophthalmology) contributed to the conception and design of the study, recruitment and follow-up of participants, contributed to the interpretation of the results and contributed to the writing/editing of the report.

Ben Burton (<https://orcid.org/0000-0001-9579-9078>) (Consultant Ophthalmology) contributed to the conception and design of the study, recruitment and follow-up of participants, contributed to the interpretation of the results and contributed to the writing/editing of the report.

Paul Knox (<https://orcid.org/0000-0002-2578-7335>) (Reader, Vision Scientist) contributed to the conception and design of the study, recruitment and follow-up of participants, contributed to the interpretation of the results and contributed to the writing/editing of the report.

Andrew J Lotery (<https://orcid.org/0000-0001-5541-4305>) (Consultant Ophthalmology) contributed to the conception and design of the study, recruitment and follow-up of participants, contributed to the interpretation of the results and contributed to the writing/editing of the report.

Sobha Sivaprasad (<https://orcid.org/0000-0001-8952-0659>) (Consultant Ophthalmology) contributed to the conception and design of the study, recruitment and follow-up of participants, contributed to the interpretation of the results and contributed to the writing/editing of the report.

Michael Donnelly (Professor, Health Services research) contributed to the conception and design of the qualitative study, management of the qualitative study, contributed to the interpretation of the results and contributed to the writing/editing of the report.

Chris A Rogers (<https://orcid.org/0000-0002-9624-2615>) (Professor, Statistics) contributed to the conception and design of the study, was responsible for the statistical analyses and statistical analysis plan, contributed to the interpretation of results and contributed to the writing/editing of the report.

Barnaby C Reeves (<https://orcid.org/0000-0002-5101-9487>) (Professor, Co-chief Investigator) co-led the conception and design of the study, contributed to the interpretation of the results, led the writing of the final report and contributed to the writing and editing of the report.

Acknowledgements

Independent members of the steering committee

We thank the independent members of the MONARCH steering committee, including our public and patient representatives for their valued contribution and oversight of the study and for their attendance at the steering committee meetings, both in person and virtually.

David Crabb (Chair), Roger Anderson, Yemisi Takoingi, Geraldine Hoad (PPI, Macular Society), Robert Harpur, Geeta Meenon, Tariq Aslam, Stevie Johnson (PPI, RNIB Northern Ireland).

Project management group

Grantholders: Dr Ruth Hogg (CI), Barnaby Reeves (co-CI), Sobha Sivaprasad, Tunde Peto, Paul Knox, Chris Rogers, Andrew Lotery, Ben Burton, Michael Donnelly, Usha Chakravarthy.

Study team and Bristol CTU support: Charlene Treanor, Sean O'Connor, Elizabeth Ward, Robin Wickens, Eleanor Gidman.

Other acknowledgements

Our thanks go to the following groups for their assistance in conducting the MONARCH study.

We extend our thanks to all the participants who took part in the study and without who the study would not be possible.

We are grateful to all the staff at the clinical sites that facilitated recruitment, training and data collection and contributed to the regular study management meetings.

We thank the team at the reading centre (Central Administrative Research Facility) who conducted the grading of the images for the MONARCH study. Thanks to Alyson Muldrew, Clare Newell, Karleigh Kelso, Vittorio Silvestri, Tunde Peto, Michelle McCaughey and David Rice.

We thank the companies and organisations who provided access to their tests for evaluation in this context. Thanks to Mark Roser and Patricia Beaton from the International Macular and Retinal Foundation for help and support with the KSJ. Thanks to Mike Bartlett and Yi-Zhong Wang from Vital Art and Science LLC for access and support with my vision track device. Thanks to Lars Frisen and Bo Frisen from Visumetrics for access and support with the Multibit device. Thanks to Novartis and Roche for access to the apps for the duration of the study.

Data-sharing statement

All available data can be obtained by contacting the corresponding author.

Ethics statement

Main REC: Northern Ireland Health and Social Care Research Ethics Committee A

Reference number: 17/NI/0235

Date of favourable ethical approval: 29 January 2018

Information governance statement

Queen's University Belfast is committed to handling all personal information in line with the UK Data Protection Act (2018) and the General Data Protection Regulation (EU GDPR) 2016/679. Under the Data Protection legislation, Queen's University Belfast is the Data Controller, and you can find out more about how we handle personal data, including how to exercise your individual rights and the contact details for our Data Protection Officer here: <https://www.qub.ac.uk/privacynotice/Research/ListofResearchPrivacyNotices/PrivacyNoticeforResearchParticipants.html>.

Disclosure of interests

Full disclosure of interests: Completed ICMJE forms for all authors, including all related interests, are available in the toolkit on the NIHR Journals Library report publication page at <https://doi.org/10.3310/CYRA9912>.

Primary conflicts of interest: Barnaby C Reeves reports former membership of: the Health Technology Assessment Commissioning Board (from January 2010 to 31 March 2016); the Health Technology Assessment Efficient Study Designs Board (October to December 2014); the Health Technology Assessment Interventional Procedures Committee B Methods Group (December 2016 to September 2022); Systematic Reviews Programme Advisory Group (Systematic Reviews National Institute for Health and Care Research Cochrane Incentive Awards and Systematic Review Advisory Group), dates unknown. He has current membership of a CTU funded by the NIHR up to 31 August 2021. He has no other competing interests. Chris A Rogers reports membership of a Clinical Trials Unit funded by the National Institute for Health and Care Research. She also reports membership of the Health Technology Assessment Funding Committee Policy Group (formally CSG) and the Health Technology Assessment Commissioning Committee. She has no other competing interests. Ruth E Hogg reports attendance at Roche Digital Health Advisory Meeting July 2019. She also received partial PhD Studentship funding from Okko Health 2021 for home monitoring of Diabetic Retinopathy. Paul Knox reports software support from Vital Art and Science Inc who produced the My Vision Track App. Andrew J Lotery reports receiving consulting fees from and owning stock or stock options of Gyroscope Therapeutics. Sobha Sivaprasad reports grants from Boehringer Ingelheim, receiving consulting fees from Boehringer Ingelheim, Novartis, Apellis, Bayer, Oculis, Oxurion, Roche and Biogen. She also received payment or honoraria from Boehringer Ingelheim and Bayer, support for attending meetings from Bayer and participation in an advisory board with Bayer. She is also a Macular Society Trustee (unpaid). She was also a former committee member of HTA Commissioning Committee from 7 August 2017 to 30 September 2021.

Publications

Ward E, Wickens RA, O'Connell A, Culliford LA, Rogers CA, Gidman EA, *et al.* Monitoring for neovascular age-related macular degeneration (AMD) reactivation at home: the MONARCH study *Eye (Lond)* 2021;**35**(2):592–600. <https://doi.org/10.1038/s41433-020-0910-4>

O'Connor SR, Treanor C, Ward E, Wickens RA, O'Connell A, Culliford LA, *et al.*, Monarch Study Group. The COVID-19 pandemic and ophthalmic care: a qualitative study of patients with Neovascular Age-Related Macular Degeneration (nAMD). *Int J Environ Res Public Health* 2022;**19**(15):9488. <https://doi.org/10.3390/ijerph19159488>

O'Connor SR, Treanor C, Ward E, Wickens RA, O'Connell A, Culliford LA, *et al.*, Monarch Study Group. Patient acceptability of home monitoring for neovascular age-related macular degeneration reactivation: a qualitative study. *Int J Environ Res Public Health* 2022;**19**(20):13714. <https://doi.org/10.3390/ijerph192013714>. PMID: 36294292.

Papers in progress

1. Diagnostic test accuracy of home-monitoring tests.
2. Uptake of digital health interventions, and inequalities in uptake, in an elderly population living with sight impairment.
3. Challenges in the evaluation of digital health interventions requiring participation in an elderly (visually disabled) population.
4. Association of morphological features of AMD lesions with lesion status classifications by ophthalmologists monitoring patients for reactivation.

References

1. Li JQ, Welchowski T, Schmid M, *et al.* Prevalence and incidence of age-related macular degeneration in Europe: a systematic review and meta-analysis. *Br J Ophthalmol* 2020;**104**(8):1077–84. <https://doi.org/10.1136/bjophthalmol-2019-314422>
2. Taylor DJ, Hobby AE, Binns AM, Crabb DP. How does age-related macular degeneration affect real-world visual ability and quality of life? A systematic review. *BMJ Open* 2016;**6**(12):e011504. <https://doi.org/10.1136/bmjopen-2016-011504>
3. Essex RW, Nguyen V, Walton R, Arnold JJ, McAllister IL, Guymer RH, *et al.*, Fight Retinal Blindness Study Group. Treatment patterns and visual outcomes during the maintenance phase of treat-and-extend therapy for age-related macular degeneration. *Ophthalmology* 2016;**123**(11):2393–400. <https://doi.org/10.1016/j.ophtha.2016.07.012>
4. Rosenfeld PJ, Brown DM, Heier JS, Boyer DS, Kaiser PK, Chung CY, Kim RY, MARINA Study Group. Ranibizumab for neovascular age-related macular degeneration. *N Engl J Med* 2006;**355**(14):1419–31. <https://doi.org/10.1056/NEJMoa054481>
5. Chakravarthy U, Harding SP, Rogers CA, Downes S, Lotery AJ, Dakin HA, *et al.* A randomised controlled trial to assess the clinical effectiveness and cost-effectiveness of alternative treatments to Inhibit VEGF in Age-related choroidal Neovascularisation (IVAN). *Health Technol Assess* 2015;**19**(78):1–298. <https://doi.org/10.3310/hta19780>
6. Gragoudas ES, Adamis AP, Cunningham ET Jr, Feinsod M, Guyer DR, VEGF Inhibition Study in Ocular Neovascularization Clinical Trial Group. Pegaptanib for neovascular age-related macular degeneration. *N Engl J Med* 2004;**351**(27):2805–16. <https://doi.org/10.1056/NEJMoa042760>
7. Islam M, Sansome S, Das R, Lukic M, Chong Teo KY, Tan G, *et al.* Smartphone-based remote monitoring of vision in macular disease enables early detection of worsening pathology and need for intravitreal therapy. *BMJ Health Care Inform* 2021;**28**(1):e100310. <https://doi.org/10.1136/bmjhci-2020-100310>
8. Busquets MA, Sabbagh O. Current status of home monitoring technology for age-related macular degeneration. *Curr Opin Ophthalmol* 2021;**32**(3):240–6. <https://doi.org/10.1097/ICU.0000000000000756>
9. Banister K, Cook JA, Scotland G, Azuara-Blanco A, Goulão B, Heimann H, *et al.* Non-invasive testing for early detection of neovascular macular degeneration in unaffected second eyes of older adults: EDNA diagnostic accuracy study. *Health Technol Assess* 2022;**26**(8):1–142. <https://doi.org/10.3310/VLFL1739>
10. Achard OA, Safran AB, Duret FC, Ragama E. Role of the completion phenomenon in the evaluation of Amsler grid results. *Am J Ophthalmol* 1995;**120**(3):322–9. [https://doi.org/10.1016/s0002-9394\(14\)72162-2](https://doi.org/10.1016/s0002-9394(14)72162-2)
11. Bittner AK, Torr-Brown S, Arnold E, Nwankwo A, Beaton P, Rampat R, *et al.* Improved adherence to vision self-monitoring with the Vision and Memory Stimulating (VMS) journal for non-neovascular age-related macular degeneration during a randomized controlled trial. *J Clin Exp Ophthalmol* 2014;**5**(1):320. <https://doi.org/10.4172/2155-9570.1000320>
12. Alster Y, Bressler NM, Bressler SB, Brimacombe JA, Crompton RM, Duh Y-J, *et al.*, Preferential Hyperacuity Perimetry Research Group. Preferential Hyperacuity Perimeter (PreView PHP) for detecting choroidal neovascularization study. *Ophthalmology* 2005;**112**(10):1758–65. <https://doi.org/10.1016/j.ophtha.2005.06.008>

13. Chew EY, Clemons TE, Bressler SB, Elman MJ, Danis RP, Domalpally A, *et al.*, Appendix 1 for AREDS2-HOME Study Research Group. Randomized trial of the ForeseeHome monitoring device for early detection of neovascular age-related macular degeneration. The Home Monitoring of the Eye (HOME) study design: HOME Study report number 1. *Contemp Clin Trials* 2014;**37**(2):294–300. <https://doi.org/10.1016/j.cct.2014.02.003>
14. Pitrelli Vazquez N, Knox PC. Assessment of visual distortions in age-related macular degeneration: emergence of new approaches. *Br Ir Orthopt J* 2015;**12**:9–15.
15. Wang YZ, He YG, Mitzel G, Zhang S, Bartlett M. Handheld shape discrimination hyperacuity test on a mobile device for remote monitoring of visual function in maculopathy. *Invest Ophthalmol Vis Sci* 2013;**54**(8):5497–505. <https://doi.org/10.1167/iovs.13-12037>
16. Wang YZ, Wilson E, Locke KG, Edwards AO. Shape discrimination in age-related macular degeneration. *Invest Ophthalmol Vis Sci* 2002;**43**(6):2055–62.
17. Bruender MC, Benjamin N, Agostini HT, Stahl A, Ehlken C. Subjective evaluation of visual acuity is not reliable to detect disease activity in different exudative maculopathies. *Graefes Arch Clin Exp Ophthalmol* 2018;**256**(9):1565–71. <https://doi.org/10.1007/s00417-018-4021-x>
18. Kaiser PK, Wang YZ, He YG, Weisberger A, Wolf S, Smith CH. Feasibility of a novel remote daily monitoring system for age-related macular degeneration using mobile handheld devices: results of a pilot study. *Retina* 2013;**33**(9):1863–70. <https://doi.org/10.1097/IAE.0b013e3182899258>
19. Pitrelli Vazquez N, Harding SP, Heimann H, Czanner G, Knox PC. Radial shape discrimination testing for new-onset neovascular age-related macular degeneration in at-risk eyes. *PLOS ONE* 2018;**13**(11):e0207342. <https://doi.org/10.1371/journal.pone.0207342>
20. Ku JY, Milling AF, Pitrelli Vazquez N, Knox PC. Performance, usability and comparison of two versions of a new macular vision test: the handheld Radial Shape Discrimination test. *PeerJ* 2016;**4**:e2650. <https://doi.org/10.7717/peerj.2650>
21. Winther C, Frisen L. Self-testing of vision in age-related macula degeneration: a longitudinal pilot study using a smartphone-based rarebit test. *J Ophthalmol* 2015;**2015**:285463. <https://doi.org/10.1155/2015/285463>
22. Reeves BC, Scott LJ, Taylor J, Hogg R, Rogers CA, Wordsworth S, *et al.* The Effectiveness, cost-effectiveness and acceptability of Community versus Hospital Eye Service follow-up for patients with neovascular age-related macular degeneration with quiescent disease (ECHOES): a virtual randomised balanced incomplete block trial. *Health Technol Assess* 2016;**20**(80):1–120. <https://doi.org/10.3310/hta20800>
23. *Internet Use Amongst Older People Subject to North South Divide*. 2015. URL: www.ageuk.org.uk/latest-press/archive/internet-use-amongst-older-people-subject-to-northsouth-divide/ (accessed 27 March 2022).
24. OFCOM. *Adults' Media Use and Attitudes Report 2014*. 2014. URL: www.ofcom.org.uk/_data/assets/pdf_file/0020/58223/2014_adults_report.pdf (accessed 27 March 2022).
25. Kobayashi LC, Wardle J, von Wagner C. Internet use, social engagement and health literacy decline during ageing in a longitudinal cohort of older English adults. *J Epidemiol Community Health* 2015;**69**(3):278–83.
26. Ward E, Wickens RA, O'Connell A, Culliford LA, Rogers CA, Gidman EA, *et al.* Monitoring for neovascular age-related macular degeneration (AMD) reactivation at home: the MONARCH study. *Eye (Lond)* 2021;**35**(2):592–600. <https://doi.org/10.1038/s41433-020-0910-4>

27. Obuchowski NA, McClish DK. Sample size determination for diagnostic accuracy studies involving binormal ROC curve indices. *Stat Med* 1997;**16**(13):1529–42. [https://doi.org/10.1002/\(sici\)1097-0258\(19970715\)16:13<1529::aid-sim565>3.0.co;2-h](https://doi.org/10.1002/(sici)1097-0258(19970715)16:13<1529::aid-sim565>3.0.co;2-h)
28. Macular Photocoagulation Study Group. Risk factors for choroidal neovascularization in the second eye of patients with juxtafoveal or subfoveal choroidal neovascularization secondary to age-related macular degeneration. *Arch Ophthalmol* 1997;**115**(6):741–7. <https://doi.org/10.1001/archophth.1997.01100150743009>
29. Pieramici DJ, Bressler SB. Age-related macular degeneration and risk factors for the development of choroidal neovascularization in the fellow eye. *Curr Opin Ophthalmol* 1998;**9**(3):38–46. <https://doi.org/10.1097/00055735-199806000-00007>
30. Whiting PF, Rutjes AW, Westwood ME, Mallett S, Deeks JJ, Reitsma JB, *et al.*, QUADAS-2 Group. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Ann Intern Med* 2011;**155**(8):529–36. <https://doi.org/10.7326/0003-4819-155-8-201110180-00009>
31. Tong A, Sainsbury P, Craig J. Consolidated criteria for reporting qualitative research (COREQ): a 32-item checklist for interviews and focus groups. *Int J Qual Health Care* 2007;**19**(6):349–57. <https://doi.org/10.1093/intqhc/mzm042>
32. Assarroudi A, Heshmati Nabavi F, Armat MR, Ebadi A, Vaismoradi M. Directed qualitative content analysis: the description and elaboration of its underpinning methods and data analysis process. *J Res Nurs* 2018;**23**(1):42–55. <https://doi.org/10.1177/1744987117741667>
33. Venkatesh V, Morris MG, Davis GB, *et al.* User acceptance of information technology: toward a unified view. *MIS Q* 2003;**27**(3):425–78.
34. Holden RJ, Karsh BT. The technology acceptance model: its past and its future in health care. *J Biomed Inform* 2010;**43**(1):159–72. <https://doi.org/10.1016/j.jbi.2009.07.002>
35. Sekhon M, Cartwright M, Francis JJ. Acceptability of health care interventions: a theoretical framework and proposed research agenda. *Br J Health Psychol* 2018;**23**(3):519–31. <https://doi.org/10.1111/bjhp.12295>
36. Ha J, Park HK. Factors affecting the acceptability of technology in health care among older Korean adults with multiple chronic conditions: a cross-sectional study adopting the senior technology acceptance model. *Clin Interv Aging* 2020;**15**:1873–81. <https://doi.org/10.2147/CIA.S268606>
37. Lyles CR, Wachter RM, Sarkar U. Focusing on digital health equity. *JAMA* 2021;**326**(18):1795–6. <https://doi.org/10.1001/jama.2021.18459>
38. Martins Van Jaarsveld G. The effects of COVID-19 among the elderly population: a case for closing the digital divide. *Front Psychiatry* 2020;**11**:577427. <https://doi.org/10.3389/fpsy.2020.577427>
39. Silva R, Cachulo ML, Fonseca P, Bernardes R, Nunes S, Vilhena N, Faria de Abreu JR. Age-related macular degeneration and risk factors for the development of choroidal neovascularisation in the fellow eye: a 3-year follow-up study. *Ophthalmologica* 2011;**226**(3):110–8. <https://doi.org/10.1159/000329473>
40. Ho CYD, Wu Z, Turpin A, Lawson DJ, Luu CD, McKendrick AM, Guymer RH. A tablet-based retinal function test in neovascular age-related macular degeneration eyes and at-risk fellow eye. *Transl Vis Sci Technol* 2018;**7**(2):2. <https://doi.org/10.1167/tvst.7.2.2>

41. Faes L, Islam M, Bachmann LM, Lienhard KR, Schmid MK, Sim DA. False alarms and the positive predictive value of smartphone-based hyperacuity home monitoring for the progression of macular disease: a prospective cohort study. *Eye (Lond)* 2021;**35**(11):3035–40. <https://doi.org/10.1038/s41433-020-01356-2>
42. Guigou S, Michel T, Merite PY, Coupier L, Meyer F. Home vision monitoring in patients with maculopathy: real-life study of the OdySight application. *J Fr Ophtalmol* 2021;**44**(6):873–81. <https://doi.org/10.1016/j.jfo.2020.09.034>
43. Sterne JAC, Savovic J, Page MJ, Elbers RG, Blencowe NS, Boutron I, et al. RoB 2: a revised tool for assessing risk of bias in randomised trials. *BMJ* 2019;**366**:l4898. <https://doi.org/10.1136/bmj.l4898>
44. Adaptive Platform Trials Coalition. Adaptive platform trials: definition, design, conduct and reporting considerations. *Nat Rev Drug Discov* 2019;**18**(10):797–807. <https://doi.org/10.1038/s41573-019-0034-3>
45. Nahen K, Benyamini G, Loewenstein A. Evaluation of a self-imaging SD-OCT system for remote monitoring of patients with neovascular age related macular degeneration. *Klin Monbl Augenheilkd* 2020;**237**(12):1410–8. <https://doi.org/10.1055/a-1271-6834>
46. Maloca P, Hasler PW, Barthelmes D, et al. Safety and feasibility of a novel sparse optical coherence tomography device for patient-delivered retina home monitoring. *Transl Vis Sci Technol* 2018;**7**(4):8. <https://doi.org/10.1167/tvst.7.4.8>
47. Song G, Chu KK, Kim S, et al. First clinical application of low-cost OCT. *Transl Vis Sci Technol* 2019;**8**(3):61. <https://doi.org/10.1167/tvst.8.3.61>
48. Yim J, Chopra R, Spitz T, et al. Predicting conversion to wet age-related macular degeneration using deep learning. *Nat Med* 2020;**26**:892–9. <https://doi.org/10.1038/s41591-020-0867-7>
49. Sekhon M, Cartwright M, Francis J. Application of a theoretical framework to assess intervention acceptability: A semi-structured interview study. *Eur Health Psychol* 2016;**18**:565.
50. Al Aufa B, Renindra IS, Putri JS, Nurmansyah MI. An application of the Unified Theory of Acceptance and Use of Technology (UTAUT) model for understanding patient perceptions on using hospital mobile application. *Enferm Clin* 2020;**30**:110–3.
51. Chao CM. Factors Determining the Behavioral Intention to Use Mobile Learning: An Application and Extension of the UTAUT Model. *Front Psychol* 2019;**10**:1652.
52. Zhou M, Zhao L, Kong N, Campy KS, Qu S, Wang S. Factors influencing behavior intentions to telehealth by Chinese elderly: An extended TAM model. *Int J Med Inform* 2019;**126**:118–127.
53. Liu H, Wu T. Estimating the area under a receiver operating characteristic (ROC) curve for repeated measures design. *J Stat Softw* 2003;**8**:1–18.
54. Bossuyt PM, Reitsma JB, Bruns DE, Gatsonis CA, Glasziou PP, Irwig L, et al. 54STARD 2015: an updated list of essential items for reporting diagnostic accuracy studies. *BMJ* 2015;**351**:h5527. <https://doi.org/10.1136/bmj.h5527>

Appendix 1 Additional tables and figures

TABLE 29 Semistructured guide for patient participant interviews

Interview schedule
1. Could you tell me about your experiences of going to the eye clinic for your regular check-ups?
2. Could you tell me about your reasons for wanting to take part in this study and to try out the tests at home?
3. What do you think about the purpose of the tests and of monitoring your vision at home?
4. What are your thoughts on monitoring at home if it meant people might have hospital appointments less often?
5. Would you be comfortable having hospital appointments less often?
6. What do you think about the idea of just monitoring at home, without hospital appointments?
7. How did you find the information and training session at the start of the study?
8. How do you feel about using technology and electronic devices like mobile phones, tablets or computers?
9. What was your experience of using technology before the study?
10. Could you tell me about your experience of using each test?
11. Did you do the tests in any particular order?
12. How did you find doing the tests once per week? Was it too often or would it be ok to test more often?
13. Did you find testing was easy to do as part of your normal routine?
14. Was there anything that stopped you from testing at home?
15. How long did it usually take to do the tests each week?
16. Did you find the tests easy to use?
17. How confident were you that you could do the tests yourself?
18. Do you think others would be able to do the tests easily?
19. How suitable was the electronic device you were given for doing the home tests?
20. Did you find the novelty of testing at home wore off during study?
21. Did you use any support or help to do the tests? Do you still use the same amount of support after testing for some time?
22. Did you see any benefits from doing the tests?
23. Was there anything enjoyable about doing the tests?
24. Did you have to change any part of your home environment to do the tests?
25. Did you prefer any of the tests to the others?
26. What are your thoughts about receiving feedback scores for one of the tests (the numbers test) and not the shape test?
27. Where there any disadvantages of doing the tests?
28. Did you experience any technical issues with the tests?
29. Did the tests work all the time as they were meant to?
30. What would you change about the tests?
31. What advice would you give people doing the tests?
32. Would you like to add anything else or ask me any questions?

TABLE 30 Final scheme for coding of interview data

Code	Definition
1	Burden/perceived amount of effort The perceived amount of effort that is required to participate in home monitoring
2	Ethicality The extent to which home monitoring is a good fit with an individual's value system
3	Intervention coherence The extent to which the participant (mis)understands home monitoring and how it works
4	Opportunity costs The extent to which benefits, profits or values must be given up to engage in home monitoring
5	Perceived effectiveness/ perceived usefulness The extent to which home monitoring is perceived as likely to achieve its purpose An individual's perception that iPods (IT system) will enhance home monitoring (job performance)
6	Self-efficacy The participant's confidence that they can perform the test(s) required; this also includes a lack of confidence
7	Perceived ease of use An individual's perception that iPods to home monitor (IT system) is free of effort
8	Attitude An individual's evaluative judgement of the target behaviour on some dimension (e.g. good/bad, harmful/beneficial, pleasant/unpleasant, also ambivalence)
9	Behavioural intention An individual's motivation or willingness to exert effort to perform home monitoring
10	Actual behaviour/ acceptance The action of undertaking home monitoring
11	Subjective norm An individual's perception of the degree to which important other people approve or disapprove of home monitoring
12	Facilitating or inhibiting conditions Perception of availability of necessary skills, resources and opportunities for home monitoring. Also, organisational and technical infrastructure to support home monitoring. In context of MONARCH, e.g. able to get a dark space for one app test which needs complete darkness, technical helpline
13	Image The degree to which an individual perceives that use of iPod home monitoring (an innovation) will enhance his or her status in his or her social system
14	Job relevance The degree to which an individual believes that the home monitoring (target system) is applicable to his or her eye health (job)
15	Output quality The degree to which an individual believes that the iPod (the system) performs his or her home monitoring (job tasks) well
16	Results demonstrability The degree to which an individual believes that the results of iPod home monitoring (using a system) are tangible, observable and communicable
17	Individual differences Individual difference variables include personality and/or demographics (e.g. traits or states of individuals, gender and age) that can influence individuals' perceptions of perceived usefulness and perceived ease of use.
18	System characteristics System characteristics are those salient features of iPod home monitoring (a system) that can help individuals develop favourable (or unfavourable) perceptions regarding the usefulness or ease of use of a system.
19	Technology apprehension The degree to which an individual experiences apprehension, or fear, when she/he is faced with the possibility of using technology (computers)
20	Technology playfulness/ comfortable with technology The opposite of technology apprehension. The degree to which an individual is open to experiment, spontaneous interact or 'give technology a go' and comfort with technology

TABLE 30 Final scheme for coding of interview data (*continued*)

Code	Definition
21	Technology persistence Not at the same level as technology playfulness but participant appears willing to try to troubleshoot or rectify problems that may arise when using the iPod
22	Perceived enjoyment The extent to which an individual perceives home monitoring to be enjoyable in its own right
23	Reported experience Participant actual experience of rather than perception of effort required to test when using the iPod
24	Voluntariness A choice rather than enforcement of iPod home monitoring
25	Technology reliability Output quality in terms of data transfer and equipment reliability
26	General health status Participant's perception of their general health status
27	Perceived threat An individual's perception of severity and susceptibility that eye health may deteriorate, e.g. AMD no longer treatable, will lose sight or progression of AMD to other eye
28	Health beliefs and concerns about eye health An individual's beliefs and concerns around AMD
29	Habit The reported experience and perception that home monitoring has/has not become habitual
30	Ageing Functional, health, sensory, cognitive and mobility changes associated with increasing age influencing home monitoring
31	Medical Services satisfaction The extent to which participants are satisfied with current healthcare services for AMD
32	Affordability of health services Refers to the affordability of health services, e.g. private health care within NHS context
33	Comfort with health services Refers to the psychological feelings of patients towards health services and hospital environment, e.g. cleanliness of hospital
34	Professionalism of health-care staff Refers to knowledge, skills and interpersonal skills of healthcare staff
35	Safety of health care Participant's perception of healthcare safety, e.g. experienced medical teams, complete medical facilities, hospital security measures
36	Waiting time Patient's perception of waiting time for appointments, treatment etc.
37	Information quality Home monitoring can provide useful information
38	Usual care vs. home monitoring The perception of an individual to choose the physical hospital environment over the 'virtual'/home monitoring hospital environment includes recommending current hospital to others and continuing to use this hospital
39	Perceived security Perceived risk or protection from the transfer, management and analysis of personal health information
40	HCP's (Doctor's) opinion (similar to subjective norm) Influence of HCPs as they are perceived to be a point of expert authority
41	Motivational intention to participate in, or seek out latest research Participant expresses an intention or gets involved in research
42	Views about training/preparation Participant describes their experience and perception of the training session or preparing themselves for home monitoring
43	Other influencing factors Participant makes references to factors not otherwise covered by codes within this framework
44	Family support (not subjective norms) Participant makes reference to the presence of family or a significant other

continued

Table 30 Final scheme for coding of interview data (continued)

Code	Definition	
45	Emotion	Participant describes stress, anxiety or experience of trauma
46	Major life event	Participant refers to major life event
47	Social context	Participant describes their living space, residential area, whether or not they live with others.
48	Reported performance	Participant describes their performance on one or more tests
49	Keep Sight Journal	Participant describes features and perception of the paper-based KSJ
50	Multibit	Participant describes features and perception of the electronic 'numbers' MBT
51	My Vision Track	Participant describes features and perception of the electronic 'bumpy circles/odd one out' My Vision Track test
52	Check-in calls for non-testers	Participant describes having been contacted by CTEU due to non-regular testing
53	Experience with eye care	Participant describes what happens or what has happened when they have been in contact with their local macular, optometry or ophthalmologist service
54	Order of test	Preferential or chronological order of tests
55	Travel	Participant describes how they get to their usual appointments, includes incurred expenses or parking difficulties

TABLE 31 Number of patients approached to participate in home-monitoring testing

Centre	Months recruiting	Recruitment target	Expected number approached ^a	Actual number approached	% predicted ^b
BEL	21	80	329	326	99.1
LIV	18	80	343	156	45.5
SOU	20	80	333	166	49.8
MOO	19	80	338	152	45.0
JAP	20	50	222	110	49.5
GLO	9	30	180	65	36.1
Overall	107	400	1745	975	55.9

a Expected number approached based on expected number to be recruited by a site. For instance, a site open for 3 months with a recruitment target of 3.5 patients per month would expect 10.5 recruited patients; as 30% of approached patients are expected to be recruited, the expected number approached should be $10.5/0.3 = 35$ patients.

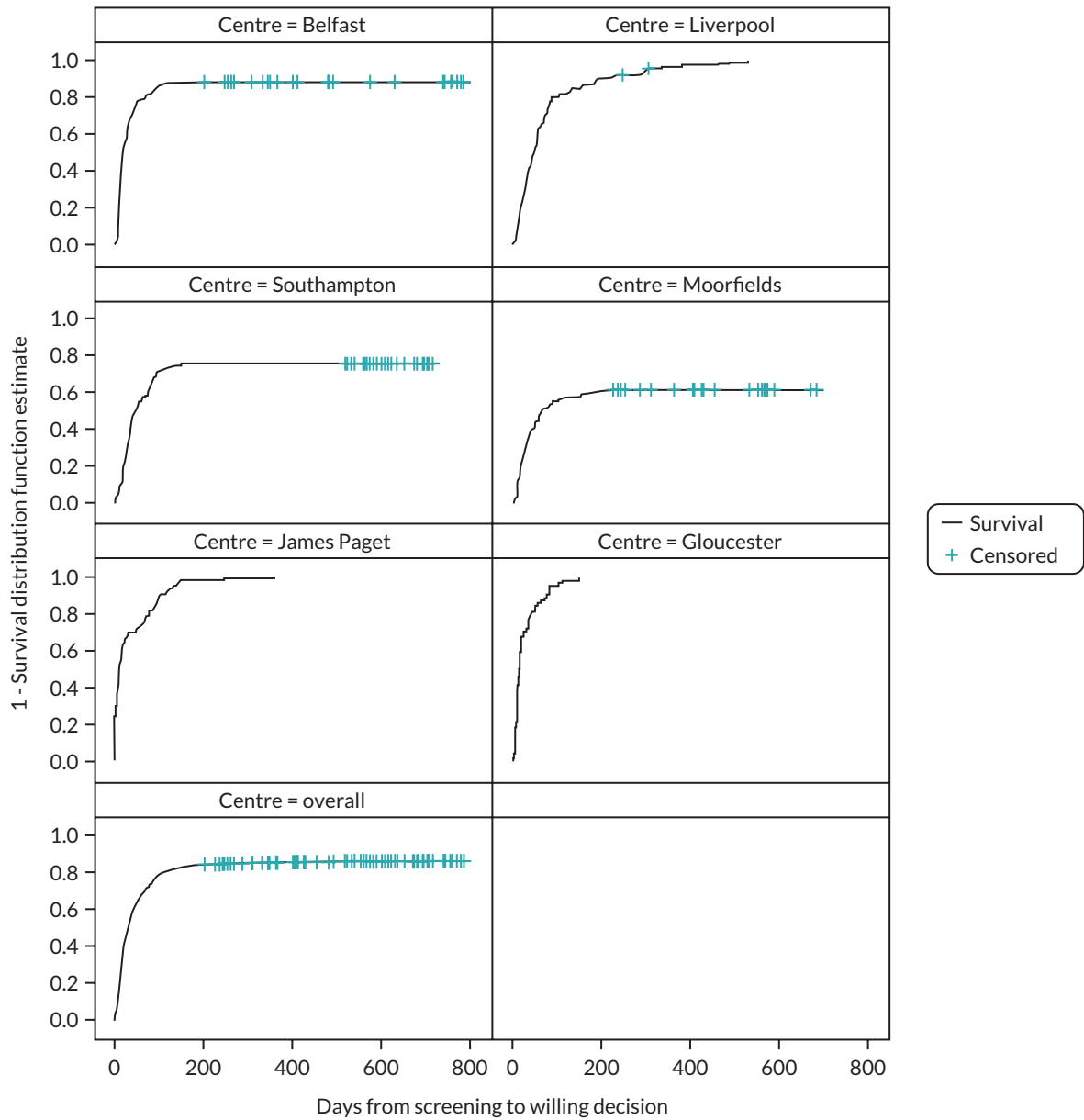
b Per cent predicted is the actual number approached as a percentage of the expected number approached.

TABLE 32 Reasons for ineligibility at screening and training

Ineligibility reason	BEL, n (%)	LIV, n (%)	SOU, n (%)	MOO, n (%)	JAP, n (%)	GLO, n (%)	Overall, n (%)
Ineligible patients at screening^a	3	8	13	7	0	0	31
< 50 years age	- (-)	- (-)	- (-)	- (-)	- (-)	- (-)	- (-)
Unable to understand English	- (-)	- (-)	- (-)	1 (14)	- (-)	- (-)	1 (3)
Not previously treated for active nAMD or currently monitored	1 (33)	6 (75)	6 (46)	7 (100)	- (-)	- (-)	20 (65)
First AMD treatment < 6 months ago	2 (67)	1 (13)	4 (31)	- (-)	- (-)	- (-)	7 (23)
Over 42 months since first treatment	- (-)	- (-)	3 (23)	1 (14)	- (-)	- (-)	4 (13)
Vision limited by other condition	1 (33)	6 (75)	8 (62)	3 (43)	- (-)	- (-)	18 (58)
Vision worse than 6/60	2 (67)	4 (50)	4 (31)	2 (29)	- (-)	- (-)	12 (39)
Surgery in eye in previous 6 months	2 (67)	2 (25)	3 (23)	- (-)	- (-)	- (-)	7 (23)
Refractive error >-6D	- (-)	- (-)	1 (8)	- (-)	- (-)	- (-)	1 (3)
Retinal or choroidal neovascularization not due to nAMD	- (-)	- (-)	1 (8)	- (-)	- (-)	- (-)	1 (3)
Ineligible patients at training^b	5	0	1	1	1	1	9
Unable to do mVT test	1 (20)	- (-)	- (-)	- (-)	1 (100)	1 (100)	3 (33)
Unable to do MultiBit (MBT) test	4 (80)	- (-)	1 (100)	1 (100)	1 (100)	1 (100)	8 (89)
Unable to complete KSJ	2 (40)	- (-)	- (-)	- (-)	- (-)	1 (100)	3 (33)
Home/personal circumstances unsuitable: BADL	- (-)	- (-)	- (-)	- (-)	- (-)	- (-)	- (-)
Home/personal circumstances unsuitable: mobile coverage	- (-)	- (-)	- (-)	- (-)	- (-)	- (-)	- (-)
Home/personal circumstances unsuitable: other	1 (20)	- (-)	1 (100)	- (-)	1 (100)	- (-)	3 (33)
Any other problems	2 (40)	- (-)	- (-)	1 (100)	- (-)	- (-)	3 (33)

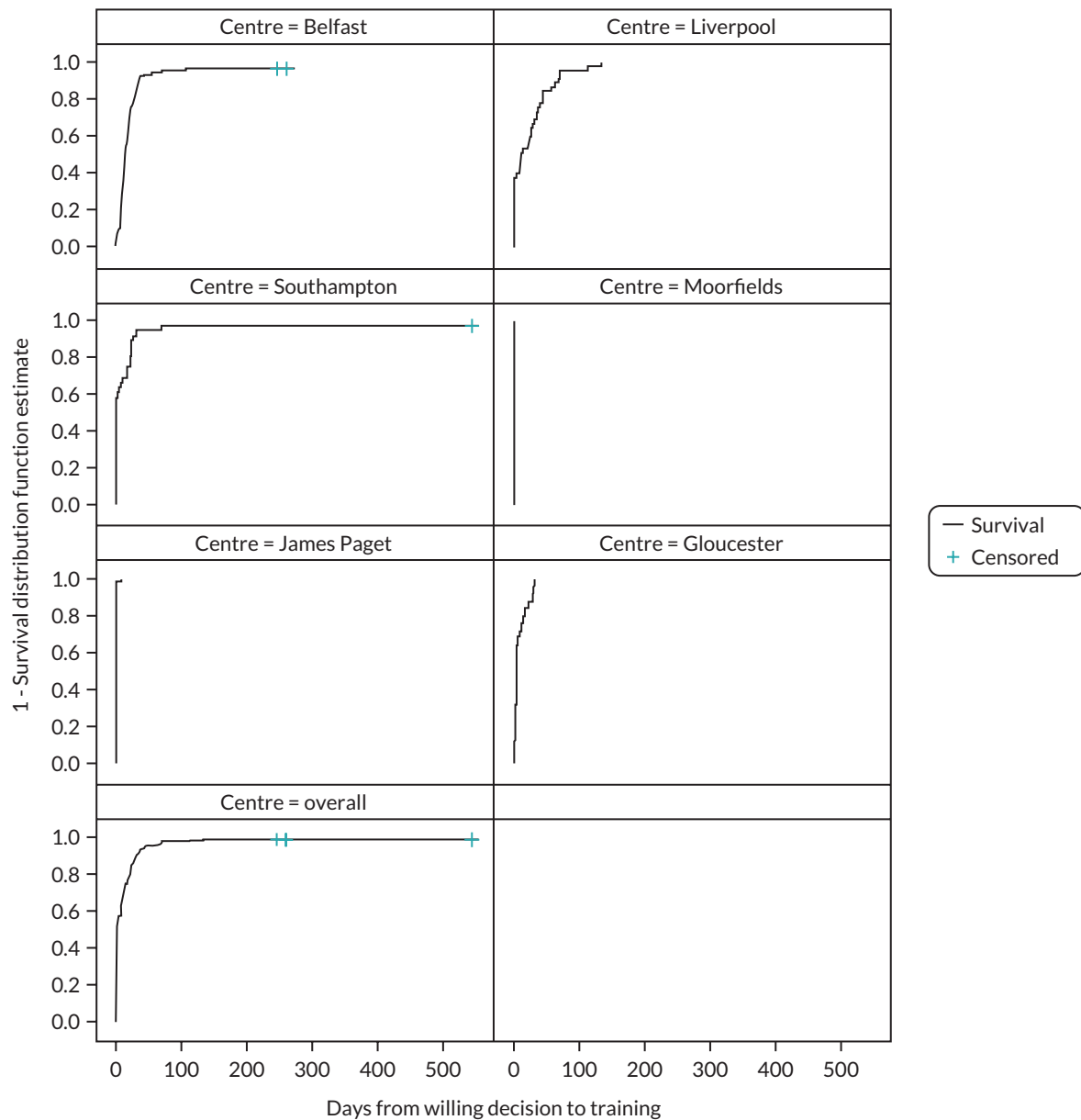
a The denominators for percentages at screening is the total number of ineligible patients (i.e. no study eye, < 50 years old or unable to understand English).

b The denominators for percentages is the total number of ineligible patients after information and training sessions are attended.



Censor ticks may indicate multiple censored observations

FIGURE 21 Time from screening to decision on willingness for eligible patients.



Censor ticks may indicate multiple censored observations

FIGURE 22 Time from willing in principle to training.

TABLE 33 Details of re-training; denominators are all consented patients

Site	Number of patients requiring retraining, n (%)
Belfast	4/82 (5)
Liverpool	0/45 (0)
Southampton	2/34 (6)
Moorfields	0/32 (0)
James Paget	11/80 (14)
Gloucester	1/24 (4)
Overall	18/297 (6)

TABLE 34 Visit rate per patient per year

Site	Number of patients with at least one follow-up, n	Total years of exposure ^a (min, max)	Total number of study visits, ^b n	Rate of visits (number of visits per patient per year)
Belfast	70	106.64	217	2.03
Liverpool	40	43.15	220	5.10
Southampton	34	53.34	141	2.64
Moorfields	29	36.47	233	6.39
James Paget	65	93.36	486	5.21
Gloucester	23	20.04	117	5.84
Overall	261	353.00	1414	4.01

a Exposure is calculated as most recent VA test at baseline to date of extract.

b The overall total (1414) does not take into account that some patients had two study eyes.

TABLE 35 Differences between intended and actual review schedules

Site	Number of visits, n	On or before intended review date, n (%)	Within 2 weeks of intended review date, n (%)	Within 4 weeks of intended review date, n (%)	> 4 weeks after intended review date, n (%)
Belfast	189	30 (15.9)	27 (14.3)	15 (7.9)	107 (56.6)
Liverpool	180	76 (42.2)	87 (48.3)	4 (2.2)	13 (7.2)
Southampton	133	36 (27.1)	33 (24.8)	17 (12.8)	47 (35.3)
Moorfields	205	151 (73.7)	36 (17.6)	7 (3.4)	11 (5.4)
James Paget	415	288 (69.4)	101 (24.3)	13 (3.1)	13 (3.1)
Gloucester	92	65 (70.7)	23 (25.0)	3 (3.3)	1 (1.1)
Overall	1214	646 (53.2)	307 (25.3)	59 (4.9)	192 (15.8)

Note

Initial follow-up visits are not included in table as the intended review date is not collected at baseline. The number of visits for Belfast includes 10 where no review schedule was given.

TABLE 36 Reasons for withdrawals from the study

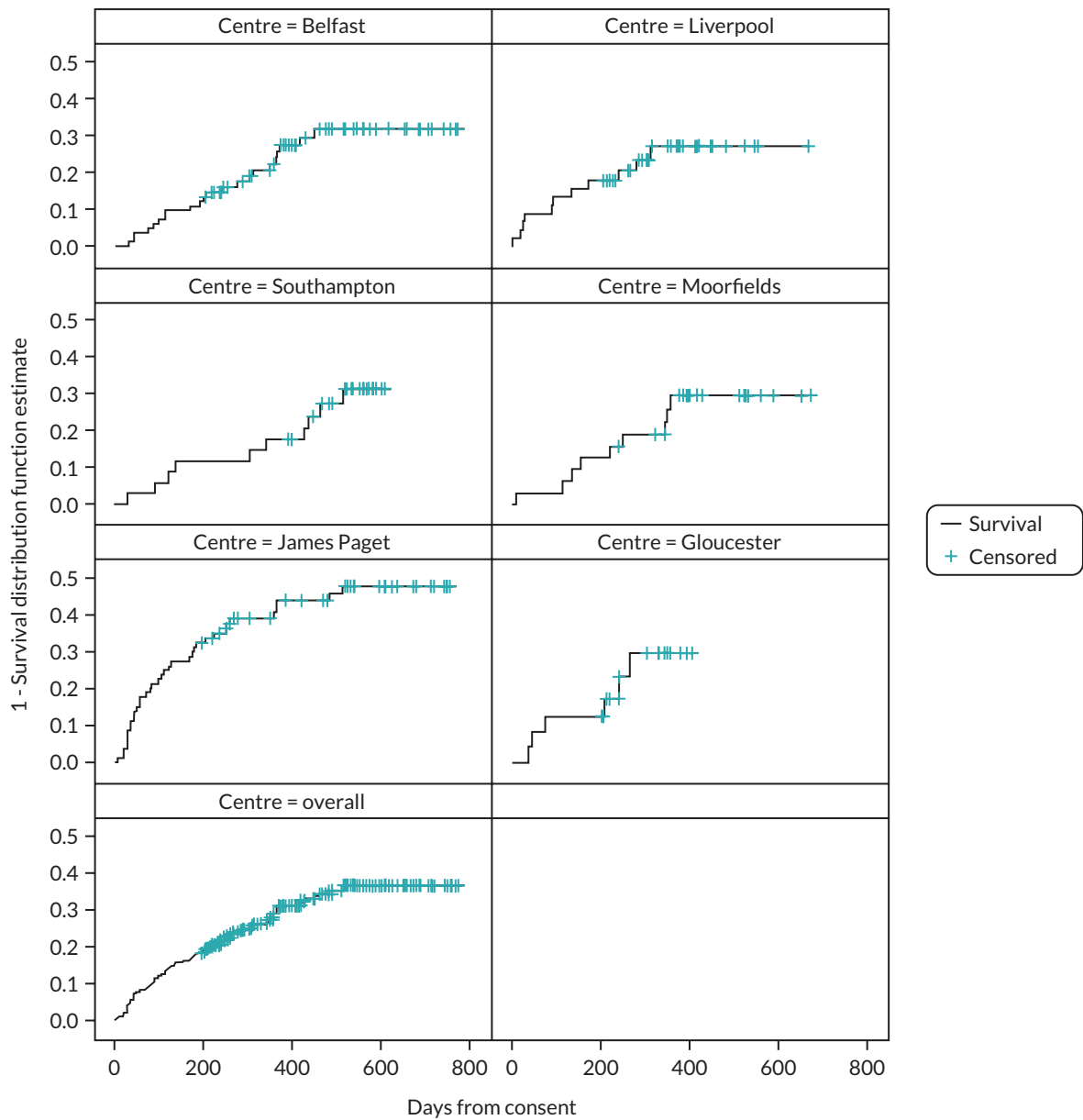
Withdrawal type	Withdrawal reason ^a	BEL (n = 82)	LIV (n = 45)	SOU (n = 34)	MOO (n = 32)	JAP (n = 80)	GLO (n = 24)	Overall (n = 297)
		n (%)	n (%)	n (%)	n (%)	n (%)	n (%)	n (%)
Overall		22 (27)	11 (24)	10 (29)	9 (28)	36 (45)	6 (25)	94 (32)
Patient's decision		21 (95)	9 (82)	8 (80)	8 (89)	36 (100)	6 (100)	88 (94)
	<i>Difficulties in operating equipment</i>	6 (29)	1 (11)	1 (13)	1 (13)	10 (28)	2 (33)	21 (24)
	<i>Home eye tests too time consuming</i>	3 (14)	1 (11)	- (-)	7 (88)	14 (39)	- (-)	25 (28)
	<i>Unhappy with mVT</i>	6 (29)	- (-)	1 (13)	2 (25)	5 (14)	- (-)	14 (16)
	<i>Unhappy with MBT</i>	10 (48)	- (-)	1 (13)	- (-)	5 (14)	- (-)	16 (18)
	<i>Unhappy with KSJ</i>	2 (10)	- (-)	1 (13)	- (-)	3 (8)	- (-)	6 (7)
	<i>Personal reasons</i>	12 (57)	7 (78)	7 (88)	5 (63)	15 (42)	4 (67)	50 (57)
	<i>Unwilling to give reason</i>	- (-)	- (-)	- (-)	- (-)	5 (14)	1 (17)	6 (7)
	<i>Other^b</i>	11 (52)	5 (56)	1 (13)	1 (13)	7 (19)	- (-)	25 (28)
Discharged from NHS review		- (-)	- (-)	1 (10)	1 (11)	- (-)	- (-)	2 (2)
Died		1 (5)	- (-)	- (-)	- (-)	- (-)	- (-)	1 (1)
Insufficient mobile coverage		- (-)	- (-)	1 (10)	- (-)	- (-)	- (-)	1 (1)
Information missing		- (-)	2 (18)	- (-)	- (-)	- (-)	- (-)	2 (2)

a Only patients' decisions have withdrawal reasons. Withdrawal reasons are not mutually exclusive.

b Other reasons include: BEL - 1 Tests too difficult, 2 Tests too difficult to see in light, 3 Medical reason, 2 Thought study only lasted for a year, 1 Made patient more aware of vision difficulties, 2 Tests too frustrating; LIV - 2 Patient unwell, 1 Tests too difficult to see in light, 1 Thought study only lasted for a year, 1 COVID-19 reasons; SOU - 1 Vision too poor; MOO - 1 Tests too frustrating; JAP - 1 Tests too difficult, 1 Tests too fiddly, 3 Tests too time consuming, 1 Equipment difficulties, 1 Vision too poor.

Note

Denominators are consented patients, except for withdrawal reasons where they are number of patients withdrawn due to a patient's decision.



Censor ticks may indicate multiple censored observations

FIGURE 23 Kaplan-Meier graphs of time to withdrawal by site.

TABLE 37 Completion of the KSJ home-monitoring test

Centre	Total patient-weeks in study	Number of booklets returned	Number of weeks fully completed ^a (%)	Number of weeks partially completed ^a (%)	Number of weeks when KSJ not completed ^a (%)	Days between completion (median, IQR) ^b
Belfast	4822	185	2830 (58.7)	525 (10.9)	1467 (30.4)	7 (7–8)
Liverpool	1979	90	1272 (64.3)	263 (13.3)	444 (22.4)	7 (7–7)
Southampton	2249	95	1402 (62.3)	172 (7.6)	675 (30.0)	7 (7–8)
Moorfields	1496	54	866 (57.9)	69 (4.6)	561 (37.5)	7 (7–8)
James Paget	4158	166	2365 (56.9)	262 (6.3)	1531 (36.8)	7 (7–7)
Gloucester	920	39	606 (65.9)	64 (7.0)	250 (27.2)	7 (7–7)
Overall	15,624	629	9341 (59.8)	1355 (8.7)	4928 (31.5)	7 (7–7)

a Week completion based on all three sections completed for all study and fellow eyes. Denominator for percentage is number of weeks patient in study.

b Where a partial or completed week has no date recorded, date of previous week is used.

TABLE 38 Frequency of mVT and MBT home-monitoring test use

Site	mVT			MultiBit		
	6–17 months, median tests per month (IQR)	18–29 months, median tests per month (IQR)	30–41 months, median tests per month (IQR)	6–17 months, median tests per month (IQR)	18–29 months, median tests per month (IQR)	30–41 months, median tests per month (IQR)
BEL	4 (3–4)	4 (2–4)	4 (2–4)	4 (2–4)	4 (2–4)	4 (1–4)
LIV	4 (2–4)	4 (3–4)	4 (2–4)	4 (1–4)	4 (2–4)	4 (2–4)
SOU	3 (1–4)	4 (3–5)	4 (3–4)	3 (1–4)	4 (3–4)	4 (3–4)
MOO	3 (1–4)	4 (2–4)	4 (0–4)	3 (2–4)	4 (2–4)	4 (1–4)
JAP	1.5 (0–4)	3 (0–4)	3 (1–4)	1 (0–4)	2 (0–4)	2 (0–4)
GLO	4 (2–4)	4 (3–5)	4 (4–5)	4 (2–4)	4 (4–5)	4 (4–5)
Overall	4 (1–4)	4 (2–4)	4 (1–4)	3 (1–4)	4 (2–4)	4 (1–4)

Note

In calculating rates, only complete calendar months are used from the date when equipment was given to the patient.

TABLE 39 Test completion for the mVT and MBT home-monitoring tests

	Belfast	Liverpool	Southampton	Moorfields	James Paget	Gloucester	Overall
MultiBit							
Patients testing (n)	77	44	34	30	57	21	263
Total number of tests ^a	5691	2587	2615	2325	2226	1228	16672
Study eyes tested ^b (n)/expected tests ^c (n)	3719/5967 (62.3%)	1516/2249 (67.4%)	1548/2499 (61.9%)	1507/2399 (62.8%)	1590/4723 (33.7%)	768/1100 (69.8%)	10,648/18,937 (56.2%)
Fellow eyes tested ^b (n)/expected tests ^c (n)	1559/2436 (64.0%)	795/1160 (68.5%)	872/1402 (62.2%)	701/1089 (64.4%)	498/1635 (30.5%)	397/647 (61.4%)	4822/8369 (57.6%)
mVT							
Total number of patients tested	78	44	34	30	74	21	281
Total number of tests ^a	6155	2595	2695	2198	2594	1245	17482
Study eyes tested ^b (n)/expected tests ^c (n)	3933/5967 (65.9%)	1567/2249 (69.7%)	1593/2499 (63.7%)	1527/2399 (63.7%)	1876/4723 (39.7%)	740/1100 (67.3%)	11,236/18,937 (59.3%)
Fellow eyes tested ^b (n)/expected tests ^c (n)	1692/2436 (69.5%)	824/1160 (71.0%)	885/1402 (63.1%)	598/1089 (54.9%)	596/1635 (36.5%)	423/647 (65.4%)	5018/8369 (60.0%)

a Includes cases where same eye is tested more than once in same day and/or week.

b Does not include instances where same eye is tested more than once in a week.

c Expected number of tests based on number of weeks in the study for each eye.

TABLE 40 Associations between different KSJ test scores

Reading comprehension	Household object		Total (%)
	Same/better (%)	Worse (%)	
Same/better	6759 (82.1)	366 (4.4)	7125 (86.5)
Worse	702 (8.5)	409 (5.0)	1111 (13.5)
Total	7461 (90.6)	775 (9.4)	8236 (100)^a
	Amsler grid		
Reading comprehension	Same/better (%)	Worse (%)	Total (%)
Same/better	6702 (79.7)	584 (6.9)	7286 (86.6)
Worse	620 (7.4)	508 (6.0)	1128 (13.4)
Total	7322 (87.0)	1092 (13.0)	8414 (100)^b
	Amsler grid		
Household object	Same/better (%)	Worse (%)	Total (%)
Same/better	7437 (85.3)	455 (5.2)	7892 (90.5)
Worse	150 (1.7)	680 (7.8)	830 (9.5)
Total	7587 (87.0)	1135 (13.0)	8722 (100)^c

a 1126 tests not completed.
b 948 tests not completed.
c 640 tests not completed.

TABLE 41 Model performance for predicting management decisions of lesion activity at thresholds defined by Youden's index for summary test scores for all index tests; primary analysis

Test	No test	MBT	mVT	KSJ
AUROC (95% CI)	0.554 (0.525 to 0.584)	0.586 (0.554 to 0.619)	0.572 (0.540 to 0.604)	0.574 (0.542 to 0.606)
Youden's index	0.09	0.14	0.12	0.13
Sensitivity	0.36	0.47	0.41	0.36
Specificity	0.73	0.67	0.72	0.77
PPV	0.66	0.68	0.68	0.70
NPV	0.44	0.46	0.45	0.45

NPV, negative predictive value; PPV, positive predictive value.

TABLE 42 Average test scores above and below thresholds defined by Youden's index for the models; primary analysis

	Below threshold		Above threshold	
	<i>n</i>	Average ^a (95% CI)	<i>n</i>	Average ^a (95% CI)
MBT	711	81.31 (79.88 to 82.75)	502	88.07 (87.03 to 89.1)
mVT	804	-0.38 (-0.39 to -0.36)	445	-0.47 (-0.49 to -0.45)
KSJ				
VA (median, IQR) ^b	860	5 (5-6)	378	5.25 (5-6)
VA worse	860	0.08 (0.07 to 0.09)	378	0.24 (0.2 to 0.27)
Amsler grid worse	860	0.06 (0.05 to 0.06)	378	0.24 (0.21 to 0.28)
Household object worse	860	0.04 (0.03 to 0.05)	378	0.16 (0.13 to 0.19)

a Average scores below/above the threshold are means for MBT and mVT, median for KSJ VA and proportions for the other three KSJ summary scores.

b VA is on a self-reported scale where 1 = worst and 6 = best.

TABLE 43 Model performance for predicting management decisions of lesion activity at thresholds defined by Youden's index for summary test scores for all index tests; sensitivity analysis 1

Test	No test	MBT	mVT	KSJ
AUROC (95% CI)	0.554 (0.524 to 0.584)	0.586 (0.554 to 0.619)	0.576 (0.542 to 0.609)	0.575 (0.542 to 0.608)
Youden's index	0.09	0.14	0.14	0.12
Sensitivity	0.36	0.47	0.46	0.47
Specificity	0.73	0.67	0.67	0.65
PPV	0.66	0.68	0.68	0.66
NPV	0.44	0.46	0.46	0.46

NPV, negative predictive value; PPV, positive predictive value.

TABLE 44 Average test scores above and below thresholds defined by Youden's index for the models; sensitivity analysis 1

	Below threshold		Above threshold	
	<i>n</i>	Average ^a (95% CI)	<i>n</i>	Average ^a (95% CI)
MBT	671	83.65 (82.30 to 85.01)	463	85.4 (83.93 to 86.88)
mVT	680	-0.36 (-0.38 to -0.35)	493	-0.47 (-0.49 to -0.45)
KSJ				
VA (median, IQR) ^b	985	5 (5-6)	117	5.5 (5-6)
VA worse	985	0.09 (0.08 to 0.10)	117	0.31 (0.26 to 0.36)
Amsler grid worse	985	0.07 (0.06 to 0.08)	117	0.35 (0.29 to 0.40)
Household object worse	985	0.04 (0.03 to 0.05)	117	0.25 (0.20 to 0.30)

a Average scores below/above the threshold are means for MBT and mVT, median for KSJ VA and proportions for the other three KSJ summary scores.

b VA is on a self-reported scale where 1 = worst and 6 = best.

TABLE 45 Model performance for predicting management decisions of lesion activity at thresholds defined by Youden's index for summary test scores for all index tests; sensitivity analysis 1

Test	No test	MBT	mVT	KSJ
Right eyes				
AUROC (95% CI)	0.611 (0.563 to 0.659)	0.602 (0.548 to 0.656)	0.646 (0.593 to 0.698)	0.636 (0.585 to 0.686)
Youden's index	0.23	0.23	0.32	0.28
Sensitivity	0.86	0.67	0.79	0.66
Specificity	0.37	0.56	0.53	0.62
PPV	0.75	0.78	0.79	0.78
NPV	0.55	0.42	0.52	0.47
Left eyes				
AUROC (95% CI)	0.5111 (0.4676 to 0.5547)	0.5767 (0.5294 to 0.624)	0.5427 (0.4961 to 0.5893)	0.5359 (0.4908 to 0.5811)
Youden's index	0.06	0.15	0.1	0.12
Sensitivity	0.26	0.49	0.57	0.17
Specificity	0.8	0.66	0.53	0.96
PPV	0.68	0.72	0.67	0.88
NPV	0.39	0.42	0.42	0.38

NPV, negative predictive value; PPV, positive predictive value.

TABLE 46 Average test scores above and below thresholds defined by Youden's index for the models; sensitivity analysis 1

	Below threshold		Above threshold	
	<i>n</i>	Average ^a (95% CI)	<i>n</i>	Average ^a (95% CI)
Right eyes				
MBT	224	74 (71.49 to 76.51)	337	91.09 (89.94 to 92.24)
mVT	179	-0.26 (-0.29 to -0.22)	404	-0.46 (-0.48 to -0.44)
KSJ				
VA (median, IQR) ^b	245	5 (4–6)	316	6 (5–6)
VA worse	245	0.18 (0.15 to 0.22)	316	0.1 (0.08 to 0.12)
Amsler grid worse	245	0.19 (0.15 to 0.23)	316	0.08 (0.06 to 0.10)
Household object worse	245	0.13 (0.09 to 0.16)	316	0.07 (0.05 to 0.09)
Left eyes				
MBT	348	91.99 (91.16 to 92.83)	270	74.22 (71.87 to 76.56)
mVT	296	-0.59 (-0.60 to -0.57)	335	-0.28 (-0.30 to -0.26)

continued

Table 46 Average test scores above and below thresholds defined by Youden's index for the models; sensitivity analysis 1 (continued)

	Below threshold		Above threshold	
	<i>n</i>	Average ^a (95% CI)	<i>n</i>	Average ^a (95% CI)
KSJ				
VA (median, IQR) ^b	952	6 (5–6)	85	5 (3–5)
VA worse	952	0.08 (0.07 to 0.10)	85	0.35 (0.27 to 0.43)
Amsler grid worse	952	0.05 (0.04 to 0.06)	85	0.45 (0.37 to 0.52)
Household object worse	952	0.03 (0.02 to 0.03)	85	0.31 (0.24 to 0.39)

a Average scores below/above the threshold are means for MBT and mVT, median for KSJ VA and proportions for the other three KSJ summary scores.

b VA is on a self-reported scale where 1 = worst and 6 = best.

TABLE 47 Model performance for predicting management decisions of a change from inactive to active lesion status for thresholds defined by Youden's index for summary test scores for index tests (except for the mVT)

Test	No test	MBT	mVT	KSJ
AUROC (95% CI)	0.554 (0.524 to 0.584)	0.586 (0.554 to 0.619)	0.576 (0.542 to 0.609)	0.575 (0.542 to 0.608)
Youden's index	0.09	0.14	0.14	0.12
Sensitivity	0.36	0.47	0.46	0.47
Specificity	0.73	0.67	0.67	0.65
PPV	0.66	0.68	0.68	0.66
NPV	0.44	0.46	0.46	0.46

NPV, negative predictive value; PPV, positive predictive value.

TABLE 48 Average test scores above and below thresholds defined by Youden's index for the models predicting management decisions of a change from inactive to active lesion status

	Below threshold		Above threshold	
	<i>n</i>	Average ^a (95% CI)	<i>n</i>	Average ^a (95% CI)
MBT	671	83.65 (82.30 to 85.01)	463	85.4 (83.93 to 86.88)
mVT	680	-0.36 (-0.38 to -0.35)	493	-0.47 (-0.49 to -0.45)
KSJ				
VA (median, IQR) ^b	931	5 (5–6)	307	6 (5–6)
VA worse	931	0.11 (0.10 to 0.13)	307	0.16 (0.13 to 0.19)
Amsler grid worse	931	0.10 (0.09 to 0.12)	307	0.14 (0.11 to 0.17)
Household object worse	931	0.06 (0.05 to 0.07)	307	0.13 (0.10 to 0.17)

a Average scores below/above the threshold are means for MBT and mVT, median for KSJ VA and proportions for the other three KSJ summary scores.

b VA is on a self-reported scale where 1 = worst and 6 = best.

TABLE 49 Model performance for predicting management decisions of fellow eye conversion for thresholds defined by Youden's index for summary test scores for index tests, primary analysis

Test	No test	MBT	mVT	KSJ
Visits	544	468	479	480
Participants	132	118	126	115
AUROC (95% CI)	0.725 (0.599 to 0.851)	0.732 (0.613 to 0.852)	0.758 (0.645 to 0.871)	0.8521 (0.794 to 0.911)
Youden's index	0.47	0.45	0.53	0.63
Sensitivity	0.71	0.76	0.82	0.88
Specificity	0.76	0.69	0.71	0.75
PPV	0.09	0.08	0.09	0.11
NPV	0.99	0.99	0.99	0.99

NPV, negative predictive value; PPV, positive predictive value.

TABLE 50 Average test scores above and below thresholds defined by Youden's index for the models predicting management decisions of fellow eye conversion, primary analysis

	Below threshold		Above threshold	
	<i>n</i>	Average ^a (95% CI)	<i>n</i>	Average ^a (95% CI)
MBT	314	95.17 (94.5 to 95.84)	154	82.93 (79.7 to 86.17)
mVT	331	-0.67 (-0.69 to -0.66)	148	-0.41 (-0.44 to -0.39)
KSJ				
VA (median, IQR) ^b	354	6 (6-6)	126	5.75 (5-6)
Household object worse	354	0.01 (0 to 0.01)	126	0.12 (0.08 to 0.16)

a Average scores below/above the threshold are means for MBT and mVT, median for KSJ VA and proportions for the other KSJ summary score.

b VA is on a self-reported scale where 1 = worst and 6 = best.

TABLE 51 Model performance for predicting management decisions of fellow eye conversion for thresholds defined by Youden's index for summary test scores for index tests, sensitivity analysis 1

Test	No test	MBT	mVT	KSJ
AUROC (95% CI)	0.725 (0.599 to 0.851)	0.754 (0.635 to 0.873)	0.759 (0.647 to 0.871)	0.895 (0.821 to 0.969)
Youden's index	0.47	0.54	0.48	0.64
Sensitivity	0.71	0.88	0.75	0.93
Specificity	0.76	0.66	0.73	0.71
PPV	0.09	0.09	0.09	0.09
NPV	0.99	0.99	0.99	1

NPV, negative predictive value; PPV, positive predictive value.

TABLE 52 Average test scores above and below thresholds defined by Youden's index for the models predicting management decisions of fellow eye conversion, sensitivity analysis 1

	Below threshold		Above threshold	
	<i>n</i>	Average ^a (95% CI)	<i>n</i>	Average ^a (95% CI)
MBT	284	96.2 (95.7 to 96.7)	158	82.4 (79.2 to 85.7)
mVT	320	-0.68 (-0.7 to -0.66)	127	-0.39 (-0.42 to -0.36)
KSJ				
VA (median, IQR) ^b	316	6 (6-6)	140	6 (5-6)
VA worse	316	0.03 (0.02 to 0.04)	140	0.12 (0.09 to 0.16)
Household object worse	316	0.02 (0.01 to 0.02)	140	0.09 (0.06 to 0.13)

a Average scores below/above the threshold are means for MBT and mVT, median for KSJ VA and proportions for the other KSJ summary scores.

b VA is on a self-reported scale where 1 = worst and 6 = best.

TABLE 53 Model performance for predicting management decisions of fellow eye conversion for thresholds defined by Youden's index for summary test scores for index tests, sensitivity analysis 2

Test	No test	MBT	mVT	KSJ
Visits	537	461	472	470
Participants	131	117	125	114
AUROC (95% CI)	0.5059 (0.3531 to 0.6586)	0.5587 (0.4155 to 0.702)	0.6012 (0.4700 to 0.7323)	0.5832 (0.4117 to 0.7546)
Youden's index	0.15	0.21	0.24	0.25
Sensitivity	0.40	0.80	0.73	0.43
Specificity	0.75	0.41	0.51	0.82
PPV	0.04	0.04	0.05	0.07
NPV	0.98	0.98	0.98	0.98

NPV, negative predictive value; PPV, positive predictive value.

TABLE 54 Average test scores above and below thresholds defined by Youden's index for the models predicting management decisions of fellow eye conversion, sensitivity analysis 2

	Below threshold		Above threshold	
	<i>n</i>	Average ^a (95% CI)	<i>n</i>	Average ^a (95% CI)
MBT	187	96.0 (95.3 to 96.7)	274	87.9 (85.9 to 89.9)
mVT	238	-0.7 (-0.71 to -0.68)	234	-0.49 (-0.51 to -0.46)
KSJ				
VA (median, IQR) ^b	288	6 (6-6)	182	5 (4-6)
VA worse	288	0.05 (0.04 to 0.06)	182	0.1 (0.05 to 0.14)
Household object worse	288	0.04 (0.02 to 0.05)	182	0.04 (0.01 to 0.06)

a Average scores below/above the threshold are means for MBT and mVT, median for KSJ VA and proportions for the other KSJ summary scores.

b VA is on a self-reported scale where 1 = worst and 6 = best.

TABLE 55 Relationship between themes and coding of interview data

Themes	The role of home monitoring		Suitability of procedures and instruments	Experience of procedures			Feasibility of regular home monitoring in usual service delivery		Impediments to home monitoring		
<i>Subthemes</i>	<i>Understanding purpose</i>	<i>Perceived impact on eye care</i>		<i>Training for home monitoring</i>	<i>Test preferences</i>	<i>Use of MBT test feedback and data</i>	<i>Frequency of home monitoring and habit formation</i>	<i>Use of ongoing support</i>	<i>Practical issues</i>	<i>Personal health and social factors</i>	
<i>Codes</i>	Relevance to eye care	Perceived usefulness	Individual preferences and use of technology	Technology apprehension	Perceived ease of use	Perceived threat of deterioration	Burden/perceived effort	Check-in calls for non-testers	Technology reliability (3)	Health status (2)	
	Threat of deterioration in vision	Job relevance	Technology apprehension	Information quality	Burden/perceived effort	Results demonstrability	Technology persistence	Family support (not subjective norm) (2)	Burden/perceived effort	Social context (2)	
	Perceived usefulness	Waiting time a concern	Individual differences – age (1)	Views on training	Actual behaviour/acceptance	Job relevance	Habit		Output quality	Individual differences – age (2)	
	Intervention coherence	Comfort with services	Attitude		Self-efficacy	Reported performance	Self-efficacy				
		Professionalism of healthcare staff	Behavioural intention to HM		System characteristics	Self-efficacy	Actual behaviour/acceptance				
		Reduced travel	Information quality		Reported experience		Family support (not subjective norm) (1)				
			Health status (1)		Technology reliability (1)						
		Image		Perceived enjoyment		Technology reliability (2)					
				Technology playfulness/comfort		Health professional opinion (similar to subjective norm)					
Note Numbers in brackets represent codes which were evident across different themes, and highlight the number of times the code emerged.											

TABLE 56 Perspectives of HCPs

HCP views	Patient theme/ subtheme	Supporting quote(s) from HCPs
Factors such as age, and lack of familiarity with technology were linked to participant's initial interest in HM and engagement with testing.	Theme 2. Suitability of procedures and instruments	'... I know how old they are before I'm approaching them and so there are misconceptions on my side because I see someone who's 80 and I think they may not be willing to participate, but they can be compared to somebody who's in their 60s or early 70s. But yes, I think older people, are just not used to an iPod'. (HCP03, Female) '... I think they split into two groups. The first is definitely the elderly, this technology is beyond them and they almost take out their phones to show you what they are accustomed to. The others are keen and interested; they like the concept'. (HCP06, Female)
Training was essential to success and to long-term engagement. It was recognised that training was information heavy and often had to be tailored based on previous experience with technology. Participants may require further training, not just a single session.	Theme 3. Experience of home-monitoring procedures Subtheme 1: Training for home monitoring	'... others are at a stage where they learn how to use technology and every piece of information is crucial. We spend a lot more time on those who are brand new users of technology because if we want a good starting point we will have more difficulties later if they are not able to remember'. (HCP02, Female) 'I think just another check in is needed, I don't think maybe the one visit is enough actually, maybe two are needed'. (HCP05, Female)
Participants preferred the MBT (which provides performance feedback ^a) and most found this to be positive (allowing self-monitoring), but there were also negatives as uncertainty over meaning of performance scores could lead to anxiety.	Theme 3. Experience of home-monitoring procedures Subtheme 3: Use of test feedback and data	'... for one of the patients that have withdrawn, one of the biggest things has been the MultiBit test, I would say, and the percentage score given at the end. I think a lot of people find that distressing because they perceive the number to be lower than what they would like, although what benchmark they are using I don't know'. (HCP05, Female) '... that's something that nearly every single participant that you speak to says is something that really distresses them (how percent scores are calculated) but nobody really knew how it was derived. To tell patients not to pay too much heed to it because it's not necessarily derived from their visual performance was really useful'. (HCP01, Female)
Home monitoring was seen as an acceptable part of patient care but may be resource intensive due to the technical support needed to support patients.	Theme 4. Feasibility of regular home monitoring in usual service delivery Subtheme 2: Use of ongoing support	'... because it is indeed involving a lot of resources, it will be probably less demanding if the patient will need to come to see the specialist every couple of months. But equally, we will need to have enough provision of support for them when it's needed and it should be straight forward access' (HCP02, Female) '... we have an understanding that a patient may experience this (connection problems) at home and despite advising them over the phone, there's not always a straight-forward solution to it'. (HCP08, Male)
Participants required additional support, through this was usually informal (delivered at follow-up visits or over the phone).	Theme 4. Feasibility of regular home monitoring in usual service delivery Subtheme 2: Use of ongoing support	'... we have to compensate when we speak with people on the phone but if we go to clinic we sometimes ask people to bring the iPod so we can offer guidance on the spot, in case it's something very simple to address'. (HCP01, Female)
A mechanism for providing 'feedback ^b ' messages (on test completion) would be beneficial (ensuring participants knew test data had transferred).	Theme 5. Impediments to home monitoring Subtheme 1: Practical issues	'... another thing fed back at every single follow up, pretty much without fail, is that they get no feedback in terms of whether their results have been received by the study team or not'. (HCP03, Female) '... it doesn't matter how much you reassure that they would be contacted if results weren't being received. They is no substitute for even just a "Thank you for completing the tests, your results have been sent" or something like that'. (HCP01, Female)

a 'Feedback' refers to information on successful test completion.

b Healthcare professionals were research ophthalmologists and nurses.

Note

'Performance feedback' refers to information on percentage scores provided on completion of the MBT.

TABLE 57 Perspectives of informal 'carers'

Views of informal 'carers'	Patient theme/ subtheme	Supporting quote(s) from informal carers
Home monitoring could be a way to reduce the frequency of clinic monitoring visits.	Theme 1. The role of home monitoring Subtheme 2: Perceived impact on eye care	<i>'I would be in favour of this self-testing if it became a permanent thing, it would save the journeys back and forward to the hospital. The fact that (partner) is prepared to shut herself away for half an hour every Friday morning and do the test, it leaves me just as somebody to answer the phone or callers to the door, it saves us a lot of time and effort.'</i> ('Carer' 10, Male)
Training was important but information recall deteriorated after home monitoring started. Partners should be involved in training so that they can support their husband/wife/partner if needed.	Theme 3. Experience of home-monitoring procedures Subtheme 1: Training for home monitoring	<i>'... so, he did need a bit of assistance and he thought he knew what he was doing after he tried it there (at the training session) but he'd forgotten a bit. If you had two people that were non-technically minded, that might be a bit of an issue, you know. if the person (a partner or carer) hadn't been to it, there would be less chance of them understanding how to help do the actual test.'</i> ('Carer' 05, Female)
Use of test feedback provided a way to monitor changes in vision. Changes in vision, including improvements in response to treatment, may be reflected in test scores.	Theme 3. Experience of home-monitoring procedures Subtheme 3: Use of test feedback and data	<i>'... but we knew it was about time in to get another injection, know what I mean? I just found it interesting that she did seem to go downhill the 2 weeks before that appointment.'</i> ('Carer' 03, Male) <i>'... other than the fact that maybe after the injection, he got higher scores.'</i> ('Carer' 01, Female)
A reduction in feedback scores could be due to deterioration in visual acuity. Patients applied their own thresholds to what was a meaningful change.	Theme 3. Experience of home monitoring procedures Subtheme 3: Use of test feedback and data	<i>'... and she (partner) is a lot more confident with it (home monitoring) now but sometimes gets disturbed at the results because there was a period where the scores were good and consistent, which suggested that there was no deterioration. Then 1 week, she came down to me and said I don't know what went wrong but the results were terrible. Instead of being over 50% it dropped to 40%. I said, maybe you weren't feeling the best, let's leave it another week and the following week it was back up to 50%.'</i> ('Carer' 09, Male)
Support for home monitoring could be provided by partners or family members if they were more familiar with using technology. Experience of testing over time may reduce the need for this support.	Theme 4. Feasibility of regular home monitoring in usual service delivery Subtheme 2: Use of ongoing support	<i>'... there were times when messages came up on the screen but all it needed was a particular button to be pressed or whatever and she wasn't sure but I was able to sort it – it was simple little things like that. Eventually she found she had the ability and the confidence to do it. It would be very rare for me to have to help her now with the actual machine.'</i> ('Carer' 02, Male) <i>'... we thought we could call in our grandchildren, if we have bother with how to work it? However, it seemed straightforward. I would be more used to technology than my husband; he would prefer to be in a book than on an iPod.'</i> ('Carer' 01, Female)

TABLE 58 Common issues reported by patients during home monitoring

Reported issues	Patient's responses to issues	Relevant constructs of technology acceptance theories	Definition of construct
Successful methods of finding a dark space when completing the MBT inhibited actual performance of the test.	Test at night-time or find/create a home area with no light (e.g. under stairs, placing a blanket or material overhead).	Opportunity Costs ^a Burden ^a Effort Expectancy ^b	Extent to which HM causes benefits or values to be given up Perceived effort required to participate in HM Extent to which HM is seen as easy to do
The need to charge the iPod device before every occasion of testing.	Devices left on charge or paper-based KSJ completed before App-based tests.	Burden ^a Effort Expectancy ^b Output Quality ^c Perceived Effectiveness ^a	Perceived effort required to participate in HM Extent to which HM is seen as easy to do Extent to which HM performs its task well Extent to which HM is likely to achieve its purpose
Mobile internet connection for transferring test data appeared to function inconsistently and participants were not aware or were uncertain whether or not test data were received by the study.	Home broadband connections used as an alternative where feasible.	Affective Attitude ^a Perceived Effectiveness ^a Perceived Risk ^b Output Quality ^c	How individuals feel about HM Extent to which HM is likely to achieve its purpose Perceived negative consequences associated with HM Extent to which HM performs its task well
Other technical issues, e.g. apps resting during tests, software updates and uncertainty among participants about how to undertake tasks such as updating software.	Use of study helpline for support or support from family members or friends	Self-efficacy ^a Effort Expectancy ^b Perceived Risk ^b Output Quality ^c	Confidence in performing HM procedures Extent to which HM is seen as easy to do Perceived negative consequences associated with HM Extent to which HM performs its task well
Difficult to establish a testing routine due to commitments including caring roles and work responsibilities.	Use of reminders and prompts including diaries, self-administered phone reminders and testing at a set time/day	Burden ^a Ethicality ^a Opportunity Costs ^a Social Influence ^b Subjective Norm ^c Individual Differences ^b Habit ^b	Perceived effort required to participate in HM Extent to which HM has a good fit with the individuals' values Extent to which HM causes benefits or values to be given up Extent to which important others believe HM should be performed Extent to which individual factors influence HM Intention to HM and continued use of HM
Feedback from tests led to anxiety if scores declined consistently and were attributed to deterioration in vision.	Consider contact-ing optometry/ ophthalmology services.	Self-efficacy ^a Intervention Coherence ^a Perceived Effectiveness ^a Performance Expectancy ^b /Job Relevance ^c Results Demonstability ^c	Confidence in performing HM procedures Extent to which HM purpose and how it works is understood Extent to which HM is likely to achieve its purpose Extent to which HM will help with eye care Extent to which results are observable and communicable

TABLE 58 Common issues reported by patients during home monitoring (*continued*)

Reported issues	Patient's responses to issues	Relevant constructs of technology acceptance theories	Definition of construct
Problems or issues with testing – if they were not easy to overcome without support, could lead to testing becoming irregular or stopping.	Use of 'problem solving' and support from study helpline or family members or friends.	Self-efficacy ^a Affective Attitude ^a Burden ^a Facilitating conditions ^b	Confidence in performing HM procedures How individuals feel about HM Effort required to participate in HM Extent to which infrastructure exists to support HM
Eye patches that were used when undertaking testing caused discomfort and were a distraction, thereby affecting test performance.	Modifications made to the eye patch.	Burden ^a Perceived Effectiveness ^a	Perceived effort required to participate in HM Extent to which HM is likely to achieve its purpose

HM, home monitoring.

a Theoretical Framework of Acceptability.^{35,49}

b Unified Theory of Acceptance and Use of Technology.^{33,50,51}

c Technology Acceptance Model.^{34,52}

Appendix 2 Summary of protocol substantial amendments

Amendment number	Updated protocol version	Updated protocol date	Brief summary of change
Substantial Amendment 001	V2.0	8 May 2018	Rewording eligibility criteria for clarification; amendment of visual acuity cut-off for eligible eyes from 6/60 to 6/24; amending eligibility criteria to specify first treatment (rather than diagnosis) must be ≥ 6 and < 42 months ago; inclusion of collection of participant self-reported data; clarifying participant withdrawal and discharge procedures; clarifying that visual acuity will be collected for both eyes, reference to 'in the better seeing eye' is deleted; correction that MiFi routers are to be provided to all participants; corrections to study schema for objectives A, C and D and objective B; clarifying eye classification for the study; clarifying that home-monitoring testing should be completed for all study and fellow eyes; amending 'non-seeing eye' to 'excluded eye'; clarifying definition of follow-up and procedures required; replacement of 'routine' with 'usual care'; providing further details on data to be collected; inclusion of collection of patient self-reported data for each follow-up visit, removal of observations of participant's home environment from plan of investigation for Objective B; clarifying co-enrolment; addition of a further reference on costs of research; updating CTEU contact details, inclusion of post-doctoral research fellow contact details.
Substantial Amendment 003	V3.0	19 March 2019	Visual acuity eligibility criterion lowered from Snellen score 6/24 to 6/60; added interviews with HCPs to Objective B; carers may be approached during a home visit with a patient participant; clarification that under certain circumstances excluded eyes can become study eyes in participants that are already taking part in the study; updated sponsor and study team details; corrected minor formatting, grammatical and typographical errors; updated ownership details of the MvT test.
Substantial Amendment 004	V4.0	1 November 2019	Any selected site may now take part in the qualitative study, not just Belfast, Moorfields and Southampton; patients who decline participation in the main study can provide verbal consent to be contacted regarding participation in the qualitative study; patients may be offered travel reimbursement for attendance of Information and Training visits; updated study management team details.

Appendix 3 Statistical analysis plan

MONARCH

Statistical Analysis Plan

Introduction

Summary of document

Scope

The statistical analysis plan (SAP) for the MONARCH study has been written in accordance with Bristol Trials Centre (BTC) standard operating procedure SOP_ST-001 and International Conference on Harmonisation (ICH) Statistical Principles for Clinical Trials E9, by Dr Eleanor Gidman, study statistician in the BTC, under the supervision of Professor Chris Rogers, statistical lead in the BTC, and covers all final statistical analyses to be performed, outlined in the study protocol found in the study master file (<https://fundingawards.nihr.ac.uk/award/15/97/02>).

Planned analyses and dissemination

The end-of-study report will be based on data collected up until the 30 September 2020 and will be disseminated to the Study Management Group when all pre-specified final analyses have been performed. An independent steering committee (SSC) will review the conduct of the study.

Background of study

Study summary

Wet AMD (nAMD) is the commonest cause of blindness in the UK. It involves new vessels growing and leaking at the back of the eye. Recent treatments for wet AMD have led to a significant reduction in the number of wet AMD patients being registered blind. However, providing prompt access to clinics for regular surveillance and treatment has proved a major challenge for the NHS. Most patients need a series of monthly injections followed by a period of regular check-up visits in case more injections are required. AMD can often flare up after a period when treatment has not been required, so check-ups are usually needed for several years.

Home monitoring to detect the need for treatment could mean that patients would not need regular hospital check-ups, allowing HES clinic appointments to be kept for patients likely to require treatment. If home monitoring indicated treatment might be required, patients could request an urgent clinic appointment. Home monitoring would be more convenient and less costly for both the patient and the NHS. The main aim of our study is to find out whether our chosen home-monitoring tests can detect when wet AMD needs to be treated as well as the surveillance tests currently carried out at hospital check-ups.

We have chosen three home-monitoring tests for which some promising results have already been reported. The tests span a range of both technical complexity and cost. The most simple and inexpensive is a paper booklet (KSJ) of self-administered 'reading tests' with space for patients to record their results on a weekly basis. The other two tests are apps that display numbers, shapes or patterns on an Apple iPod Touch patients indicate by entering information on the screen which of four shapes is the 'odd-one-out' or articulate what numbers appear briefly on the screen. Their responses will be sent to the research team through the internet.

Approximately 1620 existing patients having treatment or check-ups at participating NHS hospitals will be invited to take part in home monitoring (see Study Schema in [General considerations](#)). They will be asked to perform the home-monitoring tests weekly at home in between standard hospital check-ups over a period of 1–2 years. To ensure our results will apply to most patients needing treatment or surveillance, we will recruit patients first treated for wet AMD 6–41 months previously. Patient participants will be trained to perform the home-monitoring tests by appropriately qualified members of the local research team with experience in communicating with patients with AMD. They will also have the opportunity for refresher training throughout their participation in the study. At selected participating NHS sites, we will undertake an integrated qualitative study; a sample of patients and their carers will be interviewed to find out their experiences of performing the tests, focusing on the difficulties experienced and what could be done to make the home-monitoring tests more acceptable. In addition, HCPs involved in training and recruitment of patients will be interviewed to explore reasons why patients decline to participate in the home-monitoring study and what the barriers to uptake of such telemedicine methods might be in this population.

Study rationale

The development and implementation of care pathways for anti-VEGF treatment for a large and growing number of patients have put considerable pressure on HES. Many patients remain under regular review for several years after starting treatment. If patients could self-monitor their vision for reactivation of nAMD at home, this would be a significant advantage. Mobile phone technology allows data to be transmitted to a hospital without the need for patients to interpret test results, making home monitoring practicable.

Test accuracy of tests for self-monitoring neovascular age-related macular degeneration activity

The advent of tablet computers and mobile/wireless technology has led to the development of devices for self-monitoring of visual function in nAMD (2). The disadvantages of the standard Amsler chart have long been recognised; its sensitivity to detect the onset of nAMD has been estimated to be only 50–70% (3). Perceptual completion (4) and the inability of patients to understand the test or reliably report the results are thought to contribute to poor performance (2).

Reactivation of nAMD is more difficult to detect because some patients have distortion due to scarring and photoreceptor disorganization in the absence of disease activity; therefore, a test has to enable patients to perceive an *increase* in distortion rather than solely its *presence*. Newer technologies such as VMS grids (5), preferential hyperacuity perimetry home devices (6, 7) and shape discrimination tests (8–12) have been reported to quantify distortion more accurately than either the Amsler grid or visual acuity in clinical settings (2).

This study investigates the test accuracy of home-monitoring ‘index’ tests to detect reactivation. We are only studying index tests with supporting peer-reviewed literature and usability data; one uses paper-and-pencil and two use modern information technology, implemented as apps on an iPod Touch.

Potential inequalities in uptake

The study also aims to address the question: How do demographic, socioeconomic and visual function factors influence the uptake of home-monitoring tests for detecting active nAMD? Outcomes characterising uptake and exposures of interest are defined in *Section 2.6.2*.

A survey by Age UK in 2013 found that internet use among people aged 65 years or over varied across the UK, with a ‘north–south’ divide; more than 50% in the south (Surrey, Bedfordshire, Buckinghamshire, Suffolk and Oxfordshire) used the internet but less than a third in the north (Cumbria, Yorkshire, Hull,

Tyne and Wear).²³ With respect to smartphone use, only 20% of 65- to 74-year-olds used such a device to access the internet in 2013;²⁴ perhaps more importantly, this percentage had increased from only 12% in 2012, suggesting that the situation is changing rapidly over time. The potential importance of failure to access the internet has been highlighted by a study of men and women in the English Longitudinal Study of Ageing from 2004 to 2011;²⁵ internet use was found to be significantly 'protective against health literacy decline'.

The small percentage of regular internet and smartphone users is a potential threat to the study, especially if potential participants feel alienated by the technology and are not prepared to try out the solutions we propose. Assuming that we are able to recruit our target sample size, it is still important to determine the extent to which the technology is a barrier to consent and participation in order to project wider adoption of home monitoring in the future if it is found to have satisfactory performance. Moreover, among participants, it is possible that some tests will be easier to do for participants with limited experience of smart devices and the internet. This would be an important factor to weigh against test performance if differences in test performance were found to be small.

Aims and objectives

The aim of the MONARCH study is to quantify the performance of three non-invasive test strategies for use by patients at home to detect active nAMD compared to diagnosis of active nAMD during usual monitoring of patients in the HES.

The study has four objectives:

- A. Estimate the test accuracy of three tests to self-monitor reactivation of nAMD compared to the reference standard of detection of reactivation during hospital follow-up with OCT imaging, clinical examination and EDTRS visual acuity.
- B. Determine the acceptability of the tests to patients and carers and their adherence to home-monitoring testing regimens.
- C. Explore whether inequalities (by age, sex, socioeconomic status and visual acuity) exist in recruitment to the study and impact the ability of participants to do the tests during follow-up and the adherence of participants to weekly testing.
- D. Provide pilot data for the use of home monitoring to detect conversion to nAMD in the fellow eyes of patients with unilateral disease, compared to the reference standard of detection of conversion during hospital follow-up with EDTRS visual acuity and OCT imaging.

Professor Michael Donnelly and Dr. Charlene Treanor at Queen's University, Belfast, will complete the analysis of Objective B. The BTC will carry out study analyses in collaboration with the clinical investigators associated with Objectives A, C and D.

This SAP will cover the statistical analysis of Objectives A, C and D.

Study methods

Design

MONARCH is a non-interventional, multicentre, diagnostic test accuracy cohort study for estimating the sensitivity and specificity of home-monitoring tests in detecting nAMD activation in patients with inactive nAMD lesions.

Framework

The study will be analysed and reported in line with the reporting guidelines for studies of diagnostic accuracy.⁵⁴

Sample size

The sample size as justified in the protocol was for recruiting at least 400 participants with 2300 clinic visits. However, recruitment did not achieve that target and so the sample size will include all patients consented prior to the end of recruitment (31 March 2020) with all clinic visits undertaken prior to the end of the study (30 September 2020). For Objective C, the sample size will be expanded to include all approached patients prior to the end of recruitment. For Objective D, it will be reduced to those patients consented with at least one fellow eye.

Populations

Study populations

Inclusion criteria

A participant may enter the study if the participant is aged ≥ 50 years and has at least one study eye meeting the inclusion criteria.

A potential study eye may be included if ALL of the following apply:

- eye first treated for active nAMD ≥ 6 months ago
- eye first treated for active nAMD not more than 42 months ago
- eye currently being monitored for nAMD disease by the NHS.

Exclusion criteria

A participant may not enter study if the participant does not have at least one study eye.

A potential study eye will be excluded if ANY of the following apply:

- vision in the potential study eye worse than Snellen score 6/60, LogMar 1.04 or 33 letters
- vision in the potential study eye is limited by another eye condition other than nAMD
- surgery in the potential study eye in the previous 6 months
- refractive error in the potential study eye $> -6D$
- retinal or choroidal neovascularization in the potential study eye not due to nAMD.

In addition, a participant will be excluded if ANY of the following apply:

- inability to do one or more of the proposed tests as assessed during 'further information and training' session
- unable to understand English
- home or personal circumstances are unsuitable for home testing.

Data sources

Study data will be collected on CRFs, except for the following:

- **MultiBit home-monitoring test data:** These data will be downloaded from the MultiBit online portal.
- **mVT home-monitoring test data:** These data will be downloaded from the mVT online portal.
- **CARF image grading data:** These data will be sent from the CARF team as an Excel spreadsheet.
- **KSJ data:** These data are returned by the patient to the BTC, where it will be entered onto the study database.

Study ID will be used to link all data sources. CRF linkage to home-monitoring test data will also take account of date of home-monitoring test and date of management visit. In these cases, home-monitoring tests will be associated with management visits if their dates are before the visit in question and after the preceding visit.

Analysis populations

Objective A

All summaries and analyses will be conducted on study eyes within the consented population.

Patient study eye data will be included if the eye fulfils the following criteria:

- Patient study eye has at least one management visit with a reviewing ophthalmologist's decision on lesion presence **AND** lesion activity status **AND** patient study eye has home-monitoring test records for at least one home-monitoring test that can be associated with the interval between the management visit and a preceding hospital appointment at which lesion activity status was available (i.e. either baseline data or a preceding management visit).

Patient study eye data will not be included under the following circumstances:

- Management visits without a reviewing ophthalmologist's decision on lesion activity status will not be used in the analysis. For instance, remote visits occurring during the COVID-19 pandemic.
- Management visits without an associated OCT will not be included in the analysis.
- Home-monitoring tests that cannot be associated with a management visit will not be used in the analysis. For instance, home-monitoring tests towards the end of the study where the next management visit does not occur until after the end of the follow-up.

Note that the exclusion criteria may result in patient study eyes being assessed in one home-monitoring test model and not another if they only ever self-tested with one or other of the tests.

Objective C

Summaries and analyses will be conducted on the *consented* population when examining impact of inequalities on participants' ability or adherence in performing weekly testing.

Summaries and analyses will be conducted on the *approached* population, regardless of whether they consent, when examining impact of inequalities on participant recruitment.

Objective D

All summaries and analyses will be conducted on the fellow eyes in the consented population.

Patient fellow eye data will be included if the fellow eye fulfils the following criteria:

- an eye has been designated as a fellow eye for some observation time
- the fellow eye has at least one management visit with a reviewing ophthalmologist's decision on lesion presence **AND** lesion activity status **AND** patient fellow eye has home-monitoring test records for at least one home-monitoring test that can be associated with the interval between the management visit and a preceding hospital appointment at which lesion activity status was available (i.e. either baseline data or a preceding management visit).

Patient fellow eye data will not be included under the following circumstances:

- management visits without a reviewing ophthalmologist's decision on lesion presence or lesion activity status for the fellow eye will not be used in the analysis. For instance, remote visits occurring during the COVID-19 pandemic
- management visits without an associated OCT will not be included in the analysis
- home-monitoring tests that cannot be associated with a management visit will not be used in the analysis. For instance, home-monitoring tests towards the end of the study where the next management visit does not occur until after the end of the follow-up.

Note that the exclusion criteria may result in patients being assessed in one home-monitoring test model and not another if they only ever self-tested with one or two of the tests.

Protocol deviations

Safety population

This study does not require participants to undergo any additional investigations. Therefore, it is not possible for clinical adverse events to be attributed to study specific procedures.

There are no safety reporting procedures to be followed for this study.

Withdrawals

Participants can withdraw from the study at any time due to personal reasons, discharge from usual care or through an investigator decision. All data previously collected for a withdrawn participant will be used in the analysis, unless the participant requests that their data not be used.

Statistical analyses and report content

General considerations

Statistical analysis is the responsibility of the BTC study statistician, and all analyses will be carried out using the most recent version of Statistical Analysis Software (SAS) and/or Stata at the time of analysis.

All applicable statistical tests will be two-sided and will be performed using a 5% significance level, with the exception of tests for interactions that will be performed using a 10% significance level. CIs will be reported at the 95% level unless otherwise stated.

No formal adjustment will be made for multiple testing, but consideration will be taken in interpretation of results to reflect the number of statistical tests performed and the consistency, magnitude and direction of estimates for different outcomes.

General calculations

Unless otherwise stated, all percentages will be calculated using the total number of participants with non-missing data from the relevant population as the denominator. For categorical and binary data, all percentages will be rounded to one decimal place. For continuous measures, these will be summarised to one more decimal place than that of the data collected.

Outcomes

Primary outcome (Objective A)

The primary outcome is the classification of a study eye at a management visit as having an active or inactive disease. For the reference classification, this is the reviewing ophthalmologist's decision at a management visit about the activity status of the study eye. For index tests, this is the classification derived

from index test data captured in the interval preceding the management visit. Classifications will be derived from logistic regression model predicted probabilities where the reference classification is the outcome with the index test and, where appropriate, other measurements (patient age, patient sex, etc.) as predictors.

An additional sensitivity analysis will be performed where the reference classification will be changed to CARF-graded activity status.

Secondary outcomes

1. Lesion transition from inactive to active classification from one management visit to the next: Objective A.
2. Willing in principle to participate in study: Objective C.
3. Participation in the study: Objective C.
4. Ability of participants to weekly testing: Objective C:
 - Note: Ability is defined as the proportion of management visits for which some valid index test data are available. Only qualifying management visits will be considered.
5. Adherence of participants to perform home-monitoring tests: Objective C:
 - Note: Adherence is defined as the proportion of weeks for which some valid data for an index test are available. Only weeks between qualifying management visits will be considered.
6. Diagnosis of nAMD in a fellow eye (conversion): Objective D.

Derivation of the outcomes

Primary outcome

For all management visits

- Study eyes will be recorded as active at a management decision if *Was a nAMD lesion present?* = **Yes** and *What was the status of the nAMD lesion?* = **Active**.
- Study eyes will be recorded as inactive at a management decision if *Was a nAMD lesion present?* = **Yes** and *What was the status of the nAMD lesion?* = **Inactive**.
- Study eyes will be recorded as inactive at a management decision if *Was a nAMD lesion present?* = **No**.
- Study eyes will be recorded as uncertain at a management decision if *Was a nAMD lesion present?* = **Yes** and *What was the status of the nAMD lesion?* = **Uncertain**.

Secondary outcomes

- **Transition from inactive to active classification from one management visit to a subsequent one:**
 - *Was a nAMD lesion present?* = **Yes** AND *What was the status of the nAMD lesion?* = **Active** at management visit t_n AND
 - *Was a nAMD lesion present?* = **Yes** AND *What was the status of the nAMD lesion?* = **Inactive** at management visit $t_n - (x + 1)$, WHERE x is number of concurrent prior management visits without a management decision (see [Objective A](#)).
- **Willing to participate in the study:**
 - *Patient eligible* = **Yes** AND
 - *Willing in principle to participate in study* = **Yes** AND
 - (*Right eye classification at screening* = **Study eye** OR *Right eye classification at screening* = **Study eye**).

- **Participation in the study:**
 - Patient eligible = **Yes** AND
 - Willing in principle to participate in study = **Yes** AND
 - (Right eye classification confirmation at training = **Study eye** OR Right eye classification confirmation at training = **Study eye**) AND
 - Is patient able to do the mVT test = **Yes** AND
 - Is patient able to do the MBT test = **Yes** AND
 - Is patient able to complete the KSJ = **Yes** AND
 - Are the patient's home or personal circumstances suitable for home monitoring = **Yes** AND
 - Did the patient give consent to participate in home-monitoring testing = **YES**.
- **Ability of participants to perform home-monitoring tests:**
 - All tests: proportion of 'complete' management visits that have at least one associated valid home-monitoring test score in the prior interval.
 - Number of complete management visits with at least one valid associated home-monitoring test/number of complete management visits.
- **Adherence of participants to weekly testing:**
 - All tests: proportion of weeks for which at least one valid test score is available across all weeks that take place between 'complete' management visits.
 - Number of weeks in-between complete management visits with at least one valid home-monitoring test/number of weeks in-between complete management visits.
- **Diagnosis of nAMD in a fellow eye (conversion): Objective D:**
 - Was a nAMD lesion present? = **Yes** AND What was the status of the nAMD lesion? = **Active** at management visit t_n AND
 - Was a nAMD lesion present? = **No** at management visit $t_n - (x+1)$, WHERE x is number of concurrent prior management visits without a management decision (see [Objective A](#)).

Model predictor variables (Objectives A and D)

Electronic tests (MBT and mVT)

Home-monitoring tests were collected on multiple occasions between management decision visits and scoring outcomes differ between the electronic and KSJ tests. Both electronic tests return continuous single score outcomes.

For both electronic tests, a simple pragmatic approach of using mean test scores between visits with included management decisions (those with an OCT image) will be used as initial summary measurements for the models. However, this approach has drawbacks:

- Periods between included management visits are not fixed intervals.
- Home-monitoring test scores closer to the management visit are likely more reflective of eye health at their associated visit.
- Longer intervals may distort signals related to changing eye health that occur close to the management visits.

Due to the above, mean test scores will also be calculated using a fixed interval covering the previous 4 weeks prior to the management visit.

In the case of the mVT test, the outcome is a mean of ≥ 2 separate test runs with an associated within-test SD. Due to this, the average of the SDs will also be calculated by taking the square-root of the mean of the underlying variances for both approaches.

The time interval of the testing period in weeks and the proportion of these weeks with a valid test will also be calculated as additional summary measurements. Other summary measurements may be explored, for instance: worst index test score, most recent test score to management decision or median index test score.

KeepSight Journal

The KSJ returns three ordinal predictors for visual acuity, environmental Amsler and visual Amsler testing, each of which have a range of worse < same < better. These ordinal outcomes will be converted to -1 for worse, 0 for same and 1 for better. Medians will be used as summary measurements of 'average' performance between management visits. The time interval of the testing period in weeks and the proportion of these weeks with a valid test will also be calculated as additional summary measurements.

As for the electronic tests, median scores will also be calculating using only the immediately preceding 4 weeks prior to the management visit.

Other predictors

The following predictors are not of direct interest to the objectives, but rather are possible confounders and may be included in the models as additional independent variables:

- **Patient sex:** as recorded on the database.
- **Patient age:** calculated as age at consent.
- **Visual acuity at diagnosis:** visual acuity at diagnosis will be calculated into three categories, based on the worst eye in the study for the patient:
 - Snellen $\geq 6/18$
 - Snellen $< 6/18$ and $> 6/24$
 - Snellen $\leq 6/24$.
- **Time since study baseline:** calculated as number of days between baseline (continuous predictor) and management visit *and* as annual quarters since baseline (categorical predictor).
- **Time since previous management visit:** calculated as number of weeks between management visit and most recent management visit with an included management decision (those with an OCT image).
- **Proportion of weeks with tests:** calculated as number of weeks with a valid home-monitoring test over the total number of weeks between a management visit and most recent management visit with an included management decision (those with an OCT image).
- **Between-test SD (MBT and mVT) and interquartile range (IQR) (KSJ):** SD of complete electronic home-monitoring tests between a management visit and most recent management visit with an included management decision (those with an OCT image).
- **Average within-test SD (mVT):** calculated as the square-root of the mean of the underlying variances for each applicable test score.

Analysis of the outcomes

Objective A

Logistic regression models and AUROC comparisons will be used to assess the diagnostic test performance of the home-monitoring tests. Only data from study eyes will be analysed. All models will consider nesting of management visits within participants and, where two study eyes are present within participants, eyes. Management visits will be excluded based on the criteria given in section *Objective A*.

All models will use a binary outcome dependent variable for the primary outcome or secondary transition outcome, as assessed at a single management visit. Alongside the home-monitoring point estimate, other independent variables used will include time since first treatment for nAMD at consent strata, patient sex and patient age at consent. Random intercepts will initially be fitted for patient and study eye, with a random gradient fitted for time since study baseline within study eye. Time since study baseline will be assessed as both a continuous and a categorical variable to assess linearity within the variable. When assessed as categorical, intervals of annual quarters will be used.

Models will be built using a process of forward selection, with additional independent variable additions assessed through likelihood ratio tests. Initially, separate models will be built for each home-monitoring test. For the KSJ, median point estimates for all three ordinal predictors will be assessed within a single model. No testing of interactions is planned.

In the event of logistic regression models failing to converge, the models will be simplified by selecting a single eye in participants with two study eyes, as assessed at consent, thereby removing a level of within participant nesting. In this case, eyes will be selected at random within participants with two study eyes.

Initially, models will not include additional parameters for measurement error, but depending on initial model performance parameters may be included, such as time interval covered by testing in weeks where appropriate, proportion of weeks with a valid home-monitoring test, between-test SD/IQR, and, in the case of mVT, within-test SD. Appropriateness of including these extra terms will be assessed using likelihood ratio tests.

Weighting will not be used in the models to control for home-monitoring interval testing length due to the assumption that a longer testing interval does not necessarily imply a more precise point estimate of eye health at the time of the management decision. Instead, the impact of variable interval length will be assessed by comparing the models with complete intervals against those with intervals capped at 4 weeks prior to the management decision. Management decisions will be excluded from the capped interval model if there are no associated home-monitoring tests in the capped period.

Initially, cut-off values will be estimated for curve thresholds that maximise sensitivity and specificity as calculated by the Youden's index. However, as Youden's index assigns equal weighting to false positives and false negatives, additional cut-off values will also be explored. For instance, false negatives may be classed as more costly than false positives, and so thresholds that maximise sensitivity will be explored.

True positive rates (sensitivity), true negative rates (specificity), positive predictive values and negative predictive values will be provided for all selected cut-off values. Additionally, as cut-offs will relate to post-estimation predicted probabilities, means and SDs for home-monitoring test scores above and below the associated cut-offs will also be described.

Objective C

Linear regression models will be fitted to explore the dependent variables of ability to carry out the home-monitoring tests and adherence to the weekly testing schedule. Logistic regression models will be fitted to explore the dependent variables of inequalities in willingness to participate and inequalities in participation.

The influence of the following independent variables will be examined in the inequalities model: age, sex, centre, socioeconomic status, site and visual acuity at diagnosis. Site will be fitted as a random variable.

The same independent variables will be examined in both the ability to carry out home-monitoring tests and in the adherence models with the inclusion of the nine 'use of technology indicators'. Due to the

high number of additional binary indicators this introduces, use of technology will be explored using a process of forward selection after the other independent variables are fitted.

The IMD, as determined by participant postcode, will be used as an indicator of socioeconomic status. However, IMD is not comparable across country due to differences in calculation and IMD being a rank (ordinal) measure. Due to this, Northern Ireland IMD scores will be estimated onto the English scale using 2010 Northern Ireland IMD domain residuals.⁴⁹ All IMD scores and ranks will then be converted to quintiles, ranging from least to most deprived.

The influence of these factors will be reported as mean difference per unit change in a predictor or ORs with 95% CIs. As for Objective A, analyses will take account of the structure within the data, that is, the nesting of visits within participants where necessary.

Objective D

Analyses for Objective D will follow that outlined for Objective A, except the analyses will only include participants with a fellow eye in the study. Due to this, models will not include nesting of eye within participant, as any participant can only have one fellow eye within the study.

As for Objective A, threshold analysis and model comparison between tests will only be performed on models where coefficients for home-monitoring test predictors are statistically significant at the 5% level in predicting management visit outcomes or AUROC values are negatively affected by removing the home-monitoring test predictors from the model.

General content

Participant flow

The Consolidated Standards of Reporting Trials (CONSORT) flow diagram will be used to summarise the flow of participants through screening until follow-up throughout the course of the study and will be presented by centre. This will include the number of:

- patients approached
- eligible patients at screening
- patients willing in principle to participate
- patients attended further information and training
- eligible patients at training
- consented patients
- participants who withdraw at any time.

Confirmed serious breaches of good clinical practice will be summarised, including violations of eligibility criteria on entry into the study. The number of withdrawals of consent to the study will be presented, along with reasons for withdrawal. Any notes to file reported for each participant will be presented as a line listing.

Baseline data

Participant baseline characteristics will be tabulated using frequency and summary statistics recruiting centre and overall. Continuous variables will be summarised using the mean and SD (or median and IQR, depending on the distribution), and categorical data will be summarised as a number and percentage.

Graphical methods may also be used to summarise data where useful. Missing or unobtainable data will be detailed in the summaries. Statistical testing will be carried out on these data to address Objective C (see [KeepSight Journal](#)).

Missing data and outliers

A thorough data cleaning process will be carried out and attempts will be made to obtain any missing data by chasing until it is either received, confirmed as not available, or the study is at the analysis stage. Where data are unobtainable, all summaries will indicate how many missing results there are by time since first treatment for nAMD at consent strata.

Summaries will also be given for data excluded from Objectives A and D where management decisions cannot be associated with home-monitoring tests and vice versa (see [Objective A](#) and [Objective D](#)).

Where data points are identified as possible outliers both statistically and clinically, and are considerable in number, sensitivity analyses may be considered for all formal outcomes.

Statistical analysis plan revision history**TABLE 59** Statistical analysis plan revision history

Version number	Revision date	Justification for revision
0.1		Initial draft
0.4		Exclusion criteria added to study population in Objective A and Objective D . More detail on models and AUROC calculations in Electronic tests (MBT and mVT) .
0.5	19 February 2021	Added lesion transition as primary outcome in Primary outcome . Changed primary analysis outcome from lesion activity to lesion transition in Electronic tests (MBT and mVT) . Kept possibility of performing analysis using lesion activity as option. Change methodology for ROC calculation from that given in Liu and Wu (2003) ⁵² to inbuilt Stata function in Electronic tests (MBT and mVT) . Added additional analysis to examine length of intermanagement visit interval without reference to home-monitoring tests in Electronic tests (MBT and mVT) . Option to include this parameter as measurement error variable in home-monitoring test analysis still open in section.

EME
HSDR
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 and Care 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