NIHR | National Institute for Health and Care Research

Journals Library

Check for updates

Health Technology Assessment

Volume 29 • Issue 69 • December 2025

ISSN 2046-4924

Tele-ophthalmology-enabled and artificial intelligence-ready referral pathway for community optometry referrals of retinal disease: HERMES cluster randomised trial with a diagnostic accuracy study

Anitta Sharma, Rima Hussain, Annastazia E Learoyd, Angela Aristidou, Taha Soomro, Ann Blandford, John G Lawrenson, Gabriela Grimaldi, Abdel Douiri, Ashleigh Kernohan, Tomos Robinson, Najmeh Moradi, Christiana Dinah, Evangelos Minos, Dawn Sim, Tariq Aslam, Avinash Manna, Alastair K Denniston, Praveen J Patel, Pearse A Keane, Catey Bunce, Luke Vale and Konstantinos Balaskas

DOI 10.3310/QNDF3325

NIHR | National Institute for Health and Care Research

Health Technology Assessment

Check for updates

Extended Research Article

Tele-ophthalmology-enabled and artificial intelligence-ready referral pathway for community optometry referrals of retinal disease: HERMES cluster randomised trial with a diagnostic accuracy study

Anitta Sharma[1,2,3,4] Rima Hussain[1,3] Annastazia E Learoyd[5] Angela Aristidou[6,7] Taha Soomro[1] Ann Blandford[8] John G Lawrenson[9] Gabriela Grimaldi[1,3,10] Abdel Douiri[5] Ashleigh Kernohan[11] Tomos Robinson[11] Najmeh Moradi[11] Christiana Dinah[12] Evangelos Minos[13] Dawn Sim[3] Tariq Aslam[14] Avinash Manna[15] Alastair K Denniston[15] Praveen J Patel[1,3,10] Pearse A Keane[3,4,10] Catey Bunce[16] Luke Vale[11,17] and Konstantinos Balaskas[1,3,4,10*]

[1]Moorfields Ophthalmic Reading Centre and Artificial Intelligence Lab, Moorfields Eye Hospital NHS Foundation Trust, London, UK
[2]Optometry, Moorfields Eye Hospital NHS Foundation Trust, London, UK
[3]NIHR Biomedical Research Centre, Moorfields Eye Hospital NHS Foundation Trust, London, UK
[4]Institute of Ophthalmology, University College London, London, UK
[5]School of Population Health and Environmental Sciences, King's College London, London, UK
[6]School of Management, University College London, London, UK
[7]Digital Economy Lab, Stanford University, Stanford, CA, USA
[8]UCLIC, University College London, London, UK
[9]School of Health and Medical Sciences, City St George's, University of London, London, UK
[10]Medical Retina Service, Moorfields Eye Hospital NHS Foundation Trust, London, UK
[11]Health Economics Group, Population Health Sciences Institute, Newcastle University, Newcastle upon Tyne, UK
[12]Ophthalmology, London North West University Healthcare NHS Trust, London, UK
[13]North West Anglia NHS Foundation Trust, Peterborough, UK
[14]School of Biomedical Sciences, University of Manchester, Manchester, UK
[15]NIHR Birmingham Biomedical Research Centre, London, UK
[16]Faculty of Infectious and Tropical Diseases, London School of Hygiene & Tropical Medicine, London, UK
[17]Global Health Economics Centre, London School of Hygiene and Tropical Medicine, London, UK

*Corresponding author konstantinos.balaskas@gmail.com

Published December 2025
DOI: 10.3310/QNDF3325

This article should be referenced as follows:

Sharma A, Hussain R, Learoyd AE, Aristidou A, Soomro T, Blandford A, *et al*. Tele-ophthalmology-enabled and artificial intelligence-ready referral pathway for community optometry referrals of retinal disease: HERMES cluster randomised trial with a diagnostic accuracy study. *Health Technol Assess* 2025;**29**(69). https://doi.org/10.3310/QNDF3325

Health Technology Assessment

ISSN 2046-4924 (Online)

Impact factor: 4

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 4 and is ranked 30th (out of 174 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 NIHR127773. The contractual start date was in March 2020. The draft manuscript began editorial review in September 2023 and was accepted for publication in March 2025. 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.

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

Abstract

Background: Community-based optometrists, a major provider of primary eye care in the United Kingdom, are the main source of referrals to hospital eye services. The widespread introduction of optical coherence tomography devices in community practices provides community-based optometrists with an opportunity to identify a broader range of treatable diseases. Standard referral pathways do not effectively filter unnecessary referrals, with misclassification of urgency, and erroneous diagnoses.

Objectives: To assess the effectiveness of a teleophthalmology referral pathway between community-based optometrists and hospital eye services for retinal diseases. To measure the accuracy of an artificial intelligence decision support system for diagnosis and referral management of retinal disease.

Design: A multicentre, superiority cluster randomised controlled trial to assess the effectiveness of a teleophthalmology referral pathway. A prospective, observational diagnostic accuracy study to measure the performance of artificial intelligence decision support system. A comprehensive economic evaluation was conducted.

Settings: United Kingdom-based community optometry practices with an optical coherence tomography device and hospital eye services.

Participants: Adults requiring referral for retinal disease at the opinion of the community-based optometrists.

Interventions: Community optometry practices were randomised 1 : 1 to standard care or teleophthalmology. Referrals sent via the teleophthalmology platform were remotely reviewed by human experts based at the corresponding hospital eye services. A referral decision was provided within 48 hours. Suitable optical coherence tomography scans were solely processed by artificial intelligence decision support system (the 'Octane' model).

Main outcome measures: Cluster randomised controlled trial's primary outcome was the proportion of false-positive referrals (not required or not urgent) per arm in overall participants and in referred-only participants against an independent reference standard. Secondary outcomes included the proportion of wrong diagnosis, wrong referral urgency, false-negative referrals, safely triaged referrals for rare diseases, time from referral to consultation and treatment and cost-effectiveness of teleophthalmology. Primary outcome for the artificial intelligence study was the sensitivity and specificity of artificial intelligence referral decisions against the reference standard.

Results: Teleophthalmology significantly reduces the proportion of false-positive urgent referrals by 59% compared to standard care in referred participants. Due to the observed low event rate for false positive referrals, teleophthalmology's role for reducing false positives overall was inconclusive. No significant difference between arms for safety of referral decisions (false negatives) was found. After accounting for external factors, the time to consultation demonstrated both clinically and statistically significant benefits for the teleophthalmology arm. The time to treatment showed a clinically significant benefit.

Of 396 recruited participants, the Octane artificial intelligence model processed images contributed by 204 participants (51.5%). For referral decisions, the model showed comparable sensitivity and specificity against its own preset referral rules (rule-based reference standard) (post hoc analysis), but it showed inferior sensitivity and specificity when compared to human expert assessors making these referral decisions (clinical reference standard) (primary AI analysis). The artificial intelligence model presented challenges relating to its generalisability in a real-world evaluation context.

Limitations: Technical limitations in optometry practices, lack of ethnicity data.

Conclusions: Asynchronous teleophthalmology reduces the number of unnecessary urgent referrals, the main drivers of increasing hospital capacity pressures, provides more appropriate referral-to-treatment times and is more cost-effective compared to standard care. The Octane artificial intelligence model could not process images from 48.5% of study participants. Compared to hospital-based experts for referral decisions, Octane was less accurate at making routine and urgent referral decisions and of similar accuracy to community optometrists.

Future work: Applied health research, human–artificial intelligence interaction and artificial intelligence clinical trial design.

Trial registration: This trial is registered as ISRCTN18106677.

Funding: This award was funded by the National Institute for Health and Care Research (NIHR) Health Technology Assessment programme (NIHR award ref: NIHR127773) and is published in full in *Health Technology Assessment*; Vol. 29, No. 69. See the NIHR Funding and Awards website for further award information.

Contents

List of tables

List of figures

List of abbreviations

AI	artificial intelligence
AIC	Akaike information criterion
AI DSS	artificial intelligence decision support systems
AMD	age-related macular degeneration
AUC	area under the curve
BIC	Bayesian information criterion
BRC	Biomedical Research Centre
BRVO	branch retinal vein occlusion
CBA	cost–benefit analysis
CCA	cost–consequence analysis
CCG	Clinical Commissioning Group
CDM	critical decision method
CEA	cost-effectiveness analysis
CFP	colour fundus photograph
CI	confidence interval
CNV	choroidal neovascularisation
CO	community-based optometrist
COVID-19	coronavirus disease 2019
cRCT	cluster randomised controlled trial
CRVO	central retinal vein occlusion
CSCR	central serous chorioretinopathy
DCE	discrete choice experiment
DMC	Data Monitoring Committee
DMO	diabetic macular oedema
DNA	did not attend
eCRF	electronic case report form
EDI	equality, diversity and inclusion
EeRS	Electronic eyecare Referral Systems
ERM	epiretinal membrane
FTMH	full-thickness macular hole
GDPR	General Data Protection Regulation
GP	general practitioner
HCI	human -computer interaction
HES	hospital eye services
ICC	intracluster correlation coefficient
ICER	incremental cost-effectiveness ratio
IED	inherited retinal eye disease
IOP	intraocular pressure
MEH	Moorfields Eye Hospital NHS Foundation Trust
MO	macular oedema
MWTP	marginal willingness to pay
MXL	mixed logit model
nAMD	neovascular age-related macular degeneration
NICE	National Institute for Health and Care Excellence
NIHR	National Institute for Health and Care Research
NPT	normalisation process theory
NPV	negative predictive value
OCT	optical coherence tomography
OR	odds ratio
OWSA	one-way sensitivity analysis
PEARS	Primary Eye care Acute Referral Scheme
PI	principal investigator
PIS	patient information sheet
PM	project manager
PPI	patient and public involvement
PPIE	patient and public involvement and engagement
PPV	positive predictive value
PSA	probabilistic sensitivity analysis
PTMH	partial thickness macular hole
RCT	randomised controlled trial
ROC	receiver operating characteristic
RS	reference standard

SRC	suspected retinal conditions	VA	visual acuity
TMG	Trial Management Group	VMT	vitreomacular traction
TSC	Trial Steering Committee	VRA	vitreoretinal interface abnormalities
UCL	University College London	WTP	willingness to pay

Plain language summary

Community-based optometrists are the main source of referrals to hospital eye services in the United Kingdom. Many referrals are for problems with the retina, which is the layer at the back of the eye which allows us to see. Optical coherence tomography devices detect retinal conditions and are increasingly used by community-based optometrists; however, not all are sufficiently trained to use these machines. This leads to many inappropriate referrals and delayed access to treatment.

The HERMES study assessed the effectiveness of a teleophthalmology referral pathway between community optometrists and hospital eye services. Teleophthalmology is the review of medical information that has been electronically exchanged. Using this technology, referrals with eye scans from community optometrists were remotely reviewed by hospital-based eye specialists.

Two hundred and ninety-four participants were recruited by 26 optometry sites, of whom 158 participants were referred via the teleophthalmology referral platform and 136 participants were referred via the standard referral pathway. The teleophthalmology pathway effectively reduced the proportion of unnecessary urgent referrals by almost 60%, decreased the proportion of incorrect referral urgency by 25% and significantly reduced the proportion of incorrect diagnoses and the time to consultation. If implemented, it is likely to have lower costs and greater effectiveness.

The role of artificial intelligence to improve hospital referrals was also assessed. Artificial intelligence is a computer programme that is trained to do tasks which require human intelligence. We used an artificial intelligence model to look at eye scans and recommend if a hospital referral was required. We found that the model could not support many of the people who visit community optometry practices in England, and it was therefore used only on suitable scans from study participants. The model's referral recommendation was compared to optometrists and hospital experts, where it sometimes made different referral decisions than hospital experts but similar decisions to optometrists.

Scientific summary

Background

Ophthalmic services are challenged by an ageing demographic and the associated rise in common retinal diseases. Despite an increasing demand, there is an anticipated global shortfall of ophthalmologists. Community-based optometrists (COs) are essential in managing and referring people to hospital eye services (HES). Optical coherence tomography (OCT) devices have been increasingly installed in community practices. Although OCT expanded optometrists' scope of practice, its widespread use caused an increase in overall and unnecessary referrals. Efforts to streamline incoming referrals are vital to help alleviate pressures to HES and to improve patients' access to timely diagnosis and treatment.

This could be achieved with teleophthalmology, which is the use of medical information exchanged from one site to another via electronic communications to improve a patient's eye health. A systematic review revealed the evidence gap from well-designed, randomised controlled trials (RCTs) for validation of teleophthalmology referrals. The HERMES study assesses the clinical- and cost-effectiveness of teleophthalmology referrals compared to standard care, in reducing unnecessary hospital visits and misclassified 'urgent' referrals through a cluster RCT (cRCT).

Additionally, there have been significant developments in applying machine learning, specifically deep learning, in medical imaging for disease diagnosis. The HERMES study included a prospective observational study, reviewing artificial intelligence decision support systems (AI DSS), specifically the Moorfields-Google DeepMind model ('Octane') (Google DeepMind, DeepMind Technologies Limited, London), compared to human assessors in terms of its accuracy for retinal diagnosis and referral decisions. Its use as an AI DSS was tested for its generalisability in a diverse clinical care environment.

Objectives

To assess the effectiveness and cost-effectiveness of a digital referral pathway between community optometry and HES for referral of retinal disease enabled by a teleophthalmology platform. To measure the diagnostic and referral accuracy of the Octane AI DSS in the context of referral pathways between community optometry and HES.

Methods

Design

A multicentre, superiority cRCT to assess the clinical- and cost-effectiveness of a teleophthalmology referral pathway for retinal disease. A prospective, observational diagnostic accuracy (validation) study to measure the performance of Octane AI DSS for diagnostic and referral support. A human–computer interaction (HCI) assessment via a theoretically informed, qualitative study to explore participants' and healthcare professionals' perspectives on teleophthalmology and AI DSS. A small-scale exploratory post-implementation observational study of real-life teleophthalmology.

Settings

Community optometry practices with an OCT device and HES based in the UK.

Participants

Adults (≥ 18 years) attending an eye examination with a macular OCT scan at the participating optometry practice were recruited if there was a suspicion of retinal disease in the opinion of CO. Conditions included neovascular age-related macular degeneration (nAMD), dry age-related macular degeneration, diabetic retinopathy, macular oedema, central serous chorioretinopathy, vitreoretinal interface abnormalities, genetic eye disease and any other retinal condition not requiring an emergency referral. Participants were required to give consent. Individuals with known

retinal comorbidities in either eye triggering a referral and cases where acquisition of a good-quality OCT scan was not possible were excluded.

Interventions

Twenty-six community optometry practices (clusters) in the catchment areas of four HES sites in the UK were randomised 1 : 1 to standard care or teleophthalmology for the cRCT. Practices randomised to the control arm continued to refer participants with suspicion of retinal disease to HES using their standard method. Upon receiving informed written consent, COs uploaded their clinical findings, imaging (OCT scans), diagnosis and referral decision to the teleophthalmology referral platform.

Practices randomised to the intervention arm referred participants with suspicion of retinal disease to HES using the teleophthalmology referral platform. Human experts based at the corresponding HES reviewed every case remotely. A referral decision was provided to participants and COs within 48 hours. A hospital appointment was arranged if required. The efficiency of teleophthalmology was assessed by comparing CO and teleophthalmology recommendations against a reference standard (RS).

For the AI study, a subset of OCT scans was processed by the Octane AI DSS. The Octane AI only used the OCT scans as input without any other clinical information. Its diagnosis and referral recommendations were compared against an independent RS.

Additionally, a model-based economic evaluation was conducted, including a model-based cost-effectiveness analysis (CEA), a discrete choice experiment (DCE), and by using the results of both CEA and DCE for a cost–benefit analysis (CBA).

A HCI evaluation was conducted; this involved a combination of situated observations and semistructured interviews with healthcare professionals and participants to investigate their perspectives on teleophthalmology models of care and AI DSS.

Outcome measures

The cRCT primary outcome was the proportion of false-positive referrals (unnecessary HES visits or incorrect referral urgency) in each arm in the overall enrolled and the referred participants against the RS. Secondary outcomes included the proportion of wrong diagnosis, wrong referral urgency, false-negative referrals, uncommon referrals (rare disease) which could be safely triaged, the time from referral to consultation and to treatment and an economic evaluation to assess the costs and benefits from the NHS perspective of the teleophthalmology referral triaging pathway.

The primary outcome for the observational study was the diagnostic accuracy of the referral decision made by AI DSS when compared with the RS. Secondary measures included the diagnostic accuracy of retinal disease and referral urgency, the proportion of false-positive referrals, wrong diagnosis, the wrong referral urgency in each arm if human assessors were replaced by AI DSS, the time required by the AI DSS to process an OCT scan and its average time of end-to-end referral recommendation.

A HCI analysis using qualitative methods identified factors that facilitate the successful implementation of a digital referral platform to ensure acceptability and acceptance.

Statistical analysis

A hierarchical two-level mixed-effects model was used to calculate the required sample size. Twenty-six clusters split between arms in a 1 : 1 ratio needed to recruit 306 participants, with an average of 10 participants per cluster, to achieve 89.27% power to detect a difference in the proportion of false-positive referrals of 30%. To calculate estimates of diagnostic accuracy for the AI study, a higher total sample size of 370 participants (accounting for the drop-out rate), with 351 participants being correctly diagnosed, would produce estimates of diagnostic accuracy with a two-sided 95% confidence interval (CI).

For the cRCT primary outcome, the difference in the proportion of false-positive referrals between study arms was calculated with 95% CI, calculated using the exact binomial method adjusting for cluster. A superiority margin of 30% was used. A regression model estimating the difference in odds of false-positive referral between study arms as an odds ratio (control arm/intervention arm) was also measured. As differing referral urgency levels are of interest, false-positive urgent referrals (clinician urgently referring the patient when RS deemed an urgent referral was unnecessary) were examined using the same methods.

Diagnostic accuracy for the referral decision is reported using the sensitivity, specificity, positive predictive value (PPV) and negative predictive value (NPV), where exact binomial CIs were provided. The mean time from referral to consultation and to treatment was generated with 95% CIs calculated using the exact method for a Poisson distribution. Cox model proportional hazards regression adjusting for clustering was used to compare the time from referral to consultation between arms. A sensitivity analysis was performed following the exclusion of cases delayed by external factors from the intervention arm. The remaining secondary outcomes were examined using the same methods for the primary outcome. An additional exploratory analysis focusing on the detection and referral of nAMD was conducted.

For the AI study primary outcome, sensitivity, specificity, PPV and NPV are reported, where exact binomial CIs are provided with the exact binomial method. Summary statistics for the diagnostic accuracy of the referral decision dichotomised as urgent referral versus non-urgent and of referral urgency (urgent vs. routine referral) were computed as the primary outcome methods. Proportions for all human-assessor-related outcomes were analysed with 95% CIs calculated using the exact binomial method. The technical infrastructure details were reported in terms of time per OCT scan.

Results

One hundred and thirty-six participants were recruited to the control arm, and 158 participants were recruited to the intervention arm. Asynchronous teleophthalmology with clinician triage significantly reduced the proportion of false-positive urgent referrals by 59.5% as compared to standard care in referred participants. It lowered the proportion of false-positive referrals (for urgent/routine referrals) and of false-negative referrals. In terms of wrong referral urgency (reported as means with CIs), the standard pathway had 25.7% (14.1% to 37.3%) more incorrect referral decisions than the teleophthalmology pathway, a significant difference at the 5% level.

Both arms showed high sensitivity for detecting the need for a referral, however, specificity was significantly lower in the control arm [41.2% (18.4% to 67.1%)] versus the intervention arm [93.9% (79.8% to 99.3%)]. For urgent referrals only, there was greater accuracy in the intervention arm [sensitivity: 96.3% (81.0% to 99.9%), specificity: 99.2% (95.8% to 100%)] versus the control arm [sensitivity: 73.7% (48.8% to 90.9%), specificity: 79.5% (71.0% to 86.4%)]. There was also a significant increase in the proportion of wrong diagnoses in the standard pathway versus the teleophthalmology pathway [right eye: 19.7% (6.1% to 33.3%); left eye: 18.8% (4.3% to 33.2%)].

Focusing on nAMD, the control arm had 54.2% (30.3% to 78.0%) more false-positive referrals than the intervention arm. Additionally, the intervention arm had greater sensitivity [90.9% (70.8% to 98.9%)] and specificity [100% (97.3% to 100%)] when making a nAMD-related referral than the control arm sensitivity [57.9% (33.5% to 79.7%)] and specificity [88.9% (81.7% to 93.9%)]. The standard pathway also led to more false-positive nAMD diagnoses [53.6% (33.8% to 72.5%)] versus the teleophthalmology pathway [0, (0%)].

Following a sensitivity analysis, the mean time to consultation was significantly lower in the intervention arm [53 (51 to 55) days] versus the control arm [89 (87 to 91) days] (p = 0.039). The mean time to treatment was lower in the intervention arm [55 (52 to 57) days] versus the control arm [90 (87 to 94)] (p = 0.151). Assessment of safe triage of rare diseases showed each arm referred all suspected cases to HES.

For the AI study, 204 out of 396 participants from 17 of the 29 participating CO practices were included, where the majority of participants were from 14 sites, 201 of the 204 participants (98.5%). Two OCT device manufacturers found

in CO practices were supported, of which 30% of images were not suitable for processing by Octane due to image size or format.

When compared to human assessors (clinical RS), the sensitivity of the Octane AI model was 96.4% (92.4% to 98.7%) and the specificity was 20.0% (8.4% to 36.9%). The AI model was less accurate when deciding referral urgency for required referrals (routine vs. urgent), with wrong referral urgency by AI of 14.8% (9.0% to 22.3%) and by human assessors 2.8% (0.6% to 7.8%). Of note, human assessors considered information from OCT scans, clinical history and patient preferences when making referral decisions, whereas the AI model considered information solely from OCT scans. When recommending urgent referrals, the Octane model is less accurate than hospital-based experts in the teleophthalmology arm, with wrong referral urgency by AI of 25% (5.5% to 57.2%) and by hospital-based experts of 0%, yet the overall absolute numbers of wrong urgent referrals are small; the model is of a similar accuracy as community optometrists for urgent referrals, with wrong referral urgency by AI of 66.7% (34.9% to 90.1%) and by CO of 62.5% (35.4% to 84.8%).

In a post-hoc analysis, the Octane model showed good referral accuracy when applying its own preset referral rules (rule-based RS) for recommending a referral versus no referral. If the AI model offers a referral to a patient, there is a 100% (98.1% to 100%) probability that a referral is truly needed (PPV). If a referral is not offered by the AI, however, there is a 69.2% (38.6% to 90.9%) probability that a referral is truly not needed (NPV). Absolute numbers of no referrals being small in both pathways, however, PPV is increased and NPV is decreased due to high prevalence of participants requiring a referral.

In a post-hoc analysis, the Octane model shows good referral accuracy against the rule-based RS for recommending urgent referral, with modest reduction in referral efficiency (PPV); that is, if the Octane model offers an urgent referral to a patient, there is a 78.4% (61.8% to 90.2%) probability that an urgent referral is truly needed. For referral urgency, the Octane model chose the wrong referral urgency a comparable number of times to CO [AI: 22.7% (11.5% to 37.8%) vs. CO: 22.7% (11.5% to 37.8%)] and a greater number compared to hospital-based experts [AI: 9.4% (3.5% to 19.3%) vs. HES: 1.6% (0.0% to 8.4%)]. In terms of diagnostic accuracy for the diagnosis of retinal disease, the Octane model exhibits moderate sensitivity [68.1% (62.5% to 73.4%)] when detecting the same diagnosis as the clinical RS, which improves to 82.6% (72.9% to 89.9%) when compared to the rule-based RS.

The post-implementation substudy recruited 17 patients overall. Barriers to implementation of the Manchester Electronic eye Referral System included training gaps, inadequate communication channels among primary care, secondary care and technology suppliers and an insufficient support network.

The CEA of the economic decision model showed the HERMES pathway to have a greater effect at a lower cost than the standard pathway, meaning it was the dominant intervention. The DCE results showed the public had a greater preference for a more effective intervention that would be delivered more quickly. Importantly, they had a greater preference for obtaining a correct diagnosis compared to a reduction in the waiting time. A CBA, based on the DCE results, demonstrated a net benefit for the HERMES pathway compared with the standard pathway of £992 for every patient seen. Deterministic and probabilistic sensitivity analyses were carried out to assess the robustness of the conclusions, and in these analyses the conclusions were found not to change. These results provide strong evidence for the efficiency of the HERMES pathway compared with standard care. These data are significant as previous studies reviewing different referral interventions, including asynchronous teleophthalmology, provide limited information on this key area.

The HCI-related in-depth interview study with participants, optometrists and ophthalmologists showed that they generally attribute value in implementing teleophthalmology through improving efficiency and the ability to provide and receive feedback. The normalisation process theory analysis highlighted the need to consider multiple factors when developing and implementing teleophthalmology platforms, especially if aiming to have it adopted and normalised at a large scale.

Conclusions

Asynchronous teleophthalmology can reduce the number of unnecessary referrals, provide more appropriate referral timescales and is more cost-effective compared to standard care. Its role in reducing unnecessary referrals overall was inconclusive. When comparing the Octane AI system for referral decisions to hospital-based human expert assessors, it was less accurate for making routine and urgent referral decisions and was of similar accuracy to CO for making urgent referral decisions. The AI model presented challenges relating to its generalisability in a real-world evaluation context.

Trial registration

This trial is registered as ISRCTN18106677.

Funding

This award was funded by the National Institute for Health and Care Research (NIHR) Health Technology Assessment programme (NIHR award ref: NIHR127773) and is published in full in *Health Technology Assessment*; Vol. 29, No. 69. See the NIHR Funding and Awards website for further award information.

Chapter 1 Introduction

Ophthalmic services face the challenges of an ageing population and a rise in prevalence of common retinal diseases.[1,2] The number of new cases of late age-related macular degeneration (AMD) is expected to increase worldwide from 5.24 million in 2020 to 6.41 million by 2050, with the highest concentration in Asia (4.17 million), followed by Europe (0.65 million), North America (0.38 million) and Oceania (0.04 million).[3] In 2022–3, ophthalmology outpatient attendances in the UK reached over 8 million, the most compared to any other specialty within the NHS.[4] In January 2023, > 632,000 participants were on the ophthalmology waiting lists in England alone.[5] Despite an increasing demand, the workforce growth rate remains lower than the ageing population growth rate.[6,7] Data from the International Council of Ophthalmology showed that, although there are > 200,000 ophthalmologists worldwide, there remains an anticipated global shortfall of ophthalmologists.[7] A census report from 2022 found that over three-quarters of NHS eye units did not have enough ophthalmologists to meet the demand.[5] Workforce shortages continue to impact ophthalmology services across the UK at a greater level following the COVID-19 pandemic.[5] The rising pressures faced by the NHS hinder the ability to efficiently reduce the high levels of patient backlog. This is compounded by the inefficiency of referral processes into hospital ophthalmic services. Consequently, participants' access to diagnostic services and timely treatment are delayed, potentially resulting in a poorer prognosis.[8–10]

Community-based optometrists (COs), high street opticians, are a major provider of primary eye care in the UK, similar to European countries such as Germany[11] and North America,[12] and are the main source of referrals to HES.[13,14] A large proportion of participants diagnosed with a suspicion of retinal disease, including common conditions such as neovascular age-related macular degeneration (nAMD), are referred to HES for diagnostics and disease management.[15,16] The current process results in unnecessary referrals, erroneous diagnoses, misclassification in terms of urgency, duplication of imaging tests and delays in access to treatment.[17] A review of electronic referrals from community optometry practices to HES found just over a third were unnecessary.[18] Many factors can impact the accuracy of a referral, such as a clinician's years of experience and their ability to correctly identify clinical signs.[19,20] Models of community-based referrals for common eye (retinal) disease have been implemented in other healthcare settings (Scandinavian countries, mainly Denmark,[21,22] the USA[12]) and various community-based providers of optical coherence tomography (OCT) scanning, including pharmacies and dedicated facilities within large commercial centres.[23,24]

To further support CO and to improve patient care, a growing number of OCT devices continue to be installed in community practices.[25,26] OCT is a non-invasive imaging technique and can provide high-resolution, cross-sectional images of the posterior segment.[27] Although it provides CO with the potential to identify a broader range of treatable disease, its widespread use has led to a greater detection of asymptomatic retinal conditions and subsequent referrals.[28] While OCT is a useful diagnostic tool, accurate interpretation and management by CO are not guaranteed. The benefits of advanced imaging on patient management are limited without a good understanding of a clinical presentation and its underlying disease biology.[29] When compared to the use of colour fundus photographs (CFPs) alone, the utilisation of advanced imaging modalities (including OCT) improved the diagnostic accuracy of AMD by only 5% and led to an increase in false-positive referrals.[30] As capacity and workforce issues continue to grow, efforts to streamline incoming referrals are vital to help alleviate the pressures currently faced by HES.

Teleophthalmology for remote review of full-volume, 3D OCT scans obtained in the community by hospital-based experts and automation of the referral triaging process through artificial intelligence (AI)-assisted or AI autonomous referral pathways are key digitally enabled transformations of the referral process for common eye (retinal) conditions.

Teleophthalmology is defined as the use of medical information exchanged from one site to another via electronic/digital communications to improve a patient's eye health status. More specifically, in teleophthalmology, delivery of eye care may take several forms: through real-time teleconsultation with an eye specialist;[31] via the store-and-forward model in which digital ocular imaging is acquired and transferred via telemedical technology to remotely located (spatially and/or temporally) eye specialists;[32] or through a remote-monitoring model, which involves tracking of a patient's health data after he or she has left the hospital (e.g. smartphone-based home vision monitoring for nAMD participants).[33]

Various iterations of teleophthalmology referral models involving OCT scans have been implemented in diverse healthcare settings. A systematic review of the relevant extensive literature revealed the persistent evidence gap from well-designed, statistically powered randomised controlled trials (RCTs) for the clinical- and cost-effectiveness validation of teleophthalmology referrals.[34] Additionally, robust evidence from qualitative methods and implementation science research is of critical significance for the feasibility and scalability of teleophthalmology referrals, and this is missing from relevant literature.[35,36] This pattern is not dissimilar to the one reported in other medical specialities for telemedical healthcare models.

The HERMES superiority cluster RCT (cRCT) is an assessment of the clinical- and cost-effectiveness of teleophthalmology referrals compared to the current standard, with emphasis on reducing non-essential hospital visits and misclassified 'urgent' referrals. It focuses on improving the referral of suspicious retinal diseases from a primary/community care setting to a secondary/HES with the implementation of two digital technologies into the referral pathways: teleophthalmology and AI decision support systems (AI DSS). The teleophthalmology pathway is enabled by a custom-build, state-of-the-art, cloud-based platform with advanced ophthalmic imaging viewing functionalities, including the full 3D OCT volume scan.

There have been significant developments in the last decade of applying machine learning, specifically deep learning techniques with computer vision in medical imaging for diagnostic classification tasks.[37] This has all been in the research of and to support the use of AI DSS in diagnosing medical conditions.[37] The HERMES prospective observational real-world AI validation (diagnostic accuracy) study is aimed to generate high-quality evidence of the performance of clinical AI diagnostic assistance and referral assistance in a real-world healthcare setting. The real-world setting allows testing the performance of the AI system when it is exposed to the heterogeneity of real-world clinical care, involving variability in skills, performance, image acquisition patterns and imaging equipment. By contrast, reported validation results of clinical AI systems from in silico internal retrospective validations, or even external, out-of-sample, yet retrospective validations, are untested for performance in a prospective, real-world context. In HERMES, OCT scans originated from COs (high street opticians) and exhibited wide variability, reflective of a real-world environment. The Moorfields-Google DeepMind model ('Octane') is an AI OCT segmentation and disease diagnosis and referral recommendation model, designed to diagnose and triage participants with macular disease via OCT scans. OCT scans from the Topcon 3D OCT device (Topcon, Tokyo, Japan) were primarily used for Octane model development, but a smaller set of scans from the Heidelberg Spectralis OCT device (Heidelberg Engineering, Heidelberg, Germany) was also used. The model was, therefore, developed for a narrower use-case compared to the diversity in OCT devices and image formats encountered in real-world CO practices.

Clinical AI DSS models have been developed in recent years and were shown to have a good diagnostic accuracy against human experts in interpreting ocular imaging tests, including OCT.[38] There is a lack of evidence, however, from prospective, real-world validation studies. HERMES is the first such study on the use-case of AI DSS for referral triaging of eye disease referrals, which requires AI processing of technically complex 3D imaging (OCT). It is also the first primary care-led clinical AI study, with patient recruitment occurring exclusively at high street optician practices.[17] In parallel, human–AI interaction analysis and qualitative methods research, within the HERMES study, were used to generate evidence on the barriers and patient/clinician perceptions of AI-enabled referral triaging to inform the optimal route to real-life implementation. The results of our study, reported in this publication, provide definitive evidence on the effects of teleophthalmology for community eye referrals and produce estimates of performance accuracy for AI-enabled referral triaging in a subset of OCT devices and image formats (primarily Topcon OCT devices) encountered in CO practices.

Chapter 2 Methods

Study design

A superiority cRCT was performed to assess the effectiveness of a digital referral platform against standard care. Over a 21-month period, participants were recruited at 26 optometry practices (clusters) in the catchment areas of four HES sites: Moorfields Eye Hospital NHS Foundation Trust (MEH), University Hospitals Birmingham NHS Foundation Trust, Central Middlesex Hospital at London North West University Healthcare NHS Trust and North West Anglia NHS Foundation Trust. The selected sites covered urban, suburban and rural locations within the UK, which allowed any inferences from the study to be applicable to more of the UK population. Thirteen clusters were randomised to standard care (control arm) and 13 to teleophthalmology (intervention arm). Each cluster was able to recruit a maximum of 16 participants to the cRCT. All OCT scans and clinical vignettes from each case were transferred to the Moorfields Reading Centre to create the reference standard (RS) (diagnosis and referral recommendation). The patient management decisions made in either arm of the cRCT were assessed against a clinical RS to inform the between-arms comparison.

All OCT scans from the cRCT would be included in the AI observational prospective diagnostic accuracy study. Additionally, all cRCT sites which had reached their maximum cluster size of 16 participants before the end of the 21-month recruitment period were able to continue recruiting to the AI study in order to reach its higher recruitment target of 370 participants. The trial also allowed for up to five additional CO practices to participate for patient recruitment to contribute towards reaching the recruitment target of the AI study only. Data from one additional CO practice in the catchment area of Moorfields were processed by the Octane AI DSS. Due to limitations in generalisability of the Octane AI in a real-life clinical setting, a subset of collected OCT scans were processed by the Octane AI and included in the AI study (please see *Chapter 3, Results*). The AI DSS diagnosis and referral recommendations were compared against the clinical RS and the rule-based RS (post-hoc analysis), produced by the Reading Centre (see definitions of clinical and rule-based RSs later in the text).

During study setup, Manchester University NHS Foundation Trust and its associated catchment area were originally selected to join the cRCT. However, the site was in the process of transitioning to a regional real-life teleophthalmology referral process [Electronic eye care Referral System (EeRS)], similar to the one assessed in the intervention arm of the HERMES study. As a result, Central Middlesex and North West Anglia (Peterborough and Hinchingbrooke) were included to enable recruitment to time and target. Manchester presented, instead, an opportunity to assess enablers and barriers to real-life implementation of teleophthalmology referrals. An exploratory post-implementation substudy was planned, with limited recruitment due to local EeRS implementation challenges. We are reporting findings on the barriers to successful uptake and implementation of teleophthalmology referrals.

Patient eligibility and recruitment

Adults attending for an eye examination with an OCT scan at the participating community optometry practice were considered for participation in the study. Only individuals with a suspicion of retinal disease at the opinion of the community optometrist were recruited. Participants were required to understand the study, give consent and follow the study specific instructions. The community optometrist gave suitable participants a participant information sheet (PIS) and provided adequate time to address any queries before completing the informed written consent form.

To reflect real-world variability, no minimum experience or postgraduate credentials were required for community optometrists; supervised pre-registration optometrists were also eligible. Community optometrists provided oral consent to participate in the study and formed part of the research team.

Inclusion criteria

- Adults (≥ 18 years).
- Attended the involved community optometry practice for an eye examination and underwent a macular OCT.
- Able to give consent and understand the study.

- Able to co-operate by following study specific instructions.
- Individuals who at the opinion of the community optometrist have any suspicion of a retinal condition [including dry AMD, nAMD, diabetic retinopathy, macular oedema (MO), macular hole, epiretinal membrane (ERM), vitreomacular traction (VMT), central serous chorioretinopathy (CSCR) and genetic eye disease].

Exclusion criteria

- Individuals with known retinal comorbidities in either eye triggering a referral.
- Individuals with media opacities, inability to position or fixate or any other reason that prevents acquisition of good-quality OCT scans, at the discretion of the community optometrist.

Changes from the online protocol

As mentioned above, the following changes were made to the protocol, which were intended to facilitate the progress of recruitment and analysis to ensure that we adhere to study timelines, and these had no impact on the interpretation of our finding:

- Addition of two recruitment sites, Central Middlesex and North West Anglia (Peterborough and Hinchingbrooke), addition of exploratory teleophthalmology post-implementation substudy in the Manchester area.
- All cRCT sites which had reached their maximum cluster size of 16 participants before the end of the 21-month recruitment period were able to continue recruiting to the AI study in order to reach the higher recruitment target of 370 participants.
- Up to five additional CO practices were permitted to participate for patient recruitment to contribute towards reaching the recruitment target of the AI study only.

Ethics approval and consent

Ethical approval was obtained on 26 January 2021 from the London-Bromley Research Ethics Committee (REC 20/LO/1299). This trial was registered in the International Standard Randomised Controlled Trial Number (ISRCTN) register (ISRCTN18106677).

Informed consent

The CO at each cluster recruited eligible individuals and received written informed consent. A telephone consent form was used if consent could not be obtained in person.

Participant withdrawal

Participants were free to withdraw at any time without their standard of care being affected. Any data collected up to the point of withdrawal were included in the data analyses. All patient withdrawals were recorded on the electronic case report form (eCRF).

Randomisation and masking

Twenty-six community optometry practices were randomised 1 : 1 to standard care or teleophthalmology. A senior data manager in the research and development department at the MEH prepared a randomisation list using random permuted blocks of varying sizes with the Stata® MP, version 16 (StataCorp LP, College Station, TX, USA) statistical software. The seed number was recorded for future reference in the original randomised list.

Randomisation was performed with the unit of allocation being the cluster rather than the patient. All optometry practices were given a unique cluster identifier. Allocation concealment was at the cluster level. Allocation was conducted after confirmation of the optometry practice participation in the study. The randomisation list was sent to the trial project manager (PM) who then confirmed the randomisation allocation with the optometry practice. The name of the optometry practice, study site and date of randomisation were also recorded on the list. Community optometry practices were committed to the allocated study group for the duration of the recruitment period.

The statisticians analysing the data were masked to the management group status of the practice and patient. Unmasking only occurred following confirmation of the completion of the primary analysis.

Procedures

Cluster randomised controlled trial: teleophthalmology

Community practices randomised to the control arm continued to refer participants with suspicion of retinal disease to HES using their usual method. The standard referral methods used in the control arm included a General Ophthalmic Services (GOS) 18 form or a referral letter which was e-mailed, posted or faxed to the general practitioner (GP), or handed directly to the patient to give to the GP or to take to an accident and emergency clinic, a Rapid Access Wet AMD form and Electronic eyecare Referral Systems (EeRS). Once informed written consent was received, the COs entered their clinical findings and management decision to the teleophthalmology referral platform. The eCRF required patient demographics, baseline clinical data [visual acuity (VA), intraocular pressure (IOP), clinical signs, qualitative OCT findings], and imaging (OCT scans and optional CFP) to be entered (see *Appendix 1*). Their diagnosis and referral decision for each patient were also logged. Any participants where the CO decided not to refer and instead monitor were also recruited, and all clinical data were uploaded to the platform. All patient details were provided to the Trial Co-ordinator based at their corresponding HES site.

A diagnosis was selected from the following categories: dry AMD (drusen and macular atrophy), nAMD, diabetic macular oedema (DMO), branch retinal vein occlusion (BRVO), central retinal vein occlusion (CRVO), CSCR, inherited retinal disease (IRD) and vitreoretinal interface abnormalities (VRA), for example, ERM, full-thickness macular hole (FTMH). The referral decision was recorded as one of the following: urgent (< 2 weeks), routine (2–4 weeks), routine (4–6 weeks), routine (6–8 weeks), routine (8–12 weeks) and no referral.

Community practices randomised to the intervention arm referred participants with suspicion of retinal disease to HES using the teleophthalmology referral platform. Similar to the control arm, CO completed the same eCRF and uploaded imaging to the platform once patient consent was received. Their provisional diagnosis and referral recommendation were recorded using the same format. Human experts based at the corresponding HES reviewed every case remotely, including those where CO indicated 'no referral', and a referral decision was provided within 48 hours (tele-HES). The tele-HES recommendation was recorded on the teleophthalmology referral platform and implemented. Again, COs provided all patient details to the Trial Co-ordinator at their corresponding HES site. If required, an appointment was arranged at the corresponding HES by the Booking Centre, and both the participating CO and patient were informed of the outcome by the Trial Co-ordinator.

The findings from the first HES visit following the referral were also documented for all participants. The follow-up status was entered as either Attended, Cancelled or Did Not Attend (DNA). If attended, the date of consultation, VA, clinical signs, qualitative OCT findings, indication of any additional tests such as fundus fluorescein angiography, diagnoses and the date of first treatment (if required) were recorded.

Clinical reference standard

All patient demographic data, baseline clinical data and clinical imaging were reviewed at the Moorfields Reading Centre to create a clinical RS (ground truth) for diagnosis and referral decision for each patient. The process consisted of a review of the OCT, CFP (if available) and a clinical vignette, including VA, age, symptoms, ocular and systemic history, by an independent senior expert grader, followed by adjudication by a senior retinal specialist at the Moorfields Reading Centre. Data from the first in-person hospital visit, when available, were also considered for the RS. The efficiency of teleophthalmology was assessed by comparing both the CO and tele-HES recommendations against the clinical RS.

Observational diagnostic accuracy study: artificial intelligence decision support systems

Compatible, good quality OCT scans, suitable for processing with the Octane AI model (Topcon scans of specific size and density), were uploaded in a full-volume open-source format. OCT scans were then downloaded and processed by the Octane model within the Moorfields IT environment using local Graphic Processing Unit infrastructure. An AI diagnosis for each eye, and an AI referral recommendation for each patient was generated.

For each scan, a single AI diagnosis of either choroidal neovascularisation (CNV), dry AMD (drusen and macular atrophy), CSCR, MO, VRA and normal was given per scan. For each patient, a single referral recommendation of either urgent, routine or no referral was generated. An urgent referral was only recommended for CNV cases. No referral was recommended if no pathology was identified, and the remaining cases received a recommended routine referral.

Rule-based reference standard

As the Octane model generates a diagnosis and referral recommendation based on a set of hard-wired rules, a separate rule-based RS was also recorded for each case (ground truth for AI). Octane provides a referral recommendation of urgent referral for the retinal diagnosis of CNV; of no referral for normal cases; and of routine referral for all other retinal diagnoses, it detects. Unlike the clinical RS, where all available clinical data and imaging (OCT scans and available CFP) were reviewed, the rule-based RS was formed upon reviewing the OCT scan only. For the primary analysis of referral accuracy of the Octane AI, its referral recommendations were compared against the clinical RS, and also against the rule-based RS (post-hoc analysis). Octane had originally been assessed for referral recommendation accuracy based on AI processing of OCT scans alone compared to human experts with access to all available clinical and imaging data for each case, in a retrospective in silico validation study from NHS HES patients. Given the prospective, real-world study design of the HERMES study, involving participants seen in CO practices with an OCT device, Octane referral recommendations were compared against the clinical-based RS, formed upon reviewing the OCT scan only, for all human-assessor related outcomes. See *Appendix 2* for the equivalent diagnoses for each RS.

Outcome measures

The primary outcome measure for the cRCT was the proportion of false-positive referrals (unnecessary HES visits) and false positive urgent referrals (when routine or no referral was needed) in the standard and teleophthalmology referral pathways against the clinical RS, in enrolled participants overall and in referred participants. Secondary outcomes included the proportion of wrong diagnosis and wrong referral urgency in each arm, the proportion of false-negative referrals (participants that would have benefited from a HES review), the sensitivity and specificity in each arm against the clinical RS and the number of uncommon referrals (rare disease) which could be safely triaged in the teleophthalmology pathway. Its impact on other service delivery metrics, such as the time from referral to consultation and the time from referral to treatment for urgent maculopathies, were also assessed. Lastly, an economic evaluation aiming to assess the costs and benefits of the teleophthalmology pathway compared with the standard pathway was conducted.

The primary outcome measure for the AI study was the diagnostic accuracy (i.e. sensitivity and specificity) of the referral decision made by AI DSS, a dichotomous analysis of refer to HES and do not refer to HES, as well as refer to HES urgently, refer to HES routinely or do not refer to HES, when compared with the clinical RS and the rule-based RS. Secondary measures included the diagnostic accuracy of retinal disease and the diagnostic accuracy of Octane for referral urgency (routine or urgent referral) against the clinical RS and in a post-hoc analysis, against and the rule-based RS. The proportion of false-positive referrals (unnecessary HES visits) and the proportion of wrong diagnosis and wrong referral urgency in each arm if human assessors were replaced by AI DSS were assessed against the clinical RS.

Additionally, the technical feasibility of utilising the Octane model for real-time analysis of OCT images and its real-time operation performance were assessed by measuring the uptime and end-to-end inference speed of technical infrastructure supporting the AI DSS. The average time of end-to-end output (referral recommendation) by the AI DSS was also assessed.

The outcome measures for the pragmatic substudy were to assess the proportion of false-positive referrals (unnecessary HES visits), the proportion of false-negative referrals (participants that would have benefited from a HES review), the proportion of wrong diagnosis and the wrong referral urgency in the teleophthalmology referral pathway against the clinical RS and the intervention arm in the main cRCT.

A human–computer interaction (HCI) analysis using qualitative methods assessed the feasibility of implementation of both digital technologies. Further details on outcome measures and methods can be found in *Chapter 5*.

Data management and quality assurance

All data were handled in accordance with the General Data Protection Regulation (GDPR) and the UK Data Protection Act 2018 and the Research Governance Framework for Health and Social Care.

With regards to the confidentiality of patient records and the eCRF, identifiable patient data were not accessed outside the care team without prior consent at any stage of the project. All OCT scans were pseudonymised and no personal data were included prior to uploading the data to the digital platform. A unique identification code was assigned to each OCT scan. The log of subject codes was kept at each research centre and was not shared with the sponsor.

Active project data were stored in the dedicated, secure Reading Centre drive with appropriate back up arrangements. Access to the drive was restricted only to Reading Centre staff, where permission and access were monitored, granted and revoked on a per-user basis. Therefore, only individuals with prior authorisation could access the data.

Statistical analysis

Sample size

An audit conducted at MEH in September 2018 showed that 70% of retinal referrals were false positive.[39] A pilot study on 40 participants conducted in three optometry practices showed that this could be reduced by 60%.[39] A 95% confidence interval (CI) computed by the modified Wald Method as advised by Agresti and Coull would extend 44.6% to 73.7%.[40] There is consensus among clinicians, however, that given the savings to the NHS and benefit to participants, slightly smaller differences would be important to detect and we have powered the study to examine a reduction to 40% false referrals.

Using nQuery software version 8.3.10 (Statistical Solutions, Saugus, MA, USA), a hierarchical two-level mixed-effects model, was used to calculate the required sample size. Twenty-six clusters split between the study arms in a 1 : 1 ratio needed to recruit 306 participants overall, with an average of 10 participants per cluster, to achieve 89.27% power to detect a difference in the proportion of false-positive referrals of 30% (a drop from the current rate of 70% to the clinically relevant rate of 40%). This calculation assumed an intracluster correlation of 0.15 and the test was performed at the 5% significance level.

The same effect size assumptions applied to urgent false-positive referrals. When designing HERMES, we based our sample-size calculation on Kortuem *et al.*'s pilot data, reporting on a service evaluation without pre-planned statistical analysis and a historical control group. At the time, there were limited published data on urgent referral accuracy for suspected nAMD in community optometry. However, a prospective study of Rapid Access Referral Forms (RARFs) to a UK nAMD clinic in Scotland (Muen & Hewick, JRSM Short Rep 2011) found that 63% of urgent referrals were false positives (only 37% of referred patients actually had nAMD. The Scottish healthcare system had been an early adopter of teleophthalmology referral pathways between community optometry and HES. Reducing a 63–70% false-positive referrals to 30–40% represents a 30-percentage-point absolute difference, precisely the effect size for which HERMES was powered. Given the similarity between the RARF study's baseline of 63% and our original 70% assumption, it is reasonable to conclude that our trial was also effectively powered to detect large, clinically meaningful changes in false positive urgent referrals.

The sample size of the AI study would be reached by combining the recruited participants in the cRCT, the over-recruitment from community optometry sites within the RCT that reached their maximum cluster size of 16; the set up of up to five additional community optometry practices in the London area for recruitment into the AI study and recruitment from the pragmatic substudy combined had enabled the AI observational diagnostic accuracy study to obtain robust estimates of sensitivity and specificity. A sample size of 370 participants (accounting for the anticipated drop-out rate) with 351 participants being correctly diagnosed would produce a two-sided 95% CI with a width of 0.046.

Unit of analysis

As referral decisions are made on a per-patient basis considering the diagnoses from both eyes (if available), all referral-related outcomes were analysed using the patient as the unit of analysis. For any outcomes relating to the accuracy of diagnosing eye conditions, eyes were used as the unit of analysis instead.

Intereye correlation was accounted for as follows:

1. All analyses, where the eye was the unit of analysis, were completed twice, separately for each eye – once for the left eye and once for the right eye.
2. An intereye correlation methodology was applied using a hierarchical model with participant and optometry practice as clustering factors.

Intracluster correlation in community optometry practices

All results and tables reporting intracluster correlation coefficient (ICC) use the 'melogit' programme to calculate the intracluster correlation, report it as ICCs, and those ICCs inform the calculated difference in proportions in order to account for clustering. In the case of diagnostic accuracy results for the cRCT (secondary outcome), we use 'exlogistic' without clustering. Although the exact logistic regression did not account for clustering, the difference in proportions reported alongside the results of the exact logistic regression did account for clustering. The results are presented as unadjusted ORs and as differences in proportions, which account for clustering.

General details

Statistical analysis was completed using Stata MP v17. For both reported components, key parameters linked to each outcome are reported with 95% CIs.

Baseline characteristics for participants each in component are summarised using any relevant grouping (e.g. study arm). Continuous data were summarised using means and standard deviations (SDs) if data appeared Gaussian, or medians and interquartile ranges (IQRs). Categorical data were reported as proportions and percentages or medians (IQRs) if non-normally distributed.

Cluster randomised controlled trial: teleophthalmology

Two-sided hypothesis tests and CIs were used with a 5% significance level.

Primary analysis

For the primary outcome, two denominators determining the proportion of false-positive referrals were analysed: proportion of false-positive referrals out of all participants and proportion of false-positive referrals out of all referred participants only. The difference in the proportion of false-positive referrals between study arms (using each denominator) is calculated with 95% CI, using the exact binomial method, adjusting for cluster using arm-specific calculated ICCs. A superiority margin of 30% was used to determine the superiority of one study arm over the other. An exact logistic regression model was used to estimate the difference in the odds of false-positive referral between study arms as an odds ratio (OR) (control arm/intervention arm) was also measured. Due to convergence issues, clustering was not adjusted for in the regression, but the difference in proportions of false positives between study arms was adjusted for clustering.

Differing referral urgency levels are of interest, so an alternative definition of false-positive referral was examined. This alternative definition examined false-positive urgent referrals, defined as the assessing clinician urgently referring the patient when the clinical RS deemed an urgent referral was not necessary (patient needed routine referral or no referral). The difference in the proportion of false-positive urgent referrals was calculated using the same methods as for the primary outcome.

Although the statistical analysis plan did not prespecify a preferred denominator (all enrolled or referred participants), the false-positive rate used in the sample size was derived from a referred-only hospital dataset. Both denominators are valid; results are reported for each.

Secondary analysis

Two secondary outcomes were examined using the same methods as for the primary outcome: (1) difference in the proportion of false-negative referrals between study arms and (2) difference in the proportion of referrals with wrong referral urgency. A false-negative referral was determined as: (1) no referral when the clinical RS deemed one was needed (routine or urgent referral) and (2) routine or no referral when the clinical RS deemed an urgent referral was needed. Wrong referral urgency was defined in different ways depending on the patient population included in the analysis. When only referred participants were included in the analysis, wrong referral urgency is defined as participants with urgent instead of routine referral and vice versa. When all participants (included in the cRCT analysis) were included in the analysis, wrong referral urgency included participants with false-positive and false-negative referrals (urgent/routine referrals vs. no referral definition) as well as participants with urgent instead of routine referral and vice versa.

Diagnostic accuracy for the referral decision is reported using the following summary statistics: sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV) and area under the receiving operating characteristics (ROC) curve. Exact binomial CIs are provided for the sensitivity, specificity, PPV and NPV. The ROC curve was estimated using non-parametric methods with an asymptotic normal CI. This result does not account for clustering.

The secondary outcome of the proportion of wrong diagnoses made between study arms uses the eye as the unit of analysis. Wrong diagnosis is defined as the assessing clinician not detecting a relevant diagnosis that was detected by the clinical RS. This definition includes both: (1) the assessing clinician detecting a diagnosis not identified by the clinical RS (false positive) and (2) the assessing clinician not detecting a diagnosis when one or more was present (false negative). The same methods which were conducted for the primary outcome were used. The proportion of wrong diagnoses were examined in one eye at a time.

The time from referral to consultation and the time from referral to HES treatment were reported for each study arm. Only participants known to attend HES consultations were included in this analysis. The date of referral was the date of the baseline optometry visit. The mean time from referral to consultation and to treatment was generated with 95% CIs calculated using the exact method for a Poisson distribution. Cox regression adjusting for clustering was used to compare the time from referral to consultation between study arms. Hazard ratios with 95% CIs are reported.

The number of patients with rare diseases seen within each study arm, as well as the number of these patients receiving referrals, was reported. Proportions were not calculated due to the rarity of these conditions.

Additional analyses not specified in the published protocol

An additional exploratory analysis focusing on the detection and referral of nAMD cases was conducted. nAMD is the most consequential diagnosis as it is the main cause of urgent referrals. Audit results have shown that there are high levels of erroneous suspicion of nAMD.[28]

The following outcomes were examined for these additional analyses:

- proportion of eyes with incorrect nAMD diagnosis; unit of analysis: per eye
- proportion of false-positive referrals linked to a diagnosis of nAMD; unit of analysis: per patient
- diagnostic accuracy of the referral decision for referrals linked to a diagnosis of nAMD; unit of analysis: per patient.

In addition, a post-hoc analysis excluding any outlier community optometry practices with respect to their disproportionate contribution to false positive referrals (control arm) was conducted.

Observational diagnostic accuracy study: artificial intelligence decision support systems

As it is an observational diagnostic accuracy study, hypothesis testing needing significance levels was not used. The analysis focused on the diagnostic accuracy of the AI model as compared with the clinical and rule-based RS.

Primary analysis

The diagnostic accuracy of the referral decision dichotomised as refer (independent of urgency) versus do not refer made by AI was compared against the clinical-based RS, and, as a post-hoc analysis, the rule-based RS. The following summary statistics for diagnostics accuracy are reported: sensitivity, specificity, PPV, NPV and area under the ROC curve. Exact binomial CIs are provided for the sensitivity, specificity, PPV and NPV. The ROC curve was estimated using non-parametric methods with an asymptotic normal CI.

Secondary analysis

Summary statistics for the diagnostic accuracy of the referral decision dichotomised as urgent referral versus non-urgent (routine or no referral) was computed in the same way as the primary outcome. Summary statistics for diagnostic accuracy of referral urgency dichotomised as urgent referral versus routine referral was also computed in the same way as the primary outcome, but it excluded all participants deemed to not need a referral by either the AI model or the clinical-based RS, and the rule-based RS as a post-hoc analysis.

Summary statistics for diagnostic accuracy of the diagnosis of retinal diseases was computed for each examined eye. Six different diagnoses could be detected. The AI model was deemed to have made a relevant diagnosis if it identified the same diagnosis as the clinical-based RS. The detection of any other diagnosis by the AI model was deemed as a false positive. The detection of no diagnosis when one was present was a false negative.

All proportions for all human-assessor related outcomes were analysed with 95% CIs calculated using the exact binomial method. Human assessors indicate the clinician assessing the patient data (one clinician per participant; either CO in the control arm or the HES-based specialist in the intervention arm).

The number of false-positive referrals if the human assessors in the main cRCT were replaced with AI was determined by comparison to the clinical RS. Two definitions of false-positive referral were used following the two definitions used to dichotomise referral decision: refer (independent of urgency) versus do not refer and urgent referral versus non-urgent (routine or no referral). The number of necessary referrals as determined by the clinical RS, the number of correct (true positive) referrals and the number and proportion of false-positive and false-negative referrals if the referral decision was made by either the human assessor or the AI model were reported for the following groups:

1. all participants included in this observational diagnostic accuracy study
2. all participants within this study that are also included in the main cRCT.

The number of participants with the wrong referral urgency if the human assessors in the cRCT were replaced with AI was determined by comparison to the clinical RS. This analysis excluded all participants who were deemed to not need a referral by the clinical RS. The number of necessary routine and urgent referrals as determined by the clinical RS and the number and proportion with the incorrect urgency if the referral decision was made by either human assessors or the AI model were reported for the same two groups listed above.

The number of eyes with the wrong diagnosis if human assessors in the cRCT were replaced with AI was determined compared to the clinical RS. The number of true diagnoses as determined by the clinical RS and the number of correct (true positive) diagnoses and the number and proportion of false-positive and false-negative diagnoses if the diagnostic decision was made by either the human assessors or the AI model were reported for the same two groups listed above.

The technical infrastructure details were reported as discrete numbers. These values did not vary, so a summary of variations in time or inference speed could not be reported.

Clustering within practice was not performed for the AI study. Discussions with the clinical teams revealed that the AI DSS and the RSs that this was compared to all originated from the Reading Centre. In fact, the RSs were provided by two senior expert clinicians/graders (regardless of the optometry practice the patient came from) with adjudication by a senior retinal specialist. This meant that there were unlikely to be factors that differed by optometry practice that influenced either the AI DSS or the RSs. The OCT devices and images assessed by the AI DSS and expert clinicians/graders for the RS were subject to a quality assurance process at the Reading Centre to ensure consistency and high quality.

Furthermore, a post-hoc analysis on assessing neovascular AMD (choroidal neovascularisation) specific outcomes and the AI performance using Topcon OCT scans only was conducted.

Economic evaluation

An evaluation comparing the outcomes in terms of the costs and consequences of the teleophthalmology referral pathway and the standard care pathway was conducted.

The economic evaluation aimed to include a within-trial cost–consequence analysis (CCA) directly comparing the interventions from the trial, a cost–benefit analysis (CBA) where the consequences from the CCA were to be valued using a DCE comparing the teleophthalmology digital pathway with standard care.

Trial management

The overall management structure of the HERMES study consisted of an Executive Group (e.g.), Trial Management Group (TMG), Trial Steering Committee (TSC) and a Data Monitoring Committee (DMC).

The TMG was responsible for the day-to-day running and management of the trial. The group met to discuss the progress and examine mitigating strategies in case of issues arising.

The TSC examined the overall integrity of the study by monitoring the progress, investigating any serious adverse effects and considering the regular reports from the DMC. The TSC comprised of an independent statistician, two clinicians and a representative of the Macular Society (the Chief Investigator, study statisticians, health economists and the PM attended as observers). Some TSC members expressed a preference to have had more detailed updates during the deferral of submission period (March 2024–September 2024).

The DMC monitored the trial data to ensure the trial was implemented in accordance with the highest standards of patient safety and ethical conduct. Data on recruitment, emerging external evidence, sample characteristics and primary outcomes were monitored. The DMC included a statistician and two clinicians independent of the study team.

Sponsor

MEH was the sponsor of the HERMES trial.

Chapter 3 Results

Cluster randomised controlled trial teleophthalmology

Recruitment and participant flow

Participants were recruited between 5 July 2021 and 31 March 2023. Seventy-one optometry sites were considered for inclusion in this study; 33 were ultimately randomised in a 1 : 1 allocation, allowing for the replacement of withdrawn sites. Seven sites were withdrawn, leaving a total of 26 sites which recruited a total of 304 participants.

One hundred and thirty-six participants were recruited in the control arm, with a median of 9 participants per optometry practice (range: 5–16 participants per cluster). No participants were excluded in the control arm; thus, all were included in the analyses where the patient was the unit of analysis. A total of 267 eyes were examined as part of the referral process. One hundred and thirty-one (96.3%) participants had both eyes examined as part of the referral, while 5 participants had only one eye examined. Thirty-eight examined eyes (fellow eyes) were not reviewed to form a clinical RS as the COs did not upload the imaging data. All referred eyes were reviewed. As a result, only 229 eyes from the 136 participants in the control arm were included in the analyses where the eye was the unit of analysis.

One hundred and sixty-eight participants were recruited in the intervention arm with a median of 15 participants per cluster/optometry practice (range: 6–20 participants per cluster), of which 10 participants were excluded as they were already receiving treatment (n = 4), were ineligible for the trial (n = 4) or where relevant data were not completed (n = 2). Following these exclusions, 158 participants were included in the analyses where the patient was the unit of analysis, with a median of 12 participants per optometry practice (range: 6–16 participants per cluster). A total of 293 eyes were examined as part of the referral process. One hundred and thirty-five (85.4%) of participants had both eyes examined, while 23 participants had only eye examined. One OCT scan of the non-referred eye was not examined as the CO did not upload it to the referral platform. As a result, 292 eyes from 158 participants in the intervention arm were included in the analyses, where the eye was the unit of analysis (*Figure 1*). All referred eyes and all available data were reviewed to generate a clinical RS for both arms.

Baseline characteristics

Participants included in this study were a median of 73 years old (IQR: 61–78 years), and slightly less than half of participants were male [127/294 (43.2%)]; 69.0% of participants had existing general medical conditions (203/294), and 44.9% had known ocular history in one or both eyes (132/294 across both eyes). There were few identifiable differences between study arms except for an increase in a history of DMO in participants in the intervention arm, 9/158 (5.7%) participants across both eyes versus 1/136 (0.7%) across both eyes in the control arm. Notably, most participants in this study were recorded as being non-smokers [265/294 (90.8%)] (*Table 1*). This differs from the accepted proportion of non-smokers for a population with a median age of 73 years, which would have more individuals with a history of smoking.

Optometry practices recruited participants with an average age of 70.6 years (IQR: 68.1–73.4 years) and recruited slightly fewer men than women [median %: 42.4 (IQR: 37.5–50.0)]. Medical history and ocular history varied between clusters but followed a similar pattern (*Table 2*). These patient characteristics are consistent between optometry practices in each arm.

Basic characteristics of eyes examined by each study arm were very similar (IOP and VA). However, the assessing clinician made some different findings. COs in the intervention arm were more likely to find intraretinal fluid, where it was present in 29.5% of examined right eyes and 23.3% of examined left eyes versus 22.3% and 12.0% of eyes examined in the control arm, respectively. COs in the intervention arm were less likely to find macular atrophy (present in 3.4% of right eyes and 4.8% of left eyes vs. 16.1% and 12.8% of eyes in the control arm, respectively) but were more

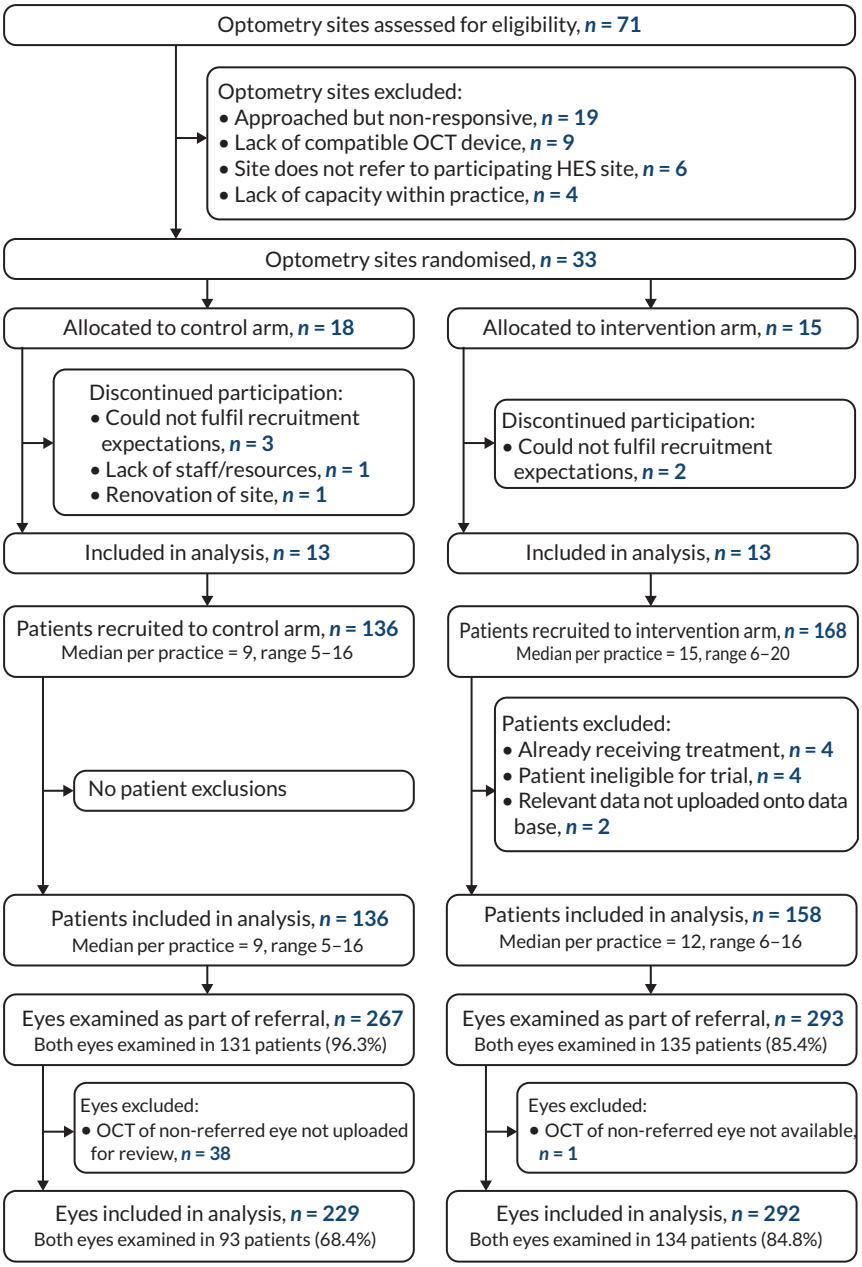

FIGURE 1 Consolidated Standards of Reporting Trials diagram demonstrating optometry site (cluster) recruitment and exclusions from the study as well as patient recruitment and exclusions.

likely to make 'other' clinical findings (present in 55.5% of right eyes and 47.3% of left eyes vs. 31.2% and 25.6% of eyes in the control arm, respectively). Overall, clinicians in the intervention arm made more clinical findings than clinicians in the control arm (*Table 3*).

Primary outcome

The proportion of false-positive referrals in each study arm

Across the 294 participants recruited to the study, 249 (84.7%) were referred across both study arms. There were more referrals in the control arm (125/136, 91.9%) than in the intervention arm (124/158, 78.5%) (p = 0.001, chi-squared test comparing arms).

TABLE 1 Patient characteristics described at the patient level for all participants and each study arm

Patient-level patient characteristics	Total, *N* = 294	Control arm, *N* = 136	Intervention arm, *N* = 158
Age (years)[a]	73 (61–78)	73 (62–78)	73 (61–78)
Male sex	127 (43.2%)	63 (46.3%)	64 (40.5%)
Smoking status			
Non-smoker	267 (90.8%)	127 (93.4%)	140 (88.6%)
Ex-smoker	17 (5.8%)	4 (2.9%)	13 (8.2%)
Current smoker	10 (3.4%)	5 (3.7%)	5 (3.2%)
Medical history			
Number of participants with medical history[b]	203 (69.0%)	91 (66.9%)	112 (70.9%)
Conditions			
Heart attack	10 (3.4%)	6 (4.4%)	4 (2.5%)
COPD	12 (4.1%)	5 (3.7%)	7 (4.4%)
Diabetes	60 (20.4%)	21 (15.4%)	39 (24.7%)
Hypertension	118 (40.1%)	52 (38.2%)	66 (41.8%)
Stroke/TIA	10 (3.4%)	5 (3.7%)	5 (3.2%)
Impaired mobility	14 (4.8%)	5 (3.7%)	9 (5.7%)
Asthma	19 (6.5%)	13 (9.6%)	6 (3.8%)
Other	100 (34.0%)	48 (35.3%)	52 (32.9%)
Ocular history			
Number of participants with ocular history[b]	132 (44.9%)	62 (45.6%)	70 (44.3%)
Conditions			
Wet AMD	23 (7.8%)	11 (8.1%)	12 (7.6%)
Dry AMD	54 (18.4%)	25 (18.4%)	29 (18.4%)
CSCR	9 (3.1%)	4 (2.9%)	5 (3.2%)
CRVO	2 (0.7%)	0 (0.0%)	2 (1.3%)
BRVO	4 (1.4%)	2 (1.5%)	2 (1.3%)
DMO	10 (3.4%)	1 (0.7%)	9 (5.7%)
IED	2 (0.7%)	1 (0.7%)	1 (0.6%)
Vitreoretinal abnormalities	28 (9.5%)	11 (8.1%)	17 (10.8%)
Other	27 (9.2%)	14 (10.3%)	13 (8.2%)
Medication for eye conditions			
Number of participants with eye medication[b]	13 (4.4%)	6 (4.4%)	7 (4.4%)
Medication type			
Prostaglandin	8 (2.7%)	5 (3.7%)	3 (1.9%)
CA inhibitors	4 (1.4%)	2 (1.5%)	2 (1.3%)
Beta-blockers	5 (1.7%)	1 (0.7%)	4 (2.5%)
AREDS	4 (1.4%)	1 (0.7%)	3 (1.9%)

TABLE 1 Patient characteristics described at the patient level for all participants and each study arm (*continued*)

Patient-level patient characteristics	Total, *N* = 294	Control arm, *N* = 136	Intervention arm, *N* = 158
Previous eye procedures			
Number of participants with previous procedures[b]	92 (31.3%)	44 (32.4%)	48 (30.4%)
Procedure type			
Cataract surgery	81 (27.6%)	39 (28.7%)	42 (26.6%)
Glaucoma surgery	4 (1.4%)	2 (1.5%)	2 (1.3%)
Eyelid surgery	3 (1.0%)	0 (0.0%)	3 (1.9%)
Other	22 (7.5%)	14 (10.3%)	8 (5.1%)

AREDS, age-related eye disease study; COPD, chronic obstructive pulmonary disease; TIA, transient ischaemic attack; CA, Carbonase Inhibitors.
a Median (IQR).
b Number of participants with medical history, ocular history, eye medications or eye procedures is across all described categories. Some participants will have multiple conditions/treatments.

TABLE 2 Average patient characteristics described at the cluster level for all clusters and clusters randomised to each study arm

Cluster-level patient characteristics	Total, *N* = 26	Control arm, *N* = 13	Intervention arm, *N* = 13
Mean age (years)[a]	70.6 (68.1–73.4)	69.6 (68.1–72.8)	71.8 (69.9–74.0)
Male sex (% in cluster)[a]	42.2 (37.5–50.0)	44.4 (38.5–50.0)	37.5 (31.3–50.0)
Medical history (% in cluster)[a]	67.7 (55.6–81.8)	66.7 (55.6–77.8)	68.8 (56.3–83.3)
Ocular history (% in cluster)[a]	44.9 (25.0–62.5)	45.5 (38.5–50.0)	40.0 (25.0–62.5)
Medication for eye conditions (% in cluster)[a]	0.0 (0.0–8.3)	0.0 (0.0–0.0)	0.0 (0.0–8.3)
Previous eye procedures (% in cluster)[a]	36.9 (18.8–43.8)	36.4 (23.1–41.7)	37.5 (16.7–43.8)

a Median (IQR).

Of all enrolled participants, there were more false positive referrals from the control arm [10/136 (7.4%)], than the intervention arm [2/158 (1.3%)], a difference of 6.1% (-4.9% to 17.1%). Out of all referrals made, there were 10 false-positive referrals in the control arm, made up of 9 unnecessary routine referrals and 1 unnecessary urgent referral, and 2 false-positive referrals in the intervention arm that were made up of 2 unnecessary routine referrals. This means that there is a false-positive referral rate of 8.0% (out of all referrals) in the control arm and 1.6% in the intervention arm, meaning that the control arm had 6.4% (−5.4% to 18.2%) more false-positive referrals than the intervention arm. This difference in proportion is insufficient to reach the superiority margin of 30% due to the unexpectedly low number of false-positive referrals in both arms(*Table 4*).

Focusing purely on urgent referrals, across the 294 participants recruited to the study, 65 (22.1%) were urgently referred across both study arms. Proportionally, there were more urgent referrals in the control arm (38/136, 27.9%) than in the intervention arm (27/158, 17.1%).

Of all enrolled participants, false positive urgent referrals were significantly higher from the control arm [24/136 (17.7%)] than the intervention arm [1/158 (0.6%)], a difference of 17.0% (10.5% to 23.5%). Out of all urgent referrals made, there were 24 false-positive urgent referrals in the control arm, where 23 participants required a routine referral instead and 1 referral was not needed and 1 false-positive urgent referral in the intervention arm which is where a routine referral was needed. This means that there is a false-positive referral rate of 63.2% in the control arm and 3.7%

TABLE 3 Clinical findings described at the eye level for all participants and each study arm

	Total		Control arm		Intervention arm	
Eye-level clinical findings	**Right eye**	**Left eye**	**Right eye**	**Left eye**	**Right eye**	**Left eye**
Number of eyes assessed	258 (87.8%)	263 (89.5%)	112 (82.4%)	117 (86.0%)	146 (92.4%)	146 (92.4%)
IOP[a]	15.1 (3.5)	15.5 (3.7)	15.6 (3.8)	16.1 (4.1)	14.8 (3.3)	15.0 (3.3)
Missing	1	2	1	2	0	0
VA (ETDRS)[b]	76 (70–85)	76 (70–85)	76 (65–85)	76 (65–85)	76 (70–85)	76 (70–85)
OCT findings						
Number of eyes with OCT findings[c]	115 (44.6%)	96 (36.5%)	53 (47.3%)	44 (37.6%)	62 (42.5%)	52 (35.6%)
Findings						
Subretinal fluid	37 (14.3%)	31 (11.8%)	18 (16.1%)	14 (12.0%)	19 (13.0%)	17 (11.6%)
Intraretinal fluid	68 (26.4%)	48 (18.3%)	25 (22.3%)	14 (12.0%)	43 (29.5%)	34 (23.3%)
Pigment epithelial detachment	42 (16.3%)	40 (15.2%)	16 (14.3%)	16 (13.7%)	26 (17.8%)	24 (16.4%)
Subretinal hyper-reflective material	26 (10.1%)	22 (8.4%)	15 (13.4%)	10 (8.5%)	11 (7.5%)	12 (8.2%)
Clinical findings						
Number of eyes with clinical findings[c]	146 (56.6%)	136 (51.7%)	59 (50.0%)	54 (46.2%)	90 (61.6%)	82 (56.2%)
Findings						
Macular haemorrhage	12 (4.7%)	12 (4.6%)	9 (8.0%)	8 (6.8%)	3 (2.0%)	4 (2.7%)
Other retinal haemorrhage	10 (3.9%)	8 (3.0%)	6 (5.4%)	5 (4.3%)	4 (2.7%)	3 (2.1%)
Exudates	15 (5.8%)	14 (5.3%)	9 (8.0%)	9 (7.7%)	6 (4.1%)	5 (3.4%)
Disc swelling	1 (0.4%)	2 (0.8%)	1 (0.9%)	1 (0.9%)	0 (0.0%)	1 (0.7%)
Macular atrophy	25 (8.9%)	25 (8.4%)	18 (16.1%)	15 (12.8%)	5 (3.4%)	7 (4.8%)
Cotton wool spot	3 (1.2%)	3 (1.1%)	1 (0.9%)	1 (0.9%)	2 (1.4%)	2 (1.4%)
Other	116 (45.0%)	99 (37.6%)	35 (31.2%)	30 (25.6%)	81 (55.5%)	69 (47.3%)

ETDRS, Early Treatment Diabetic Retinopathy Study.
a Mean (SD).
b Median (IQR).
c Number of participants with OCT findings/clinical findings is across all described categories. Some participants will have multiple findings.

TABLE 4 Proportion of false-positive referrals in each study arm

Definition of false-positive referral[a]	Control arm, *N* = 136	Intervention arm, *N* = 158	Difference in proportions[b]	Unadjusted OR[c]	Unadjusted *p*-value[c]
***Referral when not needed* (*routine or urgent referral given when not needed*)**					
Proportion of all participants	10/136 (7.4%)	2/158 (1.3%)	6.1% (−4.9% to 17.1%)	6.16 (1.28 to 58.80)	0.018
ICC	*0.492*	*< 0.001*			
Proportion of referrals	10/125 (8.0%)	2/124 (1.6%)	6.4% (−5.4% to 18.2%)	5.27 (1.09 to 50.52)	0.035
ICC	*0.472*	*< 0.001*			
***Urgent referral when not needed* (urgent referral given when routine or no referral was needed)**					
Proportion of all participants	24/136 (17.7%)	1/158 (0.6%)	17.0% (10.5% to 23.5%)	33.37 (5.28 to 1392.05)	< 0.001
ICC	*< 0.001*	*< 0.001*			
Proportion of urgent referrals	24/38 (63.2%)	1/27 (3.7%)	59.5% (41.1% to 77.8%)	42.00 (5.69 to 1901.06)	< 0.001
ICC	*0.071*	*< 0.001*			

a Two possible definitions were provided for false-positive referral. Both are compared against the clinical RS.
b Difference in proportions (control arm − intervention arm) reported with 95% CI adjusting for cluster using arm-specific calculated ICC.
c OR and unadjusted *p*-value are obtained from an exact logistic regression model not accounting for clustering. OR is presented as control arm/intervention arm.

in the intervention arm, meaning that the control arm had 59.5% (41.1% to 77.8%) more false-positive urgent referrals than the intervention arm. Using the specified superiority margin of 30%, the intervention arm can be classed as superior to the control arm at correctly identifying participants needing urgent referrals among referred participants.

Secondary outcomes

The proportion of false-negative referrals in each study arm

Across the 294 participants recruited to the study, 45 (15.3%) were not referred across both study arms. Proportionally, less participants were not provided a referral in the control arm (11/136, 8.1%) than in the intervention arm (34/158, 21.5%).

Out of participants not referred, there were four false-negative referrals in the control arm made up of three participants not referred when a routine referral was needed and one patient not referred when an urgent referral was needed. There were three false-negative referrals in the intervention arm made up of three participants not referred when a routine referral was needed. This means that there is a false-negative referral rate of 36.4% in the control arm and 8.8% in the intervention arm, meaning that the control arm had 27.5% (−2.4% to 57.5%) more false-negative referrals than the intervention arm. This difference in the proportion of false-negative referrals is not significant at the 5% significance level (based on both the CI above and the exact logistic regression described in *Table 5*).

Focusing purely on urgent referrals: across the 294 participants recruited to the study, 229 (77.9%) were not urgently referred across both study arms. Proportionally, less participants were not provided an urgent referral in the control arm (98/136, 72.1%) than in the intervention arm (131/158, 82.9%).

Out of participants not urgently referred, there were five false-negative urgent referrals in the control arm made up of one patient not referred when an urgent referral was needed and four participants given a routine referral when an urgent referral was needed. There was one false-negative urgent referral in the intervention arm where a patient was given a routine referral instead of an urgent referral. This means that there is a false-negative referral rate of 5.1% in the control arm and 0.8% in the intervention arm (0.8%), meaning that the control arm had 4.3% (−2.4% to 11.1%) more false-negative urgent referrals than the intervention arm. This difference in the proportion of false-negative referrals is not significant at the 5% significance level (based on both the CI above and the exact logistic regression described in *Table 5*).

Wrong referral urgency in each study arm

Across the 294 participants recruited to the study, 249 (84.7%) were referred across both study arms. Proportionally, there were more referrals in the control arm (125/136, 91.9%) than in the intervention arm (124/158, 78.5%). However, 48 (16.3%) were provided with the incorrect referral decision based on the clinical RS referral decision. Across all participants, an incorrect referral/wrong referral urgency was defined as any of the following: referral when none needed, no referral when one needed, routine referral when an urgent referral was needed or an urgent referral when a routine referral was needed.

Across all participants, there were 41 incorrect referral decisions in the control arm (30.2%) and seven in the intervention arm (4.4%), meaning that the control arm had 25.7% (14.1% to 37.3%) more incorrect referral decisions than the intervention arm. This difference in the proportion of false-negative referrals is significant at the 5% significance level (based on both the CI above and the exact logistic regression described in *Table 6*, $p < 0.001$).

Of those referred, 41 participants were given the wrong referral urgency based on the clinical RS referral decision (i.e. routine referral given when an urgent referral was needed, or an urgent referral given when a routine referral was needed). There were 37 referrals with the wrong referral urgency in the control arm (29.6%) and four referrals with the wrong referral urgency in the intervention arm (3.2%), meaning that the control arm had 26.4% (14.1% to 38.7%) more referrals with the wrong referral urgency than the intervention arm. This difference in the proportion of false-negative referrals is significant at the 5% significance level (based on the CI above).

TABLE 5 Proportion of false-negative referrals in each study arm

Definition of false-negative referral[a]	Control arm, *N* = 136	Intervention arm, *N* = 158	Difference in proportions[b]	Unadjusted OR[c]	Unadjusted *p*-value[c]
No referral when needed **(*no referral given when routine or urgent referral was needed*)**					
Proportion of all participants	4/136 (2.9%)	3/158 (1.9%)	1.0% (−2.5% to 4.6%)	1.56 (0.26 to 10.86)	0.836
ICC	*< 0.001*	*< 0.001*			
Proportion of non-referrals	4/11 (36.4%)	3/34 (8.8%)	27.5% (−2.4% to 57.5%)	5.60 (0.76 to 48)	0.099
ICC	*< 0.001*	*< 0.001*			
No urgent referral when needed **(no referral or routine referral given when urgent referral was needed)**					
Proportion of all participants	5/136 (3.7%)	1/158 (0.6%)	3.0% (−2.5% to 8.6%)	5.96 (0.66 to 285.19)	0.152
ICC	*0.184*	*< 0.001*			
Proportion of non-urgent referrals	5/98 (5.1%)	1/131 (0.8%)	4.3% (−2.4% to 11.1%)	6.94 (0.76 to 333)	0.106
ICC	*0.175*	*< 0.001*			

a Two possible definitions were provided for false-negative referral. Both are compared against the clinical RS.
b Difference in proportions (control arm–intervention arm) reported with 95% CI adjusting for cluster using arm-specific calculated ICC.
c OR and unadjusted *p*-value are obtained from an exact logistic regression model not accounting for clustering. OR is presented as control arm/intervention arm.

TABLE 6 Proportion of cases with wrong referral urgency in each study arm

Secondary outcome: wrong referral urgency					
Patient population[a]	**Control arm, *N* = 136**	**Intervention arm, *N* = 158**	**Difference in proportions[b]**	**Unadjusted OR[c]**	**Unadjusted *p*-value[c]**
All participants	41/136 (30.2%)	7/158 (4.4%)	25.7% (14.1% to 37.3%)	9.24 (3.90 to 25.43)	**< 0.001**
ICC	*0.105*	*< 0.001*			
Referred participants only	37/125 (29.6%)	4/124 (3.2%)	26.4% (14.1% to 38.7%)		
ICC	*0.116*	*< 0.001*			

a Wrong urgency definition depends on patient population. All participants will include those with false-positive and false-negative referrals (as determined by the clinical RS). Referred participants only include participants with urgent instead of routine and vice versa.
b Difference in proportions (control arm–intervention arm) reported, with 95% CI adjusting for cluster using arm-specific calculated ICC.
c OR and unadjusted *p*-value are obtained from an exact logistic regression model not accounting for clustering. OR is presented as control arm/intervention arm.

Sensitivity and specificity for referral decision of each study arm

Examining the overall diagnostic accuracy of each study arm, both arms have a high sensitivity for detecting the need for a referral [control arm: 96.6% (91.6% to 99.1%) vs. intervention arm: 97.6% (93.1% to 99.5%)]. This is because of the similar numbers of participants given a false-negative referral (four in the control arm vs. three in the intervention arm). However, specificity is much lower in the control arm [41.2% (18.4% to 67.1%)] than in the intervention arm [93.9% (79.8% to 99.3%)]. The lack of overlap in the CIs of these two specificities (despite a lack of adjustment for clustering) indicates a substantial difference in the specificity of each study arm. Specificity is linked to the false-positive referral rate, which suggests that more false-positive referrals are seen in the control arm than in the intervention arm.

Overall, in the control arm, if the assessing clinician offers a referral to a patient there is a 92.0% (85.8% to 96.1%) probability that a referral is truly necessary and if a referral is not offered, there is a 63.6% (30.8% to 89.1%) probability that a referral is truly not needed (according to the clinical RS). Meanwhile, in the intervention arm, if the assessing clinician offers a referral to a patient, there is a 98.4% (94.3% to 99.8%) probability that a referral is truly necessary, and, if a referral is not offered, there is a 91.2% (76.3% to 98.1%) probability that a referral is truly not needed (according to the clinical RS).

Focusing on the diagnostic accuracy of each study arm for detecting urgent referrals, the intervention arm consistently has a higher sensitivity [96.3% (81.0% to 99.9%)] and specificity [99.2% (95.8% to 100%)] than the control arm [sensitivity: 73.7% (48.8% to 90.9%), specificity: 79.5% (71.0% to 86.4%)]. As seen when examining all referrals, the difference in specificity between the two arms is sufficient that the CIs do not overlap, suggesting a substantial increase in specificity in the intervention arm compared to the control arm (despite a lack of adjustment for clustering) (*Table 7*). Specificity is linked to the false-positive referral rate, which suggests that more false-positive urgent referrals are seen in the control arm than in the intervention arm.

Overall, in the control arm, if the assessing clinician offers an urgent referral to a patient, there is a 36.8% (21.8% to 54.0%) probability that an urgent referral is truly necessary, and, if an urgent referral is not offered, there is a 94.9% (88.5% to 98.3%) probability that an urgent referral is truly not needed (according to the clinical RS). Meanwhile, in the intervention arm, if the assessing clinician offers an urgent referral to a patient, there is a 96.3% (81.0% to 99.9%) probability that an urgent referral is truly necessary, and, if an urgent referral is not offered, there is a 99.2% (95.8% to 100%) probability that an urgent referral is truly not needed (according to the clinical RS).

TABLE 7 Diagnostic accuracy of each study arm for referral decisions against the clinical RS

Diagnostic indicator[a]	Control arm, *N* = 136	Intervention arm, *N* = 158
Definition of false positive: referred when not needed		
Sensitivity	96.6% (91.6 to 99.1%)	97.6% (93.1 to 99.5%)
Specificity	41.2% (18.4 to 67.1%)	93.9% (79.8 to 99.3%)
PPV	92.0% (85.8 to 96.1%)	98.4% (94.3 to 99.8%)
NPV	63.6% (30.8 to 89.1%)	91.2% (76.3 to 98.1%)
Area under ROC curve	0.69 (0.57 to 0.81)	0.96 (0.91 to 1.00)
Definition of false positive: urgent referral when not needed		
Sensitivity	73.7% (48.8 to 90.9%)	96.3% (81.0 to 99.9%)
Specificity	79.5% (71.0 to 86.4%)	99.2% (95.8 to 100%)
PPV	36.8% (21.8 to 54.0%)	96.3% (81.0 to 99.9%)
NPV	94.9% (88.5 to 98.3%)	99.2% (95.8 to 100%)
Area under ROC curve	0.77 (0.66 to 0.87)	0.98 (0.94 to 1.00)

a Indicators come from a diagnostic test which does not account for clusters. All indicators are reported with 95% CIs.

The proportion of wrong diagnosis in each study arm

In the control arm, 42 out of 112 examined right eyes (37.5%) and 42 out of 117 examined left eyes (35.9%) were given the wrong diagnosis. While in the intervention arm, 26 out of 146 examined right eyes (17.8%) and 25 out of 146 examined left eyes (17.1%) were given the wrong diagnosis. In both eyes, an increase in the proportion of wrong diagnoses can be seen in the control arm compared to the intervention arm [right eye: 19.7% (6.1% to 33.3%); left eye: 18.8% (4.3% to 33.2%)]. This increase in wrong diagnosis is significant at the 5% significance level (based on both the CIs described above and the exact logistic regression described in *Table 8*). Accounting for clustering at the patient and optometry practice level, the odds of an eye receiving an incorrect diagnosis in the control arm was 3.77 (95% CI 1.56 to 9.12) times higher than the odds in the Intervention Arm ($p = 0.003$).

Time to consultation and hospital treatment in each study arm

Across both study arms, 222 participants (75.5%) had a hospital consultation booked (and recorded in the HERMES database). Out of those attending their hospital consultation (188 participants), the mean time to consultation from the initial optometry visit across routine and urgent referrals was 87 (85 to 88) days. The mean time to consultation was comparable between study arms [control arm: 89 (87 to 91) days; intervention arm: 84 (82 to 86) days] (comparison using survival analysis: $p = 0.836$) (*Table 9*).

Focusing on urgent referrals, the mean time to consultation from the initial optometry visit was lower in participants referred through the intervention arm [31 (29 to 34) days] compared to those referred through the control arm [53 (51 to 56) days]. However, this difference was not significant when the arms were compared using survival analysis ($p = 0.132$).

Sixty-seven participants (35.6% of those attending consultation) required treatment, with more participants requiring treatment in the intervention arm (42/98, 42.9%) compared to control arm (25/90, 27.8%). The mean time to treatment from the initial optometry visit was 85 (83 to 87) days. This time is made up of a mean time to consultation of 51 (49 to 52) days and a mean time from consultation to treatment of 34 (33 to 36) days. Overall, the mean time to treatment from the initial optometry visit was comparable between study arms [control arm: 90 (87 to 94) days; intervention arm: 82 (79 to 85) days] (comparison using survival analysis: $p = 0.665$).

Sensitivity analysis: time to consultation and hospital treatment in each study arm after exclusions

Thirty-one participants in the intervention arm experienced a delay in referral due to external factors. The reasons for these delays were:

- a delay in transfer of patient information from CO to Trial Co-ordinator at HES site ($n = 12$)
- delays largely due to Booking Centre in making an appointment (after transfer of referral request) ($n = 14$)
- issues with clinical capacity at HES sites ($n = 4$)
- a special request by the optometrist for additional independent triaging review ($n = 1$).

These delays are not considered to be related to the referral pathway. As a result, the analysis of the time to consultation/treatment was repeated without these participants to examine the best-case efficiency of the teleophthalmology referral system.

Across both study arms, 191 participants (72.6%) had a hospital consultation booked (and recorded in the HERMES database). Out of those attending their hospital consultation (157 participants), the mean time to consultation from the initial optometry visit across routine and urgent referrals was 74 (72 to 75) days. The mean time to consultation was significantly lower in the intervention arm [53 (51 to 55) days] compared to the control arm [89 (87 to 91) days] (comparison using survival analysis: $p = 0.039$).

Focusing on urgent referrals, the mean time to consultation from the initial optometry visit was still lower for participants referred through the intervention arm [23 (21 to 25) days] compared to those referred through the control arm [53 (51 to 56) days]. This difference continued to be significant when the arms were compared using survival analysis ($p = 0.047$) (*Table 10*).

TABLE 8 Proportion of wrong diagnoses made by assessing clinicians in each study arm assessing separately for each eye

Wrong diagnosis[a]	Control arm, *N* = 136	Intervention arm, *N* = 158	Difference in proportions[b]	Odds ratio[c]	*p*-value[c]
Repeated analysis with each eye					
Right eye	42/112[d] (38%)	26/146[d] (18%)	20% (6% to 33%)	2.76[c] (1.51, 5.12)	**< 0.001**[c]
ICC	*0.060*	*0.052*			
Left eye	42/117[d] (36%)	25/146[d] (17%)	19% (4% to 33%)	2.70[c] (1.47, 5.03)	**< 0.001**[c]
ICC	*0.041*	*0.155*			
Adjusted analysis combining eyes					
Across eyes	84/229[d] (37%)	51/292[d] (17%)		3.77[e] (1.56, 9.12)	**0.003**[e]
Number of participants	65/136 (48%)	44/158 (28%)			
ICC of optometry site	*0.069*	*0.092*			
ICC of participant	*0.476*	*0.276*			

a Wrong diagnosis is considered as the assessing clinician providing one or more diagnoses that is not provided by the clinical reference standard or the clinician providing no diagnoses when one is present.
b Difference in proportions (control arm – intervention arm) reported with 95% confidence interval adjusting for cluster using arm-specific calculated Intraclass correlation coefficient (ICC).
c Odds ratio and unadjusted *p*-value when examined each eye separately are obtained from an exact logistic regression model not accounting for clustering. Odds ratio is presented as control arm/intervention arm.
d Missing data come from individual eyes that could not be reviewed by the clinical reference standard (only one eye uploaded for these patients).
e Odds ratio and *p*-value when examining both eyes combined are obtained from a multilevel mixed-effects logistic regression accounting for clustering within optometry sites and within participants. Odds ratio is presented as control arm/intervention arm.

TABLE 9 Time to consultation and hospital treatment in each study arm

	Control arm, *N* = 136	Intervention arm, *N* = 158	Hazard ratio[a]	*p*-value[a]
Number with consultation booked[b]	96 (70.6%)	126 (79.7%)		
Number cancelled[c]	4 (4.2%)	9 (7.1%)		
Number DNA[c]	2 (2.1%)	19 (15.1%)		
Number attending consultation	90	98		
Time to consultation (days)[d]	89 (87 to 91)	84 (82 to 86)	0.96 (0.63 to 1.46)	0.836
Number attending consultation after urgent referral	33	25		
Time to consultation for urgent referrals (days)	53 (51 to 56)	31 (29 to 34)	0.61 (0.32 to 1.16)	0.132
Number attending consultation after routine referral	57	73		

TABLE 9 Time to consultation and hospital treatment in each study arm (*continued*)

	Control arm, *N* = 136	Intervention arm, *N* = 158	Hazard ratio[a]	*p*-value[a]
Time to consultation for routine referrals (days)	110 (107 to 112)	102 (100 to 105)		
Number needing treatment	25	42		
Time to treatment (days)[e]	90 (87 to 94)	82 (79 to 85)	0.89 (0.52 to 1.52)	0.665

a Result of Cox regression models with CIs adjusted for cluster. Hazard ratio is presented as control arm/intervention arm.
b Percentage with consultation booked is percentage of participants in each study arm.
c Percentage of participants who cancelled or DNA is percentage of participants with booked consultation.
d Only those attending a consultation are included in the analysis of time to consultations. Average time in days is presented with 95% CIs unadjusted for clusters.
e Only those attending a consultation and needing treatment are included in the analysis of time to treatment. Average time in days is presented with 95% CIs unadjusted for clusters.

TABLE 10 Time to consultation and hospital treatment in each study arm (after exclusions)

	Control arm, *N* = 136	Intervention arm, *N* = 127	Hazard ratio[a]	*p*-value[a]
Number with consultation booked[b]	96 (70.6%)	95 (74.8%)		
Number cancelled[c]	4 (4.2%)	9 (7.1%)		
Number DNA[c]	2 (2.1%)	19 (15.1%)		
Number attending consultation	90	67		
Time to consultation[d]	89 (87 to 91)	53 (51 to 55)	0.57 (0.33 to 0.97)	**0.039**
Number attending consultation after urgent referral	33	23		
Time to consultation for urgent referrals	53 (51 to 56)	23 (21 to 25)	0.48 (0.23 to 0.99)	**0.047**
Number attending consultation after routine referral	57	44		
Time to consultation for routine referrals	110 (107 to 113)	69 (66 to 71)		
Number needing treatment	25	37		
Time to treatment[e]	90 (87 to 94)	55 (52 to 57)	0.62 (0.32 to 1.19)	0.151

a Result of Cox regression models with CIs adjusted for cluster. Hazard ratio is presented as control arm/intervention arm.
b Percentage with consultation booked is the percentage of participants in each study arm.
c Percentage of participants who cancelled or DNA is percentage of participants with booked consultation.
d Only those attending a consultation are included in the analysis of time to consultations. Average time in days is presented with 95% CIs unadjusted for clusters.
e Only those attending a consultation and needing treatment are included in the analysis of time to treatment. Average time in days is presented with 95% CIs unadjusted for clusters.

Sixty-two participants (39.4% of those attending consultation) required treatment, with more participants requiring treatment in the intervention arm (37/67, 55.2%) compared to control arm (26/90, 28.9%). The mean time to treatment from the initial optometry visit was 69 (67 to 71) days. The mean time to treatment from the initial optometry visit was lower in the intervention arm [55 (52 to 57) days] compared to the control arm [90 (87 to 94) days]), although this difference was not significant when the arms were compared using survival analysis (p = 0.151).

The proportion of participants with rare disease seen and safely triaged with the teleophthalmology pathway

Eight participants with rare disease (2.7% of the cohort) were included in this study: six in the control arm and two in the intervention arm. All were referred to HES regardless of whether the assessing clinician correctly detected the rare disease (based on assessment of the OCT image and clinical examination).

Three participants had IED (two in the control arm, one in the intervention arm). All three were correctly detected by the assessing clinicians in that study arm and all were offered routine referrals. An additional patient in the control arm was wrongly diagnosed with IED, but this patient had another diagnosis 'Central vitelliform lesion and pachyoroid' (pachychoroid).

Two participants had Popper's maculopathy, both in the control arm. One patient was correctly diagnosed, while the other patient was considered to have IED. Two participants had macular telangiectasia, both in the control arm. One patient was misdiagnosed with nAMD. The other was misdiagnosed as 'odd foveal architecture'. One patient had Wyburn Mason syndrome in the intervention arm. This patient was misdiagnosed as 'nasal aspect optic disc swelling'.

Additional analysis: focusing on neovascular age-related macular degeneration (post-hoc analysis)

The proportion of false-positive referrals based on a neovascular age-related macular degeneration diagnosis

Across the 294 participants recruited to the study, 44 (15.0%) were referred due to a nAMD diagnosis across both study arms. There were 24 referrals in the control arm (17.7% of 136 participants) and 20 referrals in the intervention arm (12.7% out of 158 participants).

Out of the referrals made, there were 13 false-positive referrals in the control arm (54.2% of referrals made) with no false-positive referrals in the intervention arm, meaning that the control arm had 54.2% (30.3% to 78.0%) more false-positive referrals than the intervention arm. This difference in proportion reaches the superiority margin of 30% used for the primary outcome (*Table 11*).

Sensitivity and specificity of referral decision based on neovascular age-related macular degeneration diagnosis

Focusing on referrals that are the result of a nAMD diagnosis, the intervention arm has both a higher sensitivity and specificity for making a nAMD related referral than the control arm (*Table 12*). This effect is particularly striking when looking at the specificity of each arm. The lack of overlap in the CIs of these two specificities (despite a lack of adjustment for clustering) indicates a substantial difference in the specificity of each study arm. Specificity is linked to the false-positive referral rate, which suggests that more false-positive referrals are seen in the control arm than in the intervention arm.

Overall, in the control arm, if the assessing clinician offers a referral linked to nAMD to a patient, there is a 45.8% (25.6 to 67.2%) probability that a referral is truly necessary, and, if a referral linked to nAMD is not offered, there is a 92.9% (86.4 to 96.9%) probability that a referral is truly not needed (according to the clinical RS). Meanwhile, in the intervention arm, if the assessing clinician offers a referral linked to nAMD to a patient, there is a 100% (83.2% to 100%) probability that a referral is truly necessary, and, if a referral linked to nAMD is not offered, there is a 98.6% (94.9% to 99.8%) probability that a referral is truly not needed (according to the clinical RS).

TABLE 11 Proportion of false-positive referrals in each study arm

Definition of false-positive referral[a]	Control arm, *N* = 136	Intervention arm, *N* = 158	Difference in proportions[b]	Unadjusted OR[c]	Unadjusted *p*-value[c]
Referral when not needed					
Proportion of all participants	13/136 (9.6%)	0/158 (0%)	9.6% (1.1 to 18.0%)	23.16 (3.80 to infinity)	**< 0.001**
ICC	*0.183*	*0*			
Proportion of referrals	13/24 (54.2%)	0/20 (0%)	54.2% (30.3 to 78.0%)	29.30 (4.27 to infinity)	< 0.001
ICC	*0.227*	*0*			

a Two possible definitions were provided for false-positive referral. Both are compared against the clinical RS.
b Difference in proportions (control arm – intervention arm) reported with 95% CI adjusting for cluster using arm-specific calculated ICC.
c OR and unadjusted *p*-value are obtained from an exact logistic regression model not accounting for clustering. OR is presented as control arm/intervention arm.

TABLE 12 Diagnostic accuracy of each study arm for referral decisions linked to nAMD against the clinical RS

Diagnostic indicator[a]	Control arm, *N* = 136	Intervention arm, *N* = 158
Definition of false positive: referred when not needed		
Sensitivity	57.9% (33.5 to 79.7%)	90.9% (70.8 to 98.9%)
Specificity	88.9% (81.7 to 93.9%)	100% (97.3 to 100%)
PPV	45.8% (25.6 to 67.2%)	100% (83.2 to 100%)
NPV	92.9% (86.4 to 96.9%)	98.6% (94.9 to 99.8%)
Area under ROC curve	0.73 (0.62 to 0.85)	0.96 (0.89 to 1.00)

a Indicators come from a diagnostic test which does not account for clusters. All indicators are reported with 95% CIs.

Across both study arms, 10 participants who needed a referral linked to a diagnosis of nAMD were not given a referral for nAMD. Eight of these participants were in the control arm and two were in the intervention arm.

In the control arm, six false-negative participants out of eight were given a referral by the assessing clinician. Four participants were given urgent referrals for diagnoses of dry AMD (two participants) or MO (two participants). The other two participants were given a routine referral for dry AMD/other (other diagnosis described as a 'clinically significant cataract') or MO. The two participants not given a referral were diagnosed with dry AMD.

In the intervention arm, both false-negative participants were given a referral by the assessing clinician, one an urgent referral for other diagnosis described as 'drusen, ?acquired vitelliform lesion', the other a routine referral for a diagnosis of dry AMD.

The proportion of eyes incorrectly diagnosed with neovascular age-related macular degeneration in each study arm

According to the clinical RS, 51 eyes (9.8%) had nAMD; 27 (out of 258, 10.5%) right eyes and 24 (out of 263, 9.1%) left eyes.

Assessing clinicians in the control arm correctly identified a total of 13/24 of the diagnoses, meaning that 11 diagnoses were missed (becoming false-negative diagnoses). Meanwhile, assessing clinicians in the intervention arm correctly identified 23/27 of the diagnoses, meaning that four were missed (becoming false-negative diagnoses).

In addition to the described false-negative diagnoses, COs in the control arm misdiagnosed 15 eyes with nAMD, meaning that the control arm had a 53.6% false-positive diagnosis rate. Overall, a total of 13 participants (across both eyes) were given false-positive diagnoses of nAMD. Eleven of these participants with false-positive diagnoses received urgent referrals, with the final two participants receiving a routine (2–4 weeks) referral (*Table 13*).

There were no false-positive diagnoses in the intervention arm.

Diagnostic accuracy of a neovascular age-related macular degeneration diagnosis

The result of this decreased false-positive and false-negative diagnosis of nAMD is that the intervention arm has a higher sensitivity and specificity than the control arm (*Table 14*).

Statistical analysis of proportion of wrong diagnosis

Fifteen right eyes (5.8%) and 14 left eyes (5.3%) were given the wrong diagnosis (false positive or false negative for nAMD) by the relevant assessing clinician. Specifically, in the control arm, 13 out of 112 examined right eyes (11.6%) and 13 out of 117 examined left eyes (11.1%) were given the wrong diagnosis, while in the intervention arm, three out of 146 examined right eyes (2.1%) and one out of 146 examined left eyes (0.7%) were given the wrong diagnosis. In both eyes, an increase in the proportion of wrong diagnoses can be seen in the control arm compared to the intervention arm [right eye: 9.6% (3.2% to 15.9%); left eye: 10.4% (4.1% to 16.8%)]. This increase in wrong diagnosis is significant at the 5% significance level (based on both the CIs described above and the exact logistic regression described in *Table 15*).

Additional analysis: cRCT outcomes following exclusion of control arm site (post-hoc analyses)

There was one standard care group practice which had produced more false-positive referrals [seven (44%) of 16]. However, this was not an isolated outlier among all study-participating practices across both arms. In the intervention arm, five optometry practices would have generated multiple false-positive referrals if the optometrist's recorded recommendation was implemented (13–44% of referrals). Excluding the outlier standard care group practice with disproportionately high false positive rate reduced false-positive referral differences to 1% (95% CI −3 to 5). See *Appendix 3* for further information.

In the post-hoc analysis including only Topcon OCT data (on which the model was trained and validated) AI model performance findings showed only minimal changes.

TABLE 13 Number (proportion) of incorrect nAMD diagnosis made by each study arm

	Number of eyes	Number of true diagnoses[a]	Number of diagnoses made by clinicians	Number of correct diagnoses	Number of false-positive diagnoses[b]	Number of false-negative diagnoses[c]
Both eyes						
Control arm	229	24	28	13	15 (53.6%) (33.8 to 72.5%)	11 (5.5%) (2.7 to 9.6%)
Intervention arm	292	27	23	23	0 (0%)	4 (1.5%) (0.4 to 3.8%)

a Number of true diagnoses as determined by the clinical RS.
b Proportion of false-positive diagnoses out of the total number of wet AMD diagnoses made.
c Proportion of false-negative diagnoses out of the number of participants not diagnosed with wet AMD.

TABLE 14 Diagnostic accuracy of each study arm for diagnosis of nAMD against the clinical RS

Diagnostic indicator[a]	Control arm, *N* = 136	Intervention arm, *N* = 158
Both eyes		
Sensitivity	54.2% (32.8 to 74.4%)	85.2% (66.3 to 95.8%)
Specificity	92.7% (88.2 to 95.8%)	100% (98.6 to 100%)
PPV	46.4% (27.5 to 66.1%)	100% (85.2 to 100%)
NPV	94.5% (90.4 to 97.2%)	98.5% (96.2 to 99.6%)
Area under ROC curve	0.73 (0.63 to 0.84)	0.93 (0.86 to 0.99)

a Indicators come from a diagnostic test which does not account for clusters. All indicators are reported with 95% CIs.
b Both eyes each eye as independent regardless of the two eyes coming from the same patient for 227/294 participants

TABLE 15 Proportion of wrong diagnoses concerning nAMD made by assessing clinicians in each study arm

Wrong diagnosis[a]	Control arm, *N* = 136	Intervention arm, *N* = 158	Difference in proportions[b]	Odds ratio[c]	*p*-value[c]
Repeated analysis with each eye					
Right eye	13/112[d] (12%)	3/146[d] (2%)	10% (3% to 16%)	6.22[c] (1.65, 35)	**0.004**[c]
ICC	*< 0.001*	*< 0.001*			
Left eye	13/117[d] (11%)	1/146[d] (1%)	10% (4% to 17%)	17.97[c] (2.62, 775)	**< 0.001**[c]
ICC	*< 0.001*	*0.303*			
Adjusted analysis combining eyes					
Across eyes	26/229[d] (11%)	4/292[d] (1%)		14.13[e] (3.57, 56)	**< 0.001**[e]
Number of participants	22/136[d] (16%)	4/158[d] (3%)			
ICC of optometry site	*< 0.001*	*< 0.001*			
ICC of participant	*0.533*	*< 0.001*			

a Wrong diagnosis is considered as the assessing clinician providing one or more diagnoses that is not provided by the clinical reference standard or the clinician providing no diagnoses when one is present.
b Difference in proportions (control arm − intervention arm) reported with 95% confidence interval adjusting for cluster using arm-specific calculated Intraclass correlation coefficient (ICC).
c Odds ratio and unadjusted *p*-value when examined each eye separately are obtained from an exact logistic regression model not accounting for clustering. Odds ratio is presented as control arm/intervention arm.
d Missing data come from individual eyes that could not be reviewed by the clinical reference standard (only one eye uploaded for these patients).
e Odds ratio and *p*-value when examining both eyes combined are obtained from a multilevel mixed-effects logistic regression accounting for clustering within optometry sites and within participants. Odds ratio is presented as control arm/intervention arm.

Observational diagnostic accuracy study: artificial intelligence decision support system

Recruitment and participant flow

Participants contributing to the AI study are participants included in both arms of the cRCT and additional participants recruited for the AI study only, but not included in the RCT. The overall number of included participants decreased due to the proportion of collected OCT scans suitable for processing by the Octane AI (*Figure 2*).

Patient exclusions

Among the optometry practices recruiting to the main RCT and the four additional sites recruiting for this AI observational diagnostic accuracy study only, 204 out of 396 participants were included in the AI study. A total of 192 participants were excluded from this analysis, of whom 179 were excluded because all available OCT images were not suitable for analysis or were of insufficient quality for analysis by the Octane AI. Of the 179 patients, 90 were from the control arm and 89 were from the intervention arm. Within the control arm, 9 (10%) were not referred, 59 (65.6%) were given a routine referral and 22 (24.4%) were given an urgent referral. Within the intervention arm, 20 (22.5%) were not referred, 51 (57.3%) were given a routine referral and 18 (20.2%) were given an urgent referral. This created a total of 29 patients who were not referred (16.2%), 110 who were given a routine referral (61.5%) and 40 who were patients given an urgent referral (22.4%). Of these referrals, five patients (2.8%) were given a referral when it was not needed and five patients (2.8%) were given an urgent referral when a routine referral was needed. One patient was not given a referral when an urgent referral was needed, and eight patients (4.5%) were given a routine referral when one was not needed. In total, 141 referrals (78.8%) can be considered to be true positive (based on the definition of any referral given when one was needed), and 24 (13.4%) can be considered as true negative (no referral was given).

A total of 13 participants were excluded because of ineligibility for the study (n = 8), optometry site withdrawals (n = 3) and vital data/forms not being completed (n = 2). Although these OCT scans could not be processed by the AI model, all available OCT scans were of good enough quality to be reviewed to form the clinical RS.

Cohort characteristics

Participants recruited to this study were aged between 29 and 93 years with a median of 70 (IQR: 58.5–77) years. Slightly less than half of participants were male [85/204 (41.7%)] and the majority had at least one medical condition [145/204 (71.1%)]. Notably, the majority of participants in this study are recorded as being non-smokers [179/204 (87.7%)]. This differs from the expected proportion of non-smokers for a population with a median age of 70 years, which would be expected to have more individuals with a history of smoking.

Overall, 35.8% of participants have known ocular history with known conditions evenly distributed between right and left eyes. The most common ocular condition is dry AMD (14.7% of participants, 13.7–14.7% of eyes) followed by other unspecified conditions (8.8% of participants, 5.9–6.4% of eyes) and vitreoretinal abnormalities (5.9% of participants, 3.4–4.9% of eyes).

Only 3.9% of participants were on eye medications and 28.4% of participants had had previous eye procedures (*Table 16*).

The IOP and VA were equivalent between the right and left eyes; 50.1% of participants were presenting at the optometrist with concerns about reduced/blurry vision (3.5% had additional concerns such as ocular discomfort). Another 23.0% of participants were attending routine appointments and had no presenting concerns. The remaining presenting concerns were for a range of reasons; 48.8% (99/204) of participants had had the presenting concern for > 2 weeks. Only 15 participants (7.4%) had the concerns for < 1 week (*Table 17*).

Optical coherence tomography image collection

Optical coherence tomography images were collected for all participants who consented to HERMES. The large reduction in sample size in this study is primarily due to images being unsuitable to be processed by Octane AI. In total, 710 OCT images were collected across all participants who consented to HERMES. Only 371 images (52.3% of collected images) were suitable for the analysis by Octane. These images come from the 204 participants included in this analysis.

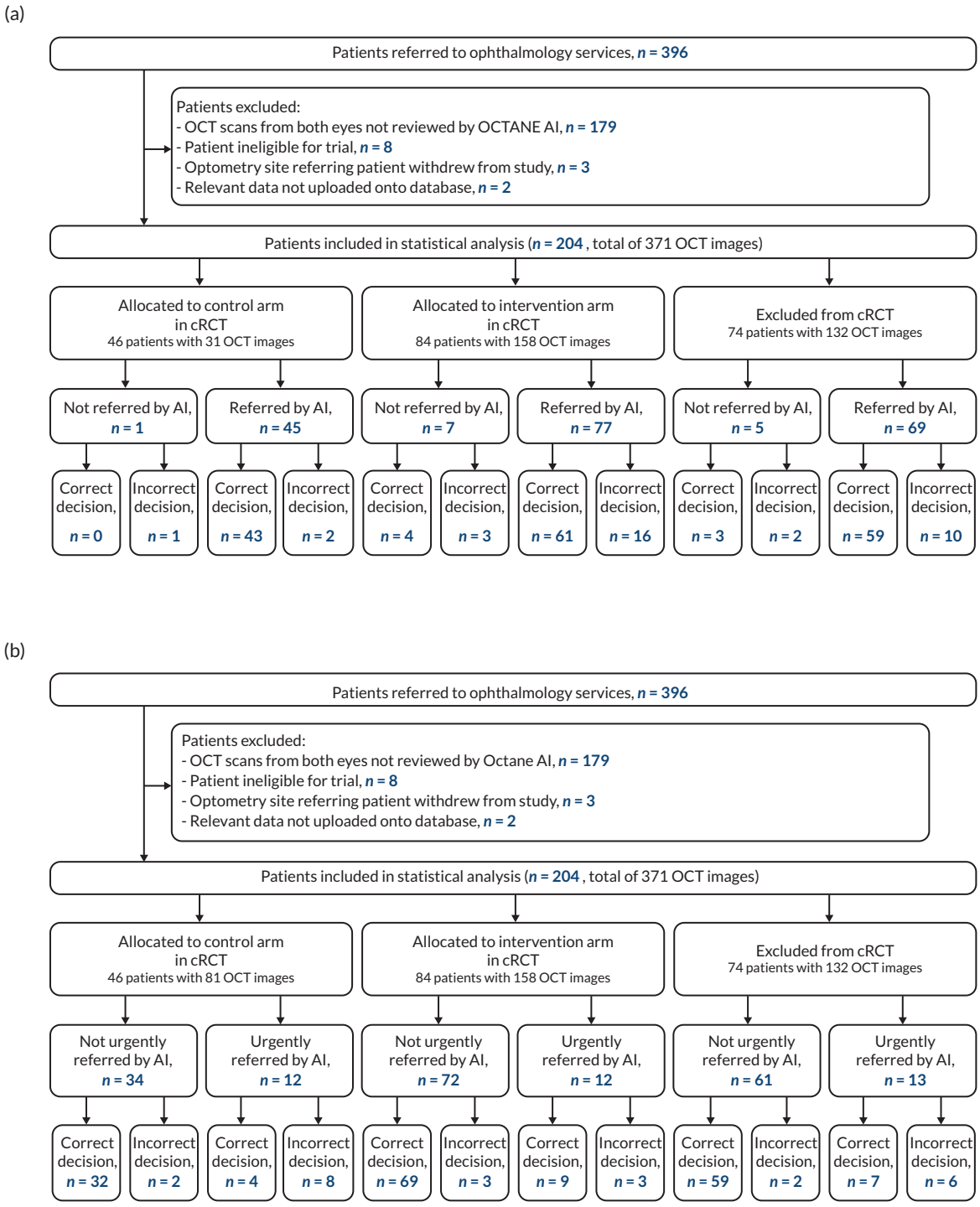

FIGURE 2 Standards for the reporting of diagnostic accuracy studies flow diagram demonstrating participant recruitment and exclusion from the AI substudy for (a) refer vs. not refer and (b) urgently referred vs. not urgently referred (routine and no referral).

Among the 204 participants who had OCT images examined by Octane AI, most of the available images have been utilised (371 out of 394 collected OCT images, 94.2%). Many participants have OCT images analysed for both eyes (167/204, 81.9%), with the rest [37 participants (18.1%)] having just one OCT image analysed by Octane AI.

The inclusion of participants in this analysis is OCT device-related, which translates to a subset of CO practices contributing to the AI study. Of the 29 community optometry practices contributing participants and OCT images to HERMES, 17 sites have recruited participants with images analysed by the Octane AI. Most of these included participants come from 14 sites [201 of the 204 participants (98.5%) included in this analysis]. Of the five OCT manufacturers used to contribute OCT images to HERMES, images that could be analysed by the Octane AI come from

TABLE 16 Patient demographics for participants included in the AI diagnostic accuracy study

	All patients, *N* = 204		**All patients, *N* = 204**		
Age (years)[a]	70 (58.5–77.0)				
Male sex	85 (41.7%)				
Smoking status			*Across both eyes*	*Right eye*	*Left eye*
Non-smoker	179 (87.7%)	Ocular history			
Ex-smoker	16 (7.8%)	Number of patients with ocular history[b]	73 (35.8%)	63 (30.9%)	55 (27.0%)
Current smoker	9 (4.4%)				
Medical history		Conditions			
Number of patients with medical history[b]	145 (71.1%)	Wet AMD	10 (4.9%)	6 (2.9%)	6 (2.9%)
		Dry AMD	30 (14.7%)	30 (14.7%)	28 (13.7%)
Conditions		Central serious chorioretinopathy	3 (1.5%)	1 (0.5%)	2 (1.0%)
Heart attack	7 (3.4%)				
COPD	12 (5.9%)	MO	13 (6.4%)	10 (4.9%)	8 (3.9%)
Diabetes	48 (23.5%)	*DMO*	*8 (3.9%)*	*6 (2.9%)*	*6 (2.9%)*
Hypertension	80 (39.2%)	Vitreoretinal Abnormalities	12 (5.9%)	10 (4.9%)	7 (3.4%)
Stroke/TIA	7 (3.4%)	Other	18 (8.8%)	13 (6.4%)	12 (5.9%)
Impaired mobility	11 (5.4%)	*CRVO*	*0 (0%)*	*0 (0%)*	*0 (0%)*
Asthma	12 (5.9%)	*BRVO*	*4 (2.0%)*	*1 (0.5%)*	*3 (1.5%)*
Other	62 (30.4%)	*IED*	*1 (0.5%)*	*1 (0.5%)*	*1 (0.5%)*
Medication for eye conditions		Previous eye procedures			
Number of patients with eye medication[b]	8 (3.9%)	Number of patients with previous procedures[b]	66 (32.4%)	56 (27.5%)	56 (27.5%)
Medication type		Procedure type:			
Prostaglandin	5 (2.5%)	Cataract surgery	58 (28.4%)	48 (23.5%)	50 (24.5%)
CA inhibitors	3 (1.5%)	Glaucoma surgery	0 (0%)	0 (0%)	0 (0%)
Beta-blockers	2 (1.0%)	Eyelid surgery	0 (0%)	0 (0%)	0 (0%)
AREDS	2 (1.0%)	Other	19 (9.3%)	15 (7.4%)	13 (6.4%)

AREDS, age-related eye disease study.
a Median (IQR).
b Number of patients with medical history, ocular history, eye medications or eye procedures is across all described categories. Some patients will have multiple conditions/treatments.

two device manufacturers, primarily Topcon (TopCon Maestro, and Topcon 3D OCT-2000) for 152 participants (74.5%), followed by Nidek (Nidek (RS-330/other) for 52 participants (25.5%).

A total of 179 participants had OCT scans which could not be processed by Octane AI. 43% (77/179) of participants' OCTs were from the Topcon device as OCTs were obtained with imaging protocols that were different to those that Octane AI was trained on. There were 75 participants with OCTs from the Heidelberg Spectralis which were not processed due to insufficient communication from the AI developer. The majority of Topcon 3D OCT-1 scans were

TABLE 17 Clinical information from the baseline optometry visit for participants included in the AI diagnostic accuracy study

	All patients, *N* = 204		All patients, *N* = 204
IOP[a]		Presenting concern	
Right eye	14.2 (3.4)	No available data	21 (10.3%)
Missing	*0 patients*	Asymptomatic	47 (23.0%)
Left eye	14.2 (3.2)	Reduced/blurry vision only	95 (46.6%)
Missing	*1 patient*	Metamorphopsia	15 (7.4%)
		Metamorphopsia and reduced vision	2 (1.0%)
VA (ETDRS)[b]		Scotoma	3 (1.5%)
Right eye	76 (70–85)	Headaches	1 (0.5%)
Left eye	76 (70–85)	Diplopia	1 (0.5%)
		Floaters and reduced vision	2 (1.0%)
Duration of presenting concern		Shadow	1 (0.5%)
< 1 week	15 (7.4%)	Scotoma and shadow	1 (0.5%)
Between 1 and 2 weeks	20 (9.9%)	Visual disturbance	1 (0.5%)
> 2 weeks	99 (48.8%)	Ocular discomfort and reduced vision	3 (1.5%)
N/A	48 (23.6%)	Diabetes related	1 (0.5%)
Unknown	21 (10.3%)	Post cataract surgery	2 (1.0%)
		Miscellaneous	2 (1.0%)
		Issues with HES care	2 (1.0%)
		Family history concerns	1 (0.5%)
		Would like new glasses	2 (1.0%)

N/A, not applicable.
a Mean (SD).
b Median (IQR).

processed apart from the wide scans with dimensions of 9x2.3x9 and 12x2.3x12. In addition, the Octane AI could not process wide-frame OCTs from the Topcon Maestro (dimensions 9x2.3x12), which represented 57 OCTs compared to the 36 standard-frame OCTS, which were processed. For Topcon Triton, Octane AI processed six OCTs with compatible dimensions (6x2.6x6 or 7x2.6x7), and not those with wide dimensions (12x2.5x9 or 12x2.6x9). Additionally, there were 20 sparse OCTs (10 b-scans) and four OCTs with small dimensions (3x2.6x3), which were not processed. There was one Nidek OCT scan from one participant that could not be processed as the image was 'unsuitable'. Optovue OCTs from 10 participants were not processed as Octane AI has not been trained on this device. Lastly, there were seven Revo OCTs that were not processed for the same reason.

Primary outcome

Diagnostic accuracy of the artificial intelligence model for referral decisions versus clinical reference standard

The AI model exhibits a high sensitivity [96.4% (92.4% to 98.7%)], detecting the need for a referral (routine or urgent) in all but six participants indicated as needing a referral by the clinical RS. However, the specificity of the AI model is very low [20.0% (8.4% to 36.9%)]; 28 participants (out of 35, 80.0%) who were deemed not to need a referral by the clinical RS would have been offered a referral by the AI model. Overall, if the AI model offers a referral to a patient, there is an 85.3% (79.5% to 90.0%) probability that a referral is truly necessary, and, if a referral is not offered, there is a 53.8% (25.1% to 80.8%) probability that a referral is truly not needed.

A secondary analysis focusing purely on urgent referrals provides a different finding for the AI model. The AI model had a reasonable sensitivity [74.1% (53.7% to 88.9%)] and a high specificity [90.4% (85.1% to 94.3%)]. Seven of the 27 participants needing an urgent referral would not have been offered one by the AI model, while an additional 17 out of 177 participants (9.6%) not needing an urgent referral would not have been offered one by the AI model. Overall, if the AI model offers an urgent referral to a patient, there is a 54.1% (36.9% to 70.5%) probability that an urgent referral is truly necessary, and, if an urgent referral is not offered, there is a 95.8% (91.6% to 98.3%) probability that an urgent referral is truly not needed.

Versus rule-based reference standard (post-hoc analysis)

The patient was used as the unit of analysis. The AI model exhibits a high sensitivity [97.9% (94.8% to 99.4%)], detecting the need for a referral (routine or urgent) in all but four participants indicated as needing a referral by the rule-based RS. The specificity is also high [100% (66.4% to 100%)], with the AI model agreeing with the rule-based RS about the nine participants who did not need a referral. Overall, if the AI model offers a referral to a patient, there is a 100% (98.1% to 100%) probability that a referral is truly necessary, and, if a referral is not offered, there is a 69.2% (38.6% to 90.9%) probability that a referral is truly not needed.

For the secondary analysis focusing purely on urgent referrals, the AI model showed moderate sensitivity [87.9% (71.8% to 96.6%)] and high specificity [95.3% (91.0% to 98.0%)] compared to the rule-based RS. Only four out of 33 participants (12.1%) needing an urgent referral according to the rule-based RS would not have been offered one by the AI model, while an additional eight participants out of the 171 participants (4.7%) not needing an urgent referral according to the rule-based RS would have been offered one by the AI model. Overall, if the AI model offers an urgent referral to a patient, there is a 78.4% (61.8% to 90.2%) probability that an urgent referral is truly necessary, and, if an urgent referral is not offered, there is a 97.6% (94.0% to 99.3%) probability that an urgent referral is truly not needed.

Secondary outcomes

Diagnostic accuracy of the artificial intelligence model for referral urgency versus clinical reference standard

In participants deemed to need a referral by both the AI model and the clinical RS only, the AI model displays similar diagnostic accuracy, as shown in *Table 18*, when looking at the urgent referrals versus routine/no referrals. The AI model has a reasonable sensitivity [74.1% (53.7% to 88.9%)] and a high specificity [87.5% (80.7% to 92.5%)]. Seven out of 27 participants (25.9%) needing an urgent referral as determined by the clinical RS would have been offered a routine referral by the AI model, while an additional 17 out of 136 participants (12.5%) needing a routine referral would have

TABLE 18 Diagnostic accuracy of the AI model for referral decisions against the clinical RS and the rule-based RS

Diagnostic accuracy for referrals[a]	vs. clinical RS, *N* = 204	vs. rule-based RS[b], *N* = 204
Routine/urgent referral vs. no referral		
Sensitivity	96.4% (92.4 to 98.7%)	97.9% (94.8 to 99.4%)
Specificity	20.0% (8.4 to 36.9%)	100% (66.4 to 100%)
PPV	85.3% (79.5 to 90.0%)	100% (98.1 to 100%)
NPV	53.8% (25.1 to 80.8%)	69.2% (38.6 to 90.9%)
Area under ROC curve	0.58 (0.51 to 0.65)	0.99 (0.98 to 1.00)
Urgent referral vs. routine/no referral		
Sensitivity	74.1% (53.7 to 88.9%)	87.9% (71.8 to 96.6%)
Specificity	90.4% (85.1 to 94.3%)	95.3% (91.0 to 98.0%)
PPV	54.1% (36.9 to 70.5%)	78.4% (61.8 to 90.2%)
NPV	95.8% (91.6 to 98.3%)	97.6% (94.0 to 99.3%)
Area under ROC curve	0.82 (0.74 to 0.91)	0.92 (0.86 to 0.98)

a All indicators are reported with 95% CIs.
b Post-hoc analysis.

TABLE 19 Diagnostic accuracy of the AI model for referral urgency against the clinical RS and the rule-based RS (patient needing referral only)

Diagnostic accuracy[a]	vs. clinical RS, *N* = 163	vs. rule-based RS, *N* = 191
Urgent referral vs. routine referral		
Sensitivity	74.1% (53.7 to 88.9%)	87.9% (71.8 to 96.6%)
Specificity	87.5% (80.7 to 92.5%)	94.9% (90.3 to 97.8%)
PPV	54.1% (36.9 to 70.5%)	78.4% (61.8 to 90.2%)
NPV	94.4% (88.9 to 97.7%)	97.4% (93.5 to 99.3%)
Area under ROC curve	0.81 (0.72 to 0.90)	0.91 (0.86 to 0.97)

a All indicators are reported with 95% CIs.

been offered an urgent referral by the AI model. Overall, if the AI model offers an urgent referral to a patient, there is a 54.1% (36.9% to 70.5%) probability that an urgent referral is truly necessary, and, if an urgent referral is not offered, there is a 94.4% (88.9% to 97.7%) probability that an urgent referral is truly not needed and a routine referral should be offered instead.

Versus rule-based reference standard (post-hoc analysis)

Compared with the rule-based RS, the AI model has a moderate sensitivity [87.9% (71.8% to 96.6%)] and a high specificity [94.9% (90.3% to 97.8%)] (*Table 19*). Only 4 out of 33 participants (12.1%) needing an urgent referral as determined by the rule-based RS would have been offered a routine referral by the AI model, while an additional 8 participants out of 158 (5.1%) needing a routine referral would have been offered an urgent referral by the AI model. Overall, if the AI model offers an urgent referral to a patient, there is a 78.4% (61.8% to 90.2%) probability that an urgent referral is truly necessary, and, if an urgent referral is not offered, there is a 97.4% (93.5% to 99.3%) probability that an urgent referral is truly not needed and a routine referral should be offered instead.

Number of false-positive referrals if human assessors were replaced with artificial intelligence

For the participants included in this substudy and included in the main cRCT (*n* = 130), a total of 109 referrals (routine or urgent) were made by the human assessors (between study arms). Three (2.8%) of these referrals were false

positives. If the AI model replaced the human assessors, a higher total of 122 referrals would have been made, of which 18 (14.8%) would have been false positives (*Table 20*).

Comparing CIs, the false-positive rate for referral (routine and urgent) is higher for the AI model (14.8% (9.0% to 22.3%) compared with human assessors [2.8% (0.6% to 7.8%)], with no overlap of CIs.

In a secondary analysis, focusing purely on urgent referrals, a total of 27 urgent referrals were made by the human assessors (between study arms). Ten (37.0%) of these referrals were false positives. If the AI model replaced the human assessors, 24 urgent referrals would have been made, of which 11 (45.8%) would have been false positives.

The high number of false-positive urgent referrals made by the human assessors is made entirely for participants in the control arm of the cRCT. Comparing CIs, the AI model has a similar false-positive rate for urgent referrals compared to the control arm [AI: 66.7% (34.9% to 90.1%) vs. assessors in control arm: 62.5% (35.4% to 84.8%)], with overlapping CIs.

The AI model is more inaccurate than human assessors when making a referral decision for required referrals (routine and urgent), and it has a similar accuracy when making an urgent referral decision to assessors in control arm.

Number of wrong referral urgency if human assessors were replaced with artificial intelligence

In the participants included in this substudy and included in the main cRCT (n = 130), 90 required routine referrals and 18 required urgent referrals as determined by the clinical RS. Thus, a total of 108 of the 130 participants required a referral. Within these 108 participants, human assessors assigned the incorrect referral urgency to 11 (10.2%) of participants: 10 by community optometrists in the control arm (out of 44 necessary referrals, 22.7%) and one by hospital-based experts in the intervention arm (out of 64 necessary referrals, 1.6%). If the AI model replaced the human assessors, 16 (5.4%) out of the 108 necessary referrals would have been referred incorrectly (routine when it should be urgent *or* urgent when it should be routine). The AI recommended the wrong referral urgency the same number of times as assessors in the control arm [AI: 22.7% (11.5% to 37.8%) vs. community optometrists in control arm: 22.7% (11.5% to 37.8%)]. The proportion of wrong referral urgencies made by the AI model was higher than that seen in the intervention arm [AI: 9.4% (3.5% to 19.3%) vs. hospital-based experts in intervention arm: 1.6% (0.0% to 8.4%)] (*Table 21*).

Diagnostic accuracy of the artificial intelligence model for the diagnosis of retinal disease versus clinical reference standard

As this outcome examines diagnostic accuracy, it was completed using the eye and not the patient as the unit of analysis. The AI model exhibits moderate sensitivity [68.1% (62.5% to 73.4%)] and specificity [67.1% (55.4% to 77.5%)] when detecting the same diagnosis as the clinical RS. The AI model correctly detected a specific retinal diagnosis in 201 images (54.1% of diagnosed images). In 25 images, the AI model detected a different diagnosis than the clinical RS, while in 60 images, the AI model detected a pathology but could not determine its nature in order to make a diagnosis (unclassified). The AI model provided a diagnosis of normal when the clinical RS indicated a specific retinal diagnosis in 34 images.

For each OCT image/eye, if the AI model makes a diagnosis, there is an 88.9% (84.1% to 92.7%) probability that this diagnosis agrees with a diagnosis made by the clinical RS. If the AI model makes no diagnosis, there is a 60.0% (48.8% to 70.5%) probability that there truly is no diagnosis needed (according to the clinical RS).

Versus rule-based reference standard (post-hoc analysis)

The AI model exhibits moderate sensitivity [74.0% (68.5% to 79.0%)] and good specificity [82.6% (72.9% to 89.9%)] when detecting the same diagnosis as the rule-based RS (*Table 22*). The AI model displays higher sensitivity and specificity when compared with the rule-based RS than with the clinical RS, with sensitivity [68.1% (62.5% to 73.4%)] and specificity [67.1% (55.4% to 77.5%)]. Two hundred and twenty-six images (60.9%) were diagnosed with a specific retinal condition by the rule-based RS. The AI model detected relevant pathologies in 211 images (93.4% of diagnosed images) and so correctly diagnosed these images. In 15 images, the AI model gave a diagnosis that did not match that given by the rule-based RS, while in 60 images, the AI model detected a pathology but could not determine its nature

TABLE 20 Number (proportion) of false-positive referrals if human assessors were replaced with AI

	Number of patients	Number of necessary referrals[a]	Total number of referrals made by…		Number of correct referrals made by…		Number of false-positive referrals made by … [b]	
			Human assessors	AI algorithm	Human assessors	AI algorithm	Human assessors	AI algorithm
Routine/urgent referral vs. no referral								
All patients	204	169	–	191	–	163	–	14.7% (10.0 to 20.5%)
cRCT[c]	130	108	109	122	106	104	2.8% (0.6 to 7.8%)	14.8% (9.0 to 22.3%)
Control arm	46	44	44	45	43	43	2.3% (0.1 to 12.0%)	4.4% (0.5 to 15.1%)
Intervention arm	84	64	65	77	63	76	3.1% (0.4 to 10.7%)	20.8% (12.4 to 31.5%)
Urgent referral vs. routine/no referral								
All patients	204	27	–	37	–	20	–	46.0% (29.5 to 63.1%)
cRCT[c]	130	18	27	24	17	13	37.0% (19.4 to 57.6%)	45.8% (25.6 to 67.2%)
Control arm	46	6	16	12	6	4	62.5% (35.4 to 84.8%)	66.7% (34.9 to 90.1%)
Intervention arm	84	12	11	12	11	9	0%	25.0% (5.5 to 57.2%)

a Number of necessary referrals as determined by the clinical RS.
b Proportion of false-positive referrals out of the total number of referrals made by either human assessor or AI algorithm (as applicable) with 95% CI.
c Only patients included in the main cRCT study (after AI study exclusions)

TABLE 21 Number (proportion) of referrals with wrong urgency if human assessors were replaced with AI

	Number of patients	Number of necessary referrals[a]		Number of referrals made by…					
				Human assessors		AI algorithm		Number of referrals with wrong urgency made by … [b,c]	
		Routine	Urgent	Routine	Urgent	Routine	Urgent	Human assessors	AI algorithm
All patients	204	142	27	–	–	154	37	–	14.2% (9.3 to 20.4%)
cRCT[d]	130	90	18	82	27	98	24	10.2% (5.2 to 17.5%)	14.8% (8.7 to 22.9%)
Control arm	46	38	6	28	16	33	12	22.7% (11.5 to 37.8%)	22.7% (11.5 to 37.8%)
Intervention arm	84	52	12	54	11	65	12	1.6% (0.0 to 8.4%)	9.4% (3.5 to 19.3%)

a Number of necessary referrals as determined by the clinical RS.
b Referrals with wrong urgency only includes necessary referrals incorrectly chosen as routine or urgent (false-positive/negative referrals are not included).
c Proportions of referrals with wrong urgency out of the total number of necessary referrals (routine + urgent referrals).
d Only patients included in the main cRCT study (after AI study exclusions).

TABLE 22 Diagnostic accuracy of the AI model detecting a diagnosis against the clinical RS and the rule-based RS

Diagnostic accuracy for diagnosis[a]	vs. clinical RS, *N* = 371	vs. rule-based RS, *N* = 371
Sensitivity	68.1% (62.5 to 73.4%)	74.0% (68.5 to 79.0%)
Specificity	67.1% (55.4 to 77.5%)	82.6% (72.9 to 89.9%)
PPV	88.9% (84.1 to 92.7%)	93.4% (89.3 to 96.2%)
NPV	35.2% (27.4 to 43.5%)	49.0% (40.6 to 57.4%)
Area under ROC curve	0.68 (0.62 to 0.74)	0.78 (0.74 to 0.83)

a All indicators are reported with 95% CIs.

TABLE 23 Diagnostic accuracy of the AI model for diagnosis of CNV against the clinical RS and the rule-based RS

Diagnostic accuracy for CNV[a]	vs. clinical RS, *N* = 371	vs. rule-based RS, *N* = 371
Sensitivity	94.7% (82.3 to 99.4%)	95.3% (84.2 to 99.4%)
Specificity	96.4% (93.8 to 98.1%)	97.9% (95.7 to 99.1%)
PPV	75.0% (60.4 to 86.4%)	85.4% (72.2 to 93.9%)
NPV	99.4% (97.8 to 99.9%)	99.4% (97.8 to 99.9%)
Area under ROC curve	0.96 (0.92 to 0.99)	0.97 (0.93 to 1.00)

a All indicators are reported with 95% CIs.

in order to make a diagnosis. In a single image, the AI model did not detect any pathology despite the presence of a diagnosis by the rule-based RS.

For each OCT image/eye, if the AI model detects a diagnosis, there is a 93.4% (89.3% to 96.2%) probability that this diagnosis agrees with a diagnosis made by the rule-based RS. If the AI model detects no diagnosis, there is a 49.0% (40.6% to 57.4%) probability that there truly is no specific retinal diagnosis present (according to the rule-based RS).

Choroidal neovascularisation as an individual diagnosis (not accounting for any other diagnoses present, post hoc analysis)

Versus clinical reference standard

The AI model has a high sensitivity [94.7% (82.3% to 99.4%)] and a high specificity [96.4% (93.8% to 98.1%)] for detecting CNV compared with the clinical RS; two of the 38 eyes (5.3%) with CNV (clinical RS) were not diagnosed with CNV by the AI model. An additional 12 out of the 333 eyes (3.6%) without CNV (clinical RS) were diagnosed with CNV by the AI model. Overall, if the AI model detects CNV as present, there is a 75.0% (60.4% to 86.4%) probability that CNV is truly present. If the AI model detects no CNV, there is a 99.4% (97.8% to 99.9%) probability that CNV is truly not present (compared to the clinical RS) (*Table 23*).

Versus rule-based reference standard (post-hoc analysis)

The AI model has a high sensitivity [95.3% (84.2% to 99.4%)] and specificity [97.9% (95.7% to 99.1%)] compared with the rule-based RS. Only two of the 43 images (4.7%) diagnosed with CNV by the rule-based RS were not diagnosed with CNV by the AI model, while an additional two images out of the 328 images (2.1%) not diagnosed with CNV according to the rule-based RS were diagnosed with CNV by the AI model. Overall, if the AI model detects that CNV is present, there is an 85.4% (72.2% to 93.9%) probability that CNV is truly present. If the AI model detects no CNV, there is a 99.4% (97.8% to 99.9%) probability that CNV is truly not present (according to the rule-based RS) (*Table 23*).

Number of wrong diagnoses if human assessors were replaced with artificial intelligence

In the participants included in the main cRCT (*n* = 239), a total of 176 diagnoses were made by the human assessors. Nineteen (10.8%) of these diagnoses were false positives. An additional 26 eyes (42.6% of those not receiving a diagnosis) should have been diagnosed according to the clinical RS (false negatives). If the AI model replaced the human assessors, 153 diagnoses would have been made, of which 18 (11.8%) would have been false positives. Nineteen eyes (40.0% of those not receiving a diagnosis by the AI) should have received a diagnosis according to the clinical RS (false negatives) and an additional 37 images could not be classified by the AI model but had a retinal diagnosis by the clinical RS. These combined 56 false-negative diagnoses mean the AI model has a combined false-negative rate of 65.1% (*Table 24*).

The false-positive rate for retinal diagnosis, defined as a retinal diagnosis provided by Octane that is different to the retinal diagnosis provided by human assessors, is comparable between the AI model and human assessors across both study arms of the cRCT on the basis of CIs. The false-negative rate, defined as no diagnosis (normal eye) or unclassified diagnosis by Octane, when human assessors have provided a specific diagnosis, is higher for Octane compared to human assessors in both study arms of the cRCT (*Table 24*). Accounting for unclassified images does increase the false-negative rate of the AI model but is not enough to create significant differences when comparing CIs.

Artificial intelligence–infrastructure-related outcomes (seconds and minutes per optical coherence tomography volume)

In its current configuration, Octane inference time per OCT volume is approximately 40 seconds on average, though this may vary depending on the deployment environment and OCT type.

End-to-end processing time from the receipt of OCT scan to referral recommendation is, on average, 8 minutes per OCT volume.

Observational pragmatic substudy

The regional commissioned teleophthalmology referral pathway in Manchester encountered challenges during deployment, resulting in limited uptake and use. Data were collected from a total of three optometry sites that were recruited in that catchment area. Of these, one practice did not contribute to recruitment.

Seventeen patients from two community optometry practices were recruited overall from the Manchester real-life teleophthalmology pathway. Of these, according to the clinical RS, four patients required an urgent referral, 12 required a routine referral and one required no referral. Interestingly, 10 patients were referred urgently by the community optometrists for suspicion of neovascular AMD, four were referred routinely and three were not referred. Six out of the 10 urgent referrals by community optometry were downgraded to routine referrals by the hospital-based clinicians, who reviewed the referrals remotely. Two patients were referred with suspicion of nAMD, three with MO (which does not warrant an urgent referral, other than for patient-specific reasons at the clinician's discretion) and one with retinal haemorrhages and roth spots (indicating a likely non-urgent systemic origin). One suspect nAMD case was diagnosed as CSCR by the hospital-based clinicians, and one as MO, both requiring routine referrals. The small number of recruited patients would not allow providing sufficient insight into the safety and efficiency of a real-life teleophthalmology setting, in the iteration implemented in Manchester.

However, as indicated above, the collected data were still valuable and used in the AI substudy.

Although HERMES was not designed to perform a formal qualitative methods analysis to identify barriers to sustainable deployment at the scale of the Manchester teleophthalmology pathway, the reported technical and operational challenges included:

Insufficient training and support of community optometry practitioners in the use of the teleophthalmology referral system.

TABLE 24 Number (proportion) of wrong diagnosis if human assessors were replaced with AI

			Total number of diagnoses made by…		Number of correct diagnoses made by…		Number of wrong diagnoses made by … [b]	
	Number of images	Number of true diagnoses[a]	Human assessors	AI algorithm[c]	Human assessors	AI algorithm	Human assessors	AI algorithm
All patients	371	283	–	226 + 60	–	201	–	32.1% (27.4 to 37.1%)
cRCT[d]	239	183	176	153 + 37	157	135	18.8% (14.1 to 24.4%)	31.0% (25.2 to 37.2%)
Control arm	81	58	52	52 + 16	45	46	24.7% (15.8 to 35.5%)	32.1% (22.2 to 43.4%)
Intervention arm	158	125	124	101 + 21	112	89	15.8% (10.5 to 22.5%)	30.4% (23.3 to 38.2%)

a Number of true diagnoses as determined by the clinical RS.
b Proportion of false-positive diagnoses out of the total number of diagnoses made by either human assessor or AI algorithm (as applicable) with 95% CI.
c The number of diagnoses made by the AI algorithm is presented as the number of diagnoses (larger number) plus the number of unclassified images (smaller number).
d Only patients included in the main cRCT study (after AI study exclusions).

Lack of training in the management of diverse and complex image (OCT) files by the community optometrists, to allow transfer of full-volume OCT scans. This is a key technical requirement for a safe and efficient teleophthalmology triaging pathway.

Limited communication between community optometry and HES, leading to occasional confusion. Clear lines of communication, support and escalation were not established by the technology supplier and stakeholders.

Hesitation of community optometry practices in using the teleophthalmology pathway, even when available to them, because of the above listed challenges.

Chapter 4 Economic evaluation

Introduction

An economic evaluation is described as a comparative analysis of two or more alternatives in terms of their costs and consequences.[41] For the HERMES trial, an evaluation comparing the outcomes in terms of the costs and consequences of the HERMES teleophthalmology pathway and the current standard care pathway was carried out. The analysis was done from the NHS perspective; however, to account for the fact that national commissioned schemes for the reimbursement of OCT scanning in the community are not established at present, and in order to account for this additional cost pressure to patients, we assumed a higher sight test fee in the HERMES pathway for the additional provision of the OCT, which was based on the fees that are paid for diagnostic services in the Scottish general ophthalmic services and the Welsh Primary Eye care Acute Referral Scheme (PEARs) systems.

The economic analysis for the HERMES Studies had two workstreams; the first was an economic evaluation decision model comparing the HERMES teleophthalmology pathway with standard care based in community optometry. The second workstream utilised a DCE to assess the preferences of the general public to understand what facets of the diagnostic pathway were the most important to them as a potential user of the service. The results of the DCE workstream were then incorporated into the economic evaluation decision model to produce a CBA, where both the costs and benefits were expressed in monetary terms.

Economic evaluation decision model

The model that was designed to assess the costs and benefits took the form of a decision tree model, designed in Microsoft Excel® (Microsoft Corporation Microsoft Excel, 2018) (Microsoft Corporation, Redmond, WA, USA). The model took the form of a cost-effectiveness analysis (CEA) and was conducted from the perspective of the healthcare service. The time horizon was until referral to secondary care for care of any suspected eye disease identified.

Objective

The purpose of this economic evaluation decision model was to compare the cost and benefits of the HERMES pathway compared with the standard optometry-based pathway in the correct identification of retinal disease.

Clinical pathway

In the standard care referral pathway, COs referred their participants for suspicious retinal disease to a HES via posting or e-mailing the referral letter to the GP or instructing the patient to give the referral to their GP. Electronic referral pathways are also used, where referrals are directly sent from CO to a HES. Participants who are urgently or routinely referred are expected to be examined within 2 weeks and 3 months, respectively.

In the HERMES teleophthalmology pathway, the clinical data and imaging (OCT scans and CFP) were uploaded to a referral platform. A hospital-based eye specialist reviewed the referral and provided a management decision to the CO and patient within 48 hours. If required, the corresponding NHS site arranged an appointment.

Model structure and assumptions

The structure for the economic decision took the form of a CEA in the form of the decision tree. Each arm of the decision tree represented one of the care pathways that could be utilised to correctly diagnose the number of correctly identified cases of eye disease in the pathway. The model structure is shown in *Figure 3*.

The model was subject to the following key simplifications:

- Due to the heterogeneity of the different diseases which are considered for inclusion in the HERMES cluster trial, the individual eye diseases are not included.

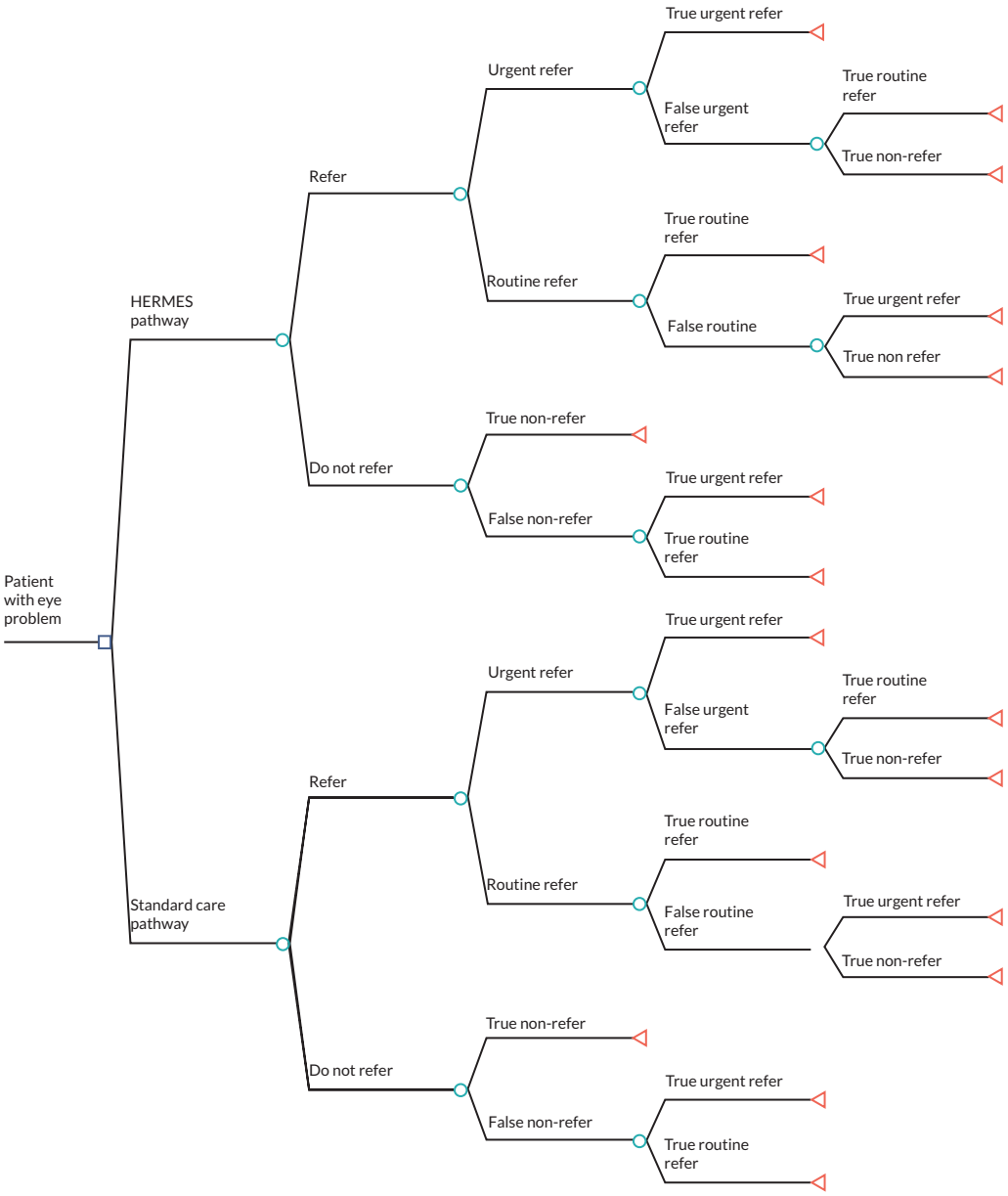

FIGURE 3 Model structure for the comparing HERMES diagnostic pathway.

- Until the point of diagnosis, all the referrals have the same pathway.
- All examinations that are included in the economic decision model are costed as NHS-funded eye examinations, although some of the participants in the standard care pathway may have paid for their eye examination out of pocket.

Costing data

Cost data were derived from published sources, and costs were presented in UK GBP (2022). The hospital outpatient costs were derived from NHS reference costs 2021–2.[42] COs in the standard care pathway were costed using the current costs of an NHS eye examination which was derived from the 2023 reimbursement rate for an NHS sight test fee.[43] The costs for the optometrist in the HERMES teleophthalmology pathway was estimated based on reimbursement fees paid to other optometrists who are working in shared care schemes in other parts of the UK, especially the fees for the Scottish supplementary eye examination and the Welsh PEARS.[44,45]

In addition to the sight test fee, the cost of the HERMES pathway included the cost of staff time needed for analysing and interpreting the information that was uploaded on to the HERMES platform. This information was provided by the expert opinion of the trial team who designed the HERMES pathway. At present, the software required to run the platform does not have an associated licensing cost. Therefore, this was not included in the base-case analysis. However, a per-patient software cost was examined as part of the deterministic sensitivity analysis as described below.

Effectiveness data

The effectiveness data were derived from the results of the randomised trial, specifically the number of individuals in each arm who had specific referral outcomes. From these data, the probabilities used to populate the economic evaluation decision model were derived. The probabilities for a number of different referral outcomes are detailed in *Table 25*.

The key referral events from the trial were summarised and probabilities were derived as to the likelihood of the different referral outcomes to be used in the economic decision model. The primary model outcome was the incremental cost per case correctly referred (i.e. the difference in cost between arms divided by the difference in the number of true positives and true negatives between arms).

TABLE 25 Summary of probabilities of referral outcomes used in decision model

Probability	Description
p_1 – probability of referral	The likelihood that the patient is referred to HES at any urgency
$(1-p_1)$ – probability of non-referral	The likelihood that the patient is not referred to HES at any urgency
p_2 – probability of urgent referral	The likelihood that the patient is referred to HES as for an urgent appointment
$(1-p_2)$ – probability of routine referral	The likelihood that the patient is referred to HES as for a routine appointment
p_3 – probability of true-positive urgent referral	The likelihood those who are referred for urgent appointments have been referred at the correct urgency
$(1-p_3)$ – probability of false-positive urgent referral	The likelihood those who are referred for urgent appointments have not been referred at the correct urgency
p_4 – probability of false-positive urgent referral being true routine referral	The likelihood of those who were incorrectly referred for urgent appointments should have been referred for a routine appointment
$(1-p_4)$ – probability of false-positive urgent referral being true non-referral	The likelihood of those who were incorrectly referred for urgent appointments should not have been referred at all
p_5 – probability of true-positive routine referral	The likelihood those who are referred for routine appointments have been referred at the correct urgency
$(1-p_5)$ – probability of false-positive routine referral	The likelihood those who are referred for routine appointments have not been referred at the correct urgency
p_6 – probability of false-positive routine referral being true urgent referral	The likelihood of those who were incorrectly referred for routine appointments should have been referred for an urgent appointment
$(1-p_6)$ – probability of false-positive routine referral being true non-referral	The likelihood of those who were incorrectly referred for routine appointments should not have been referred at all
p_7 – probability of true non-referral	The likelihood that the patient is correctly not referred to HES
$(1-p_7)$ – probability of false non-referral	The likelihood that the patient is incorrectly not referred to HES
p_8 – probability of false non-referral being true urgent referral	The likelihood of those who were not referred to HES should have been referred urgently
$(1-p_8)$ – probability of false non-referral being true routine referral	The likelihood of those who were not referred to HES should have been referred routinely

Base-case analysis

The base-case analysis was carried out from the NHS perspective, with a time horizon that covered the period from the patient presenting at the optometry service until diagnosis. As this time horizon was < 1 year, no discounting was applied.[46] The costs were represented by the outpatient appointments, the costs of the sight test, the costs of the HERMES pathway (in the interventional arm) and the costs of the outpatient appointments needed to diagnose the patient to begin their management. In the case of false-positive urgent referrals that should have been referred routinely to clinic, the cost of two outpatient appointments was included: one to represent the initial emergency referral appointment, and one to refer them into the appropriate routine outpatient clinic.

Sensitivity analysis

One-way sensitivity analysis (OWSA) was carried out to assess the impact of individual parameters on the conclusions of the economic decision model. In this instance, the values for costs were increased and decreased by 20% to create an upper and lower bound. Other parameters were varied using pragmatic assumptions, for example, the time that each consultant reviewed the images was halved and then doubled to represent cases of differing complexity. The potential of software licencing costs (if e.g. the software was run commercially) was varied between no costs and a cost of £10 per grading as an upper bound. These variations in the parameters were then used to assess the impact of change on conclusions. The list of parameters that were varied to assess the impact on the results are listed in *Table 26*.

The results of the OWSA were reported in the form of a tornado diagram. A tornado diagram shows which of these parameters had the greatest impact on the overall results. As we are estimating two different incremental cost-effectiveness ratios (ICERs), the tornado diagrams were produced, one for each ICER measure. The far left and right values of the horizontal bars indicate the ICER values associated with the lower and upper values for each parameter according to the legend colour-coding.

In addition to the deterministic OWSA, a probabilistic sensitivity analysis (PSA) was also carried out. Where an OWSA assessed the impact on the cost-effectiveness of individual parameters when changed one at a time, a PSA assessed the impact on estimates of incremental effect and incremental cost of uncertainty in all parameters when varied simultaneously. Uncertainty surrounding each parameter was characterised by defining a distribution for that parameter. This was represented by a beta distribution around the probabilities for referral outcomes and the gamma distributions around costs.

Monte Carlo simulation was used to draw an estimate for each parameter and estimate the incremental costs and incremental effects. This process was repeated for 1000 times and was used to present costs and effect plots.

TABLE 26 Variable assessed in deterministic sensitivity analysis

Variables in deterministic sensitivity analysis
Consultant ophthalmology appointment (hospital cost)
Community optometrist fee
Time spent on analysing HERMES information by optometrist
The application of a software licence cost
Time spent on analysing HERMES information by consultant ophthalmologist
Probability of true-positive emergency referral – standard practice
Probability of true-positive emergency referral – HERMES

The results for the PSA are displayed on a cost-effectiveness plane, and the costs and effects were plotted and calculated from each individual PSA distribution. The cost-effectiveness plane shows the spread of likely costs and effects and how they are linked together. The north-east quadrant represents more cost and greater effect, the north-west represents greater cost and less effect (where the treatment would be dominated), the south-west quadrant represents less effect at less cost and the south-east quadrant represents less and costs at greater effect (new treatment dominates).

Changes from protocol

The economic evaluation methods in the protocol (version 1.6)[17] was described as a within-trial CCA. However, due to the heterogeneity of service provision that was introduced by the pragmatic substudy, it was decided to adopt a modelling-based approach on provision that the telehealth service would be adopted. This enabled modelled estimates of cost-effectiveness and cost consequences to be estimated using the effectiveness data from the trial and resource and costs data from published sources.

The results of the economic decision model are based on the cRCT results, which focus on the telehealth intervention, which was separate from the diagnostic accuracy study. As such, the costs and the benefits of the AI diagnostic accuracy study are not included at this stage. The CCA (which can be thought of as a disaggregated CBA) results are integrated into the CBA. In the CBA, the outcomes that were to be included in the CCA are reported (see *Tables 37* and *38*), but these are then valued and integrated into a single metric, which we would argue is easier to interpret.

Results

Costs

The unit costs utilised for each arm are described in *Tables 27* and *28*. *Table 27* describes the sources of the unit costs that were used to cost each arm of the pathway. *Table 28* describes the costs attributed to the different referral outcomes. These costs came from the data detailed in *Table 27*.

The highest costs are associated with a referral that is wrongly classified as urgent but still requires a routine appointment. This represents the need for care, but the opportunity cost associated with an inappropriate use of the urgent care appointment.

TABLE 27 Summary of intervention costs

Intervention costs							
Standard care				**HERMES**			
Item	**Unit cost (£)**	**Notes**	**Source**	**Item**	**Unit cost (£)**	**Notes**	**Source**
Optometrist fee	23.14	Optometry sight test	Department of Health[43]	Optometrist fee	39.14	Based on cost for supplementary examination for new symptoms. The PEARS band 2 exam has a similar value	Scottish Government[45]
Ophthalmology outpatient costs	147	Consultant-led ophthalmology appointment	NHS cost collection 2021–2[42]	HERMES intervention	14.37	Based on 10 minutes of optometrist analysis time and 2 minutes of consultant time	Expertise from the trial team and salary costs from PSSRU costs[47]
				Ophthalmology outpatient costs	147	Consultant-led ophthalmology appointment	NHS cost collection 2021–2[42]

PSSRU, Personal Social Services Research Unit.

TABLE 28 Cost of each referral outcome

Referral outcome	Standard care (£)	HERMES (£)
Correct no referral	23.14	39.14
Correct urgent referral	170.14	200.51
Incorrect urgent referral, should have been routinely referred	317.14	347.51
Incorrect urgent referral, should not have been referred	170.14	200.51
Correct routine referral	170.14	200.51
Incorrect routine referral, should have been urgently referred	170.14	200.51
Incorrect routine referral, should not have been referred	170.14	200.51

Effectiveness results

The results of the cRCT were used to derive probabilities for the likelihood of the different referral events in each arm. The results are summarised in *Table 29*. The key difference between the HERMES and standard care pathways is the difference in the likelihood of false positives due to the higher likelihood of correct urgent referrals in the HERMES pathway compared with the standard care pathway.

Cost-effectiveness results

The results of the base-case cost-effectiveness are shown in *Table 30*.

The expected total cost for the HERMES pathway is £167 compared with £183 for the standard care pathway, resulting in a lower cost of £16 per case for the HERMES pathway. The rate of correct diagnosis is also greater for the HERMES pathway compared with the standard care pathway, giving the HERMES pathway a 0.25 greater incremental correct rate of correct diagnosis. As the HERMES pathway has a greater effect at a slightly lower cost, this means it is, on average, the dominant pathway in the base case.

One-way sensitivity analysis

To assess the robustness of the conclusion of the economic decision model, an OWSA was carried out, varying specific parameters to assess their impact on the study conclusions, and the factors that were varied were summarised in *Table 31* and the results are shown in *Table 31* and *Figure 4*.

As described in the *Methods* section, the parameters were varied by a set amount to create an upper and lower bound value, and these values were changed to assess their impact on the ICER. None of the variation of these parameters had a significant impact on the ICER, and subsequently, the conclusions. In each scenario, the HERMES pathway continues to be less costly and more effective than the standard care pathway.

Probabilistic sensitivity analysis

The results of the PSA are shown in *Figure 5*.

The majority of the points on the cost-effectiveness plane (where each point represents a single iteration of the Monte Carlo simulation) are associated with less cost and greater effect, with this being the case in 98% of the iterations in this sample. Both the deterministic and PSA demonstrate a level of certainty as the cost-effectiveness of the HERMES pathway in the correct diagnosis of the urgency of retinal pathology in this context.

Discrete choice experiment

In addition to the economic decision model that was undertaken to assess the cost-effectiveness of the HERMES pathway, a DCE was conducted to understand the preferences of members of the public about what is important in

TABLE 29 The probabilities of different referral events in each arm

Standard care				
Probabilities	**Number of events (*n*/*N*)**	**Value**	**Lower bound**	**Upper bound**
p_1 – probability of referral	125/136	0.92	0.86	0.96
(1 – p_1) – probability of non -referral	11/136	0.08	0.04	0.14
p_2 – probability of emergency referral	38/125	0.30	0.22	0.39
(1 – p_2) – probability of routine referral	87/125	0.70	0.61	0.78
p_3 – probability of true-positive emergency referral	14/38	0.37	0.22	0.54
(1 – p_3) – probability of false-positive emergency referral	24/38	0.63	0.46	0.78
p_4 – probability of false-positive emergency referral being true routine referral	23/24	0.96	0.79	1.00
(1 – p_4) – probability of false-positive emergency referral being true non-referral	1/24	0.04	0.00	0.21
p_5 – probability of true-positive routine referral	75/87	0.86	0.77	0.93
(1 – p_5) – probability of false-positive routine referral	12/87	0.14	0.07	0.23
p_6 – probability of false-positive routine referral being true emergency referral	3/12	0.25	0.05	0.57
(1 – p_6) – probability of false-positive routine referral being true non referral	9/12	0.75	0.43	0.95
p_7 – probability of true non-referral	7/11	0.64	0.31	0.89
(1 – p_7) – probability of false non-referral	4/11	0.36	0.11	0.69
p_8 – probability of false non-referral being true emergency referral	1/4	0.25	0.01	0.81
(1 – p_8) – probability of false non-referral being true routine referral	3/4	0.75	0.19	0.99
HERMES				
Probabilities	**Number of events (*n*/*N*)**	**Value**	**Lower bound**	**Upper bound**
p_1 – probability of referral	124/158	0.78	0.71	0.85
(1 – p_1) – probability of non -referral	34/158	0.22	0.15	0.29
p_2 – probability of emergency referral	27/124	0.22	0.15	0.30
(1 – p_2) – probability of routine referral	97/124	0.78	0.70	0.85
p_3 – probability of true-positive emergency referral	26/27	0.96	0.81	1.00
(1 – p_3) – probability of false-positive emergency referral	1/27	0.04	0.00	0.19

continued

TABLE 29 The probabilities of different referral events in each arm (*continued*)

HERMES				
Probabilities	**Number of events (*n*/*N*)**	**Value**	**Lower bound**	**Upper bound**
p_4 – probability of false-positive emergency referral being true routine referral	1/1	1.00	.	.
(1 – p_4) – probability of false-positive emergency referral being true non-referral	0/1	0.00	.	.
p_5 – probability of true-positive routine referral	94/97	0.97	0.91	0.99
(1 – p_5) – probability of false-positive routine referral	3/97	0.03	0.01	0.09
p_6 – probability of false-positive routine referral being true emergency referral	1/3	0.33	0.01	0.91
(1 – p_6) – probability of false-positive routine referral being true non-referral	2/3	0.67	0.09	0.99
p_7 – probability of true non-referral	31/34	0.91	0.76	0.98
(1 – p_7) – probability of false non-referral	3/34	0.09	0.02	0.24
p_8 – probability of false non-referral being true emergency referral	0/3	0.00	.	.
(1 – p_8) – probability of false non-referral being true routine referral	3/3	1.00	.	.

TABLE 30 The results of the base-case analysis comparing the HERMES pathway with the standard care pathway (monetary values rounded to nearest £)

Strategy	Expected total cost (£)	Incremental costs (£)	Correct diagnosis	Incremental correct diagnosis	ICER
HERMES	167	−16	0.96	0.25	Dominant
Standard care	183		0.71		

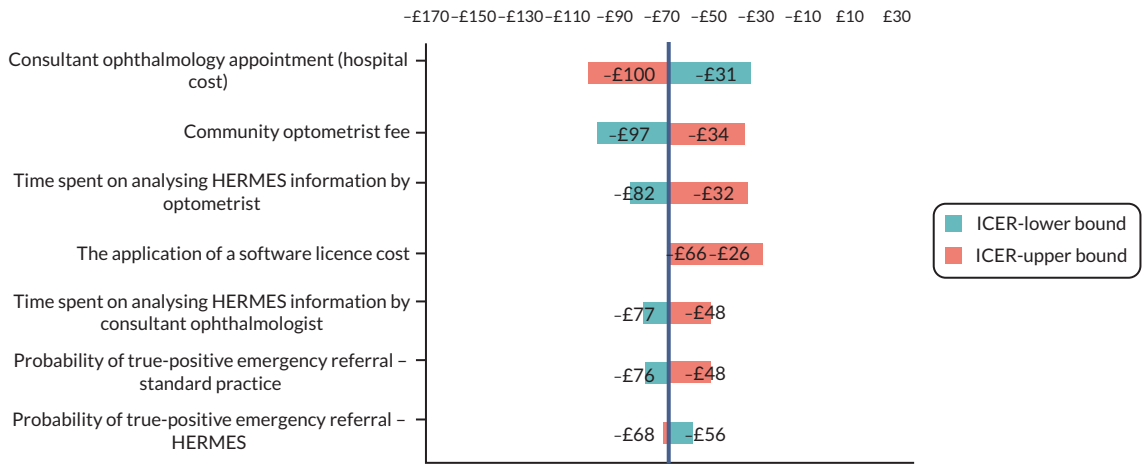

FIGURE 4 Tornado diagram displaying the results of the OWSA for the CEA (monetary values rounded to nearest £).

such a service. As the population grows older, the likelihood of developing a problem with their eyes, particularly in later life, is set to increase.[48] As such, everyone's views on such a service are valuable.

A DCE is an attribute-based survey method for measuring benefits in which participants choose between hypothetical scenarios (choice sets) with two or more alternatives.[49] In this instance, different levels and attributes were used to understand the level of preference that individuals have towards the various facets of the diagnostic pathways being compared in the HERMES study. One specific attribute that was included in the DCE was a cost attribute to measure how the cost of the diagnostic pathway to the healthcare system affects an individual's preference. The results of the DCE were incorporated into the economic evaluation model to produce a CBA.

Objectives

The aims of the DCE component were as follows:

- to understand the strength and direction of preference for the different potential outcomes in the HERMES and standard care pathways
- to derive a willingness-to-pay value (WTP) from the DCE.

Levels and attributes

The levels and attributes of the DCE were derived from the key features of the diagnostic pathway and the comparators. They were designed in conjunction with clinical expertise from those involved with the design of the diagnostic pathway. The key elements that were considered to be impacted by the telehealth system were:

- the likelihood of a problem not being identified when it presents to a healthcare practitioner (false negative)
- the likelihood of an unnecessary hospital visit for someone who does not need further care (false positive)
- the waiting times for both an urgent and non-urgent referral to a hospital eye specialist.

This resulted in the derivation of five attributes (*Table 32*), the levels for each attribute are summarised in *Table 33*.

In the DCE survey, potential respondents were presented with varying combinations of the attributes and levels, which are varied and presented as a series of choices. Potential respondents were presented with varying combinations of the attributes and levels, as shown in *Table 33*. As there were three attributes with five levels and two attributes with three levels, there were 1125 possible scenarios (the full factorial design) that could be presented to participants. A full factorial design was not possible, as this involves more scenarios that participants could reasonably be asked to answer. Therefore, the DCE used an efficient design to estimate the parameters while keeping the standard errors as low as possible, specifically a d-efficient design was used. This design seeks to minimise the d-error (a measure that describes

TABLE 31 One-way sensitivity analysis results comparing the HERMES pathway with the standard care pathway (monetary values rounded to nearest £)

No.	Variable	ICER-lower bound (£)	ICER-upper bound (£)	Range (£)
1	Consultant ophthalmology appointment (hospital cost)	−31	−100	68.98
3	Community optometrist fee	−97	−34	62.68
4	Time spent on analysing HERMES information by optometrist	−82	−32	50.27
2	The application of a software licence cost	−66	−26	40.03
5	Time spent on analysing HERMES information by consultant ophthalmologist	−77	−48	29.06
6	Probability of true-positive emergency referral – standard practice	−76	−48	28.71
7	Probability of true-positive emergency referral – HERMES	−56	−68	11.45

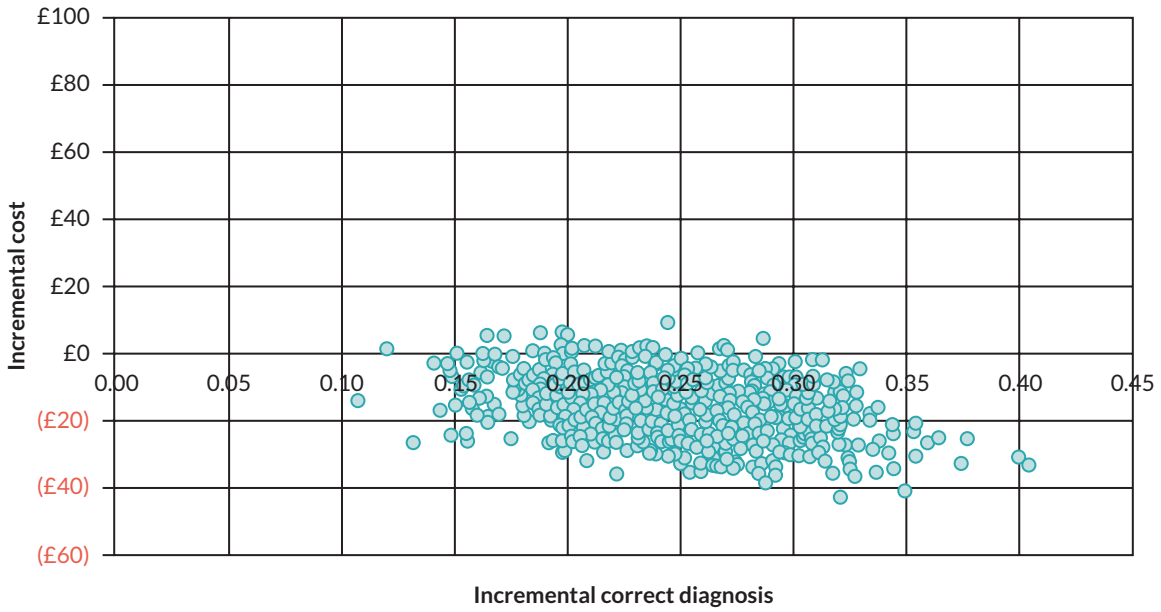

FIGURE 5 Displays the cost-effectiveness plane for HERMES pathway compared with the standard care pathway.

TABLE 32 Description of the DCE attributes

Characteristic	Summary
Number of people who are sent to the hospital eye department unnecessarily	This describes the number of people who were sent for tests at the hospital eye department but who did not need to be sent there as they could have been managed by their optician, GP or pharmacist
Number of people who should have been sent to the hospital eye department but were not	This describes how many people had an eye disease that needed urgent treatment that was not identified. For these people, the delay in treatment for their eye disease could result in their disease worsening before any treatment is administered
Time between being referred for an urgent appointment at the hospital and getting that urgent appointment	This is the time that occurs between being referred by the service for an *urgent* consultation appointment at the hospital eye department until being seen for that urgent appointment
Time between being referred for a non-urgent appointment at hospital and getting that non-urgent appointment	This is the time that occurs between being referred by the service for a *non-urgent* consultation appointment at the hospital eye department until being seen for that non-urgent appointment
Cost to you	The one-off cost paid to attend the service. We know that you do not have to pay for treatment in the NHS but are asking you to imagine a hypothetical situation where you do pay an out-of-pocket charge

how well a design can extract information) and the design with the lowest d-error is chosen.[50–52] The DCE questionnaire was designed in the NGene (ChoiceMetrics, Sydney, NSW, Australia) software program.[50]

A survey was then designed incorporating an introduction and explanation, some initial screening questions, the DCE task and some demographic information. The survey was structured to present each participant with nine choices. There were 12 potential blocks of choices that the participant could have been presented with. These blocks were chosen to make sure that the number of choice tasks could be a reasonable amount for a participant to do while covering all the necessary choices from the d-efficient design.

One of the nine choices presented to each participant was included to explore that participants were engaging with the questions set out in the DCE. This took the form of a dominance question, where the respondent was presented with the very best outcomes for every attribute compared to the very worst. This checks that the participants understand the

TABLE 33 Levels and attributes for the DCE

Attribute	Definition	Levels
Number of people who are sent to hospital eye department unnecessarily	This describes the number of people cared for under this service who were sent for tests at the hospital eye department but who did not need to be sent there as they could have been managed by their optician, GP or pharmacist	1 person out of every 100
		5 people out of every 100
		10 people out of every 100
		25 people out of every 100
		50 people out of every 100
Number of people who should have been sent to the hospital eye department but were not	This describes how many people had eye disease that needed urgent treatment from the hospital eye department, which was not identified, and they were told they did not need further treatment. For these people, the delay in treatment for their eye disease could result in their disease worsening before any treatment is administered	1 person out of every 100
		5 people out of every 100
		10 people out of every 100
		15 people out of every 100
		20 people out of every 100
Time between being referred for urgent treatment and getting that urgent treatment	The time that occurs between being examined and referred by an eye care service for *urgent* treatment until being seen at the hospital eye department to get that urgent treatment	2 weeks
		4 weeks
		6 weeks
Time between being for non-urgent treatment and getting that non-urgent treatment	The time that occurs between being examined and referred by an eye care service for *non-urgent* treatment until being seen at the hospital eye department to get that urgent treatment	8 weeks
		10 weeks
		12 weeks
Price to access the service	The one-off cost paid to attend the service	£0
		£20
		£30
		£60
		£150

task, as there would be no reason to select the choice set containing the worse outcomes over the choice set containing the better ones.

Carrying out the survey

After the pilot phase of the study, the main data collection phase began. The questionnaire survey was carried out by the survey company Dynata.[53] The survey company carried out internal checks to ensure data quality (such as question completion time). Participants were compensated for their time by points earned for participation, which can be exchanged for rewards.

As noted above, the survey began by asking participants basic demographic questions, including age, gender and geographic location. The participants were then given the background and context of the survey and were asked to complete the DCE task. Finally, at the end of the survey, further demographic information was collected, including ethnicity, educational attainment, caring responsibilities and experience with eye disease. The participants were also asked about their approximate household income. To do this, they were asked to select from brackets ranging from < £10,000 to > £150,000. This is key for assessing the results of the cost attribute to understand the spread of income for those taking part in deriving the WTP values.

Inclusion criteria and sample

The inclusion criteria for participants of the study included being over 18 years of age and able to give informed consent to participate in the study. The sampling strategy was purposive to include a representative sample of the English population with regards to age, gender and ethnicity.

A sample of approximately 700 participants was sought to address the research aims for this survey. The sample size for DCEs was based on the number of true choice tasks (which is not something that can be known prior to undertaking the research task). Some analysts undertake a rule-of-thumb approach based on attributes and levels,[54] but, in this instance, a sample of 700 was considered in order to ensure that sample was as large as practically possible, which has been utilised effectively in other DCE studies.[55]

Piloting

Piloting the survey took part in two phases; the first phase piloted the survey within the Newcastle University Population Health Science Institute. This included academic and non-academic members of staff, with a view to checking whether the language was clear and the task was understandable. This led to the development of *Table 32*, with the explanation of the attribute to the participant.

The second phase of piloting took place when the survey was transferred to the online platform where it would be administered to participants. This second stage of piloting used approximately 10% of the targeted survey participants. This element of piloting was followed by an interim analysis that allowed the data to be checked for completeness. If there proved to be any problems with the data completion or other data issues, then the survey could be retracted temporarily from the platform to make any necessary amendments.

Data analysis and interpretation

The first stage of data analysis was to produce descriptive statistics for the demographic data. This included gender, age, geographical region, whether they had caring responsibilities or previous eye problems and household income.

The second stage of analysis was to include the respondent's preferences for each of the different attributes using a mixed logit model (MXL).[56] Different choice models were obtained from different assumptions about the distribution of the random terms. The final model selection was determined based on model fit, as measured by several measures, including the log-likelihood ratio, McFadden's pseudo-R^2, the Akaike information criterion (AIC) and the Bayesian information criterion (BIC) (where the lower the AIC and BIC, the better the model fit).

Marginal rates of substitution were then calculated to assess the trade-offs between the different attributes. Marginal willingness-to-pay (MWTP) was used to estimate the different WTP values between the attributes.[49] All data analyses took place using the Stata software program (StataCorp, Stata Statistical Software. 2019.

Ethics and dissemination

Ethical approval was sought for this study from the Newcastle University Ethics Committee. The project has received ethical approval from the Ethics Committee at Newcastle University reference number 36993/2023.

All data that were received from the survey company were completely anonymised, and participants were only known by an identification number. All data received were kept in a secure, password-protected university server. Compliance with GDPR was maintained throughout the project.[55]

Results

In total, 711 participants took part in the survey. The initial pilot study included 81 participants that demonstrated good levels of completion and the survey continued to the full sample. The results of the demographic information are shown in *Table 34*.

TABLE 34 Demographic details of the DCE sample

Variable	*N*	%
Gender		
Male	341	48
Female	367	52
Prefer not to say	3	0
Age (years)		
18–24	73	10
25–34	118	17
35–44	115	16
45–54	119	17
55–64	112	16
65 +	95	13
Region		
East Midlands	62	9
East of England	79	11
London	109	15
North East	33	5
North West	93	13
South East	117	16
South West	74	10
West Midlands	74	10
Yorkshire and the Humber	70	10
Ethnicity		
White	596	84
Mixed	18	3

TABLE 34 Demographic details of the DCE sample (*continued*)

Variable	*N*	%
Asian	61	9
Black	27	4
Arab/other	5	1
Prefer not to say	4	1
Caring responsibilities		
Yes	182	26
No	523	74
Prefer not to say	6	1
Education		
None	2	0
Secondary school	151	21
College/sixth form	199	28
University	339	48
Other	18	3
Prefer not to say	2	0
Household income		
< £10k	38	5
£10–20k	91	13
£20–30k	133	19
£30–40k	115	16
£40–50k	72	10
£50–60k	50	7
£60–70k	62	9
£70–80k	23	3
£80–90k	24	3
£90–100k	19	3
£100–150k	33	5
> £150k	9	1
Prefer not to say	42	6
Previous eye problem		
Yes	176	25
No	533	75
Prefer not to say	2	0
Device		
Desktop/laptop	598	84
Tablet	113	16

The sample was targeted to be representative of the English population with regards to age, gender and ethnicity. Most participants did not have a previous history of eye disease that required a visit to HES or have any caring responsibilities. The results of a MXL are described in *Table 35*.

The attribute-specific constant, described in *Table 35*, indicates that there was no bias for the left-hand side of the screen on which the questionnaire was presented (this is important for English readers who read from left to right), which is a good sign of the participant. In addition, the dominance question was answered correctly by 93% of the participants. This indicates that the participants were attentive and understood the task. The data were of sufficient quality after the initial 10% pilot phase that the study was deemed fit to continue.

The results from this MXL demonstrate that respondents prefer less false positives, less false negatives, less waiting times and less cost for the diagnostic pathway. The preference is stronger for the avoidance of a false negative compared to a false positive. The preference is stronger for less waiting time for an urgent referral than a non-urgent one. Given the nature of these attributes, this makes intuitive sense and indicates that the DCE is performing as expected.

The results from a further MXL 'WTP space' are shown in *Table 36*. In this model, the values that were chosen for the cost attribute can be used to derive a WTP per unit for the other attributes. This WTP value can be used as a measure of the strength of preferences for different ways that care can be provided.

The results of the WTP analysis from the DCE are described below:

- Respondents were willing to pay £16.22 for a 1% decrease in the false-positive rate.
- Respondents were willing to pay £55.17 for a 1% decrease in the false-negative rate.
- Respondents were willing to pay £1.95 for a 1-week decrease in the wait for urgent care.
- Respondents were willing to pay £1.04 for a 1-week decrease in the wait for non-urgent care.

The strongest preference demonstrated by the participants is for less false negatives, followed by less false positives. A 1% reduction in false-negative rates is considered to be over three times more important than a 1% reduction in the false-positive rate. As these data show, while participants also prefer shorter waiting times (more so for urgent referrals), very much stronger preferences were expressed for the reduction of false positives and false negatives.

TABLE 35 Participant preferences for the levels and attributes

Attribute	Coefficient	Standard error	95% CI (lower)	95% CI (upper)
False positive (%)	−1.673	0.133	−1.933	−1.413
False negative (%)	−5.114	0.353	−5.805	−4.423
Wait for urgent care (weeks)	−0.196	0.013	−0.223	−0.170
Wait for non-urgent care (weeks)	−0.088	0.013	−0.115	−0.062
Cost	−0.025	0.001	−0.028	−0.023
Attribute-specific constant	0.071	0.037	−0.002	0.144
Log-likelihood	−2868.389			
Observations	11,376			
Participants	711			

TABLE 36 Participant preferences for the levels and attributes, including WTP values

Attribute	Coefficient	Standard error	95% CI (lower)	95% CI (Upper)	WTP (£)	95% CI (lower)	95% CI (upper)
False positive (%)	−65.975	6.538	−79.035	−53.317	−16.22	−19.43	−13.11
False negative (%)	−224.419	18.148	−261.776	−187.811	−55.17	−64.37	−46.18
Wait for urgent care (weeks)	−7.921	0.711	−9.312	−6.453	−1.95	−2.29	−1.59
Wait for non-urgent care (weeks)	−4.234	0.629	−5.485	−3.031	−1.04	−1.35	−0.75
Cost	−4.067	0.054	−4.172	−3.959	N/A	N/A	N/A
Attribute-specific constant (Alt A)	2.730	1.753	−0.600	6.259	N/A	N/A	N/A
Log-likelihood	−2992.241						
Observations	11,376						
Participants	711						

Cost–benefit analysis

As noted above, the DCE elicited preferences from the general public as to the order and strength of their preference with regards to the different facets of the HERMES pathway. Using the cost attribute, these preferences have also been expressed in monetary terms. Using the valuation of the false-positive and false-negative outcomes, we can re-express the economic evaluation as a CBA. A CBA is a type of economic evaluation where both the costs and the benefits are expressed in the same, normally monetary terms.[57]

In the DCE, a WTP value for both a 1% decrease in the false-positive and false-negative rates was generated. If we assume transitivity with regards to the outcomes, then these WTP values also represent a corresponding 1% increase in the true-positive and true-negative outcomes. As such, the increase of the detection of true positives and true negatives can be used to help generate the net monetary benefit for diagnostic accuracy benefits of each pathway. The MWTP (i.e. the amount people are willing to pay for each measure of additional benefit) of the diagnostic accuracy outcomes is described in *Table 37*.

Marginal willingness to pay for the HERMES pathways compared with the standard care

Table 37 shows that, although there is a difference in the rates of the detection of true positives that results in net benefit for the HERMES pathway, the greatest amount of net benefit is driven by the higher rates of detection of true negatives in the HERMES pathway compared with the standard care pathway.

The DCE also included waiting times attributes, and the WTP for reductions in waiting time can also be estimated and included in the CBA. Data reported in *Table 9*, and reproduced in *Table 38*, report the waiting times associated with the HERMES and standard care pathways. As these values are reported in days, the values were divided by seven to express them in weeks to correspond with the DCE. The WTP values described in the DCE study are for being seen for a consultation 1 week sooner. To use these data in the CBA, the differences in waiting times between the two pathways were calculated. The HERMES pathway was approximately 3 weeks faster for urgent referrals and 1 week faster for routine referrals. Multiplying these values by their respective WTP values resulted in a small additional net benefit of £7 for the HERMES pathway (*Table 39*)

Considering both the incremental WTP for improved diagnostic accuracy and for reduced waiting associated with the HERMES pathway along with its small cost saving, the mean net benefit of the HERMES compared to the standard care pathway was £992. That is, for every patient seen by the HERMES compared to the standard care pathway, we gain £992.

To assess the robustness of the conclusion of the CBA, an extreme value sensitivity analysis was carried out. In this analysis, each of the probability parameters in the economics model were changed so that they took the plausible least favourable value for the HERMES pathway. This meant that, in the model, we used the lower bounds of the CIs associated with diagnostic performance and the upper bounds for the CIs for waiting time. The CIs associated with WTP values referenced in *Table 35* do not cross zero, meaning that changing these values would also not result in a negative net benefit for the HERMES pathway. This demonstrates that even when using these extreme values, the HERMES pathway still results in a positive net benefit, demonstrating a very high certainty in the results of the base-case analysis, which is presented in *Table 36*.

Discussion of economic evaluation

This economic evaluation aimed to assess the costs and benefits of the HERMES pathway compared with the standard optometry-based pathway. This was carried out in a number of ways, including a model-based CEA, a DCE and using the results of both for a CBA. In both analyses, the HERMES pathway has a lower mean cost and greater mean effect. The deterministic and probabilistic sensitivity CEA analyses provide strong evidence that this result holds. The results of the CEA are consistent with the CBA, which found very strong evidence (the results held even under extreme assumptions) favouring the HERMES pathway compared with the standard care pathway.

TABLE 37 Participant preferences for the levels and attributes, including WTP values for false-positive and false-negative rates (monetary values rounded to nearest £)

Strategy	WTP per 0.01 improvement of false-positive rate (£)	True-positive rate	WTP true positive (£)	Incremental WTP for greater true-positive rate (£)	WTP per 0.01 improvement of false-negative rate (£)	True-negative rate	WTP true negative (£)	Incremental WTP for greater true-positive rate (£)	Overall WTP (£)	Incremental WTP for correct diagnosis (£)
HERMES	16	0.76	1232	170	55	0.20	1082	798	2314	969[a]
Standard care		0.65	1061			0.05	284		1345	

a Note this amount is not fixed but varies as the relative mix of true positives and negatives in any given estimate of the correct diagnosis rate varies.

TABLE 38 Participant preferences for the levels and attributes including WTP values for waiting times (monetary values rounded to nearest £)

Strategy	WTP per 1-week improvement in urgent referral times (£)	Average waiting time urgent referrals (weeks)	Incremental reduction in urgent waiting times (weeks)	Incremental WTP faster urgent referral times (£)	WTP per 1-week improvement in routine referral times (£)	Average waiting time routine referrals (weeks)	Incremental reduction in routine waiting times (weeks)	Incremental WTP faster routine referral times (£)	Overall WTP for faster referral times (£)
HERMES	2	4.46	−3.15	6	1	14.63	−1.12	1	7
Standard care		7.61				15.74			

TABLE 39 Participant preferences for the levels and attributes including WTP values (monetary values rounded to nearest £)

Strategy	Expected total cost (£)	Incremental costs (£)	Incremental WTP for diagnostic accuracy	Incremental WTP for reduced waiting times	Overall net benefit for HERMES pathway
HERMES	167	−£16	969	7	992
Standard care	183				

Comparison with existing literature

There have been previous economic evaluations that have assessed telehealth interventions with the aim of population screening to identify eye disease.[58] However, there has been less evaluation of the use of telehealth interventions in areas such as referral refinement or management pathways for eye disease. This is the first study which evaluates the HERMES pathway, which aids in classifying the urgency of referrals from primary care. Although there has previously been a greater emphasis on studies for screening and disease identification, there have been some previous studies evaluating services to reduce the burden of unnecessary referrals to the HES in the UK. For example, a retrospective economic analysis of a minor eye condition scheme also found that patient volume could be lowered at a lower service cost.[59] The intervention does differ from the HERMES pathway, as it is based in the community and there is less of a focus on the use of telehealth, but it did find that referral volume could be lowered at a lower cost by utilising a referral refinement scheme. A further recent study also provided preliminary evidence of lower costs and better patient outcomes for a telehealth intervention in post-surgical cataract care,[60] meaning that telehealth can be a promising strategy to maximise the use of healthcare resources in ophthalmology.

Patient preferences about pathways

The DCE component of the study measured patient preference, specifically deriving a monetary value for the facets of the diagnostic pathways that are most important to them. Interestingly, the diagnostic accuracy of the pathway was much higher than the values related to waiting times for consultations, demonstrating that the avoidance of any missed pathology, followed by the avoidance of an unnecessary hospital visit are the factors that are valued highest by a general population sample. This adds to the knowledge base about what a large sample of individuals value about their eye care services.

Strengths and limitations

There are several strengths and limitations associated with the economic evaluation presented here. Strengths include its novelty in the use of a telehealth intervention in this context. In addition, the outcomes of the economic evaluation are expressed in different ways (a CEA and a CBA) and the results were consistent. Another strength of the economic evaluation was that factors that may be key to future implementation, such as software licensing or clinical analysis time, are included in deterministic sensitivity analysis and do not change the overall conclusion of the study.

The strengths of the DCE approach that was undertaken include a large representative sample of the English population to represent the potential recipients of the diagnostic pathway. The DCE also had good completion rates, and (given the results or the dominance and attribute constant for the left hand of the screen) it can be attributed that participants were attentive to the task. The results also made logical and intuitive sense, with no outcomes that would result in worse clinical performance or longer waiting times being preferred over better outcomes.

One of the key limitations of the economic decision model was the limited time horizon of the model, as the outcomes were only considered up to the point of diagnosis. Subsequent visual outcomes, impacts on health-related quality of life and longer-term costs were not considered. However, given the diagnostic finding and assuming effective and efficient treatments are used, then it would be implausible that the inclusion of the later effects could change the conclusions. However, future research could explore this further by considering a longer time horizon, including the specific health outcomes associated with the referral decisions.

Another limitation in the economic model is the lack of quantification of the impacts of false negatives in the model beyond the initial diagnostic phase. In terms of the conclusion that can be drawn from the CEA, the HERMES pathway is highly likely to both lower cost and have a higher efficacy than the standard care pathway. Furthermore, the numbers of false negatives across both arms of the trial were small (one urgent referral and three routine referrals in total for the 294 participants in the HERMES trial). Given the small number of events, the impact of the treatment costs of these individuals on mean costs would be expected to be small.

With respect to the DCE, a limitation could be a lack of specific preferences from those with a history of macular disease or with experience of caring for someone with macular disease. These individuals may have an interesting perspective on waiting times or efficacy compared to those of the general public. Future research could investigate the preferences of these individuals to understand how they compare to the larger sample of the general public.

Conclusion

The HERMES pathway is highly likely to be a cost-effective strategy to identify and prioritise urgent eye disease compared with the standard care pathway. It is highly likely to have a lower cost and greater effectiveness. The general public also expressed a strong preference for the improvement in clinical outcomes and lower waiting times that the HERMES pathway can provide.

Chapter 5 Human–computer interaction evaluation

This is a qualitative study to explore the acceptability, barriers and facilitators of the implementation of new teleophthalmology technologies between community optometry practices and HES.

Introduction

Primary eye care in the UK is mainly delivered by community optometry practices.[17,62] If participants are suspected of having a retinal condition, non-urgent referrals to hospital eye services (HES) are typically processed by their GP based on recommendations from the CO. Thus, optometrists are often not involved in making direct referrals using electronic referral platforms or are informed of outcomes. The additional step can reduce the specificity of clinical details included in the referral, as GPs are not specialists in eye care and rarely have the time or expertise to undertake eye examinations.[63] For several reasons, including concerns over capacity, changes in practice due to the COVID-19 pandemic and the high number of referrals to HES, there is a need for disruptive changes in the optometric referral pathways for suspected retinal conditions (SRC).[64]

Teleophthalmology can have several benefits in the context of triage. For example, in one scoping review, teleophthalmology was found to contribute to reducing face-to-face appointments with ophthalmologists by 16–48% through reducing inappropriate and unnecessary referrals.[65] Similarly, implementing remote retinal imaging-based referrals reduced the waiting time for participants to see an ophthalmologist from 14 weeks to 4 weeks.[65] Teleophthalmology has been found to improve elderly participants' access to specialist eye care and reduce workload on specialist centres and unnecessary visits.[66] Participants also reported high levels of satisfaction with teleophthalmology services due to reduced cost and time of travel as well as increased accessibility to services.[67] A recent systematic review has also emphasised the potential of teleophthalmology to serve as an alternative eye care delivery model by demonstrating its feasibility and cost-effectiveness for the management of various eye conditions in several countries, including the UK.[68] Additionally, in recent years, advances in AI, particularly in deep learning, hold great promise for expanding the use of teleophthalmology.[69–72] Deep learning can improve referrals by identifying participants who are more likely to develop a specific condition and require urgent care or frequent follow-ups, increasing participants' access to appropriate eye care.[70–72]

The non-urgent referral pathway with and without teleophthalmology is presented in *Figure 6*. A recent systematic review reported that teleophthalmology and digital referrals could reduce waiting time, costs and unnecessary referrals. It also noted that teleophthalmology could lead to earlier detection and diagnosis[73] and as such is an underutilised resource for HES. This review was based on reviewing referrals, not people's first-hand experiences. Therefore, there is a need to understand how implementation would affect users in practice.

The growing use of OCT scanning has increased the number of referrals to HES;[28,74] therefore, teleophthalmology has been suggested to reduce unnecessary referrals, manage growing capacity concerns and potentially manage increasing workloads by reviewing referrals before participants are seen in clinics.

Despite these promising findings, triaging referrals via teleophthalmology has been limited in practice. For example, during the COVID-19 pandemic, a period associated with increased adoption of telehealth applications,[75] primary care optometrists were less willing to adopt teleophthalmology in the context of referrals.[76] Although the study did

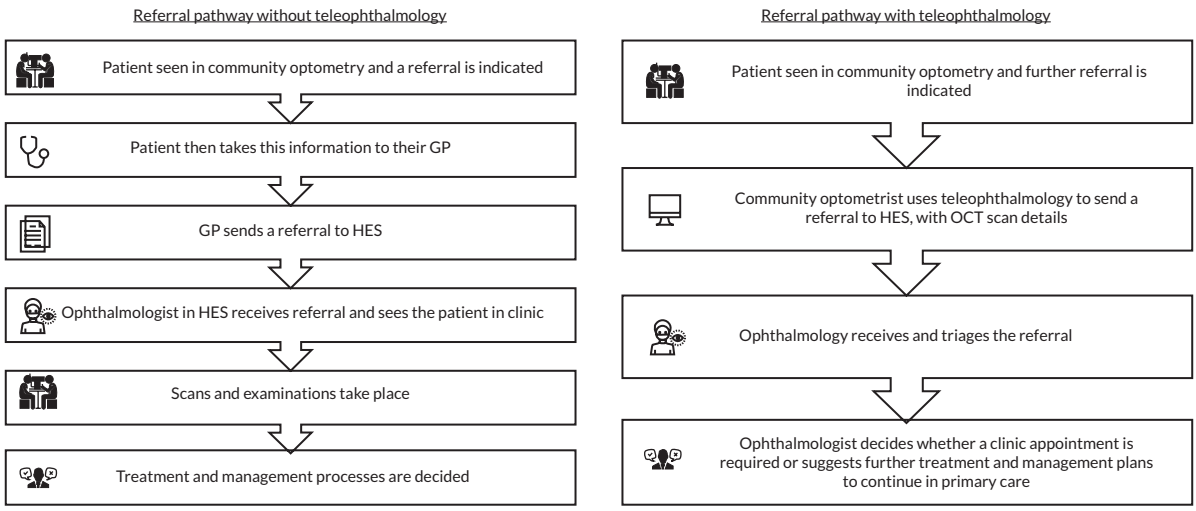

FIGURE 6 Typical referral pathway with and without the inclusion of teleophthalmology (images from www.freeicons.io).

not explore in-depth reasons for this limited adoption, this finding is not surprising. Generally, implementing digital health interventions in practice is acknowledged to be complex due to the multiple components that should be considered during implementation.[77–80] These include professionals' and participants' acceptance of the technology, staff training and education, changes in staff roles and practices, the organisational culture, capacity and readiness to accept innovations and the wider context (e.g. policy and regulations).[77–79] The application of deep learning models in ophthalmology referrals also brings with it a new set of challenges. For example, there are risks related to data security and privacy, as well as potential harm from false-negative diagnosis, that may impact the implementation and acceptance of deep learning models for clinical image classification.[69,71,72]

We present findings from a study linked with the cluster randomised controlled trial evaluating the effectiveness of the teleophthalmology platform.[17] Those in the intervention arm of the trial were using teleophthalmology to refer participants to HES, while those in the control arm were using their regular referral pathways. The differences between the pathways followed by practices in the intervention and control arms are presented in *Figure 7*. We report the novel qualitative findings of participants' and healthcare professionals' experiences to understand the practical implications of implementing teleophthalmology.

Aims and objectives

Previous research on the facilitators and barriers of teleophthalmology implementation has mainly focused on diabetic retinopathy screening,[81–83] with limited research focusing on the referral process between CO and HES for other retinal conditions. Therefore, this study aimed to assess participants' and healthcare professionals' acceptance of, and barriers and facilitators for, the adoption of two innovative digital technologies supporting referral pathways between CO and HES. These are a teleophthalmology platform and the AI DSS.[38] A HCI approach was used to understand professionals' and participants' interactions with the proposed technological solutions as well as the contexts in which these technologies are implemented. Five research objectives addressed the overall aim of this study:

1. to understand current workflows and practices of staff and participants in community optometry and HES to identify key user requirements for teleophthalmology tools from the perspectives of both groups
2. to understand workflows and practices of staff and participants in community optometry practices and HES with already established teleophthalmology pathways to identify technical, logistical and human factors affecting the implementation of teleophthalmology in practice

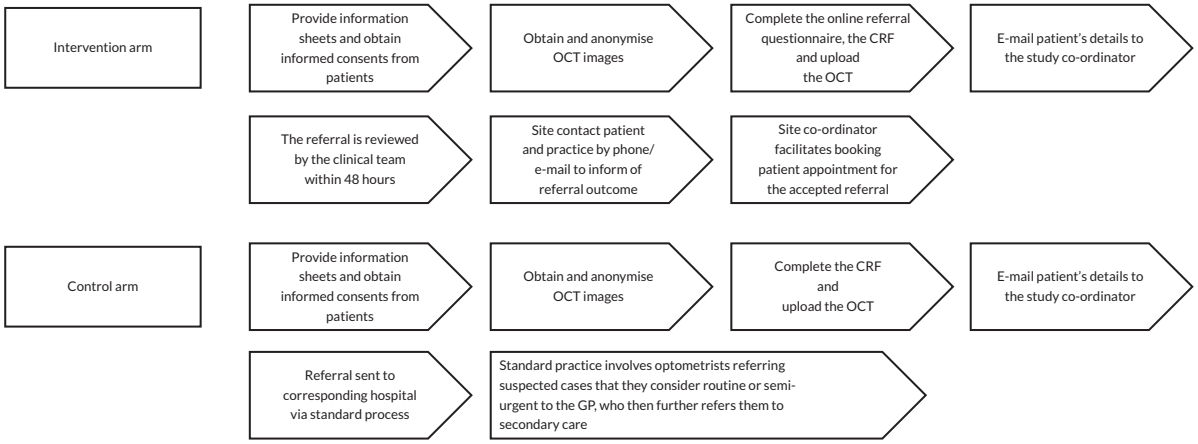

FIGURE 7 Pathways followed by the intervention and control arms in the HERMES trial. CRF, case report form.

3. to identify factors that shape professionals' and participants' attitudes to, and trust in, AI and how to present information in ways that instil appropriate confidence
4. to understand whether and how work practices are likely to change following the adoption of AI
5. to identify factors that ease the deployment of a digital referral platform to ensure acceptability and acceptance by all user groups and to understand the adoption process.

Methods and analysis

A theoretically informed, qualitative study was conducted to explore participants' and healthcare professionals' perspectives on teleophthalmology models of care and AI DSS. A combination of situated observations and semi-structured interviews with healthcare professionals and semistructured interviews with participants was conducted. This enabled us to understand the contexts in which new technologies are being implemented, focusing on understanding workflows, practices and user requirements as well as identifying potential barriers and facilitators to implementation. It enabled us to gain an in-depth understanding of staff and participants' expectations and experiences with the implementation of new technologies.

Participant selection and recruitment

Purposive sampling was applied to recruit participants who were representative of relevant patient and professional groups. This type of sampling is used to select participants who are most likely to produce valuable data.[84] Selection criteria for patient participants were:

- able to communicate in English, understand the study and give informed consent
- adults (≥ 18 years) attending the involved community optometry practices who underwent an OCT scan
- individuals who, in the opinion of the CO, have any suspicion of a retinal condition (including dry AMD, nAMD, diabetic retinopathy, MO, macular holes, ERMs, CSCR and genetic eye disease).

Participants with retinal conditions who are not routinely visualised or diagnosed using an OCT scan or those with conditions that prevent acquisition of good-quality OCT were excluded. This included peripheral retinal comorbidities such as peripheral retinal degeneration, retinal tear, retinal detachment, peripheral retinochoroidal tumours, Coat's disease, retinopathy of prematurity, familial exudative vitreoretinopathy and sickle-cell retinopathy.

Participants who had undergone an OCT and, in the opinion of the CO, had any suspicion of a retinal condition were invited to participate in an interview. Potential patient participants were invited to participate following their consultation at a participating community optometry practice. The optometrist explained the study to potential

participants, highlighting its purpose, possible advantages and disadvantages and what it entailed. Potential participants were given sufficient time to think about their participation and ask questions about the study. The researcher later called potential participants to obtain their decision to participate and book a provisional interview date for those who agreed to participate. Interviews were conducted at the optometry practice where the participant was recruited, or via telephone or video conferencing.

Professional participants included COs and clinicians (medics or specialist optometrists) with a minimum of 2 years' experience of independent practice in retinal clinics in HESs. Some of the participants' characteristics (e.g. their level of experience) were monitored during recruitment to ensure that diverse views were included in the sample.

Participants were recruited from three settings: (1) community optometry clinics in the control arm (pre-transitioning to teleophthalmology); (2) community optometry clinics in the intervention arm (post-transitioning to teleophthalmology) and (3) HESs. These settings helped us understand and compare experiences and work practices before and after implementing the new teleophthalmology technologies as well as identify barriers and facilitators during their implementation.

Sixteen community optometry practice managers (10 from the intervention arm and 6 from the control arm) and 7 ophthalmologists [(4 principal investigators (PIs) and 3 clinical fellows)] were invited via e-mail to participate in the study. The e-mail invitation included details about the aim of the study, what their participation would entail, the participant information sheet (PIS) and a copy of the informed consent form. Non-responders were followed up with an e-mail and/or a phone call to answer any questions they had about the study.

No restriction was applied on the level of experience of community optometrists, as it was important to include views from optometrists with various levels of experience. Ophthalmologists recruited to the study were all triaging referrals received from community optometrists via the HERMES teleophthalmology platform. Those were the local PIs of the HERMES trial at four sites or clinical fellows designated by the local PIs to triage referrals at those sites. Ophthalmologists had to have a minimum of 2 years' experience of independent practice in retinal clinics in HES.

Seventeen participants were recruited from community practices, of which 14 were from clinics in the intervention arm and 3 were from clinics in the control arm. These sites were affiliated with four NHS Trusts: MEH, Central Middlesex Hospital at London North West University Healthcare NHS Trust, University Hospitals Birmingham NHS Foundation Trust and North West Anglia NHS Foundation Trust.

A total of 24 health professionals were recruited to the study: 18 optometrists and 6 ophthalmologists. The 18 optometrists were recruited from 14 practices – 9 from the intervention arm and 5 from the control arm. This group included 11 practice managers or lead optometrists and 7 optometrists identified by their managers. The six ophthalmologists consisted of three PIs and three clinical fellows, recruited from four hospitals.

Design of observations and interviews

Observations

The aim of the observations was to gain familiarity with the contexts in which the two innovative technologies could be implemented. In particular, we aimed to establish an understanding of current practices and work routines and identify any differences in the workflows between practices. Understanding what people do and how they work in real life is a core focus for normalisation process theory (NPT). Additionally, findings from the observations helped to set the context for the semistructured interviews with healthcare professionals.

Observations focused on clinician–patient interactions around the diagnostic and referral processes. The researcher took field notes on the workflow, how referral decisions are made and communicated to participants about the clinician

interaction with the new teleophthalmology platform and any facilitators or barriers experienced during the interaction. To facilitate capturing these data, the flow and sequence work models from contextual design were used.[85] The flow model describes communication and co-ordination patterns that are important to accomplish the work, while the sequence model represents the detailed steps that people take to accomplish the tasks and the problems that they may encounter while doing them.

Interviews

The aim of the interviews was to gain an in-depth understanding of the expectations, perceptions and experiences of participants and health professionals with the teleophthalmology platform. Interviews were semistructured, allowing us to address the study aim while also following up on new insights as they emerged.[86] Two approaches were used to conduct the semistructured interviews with healthcare professionals: contextual inquiry interviews and critical decision method (CDM).

Contextual inquiry is a method commonly used in the HCI field to gain a deep understanding of users' work practices.[85,87] It is based on the premise that users are tacitly aware of their own work practices as they are immersed in their everyday activities. To understand their actions and reveal their motivations, intents and strategies, it is important to observe and speak to them in the context in which they perform their day-to-day activities. In other words, contextual inquiry involves conducting observations and following them up with questions to understand the work at hand.[86] In this study, contextual inquiry with healthcare professionals complemented the observations made in HESs and optometry practices.

The CDM, originated from the critical incident technique, is a cognitive task analysis approach used to elicit expert knowledge.[88] The CDM focuses on a retrospective analysis of critical incidents experienced by the interviewees.[89] In the context of HCI studies, critical incidents can include events when the technology failed or the system experienced particular demands.[86] The CDM uses a set of techniques to minimise recall biases and aid the interviewees to recall critical decisions as accurately as possible.[89] For example, the technique involves probing the interviewee to identify and describe a specific critical incident or incidents from beginning to end.[88] The researcher then composes a decision timeline and employs probe questions that allow the interviewee to provide corrections or more details.[88] The interviewee is also asked 'what–if?' questions to understand what might have happened differently. In this study, critical incident interviews were conducted with healthcare professional participants in the intervention arm to gain a deep understanding of their perceptions and experiences with the teleophthalmology platform as well as to explore barriers to its implementation in practice (e.g. when the platform failed and reasons for that).

A semistructured topic guide was used in all interviews and included questions related to the research topic and NPT. The topic guide was tailored to each group (participants and healthcare professionals in the intervention and control arms) as well as to suit the approach employed (contextual inquiry and CDM). The interview procedure followed the five steps to conduct HCI semistructured interviews.[86] Step 1 (opening the conversation) aimed to put the participants at ease and assure them they have the desired knowledge and expertise. Step 2 (introducing the research) introduced the topic and ensure that participants were aware of the purpose, reaffirming their confidentiality and right to withdrawal and requesting permission to record the interview. Step 3 (beginning the interview) gathered contextual information about the participant, such as their role, technology use and prior experience. Step 4 (during the interview) aimed to gain in-depth information about the topic under investigation. NPT components (coherence, cognitive participation, collective action and reflective monitoring) informed the questions in this step. Questions about coherence focused on participants' expectations from the teleophthalmology platform as well as its perceived benefits and barriers. Questions based on cognitive participation explored participants' engagement with the teleophthalmology platform and the issues they faced when using this new technology. Questions about collective action focused on participants' views on the impact of the teleophthalmology platform on eye care and practice as well as the changes that may be required to integrate this new technology in routine practices. Questions based on reflective monitoring explored participants' perspectives on how the teleophthalmology platform should be implemented in the future. For the AI DSS, issues around the 'black box' phenomenon, as well as the optimal place in the care pathway, confidence and trust were investigated. Probes such as anonymised screenshots from the digital referral platform and illustrative prototypes from the AI DSS were used to support the exploration of the themes. Step 5 (closing the interview) provided the participant

with an opportunity to express more thoughts and thanked them for their contribution to the study and the design of the technology. All interviews were audio-recorded, with participants' permission, and were transcribed verbatim.

Interviews took place between December 2021 and May 2022. The majority were conducted via videoconference, and the rest were conducted via telephone. Verbal consent was obtained from participants and was audio-recorded before the interview. All interviews were audio-recorded, with participants' permission. The average duration of the interviews was 36 minutes. Five interviews were transcribed verbatim by the study researcher. The remaining interviews were transcribed using Scrintal Software (Scrintal Labs, Stockholm, Sweden) and were checked for accuracy and anonymized by the researcher.

Data analysis

A combination of inductive and deductive thematic analyses were used, following Braun and Clarke's guidance on conducting a thematic analysis.[90] Two independent analyses were conducted; one was inductive and compared the professionals' and participants' perspectives; the second was structured according to the constructs within NPT, following the analysis method described by May *et al.*[91]

Inductive thematic analysis of participants' and professionals' data

The transcripts were coded using NVivo (QSR International, Warrington, UK) by an independent researcher (DP) who did not conduct the interviews, using inductive thematic analysis methods.[90,92]

Analysis started with the lead analyst (DP) familiarising herself with the data by listening to the audiotapes and reading transcripts and field notes. An open approach was followed at the start of the coding, where data from the first few transcripts and field notes were open-coded line by line, enabling interesting codes and insights to emerge from the data. NVivo V.20 software was used to manage the data analysis.

Thirty-five codes were initially defined and discussed with the research team. These were then refined and categorised into overarching themes. We present the three main themes that explore experiences of teleophthalmology for referrals for SRC.

Data analysis informed by normalisation process theory

A second data analysis based on NPT[93,94] was led by SA, who had led the data gathering. The recently published NPT coding book was used to structure the data analysis.[91] The coding framework consists of three main domains (contexts, mechanisms and outcomes) that includes 12 primary constructs (strategic intentions, adaptive execution, negotiating capacity, reframing organisational logics, coherence building, cognitive participation, collective action, reflexive monitoring, intervention performance, relational restructuring, normative restructuring and sustainment). The coding framework also includes 16 subconstructs related to constructs in the mechanism domain. Data analysis started by reading the interview transcripts several times to become familiar with the full data set. Data were then coded deductively but broadly to the four primary constructs that informed the interview topic guide (coherence, cognitive participation, collective action and reflective monitoring). These constructs align to the four primary constructs that fall under the mechanisms domain in the NPT coding book. Transcripts were reread to explore the relevance of the data to the contexts and outcomes domains, as the interviews did not specifically focus on these domains. Four primary constructs were found to be relevant, and data were coded to these constructs: strategic intention and adaptive execution from the contexts domain, and intervention performance and sustainment from the outcomes domain. Transcripts were then reread and coded in more depth to subconstructs, where data supported that level of detail. For example, in coherence, data were coded to the internalisation and differentiation subconstructs. Primary and secondary constructs identified in this study are summarised in *Table 40*. Data that did not fit the framework were coded separately and grouped under a theme called miscellaneous. These data were not included in the final analysis as they were not relevant to the study questions. The coding strategy was led and implemented by SA and discussed regularly with AB. In total, three themes (contexts, mechanisms and outcomes) and eight subthemes (strategic intention, adaptive execution, coherence, cognitive participation, collective action, reflective monitoring, intervention performance and sustainment) are presented in *Chapter 5*, *Results*.

TABLE 40 Primary and secondary NPT constructs identified in this study[a]

Domain	Primary constructs	Secondary constructs
Contexts – events in systems unfolding over time within and between settings in which implementation work is done	Strategic intention – how do contexts shape the formulation and planning of interventions and their components?	
	Adaptive execution – how do contexts affect the ways in which users can find and enact workarounds that make an intervention and its components a workable proposition in practice?	
Mechanisms – motivate and shape the work that people do when they participate in implementation processes	Coherence – participants contribute to enacting intervention components by working to make sense of its possibilities within their field of agency	Differentiation – how do people distinguish interventions and their components from their current ways of working?
		Internalisation – how do people construct the potential value of interventions and their components for their work?
	Cognitive participation – participants contribute to enacting intervention components through work that establishes its legitimacy and that enrols themselves and others into an implementation process	Enrolment – how do people join in with interventions and their components?
		Activation – how do people continue to support interventions and their components?
	Collective action – participants mobilise skills and resources and make a complex intervention workable	Interaction workability – how do people do the work required by interventions and their components?
		Relational integration – how does using interventions and their components affect the confidence that people have in each other?
	Reflexive monitoring – participants contribute to enacting intervention components through work that assembles and appraises information about their effects and utilizes that knowledge to reconfigure social relations and action	Individual appraisal – how do people individually assess interventions and their components as worthwhile?
Outcomes – the practical effects of implementation mechanisms at work	Intervention performance – what practices have changed as the result of interventions and their components being operationalized, enacted, reproduced, over time and across settings?	
	Sustainment – how interventions and their components can be incorporated in practice?	

a Descriptions provided as per the NPT codebook by May *et al.*[93]

Ethics and dissemination

Health Research Authority (Health Research Authority, London, UK) and Health and Care Research Wales ethical approvals have been obtained from London-Bromley Research Ethics Committee (Rec ref number: 20/LO/1299). Participant information sheets were provided to all potential participants. Written or audio-/video-recorded informed consent was obtained from all participants before they took part in the study. All interviews were conducted at a time and place convenient to participants. Participants were reminded of their rights to withdrawal from the study to avoid negative consequences on their work or the care they receive.

All data are handled following the GDPR, UK Data Protection Act 2018 and the Research Governance Framework for Health and Social Care. Participants' anonymity and confidentiality are maintained during the study. Interviews were conducted using encrypted audio recorders, and recordings were removed from the portable device permanently as soon as they are transferred to an access-restricted folder on the university home drive. People transcribing the interviews were subject to a nondisclosure agreement. Field notes and interview transcripts were pseudonymised.

Results: inductive analysis

We present the results from the two separate analyses in separate sections. There is, unsurprisingly, some overlap between the outcomes of the different analyses.

For the inductive analysis, we present our findings under three themes, Efficiencies of teleophthalmology, Teleophthalmology enables feedback and Concerns about teleophthalmology. We found most participants were optimistic about the implementation of teleophthalmology in the optometric referral pathway due to the efficiencies the platform would enable. All participants expressed needing feedback during the referral process to improve care and highlighted some concerns.

Efficiencies of teleophthalmology

All welcomed teleophthalmology due to its ability to improve patient and clinician experiences. There is regional variation in referral pathways depending on the specific condition; GPs typically process routine referrals for SRC. However, it was reported that GPs were not always suited to create referrals due to their limited skillset in specialist eye care. Participants often described GPs as the unnecessary 'middleman'. Participants reported being keen to be referred directly by their optometrist and felt this would reduce their waiting time to hear back from HES if referrals were processed directly.

> *If the optician can do the referral directly rather than you know, the optician telling me you'll need to go and see your GP who will refer you (…) I would be more comfortable because they [optometrist] know what they're doing whereas the GP is just saying you are alright then, if your optician told you that, then I'll send you.*
>
> *Patient 11*

In the HERMES study, optometrists used teleophthalmology to refer participants to HES, enabling a quicker referral process direct to triaging ophthalmologists. They shared that teleophthalmology could improve patient satisfaction and help relieve hospital capacity pressures by reducing unnecessary hospital visits.

> *I think the whole teleophthalmology thing will improve patient satisfaction and it makes life a lot easier, less participants, elderly participants having to find transport to the hospital, and being dilated once in the optometry practice and then being dilated again, back at the hospital and patient transport having to be arranged, so overall good, big saving of cost and finance and less crowded waiting rooms at the hospital.*
>
> *Optometrist 18*

Ophthalmologists shared that the teleophthalmology platform introduced uniformity to the referral process by requiring the same data fields to be completed for each patient, which was easy for the referring optometrist. This enabled referrals to be triaged and reviewed more quickly: referral decisions could be made promptly, and the appropriate triaging decision could be made regarding the indication and the urgency for a hospital visit.

> *It's more informative because, as you know, the platform has the questions with the tick box, um, on the optician findings which [is] not always involved in a classic referral proforma. And obviously, it has the imaging as well, which helps us to make a decision very quickly.*
>
> *Ophthalmologist 1*

As mentioned, the key benefit of the teleophthalmology platform is the ability to review and triage patient referrals without them having to attend a face-to-face appointment. All participants recognised and shared the advantage of saving time and resources by using teleophthalmology.

Teleophthalmology enables feedback

Some participants reported that they were happy for teleophthalmology to be used to assess their referrals as they would not want to attend HES if not required. Still, they expected feedback to explain the reason for not being seen for a face-to-face appointment. This was not always provided. Some participants reported dissatisfaction with their referral experience due to the lack of communication. This was also shared in the context of not receiving feedback promptly. Participants expected to hear about their referral decision more promptly through the teleophthalmology process, which increased concerns over their eye health when this expectation was not satisfied. In these cases, participants wanted to be seen or told directly and promptly why an appointment was not required and not to be kept waiting in uncertainty. Teleophthalmology can overcome this expectation discrepancy through accurate information presentation and timely and clear feedback.

> *If the participants were to get a letter or some form of communication from the hospital or the specialist to reassure them that your case has been looked at and this is what has been concluded, I think that would be enough to put somebody's mind at ease.*
>
> *Patient 10*

Optometrists stated that one of the significant benefits of teleophthalmology was the ability to receive feedback from ophthalmologists. Optometrists reported that when participants were referred to HES in the absence of teleophthalmology, participants would often return to them seeking more information and advice about their care; therefore, it was important for optometrists to be involved in the referral pathway and remain informed of their participants' management plans. Optometrists shared that many participants were not a reliable source of information about their eye health/treatment, which could affect future care or monitoring they provided.

> *Once we refer the patient, we don't actually know then what is happening thereafter, unless we chase the patient or, our participants are quite loyal, so they would, we would see them a year later, we will say or remember, we referred you last time, what happened? (…) So, it's actually we're basing it of what the patient is then telling us, so we are actually getting like the second story through the patient rather than the actual clinical information.*
>
> *Optometrist 12*

By receiving feedback, optometrists can also verify whether their referrals were appropriate and audit themselves to improve the quality of their referrals.

> *Because if we keep referring something that we think is urgent, but [the] ophthalmologist tells us this is not urgent, and if you learn by that, that's going to help you, you see. Right now, there's no feedback. (…). But if I got feedback from the ophthalmologist that saw the patient and I will know for next time when I see that similar sort of situation that well, actually this isn't urgent.*
>
> *Optometrist 10*

Having a system where the community-based clinic is connected to the HES was also seen as a great benefit for ophthalmologists. They welcomed being able to provide feedback to the referring optometrists, especially to enable the sharing of referral decisions directly and concurred with the need to provide feedback on the referral quality to improve future referrals. It was suggested that the teleophthalmology system should send referral replies to the referring optometrist, patient and GP so that all are informed of the outcome.

Concerns about teleophthalmology

There were some concerns about using technology to manage patient referrals. Some participants still wanted the reassurance of seeing a clinician rather than having their referral decision and notification completed remotely. Seeing someone face to face provided the holistic care some participants reported wanting and addressed their worries and anxieties.

> *I think [if] you don't get a chance to see the patient yourself, there is something about looking at data visually transactionally, that is fine, but there is also something about talking to the participants about how they're feeling and how they're coping with things.*
>
> *Patient 10*

Optometrists were mainly concerned with the practicalities of implementing a new system into their workflow. This included concerns over training to use the equipment, the reliability of network connectivity and equipment costs that some smaller practices may not be able to bear as well as remuneration for their time for taking on additional roles. Some also reported that completing a referral on the teleophthalmology platform took time.

> *The barriers would be cost, because this is all based on the information that is being sent from an OCT device, yeah, as part of the process of referral, it's not just from a letter, so when it comes to having the equipment, that's an immediate barrier. And having the right remuneration for the equipment.*
>
> *Optometrist 2*

Ophthalmologists also shared these concerns; however, they were positive towards the ability of teleophthalmology, enabling them to use their time more efficiently.

Results from normalisation process theory analysis

We present the results of this analysis, which focused on healthcare professionals' data, structured according to the constructs of NPT – namely, context, mechanisms and outcome.

Context

Strategic intentions

Most participants recognised the problems that faced the referral processes and suggested various approaches to overcome these issues.

One of the referral problems mentioned by many optometrists is the GP referral pathway. Many participants believed that sending the referrals to GPs instead of directly sending them to secondary care is inefficient in terms of time and is a significant hindrance. They felt that GPs lack the necessary knowledge or the right equipment to inform the referral decision to secondary care. Inefficiencies in the processing of referrals was also reported by many participants.

Cutting out the so-called 'middleman' and directly referring participants to secondary care was suggested by many participants as a way to improve the efficiency of the referral process.

> *I'll be asking the patient what's happened, so did you get the, they're like, oh, I never heard anything. And then when we look into the bottom of it, it hasn't been actioned at the GP, either the letter has been, and so many times participants have taken the letters and they had to come back several times saying, oh, it's been lost at the surgery, in my opinion, avoiding that route and going directly to a hospital or a triage services is probably would be better.*
>
> *Optometrist 10*

Optometrists also suggested providing more direct pathways to refer SRC as a potential way to improve referral efficiency. Currently, the only direct pathway that was commonly used by participants is the rapid access pathway to refer participants with suspected wet AMD. However, practices differed in how they referred the participants, with some using e-mails or online forms, whereas others used faxes. Some optometrists also mentioned that there is a lack of clarity on the pathways that should be used as trusts and Clinical Commissioning Groups (CCGs) have different pathways and protocols, suggesting the need to standardise the referral pathways across optometry practices.

> *Well, I think more of these pathways are useful. And then more, um, better links between the ophthalmologists and the optometrists in the community. Um, that would be helpful, but there are some more pathways, and they are trying. So but*

> *I think it's so random between practices that it would be nice if it was standardised. You know, every little borough has their own little way of doing things. I work in two different places, you know, and I and I worked on [name of an area], I don't anymore. But I was. And you know, all the pathways are so different wherever you go. It would be nice if there was some sort of standardisation because otherwise it's a postcode lottery all the time.*
>
> *Optometrist 9*

Most ophthalmologists, on the other hand, believed that the quality of the referrals by optometrists needs to improve. Having insufficient details on referrals and unclarity of the attached OCT images were two of the main issues mentioned, making it difficult for ophthalmologists to triage referrals appropriately. Referring participants unnecessarily to secondary care was another important issue mentioned by ophthalmologists, with one of the ophthalmologists mentioning that around 60% of the referred nAMD cases are considered as false referrals. Some ophthalmologists identified limited knowledge and experience of some optometrists as potential reasons for this, identifying the need to train optometrists to improve the quality of the referrals.

> *There is an issue which is a knowledge issue which is we get referred, uh, participants that do not need to be seen in the hospital setting. And this is a problem that we have tried to address, with sort of teaching sessions to the community, and we had a successful one, um, last year.*
>
> *Ophthalmologist 5*

Many optometrists recognised the need to improve their decision-making in handling referral cases; however, they identified issues that act as barriers. Not receiving feedback from secondary care and ophthalmologists regarding the referred cases was identified by most optometrists as a major issue that needs to be addressed. It was described as a 'one-way communication system' and a 'pain of our lives', making some optometrists feel unvalued in the system. Two types of feedback were identified in the interviews that could potentially help optometrists improve their overall practice and care to participants. The first is feedback from ophthalmologists. Most optometrists identified the need to know more about the ophthalmologists' assessment of their referral to help them improve their future referral decisions and avoid sending participants to hospital unnecessarily. The second is regarding the status of the referral, that is whether the referral was received or not.

> *But in 35 years, I could probably count on one hand the number of letters I've got back from the consultant. And that, I think, has to change*
>
> *Optometrist 16*

> *[B]ut I'm not going to hear back from that. I'm not going to hear back from that. So were those angles narrow? Did I do the right thing? Or was that, uh, was that a false positive referral that I did, um, you know, how would I know, because then the next time I'm in that situation, I can improve that. And for me, it's all about improving it for the future*
>
> *Optometrist 5*

The limited knowledge or experience of some optometrists, particularly junior ones, is another factor that could affect the quality of the referrals.

Some optometrists reported that optometrists with less experience sometimes prefer to refer any suspected case to secondary care, as they were concerned about their registration. The wide availability of OCT machines in optometry practices also meant that a large number of OCT images are produced. Although many optometrists and ophthalmologists recognised the importance of OCT in eye examinations, a few optometrists believed this could result in overburdening the hospital, particularly when many optometrists have limited knowledge in interpreting OCT results.

> *So I imagine that the ophthalmologists that are getting an awful lot of what they call it victims of multi-imaging technology have you come across it, you know, basically, these people have been getting false positive, referrals. They're getting a false referral because the clinician is covering their back. Yeah, basically the same. I don't know what that picture says, therefore, I'm going to send it*
>
> *Optometrist 14*

Overall, it was evident from the interviews that participants are open to new approaches and solutions to improve the efficiency of the referrals process and the quality of their referrals.

Adaptive execution

Optometrists and ophthalmologists reported several solutions that are currently implemented to overcome issues with the referral process.

Having a direct link with ophthalmologists via e-mail or by observing them in clinic has been identified by many optometrists as one of the effective ways to improve the quality of the referrals. This is because optometrists can use this pathway to consult with the ophthalmologists regarding the urgency of the referral and receive direct feedback regarding the case. Some optometrists also have access to group chats with fellow optometrists, which they use to get input regarding suspected cases. In their view, it has the advantage of providing them with a quick way to get feedback regarding cases they are doubtful about, particularly in the absence of links with clinicians in hospital. Most ophthalmologists also recognised the importance of having direct links with optometrists, with some taking the initiative to write back to optometrists regarding the clinical findings, as they considered it as an opportunity for educating optometrists. However, one of the barriers that ophthalmologists faced is the dynamic nature of the optometrists' community with continuous locum movement and newly graduated optometrists joining in. This is the same barrier that they faced when they started training programmes to educate community optometrists on pragmatic ways to interpret OCT images.

> *But by giving the feedback to them to state, these are the reasons why I think this is not urgent would be a learning process. And, uh, it has worked in some ways, but because there is quite a bit of a locum bank of opticians who come through, it has been an issue.*
>
> *Ophthalmologist 6*

Another solution that has been implemented in several practices is using online systems to refer participants instead of the paper referral. According to some optometrists, these systems help in improving the efficiency of the referral process, as it allows them to upload referral letters and attach images. This way, they are reassured that their referrals 'are not lost in the system'. However, a few optometrists reported not using these online systems due to not getting feedback about their referred case and because sometimes GPs are not registered on these systems.

> *Not all GPs are registered on the platform yet. So we basically type in the name, if their name comes up, then we can send it via [the name of the online referral system], if it's not on there would still have to do the traditional method of writing it and sending it with the GP or posting it to the GP ourselves, uhm.*
>
> *Optometrist 5*

Some practices also had access to central referral hubs. These hubs act as a single point of access to secondary care. It involves hospital-level optometrists or ophthalmologists triaging the referral and deciding on its urgency. It might also involve repeating some of the tests like OCT. Generally, optometrists who have sent referrals through these hubs thought that it was a positive step and could improve the accuracy of referrals. These systems bypass the GP, although it notifies them, which is another advantage of these systems from optometrists' point of view.

Overall, there were several ongoing initiatives to improve the efficiency and quality of the referrals. Optometrists and ophthalmologists who were involved in the implementation of these solutions were generally positive about them. However, except for the aforementioned rapid access pathway for nAMD suspected cases, there was not a large-scale solution that all practices implemented or had access to.

Mechanisms

Coherence

There are some common characteristics that were mentioned by optometrists, particularly those in the control group, which they expect to see in a teleophthalmology technology to improve the overall referral process. One of these characteristics is the ease of use and time efficiency of the platform where they can upload the OCT images and

the referral letter or information into the system, ideally on the same screen. The information on the referral is also expected to be streamlined to the currently used referral system in the UK, as any additional step could be considered as an added burden to their routine. Additionally, the system should allow them to have direct communication with ophthalmologists; however, optometrists' expectation on the means of communication varied based on their prior exposure to teleophthalmology platforms. Those with previous experience expected the communication to be online or via e-mail, whereas those with limited experience believed it would involve phoning up ophthalmologists as the term teleophthalmology implies. Optometrists expected that having a direct link with ophthalmologists in a teleophthalmology platform would provide them with the opportunity to consult with them regarding the referral decision. They also expected that they would receive feedback via the teleophthalmology platform, particularly about the ophthalmologist's clinical decision, as would be included in a discharge letter.

> *Yeah, yeah, maybe it's just an ideal thing would be a good and easily accessible user friendly system or platform that is time effective that where the optom can upload or communicate with ophthalmologists or other experts nurses, etcetera and get maybe sometimes feedback, ideally, within a few days if there's uncertainty or queries or if, you know, making referrals to the right place in a good time.*
>
> *Optometrist 18*

Several perceived benefits of teleophthalmology were identified by optometrists. Many believed that receiving immediate feedback from ophthalmologists via teleophthalmology would reduce unnecessary referrals to hospitals as it would help them improve their practice and avoid mistakes. They also believed that they would be valued and trusted more by ophthalmologists, especially when they diagnose correctly. Several optometrists believed that implementing teleophthalmology would improve participants' experience of the referral as it would avoid them going to hospital unnecessarily, alongside keeping them informed about their care. Saving on NHS cost by bypassing GP and reducing participants' unnecessary visits was thought to be another potential benefit by some optometrists.

> *Well, it's just a case of, knowing what you're doing, where we're going, of course, the finesse is coming to it, and to reassure participants that they are being seen by ophthalmologist, and to give us credit that you know we have diagnosed correctly and we are referring in a valid and timely manner. Um, as I say, sometimes we do waste GPs' time by doing the referral to them, we are sending the referral to them when we could do it directly and more efficiently, so to waste and stop wasting national health service time.*
>
> *Optometrist 1*

A few barriers were identified by optometrists, which they thought may face the implementation of any teleophthalmology technology. Many optometrists believed that the poor IT infrastructure of some community optometry practices might be a potential barrier. For example, many practices currently use old systems or equipment in their practices and hence they do not have the necessary hardware or software to operate new teleophthalmology platforms. Also, lack of interoperability between the different systems used in practices, particularly those used to manage the OCT machines and the patient details, could pose challenges for implementing teleophthalmology platforms. Some optometrists also thought that the IT skills of optometrists, especially older ones, might be a potential barrier. Cost of the OCT machine and not receiving the appropriate remuneration were perceived by some optometrists as potential barriers for implementing teleophthalmology, particularly in less well-off areas.

Cognitive participation

Most optometrists were ready to engage with the HERMES teleophthalmology platform as they were familiar with teleophthalmology platforms or online referral systems, and it also addressed their desire to overcome referral issues. All optometrists in the intervention arm received training on using the HERMES teleophthalmology platform, which was supplemented with flow charts that explained the steps they need to follow when referring a patient via the platform. Therefore, in theory, all optometrists were clear on how to refer a patient through the teleophthalmology pathway. However, in practice, several optometrists faced difficulties in following the process described in the flow chart, particularly for the first few participants. The main issues raised by optometrists regarding the process of referral were specific to conducting research in practice. For example, the majority of optometrists found that explaining the study and consenting participants were time-consuming, describing it as 'tedious', 'the longest part' and 'long-winded'. Anonymizing the OCT scan was another issue that was found laborious by many optometrists. However, some

optometrists recognised that these steps are specific to the research context and are not included in real practice. In addition to the issues related to consenting participants and anonymisation of data, many optometrists found filling the questionnaire to be time-consuming, with some estimating the time to be around 30 minutes. Another issue raised by many optometrists was that the questionnaire included repetitive questions. Optometrists were also asked to answer each question on the questionnaire, although some were irrelevant to the referred case.

> *But I think the whole process from the questionnaire to the, um, take the images twice and then sending them across, you know, that is cumbersome. It's a laborious process*
>
> *Optometrist 11*

The support provided by the PM was essential to those who faced any difficulty in the process, particularly those with limited information technology (IT) skills. Sharing feedback with the PM regarding the long process in recruiting participants, the repetitiveness of the questionnaire and the added administration burden to optometrists' daily work resulted in a few changes to the process, such as telephone-consenting participants and pre-filling the questionnaire.

Collective action

Optometrists employed different strategies to incorporate doing the referral on the platform in their existing practice. One of the main changes was adding more time to their regular eye examination appointments, particularly for those participants who they believed might be a retinal referral. Many optometrists completed the referral when the patient was not present, taking advantage of their ability to access the platform remotely. This included completing the referrals before or after clinics, during their lunch breaks or on days off. Another change that was implemented in practice was dividing the eye examination and the referral process. For example, a few optometrists mentioned that when spotting a referable case to HERMES, they would ask the patient to do the OCT test and consent on another visit. It was also an opportunity to provide participants with more time to think about the study. Other optometrists prioritised doing the OCT for their participants when they were constrained with time, particularly if they believed the referral was urgent. Some optometrists depended on their colleagues to help them with some of the steps such as anonymisation of the participants' details. Those with less experience in referring participants focused on consenting participants and completing the online questionnaire; however, the final referral decision was jointly made with a senior member of staff.

> *It is just, it is just time I think, I added time, I [not clear, 18:38] folder for all the consent forms and the info sheets and everything. I've had to work from home and log in and just to complete it because, you know, some of these referrals need to be done soon, quickly you know we can't wait for some time. You know two or three working days later, so it's going to be done quickly. So had to work overtime in that sense to sort of action it and get it done.*
>
> *Optometrist 6*

Ophthalmologists received an e-mail notification when a referral had been uploaded on the platform, alongside an e-mail from the study co-ordinator. Most ophthalmologists actioned the referral as soon as they were available to access the platform and review the referral. This depended on their clinic timings and the time of the day they received the referral. Like optometrists, ophthalmologists were able to access the platform remotely and review some of the referrals. However, some OCT images could only be viewed in a hospital setting, which according to one ophthalmologist is 'quite frustrating'. Additionally, there were some variations in how referrals were managed in different sites. In some cases, the referrals were triaged based on whoever saw them first, while in other sites, a specific person was assigned to review all referrals, with another person covering for them during their absence.

Reflective monitoring

There were some common factors that optometrists use to judge the value of the teleophthalmology platform. Receiving feedback from secondary care regarding their referred case was one factor that several optometrists valued. However, they reported inconsistencies regarding feedback received via this pathway. Some optometrists received regular feedback regarding the cases referred through the platform. They were generally satisfied about it, given that, in usual practice, they do not normally receive feedback about referred cases. It also provided them with reassurance that what they had referred was correct. However, in some cases, optometrists would have liked a more detailed feedback regarding the case. Conversely, some optometrists reported not receiving feedback on their referred cases. In addition

to the importance of feedback for their professional development, optometrists required this feedback to inform participants regarding the status of their referral.

> *That would be a critical factor for us that there was, uh, very clear feedback both to us and the patient. I've not really had confirmation of any of the sort of provisional diagnosis that we've made on the HERMES, whether we refer them correctly or not. And that would be very useful for us. That's very useful for learning. Because generally if you say to someone you have a problem, it would worry them, even if it's not serious, it still concerns people, so telling the participants that they will be contacted in a couple of days give them this reassurance.*
>
> *Optometrist 7*

The speed of the referral and that a referral decision would be made within 48 hours through the teleophthalmology platform was considered to be advantageous by many optometrists. Many optometrists convinced participants to send their details through the HERMES platform based on the speed of this referral route compared to the traditional pathway. Therefore, some optometrists felt disappointed when the process took longer than expected, which led them to question the value of this referral pathway. Additionally, a few believed that this advantage is because of the study environment, and this might not be the case when teleophthalmology is implemented in practice, as the volume of referrals will be larger.

Ophthalmologists in general had positive experience triaging referrals via the platform. Most ophthalmologists found the platform easy to use and straightforward. Time was an important factor that ophthalmologists used to assess their experience triaging referrals on the platform. The majority believed the time was reasonable. Although a few ophthalmologists believed it took longer than usual practice because of the need to review OCT images, this was considered to be acceptable as it helped them achieve high-quality decisions. Despite their overall positive experience, ophthalmologists suggested a few points to improve their interaction with the platform. One suggestion was adding a question regarding the symptoms and their duration, which would help them decide on the urgency of the referral. Another suggestion was to get a notification to triage the referral only when the referral is completed on the system, as they are currently getting a notification for each step performed by the optometrist. Improving the way a specific OCT machine interfaces with the platform was another point mentioned. However, suggestions to improve the platform were viewed as minor, while addressing the barriers that optometrists may face when using the platform was considered to be more significant.

Outcome

Intervention performance

The views on the intervention performance came largely from sites that had referred several participants to the HERMES study. Ophthalmologists' views were mixed regarding the quality of the referred cases using the teleophthalmology platform as compared to the traditional pathway. Some ophthalmologists believed that the quality of their decisions using the teleophthalmology platform was better because they had the required information to triage effectively and that the quality of the referrals received via the platform improved compared to what they usually receive from community optometrists. This is because the referrals received via the platform had the optometrists' provisional diagnosis, whereas in usual practice, optometrists are not required to propose a diagnosis. However, a few ophthalmologists thought this might be because of the study environment and that optometrists felt their decisions would be scrutinised. A few ophthalmologists also thought that some information in the referrals were missing and that optometrists are still sending cases that should not be referred.

> *[T]hey are more cautious of referring let's say silly things, for example, and unexplained, visual loss without any scan, uh, without the pathology and OCT scan or at a photograph. So they don't refer it, of course, to the medical retina. Uh, but still, you refer, you have cases that they have just a single, uh, epiretinal membrane or very mild lamela hole that cannot progress very quickly and possibly doesn't need to be seen at the secondary care environment unless it shows progression through the time.*
>
> *Ophthalmologist 4*

Optometrists' views on the impact of using the teleophthalmology platform on their usual practice were also mixed. Some optometrists believed that using this pathway improved the efficiency of the referral process, given the speed of the review process. A few optometrists believed that using this pathway increased confidence in their referral decisions and alleviated some of the pressure they experienced, particularly when those referrals were accepted. Conversely, several optometrists did not think that using the teleophthalmology pathway had an impact on what they would normally refer. They also felt that the main change in their practice was adding time to their normal routine. In terms of relationship with secondary care, some optometrists believed that using this pathway had consolidated existing relationships with ophthalmologists as they had regular communication and updates regarding the referred cases. However, other optometrists did not believe it had an impact on their relationships with ophthalmologists due to existing good relationships or limited contact with secondary care.

> *I mean, it's a referral, ok, it makes no difference, ok, I see a problem, I discuss it with the patient and say listen I'm not happy with it, I think we need to see a specialist about it, it's, when I realized that something has to be done, it's not oh no what I do, it's just knuckle down and do it.*
>
> *Optometrist 3*

> *[B]esides learning to anonymize the scans and, you know, setting up the obvious platforms on the computer, I don't think we've done anything different.*
>
> *Optometrist 9*

Sustainment

Some factors were identified that could contribute to the sustainability or the incorporation of the platform in practice. Some of these factors were related to the platform, such as ease of use, the time to complete the online referral and offering direct communication between optometrists and ophthalmologists. These factors were similar to those the optometrists expected in any teleophthalmology platform implemented in practice. Some ophthalmologists believed that the platform should be incorporated in the participants' electronic records. The platform should also allow them to send feedback to participants in addition to optometrists. Some factors were related to the optometrists' skills. For example, some optometrists and ophthalmologists believed that providing optometrists with hands-on training on how to interpret the OCT images would improve the quality of their referrals via the platform. The limited IT skills of optometrists, particularly older optometrists, were often mentioned as a potential barrier and should be addressed when implementing the platform. Other factors were largely related to finances. Providing remuneration to the optometry practice to purchase OCT machines was considered as an important factor to facilitate the adoption of teleophthalmology platforms in practice. Persuading CCGs and NHS to invest in community optometrists and new direct pathways as well as contracting out services in the community were suggested as other factors that could aid the integration of teleophthalmology pathways in practice. Increasing participants' awareness of the importance of doing an OCT in their eye examination was also considered to be essential to implement the teleophthalmology pathway in practice.

> *Very difficult for me. Uh, first of all, you have to persuade all managers, okay. Not only the managers, but also the CCGs to, uh, the CCGs to have this platform on. The second thing is, uh, this is one point from the secondary care, so we were getting involved in each county to go and persuade them to pass through, into the national system as we have passed the test. But on the other side now of the optometrists, I mentioned, uh, also some problems that we face, and there are big changes of optometrists like [name of high street optometry practices] that they might not accept this and also, uh, the representatives of them in the CCGs. Yeah, So it's not because the clinicians don't like or optometrists don't like. Sometimes it's a big it's a big game behind that.*
>
> *Ophthalmologist 4*

Discussion

We present the discussion as related to each of the analyses presented above.

Inductive analysis of the perspectives and experiences of participants and clinicians

While previous work has focused on the efficacy and efficiency of teleophthalmology platforms through reviewing referrals,[28,95] we report insights based on experiences from participants, optometrists and ophthalmologists to validate previous findings on perceptions of using teleophthalmology for SRC.

All participants recognised the value of implementing a teleophthalmology system into their ophthalmic care pathway due to its potential to improve patient care and health services efficiencies. This is supported by others,[22] who found through a review of referrals that teleophthalmology can reduce unnecessary HES visits and significantly impact patient anxieties.[96] Our study has shown this in practice, with many participants sharing that they would not want to attend HES if not required.

Both optometrists and ophthalmologists reported that teleophthalmology's significant advantage is the ability to electronically refer participants directly from optometrist practices to HES, which can significantly reduce the waiting time for participants. The ability to triage referrals electronically also enabled ophthalmologists to provide replies and feedback to the referring optometrist via the teleophthalmology platform, which they greatly valued.

Implementing teleophthalmology into the eye care pathway would remove the burden on GPs of having to process patient referrals, but they must be informed of such referrals. GPs have, in principle, supported the suggestion of optometrists referring participants directly,[95,97] and we found that participants and optometrists would support this change in practice. There is great value in involving optometrists in the referral process, as it has been reported that this could reduce unnecessary referrals by approximately half.[28]

The overarching theme shared by all participants that substantiates many of the benefits of the teleophthalmology platform is the potential of the platform to facilitate the provision of feedback. The importance of receiving timely feedback in eye care, in general, has been reported by others,[63,98] and it was essential for participants to alleviate their concerns over their eye health. Harvey *et al.* specifically outlines the factors that could affect the provision of feedback, and a key implication of their work is the call for technology to support this provision.[98] Thus, we found that teleophthalmology could overcome concerns in the optometric referral pathway.

While feedback can keep both participants and optometrists informed,[99] it can also improve future referral quality through open conversations between referring clinicians. Our results concur as we report optometrists greatly value receiving referral replies directly from ophthalmologists to remain informed of their participants' care and audit their referrals. Additionally, previous research has highlighted that the lack of communication between optometrists and ophthalmologists can be problematic;[100,101] therefore, implementing a teleophthalmology platform could help to overcome this.

Potential barriers raised by some optometrists were the initial setup costs, which include time, training and financial costs. While others have reported this should be considered,[73] we found this to be a key concern in practice. According to the current General Ophthalmic Services contract (2023), OCT scans are not a contracted service in optometric care.[97] Therefore, there is a need to ensure optometrists are appropriately remunerated for providing this service. Optometrists would also need to have appropriate access to NHS e-referral systems to expediently refer participants to the HES system. Others have begun to explore the cost-effectiveness of implementing teleophthalmology systems;[58] further work is needed to establish health–economic benefits concerning SRC that were raised in our study.

To successfully implement teleophthalmology into the optometric referral pathway, there needs to be an investment to enable parity among optometry practices to support new technology setup. While the implementation of teleophthalmology was perceived to initially increase the workload of optometrists to process referrals and ophthalmologists to triage referrals, with the correct remuneration, there is significant potential to relieve pressures on stretched eye care services.

Understanding teleophthalmology implementation in terms of normalisation process theory

Using NPT as a theoretical approach, our study highlighted several factors that could influence the adoption of teleophthalmology platforms in the referral process by optometrists and ophthalmologists or, alternatively, lead to the abandonment of these platforms.

One of the findings of our study is that most optometrists recognised the value of using the teleophthalmology platform and were generally ready to engage with the platform. For example, they identified reducing unnecessary referrals to secondary eye care and improving participants' experience of the referral as potential benefits of using the platform, consistent with existing literature.[65,67] Understanding the value of the new technology and its potential benefits is often an important facilitator in accepting new technologies.[102,103] Similar studies show health professionals' openness to new technologies improving workflows and participants care.[104,105] However, despite willingness, occasional shortcomings in meeting support expectations can serve as a barrier.[104] Optometrists in our study expected teleophthalmology platforms to facilitate communication with ophthalmologists, providing an opportunity to discuss referred cases and enhance the quality of their future referrals. However, feedback sometimes fell short of their expectations, leading many to feel that the platform did not improve their referral cases. Ophthalmologists, on the other hand, generally had a more positive experience when using the teleophthalmology platform, noting improved decisions due to having the necessary information to triage effectively. These findings suggest that the teleophthalmology platform may have been developed with a preference for ophthalmologists' perspective to address over-referral issues, potentially overlooking optometrists' viewpoints. They also underscore the importance of identifying key stakeholders and involving them during the early stages of technology design and development before the implementation stage.[106–108] This approach can help in identifying user expectations and designing technology that meets them, enhancing its potential for future adoption.

Another important factor that may affect the experience of health professionals is linked to the technology, particularly the additional time it adds to their current work routine. In our study, this may have contributed to the relatively positive experience that ophthalmologists had when interacting with the platform, with the majority of them finding the additional time to be reasonable. By contrast, many optometrists found filling out the questionnaire on the platform to be relatively time-consuming, which might have led to a less-positive experience with the platform. This is an important consideration when developing the teleophthalmology platform as health professionals' time constraints could be a barrier to incorporating new technologies into their work routines, especially if the technology is time-consuming or increases their workload.[102,105,109] However, the results from ophthalmologists demonstrate that the additional time was acceptable because it contributed to an improvement in their decision-making process. This suggests that health professionals might be willing to incorporate some extra time into their usual practices if it leads to an overall improvement in patient care. In other words, the additional time should provide value to them rather than merely adding administrative burden.

In our study, individual factors played a critical role in either facilitating or inhibiting the adoption of the teleophthalmology platform. Among optometrists, there were differences in how they adjusted their work routines, with some prioritising conducting the OCT scan and others dividing the eye examination differently. Additionally, some optometrists, especially those with limited IT skills, required extra support to complete the online referral form. The success of technology implementation may, arguably, depend on how well these differences are considered when designing, developing and implementing the technology. In our study, the HERMES trial manager played a crucial role during the implementation stage by providing continuous support to optometrists. Providing training and support to potential users of the digital health technology are widely recognised as essential for successful implementation.[102,103,109] This assistance can help health professionals overcome barriers related to unfamiliarity with the technology and equip them with the necessary skills for its use.[110] Offering training and support to health professionals can also help them in integrating technology into their work routines.[103] However, one might question the feasibility of providing the level of intensive support observed in our study in a non-trial setting, particularly considering that many digital health interventions are implemented in resource-constrained settings. One approach to address this issue is to identify individuals who require support and assess the extent of support needed, taking into account the technology complexity and available resources. As demonstrated in our study and in other research,[109] health professionals have varying support needs and not everyone requires an intensive level of support or training.

Our study emphasised the importance of understanding the implementation context of the digital health intervention to identify facilitators and barriers to its implementation. For example, the study revealed that the issue of referrals is well recognised, and both optometry practices and secondary care were keen to solve them by implementing various solutions. However, most solutions were at a small scale, and except for the rapid access pathway to refer nAMD cases, there was no large-scale digital health intervention or referral platform that all practices implemented or had access to.

This highlights the complexity associated with implementing digital health interventions and that factors unrelated to the individual motivation to solve the problem or readiness to use of the technology must be considered for successful large-scale implementations of such interventions. Participants in our study mentioned key factors that may act as barriers to integrating the teleophthalmology platform into their usual practice, including the cost of the OCT machine, poor IT infrastructure, lack of interoperability between OCT machines and optometry practice systems and the need for investment in community optometry services. Some of these factors are already identified in the literature as potential barriers to implementing healthcare technologies.[105,110–112] Additionally, our understanding of the context revealed that implementing the platform alone might not result in the desired improvement in the referral quality. Initiatives like training optometrists on interpreting OCT images may be necessary to achieve a significant improvement in the quality of the referral. This finding aligns with the research conducted by Ramchandran *et al*.,[105] who identified the importance of hands-on training in eye health as a key factor from the staff perspective to facilitate engagement with teleophthalmology to improve diabetic retinopathy surveillance. Therefore, to achieve a large-scale implementation of the teleophthalmology platform, addressing these organisational factors is essential. Otherwise, there is a risk of limiting its implementation to a small scale or even potential abandonment of the technology, which is a common risk associated with digital health interventions.[77,112–114]

Another important consideration is the potential impact of research processes on participants' engagement with digital health technology. In our study, some optometrists found research processes, such as consenting participants, to be time-consuming. This may have acted as a barrier to their engagement with the research and the digital health intervention. Previous research supports this observation,[114] indicating that research processes can lead potential participants to decline involvement in interventions. Given the equal importance of involving end-users and complying with these research processes, it is important to find ways that strike a balance between the two. One potential solution identified in our study is obtaining participants' consent for the HERMES trial via telephone which seemed to be well received by optometrists.

Additionally, our findings, consistent with Perski and Short,[114] highlight the importance of differentiating between barriers that stem from the technology itself and those that arise from the research process. For instance, optometrists' experiences with teleophthalmology platform might have been more positive if the research processes were time-efficient and aligned with their needs. As described above, the HERMES teleophthalmology platform was implemented in the context of a trial, with specific steps due to the trial nature. Optometrists' experiences may be more positive in real-life implementation since these trial-specific steps will not be necessary. *Figure 8* presents the anticipated pathway of the teleophthalmology platform in the context of referrals if implemented in real life.

Strengths and limitations

Through the use of in-depth interviews, we were able to elicit the lived experiences of those involved in the study, many of whom had direct experience with the teleophthalmology platform. This proved valuable in comparing and contrasting their experiences, revealing differences and identifying barriers faced by different groups. We were also able to recruit participants across the stakeholder group, thus providing multiple perspectives on the teleophthalmology pathway. Optometrists and ophthalmologists provided their perspectives on using the platform and the practicalities involved in use, while participants reported on the personal impacts of navigating their eye health journey through teleophthalmology. These diverse perspectives have enabled us to corroborate and extend existing understanding of the practical implications of implementing teleophthalmology. However, it is noteworthy that some of the participants

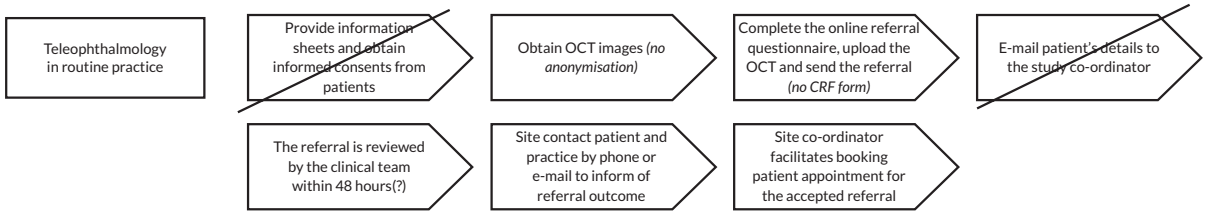

FIGURE 8 Anticipated teleophthalmology pathway in routine practice. CRF, case report form.

had prior experiences with teleophthalmology platforms which could have influenced their experience with the platform under investigation.

Study limitations include the influence of participation bias. The participants who chose to participate in our research may have different views from those who did not participate, which have not been captured in our study. We also recognise that the ophthalmologists who chose to participate were involved in the HERMES study itself, which led to an increased knowledge of teleophthalmology. Future work should endeavour to recruit a more diverse sample of participants to capture broader views on the experiences of teleophthalmology.

Regarding sustainability, a limitation is that, given the short time the platform was implemented at the time of the interviews, most of the identified factors were related to the perceptions of optometrists and ophthalmologists regarding what would sustain the intervention rather than what was actually contributing to its sustainment. Also, because this qualitative study took place within the context of a clinical trial rather than routine deployment, participants' perceptions may have been coloured by the experience of participating in a trial as well as that of implementing teleophthalmology.

Conclusions

Implementing teleophthalmology into the optometric referral pathway has numerous benefits, as outlined by all participants. Through our in-depth interview study with participants, optometrists and ophthalmologists, we found that they generally report great value in implementing teleophthalmology through improving efficiency and the ability to provide and receive feedback. Participants were satisfied if their referrals were reviewed with teleophthalmology to reduce the possibility of unnecessary HES visits if this was clearly and efficiently communicated back to them. Optometrists felt they were better suited than GPs to write and process patient referrals and would feel more valued if they were more directly involved in the pathway. Finally, ophthalmologists were pleased with a system enabling them to manage their caseloads more efficiently. Further efficiencies teleophthalmology can promote include removing the burden on GPs, the time participants wait to be seen by HES, the time it takes for ophthalmologists to review and provide referral replies and, finally, the overarching benefit to all participants of being involved and receiving feedback.

The NPT analysis highlighted the need to consider multiple factors when developing and implementing teleophthalmology platforms, especially if we want to see the technology to be adopted and normalised at a large scale. These factors relate to the technology, such as the time it takes to complete the referral on the platform, to the individuals, such as their various levels of IT skills, and organisational considerations like cost of equipment, IT infrastructure and training and investing in community optometry practices. Additionally, we emphasised the importance of involving key stakeholders in the early stages of technology design and development, as well as during implementation, to ensure that the developed technology matches their expectations of support. The importance of designing and implementing research processes that are time-efficient can also be considered as key for an active and successful engagement of health professional participants with the digital health intervention. Overall, NPT proved to be a useful approach to untangle some of the complexities associated with the implementation of the HERMES teleophthalmology platform.

Acknowledgements

We are grateful to all the participants, optometrists and ophthalmologists who participated in this study for sharing with us their valuable experiences.

Chapter 6 Discussion

Cluster randomised controlled trial

There is an ever-increasing workload for hospital ophthalmic services worldwide, with shifting demographics creating an increased demand for treating common age-related retinal conditions such as nAMD.[1,2] A large majority of referrals to hospital ophthalmic services arise from primary eye care, including community optometry services/high street opticians in many developed countries.[11,12,14,21] In the UK healthcare setting, it has been reported that 72% of hospital referrals originate from CO.[13] There is a paucity of high-quality evidence reviewing the benefits of telehealth in the form of clinical studies.[115] In the case of asynchronous teleophthalmology, where imaging data and clinical information are stored and reviewed at a later time point, previous studies have shown that virtual review of eye disease referrals can reduce unnecessary referrals from 34% to 80.5% to HES for general ophthalmology and retinal conditions.[34] Asynchronous teleophthalmology is well suited for the efficient teleophthalmology triaging of referrals for retinal disease, and, particularly, for urgent retinal referrals of common, vision-threatening, eye conditions, such as nAMD. The main reason for this is that retinal disease referrals are highly dependent on the availability of complex, three-dimensional eye scans (OCTs) for remote, teleophthalmology review by hospital-based experts. We are only aware of two studies where the OCT was reviewed for asynchronous teleophthalmology referrals of retinal disease, one of which assessed referrals from CO. Kern *et al.* were able to show a 52% reduction in unnecessary referrals using this novel referral pathway.[28]

In HERMES, teleophthalmology significantly reduced unnecessary urgent hospital referrals by 17% compared with standard care. Among referred participants only, teleophthalmology reduced unnecessary urgent referrals by 59%, meeting the prespecified 30% superiority margin. Overall unnecessary referrals (routine and urgent combined) were reduced by 6% in the teleophthalmology versus standard care groups comparison, both in the overall enrolled and in the referred participants. However, the total number of false positive cases observed was small. Teleophthalmology has the potential to reduce unnecessary urgent referrals, particularly for suspected neovascular AMD—a major pressure point for HES. Despite the statistically significant observed 6% reduction in overall unnecessary referrals with teleophthalmology, the absolute numbers were small introducing some statistical fragility. Additionally, after excluding the outlier control group optometry practice from the analysis, between arms difference decreased to 1% and lost significance. We are therefore adopting the conservative interpretation of these findings, i.e., the effect of teleophthalmology in reducing unnecessary referrals overall was inconclusive.

The observed event rate of false positive (unnecessary) referrals made by community optometrists overall, both in the control group optometry practices (standard care) and in the intervention group optometry practices (teleophthalmology) was lower than estimated. Early recruitment patterns suggested optometrists across both groups adopted greater caution with borderline cases, probably due to awareness of scrutiny (a Hawthorne effect), indicating the trial environment promoted more considered referrals. Although beneficial clinically, this behaviour might reflect a temporary behavioural shift. Additionally, post-COVID-19 upskilling initiatives in community optometry probably improved practitioner competence, contributing to enhanced referral accuracy.

The low event rate (unnecessary referrals by community optometrists overall irrespective of urgency level) led to few non-referred cases, with two implications. First, it limited statistical power to detect overall referral differences, causing absolute and relative measures (ORs) to diverge; therefore, we prioritised absolute differences in our primary interpretation. Second, for urgent referrals—when optometrists are more cautious due to the high risk of missed diagnoses—the Hawthorne effect is less likely to occur, allowing a more robust assessment of teleophthalmology. In this context, the reduction in urgent referrals was significant and met the superiority threshold. Because urgent referrals for conditions such as neovascular AMD are less affected by behavioural bias, they better reflect the true potential effect of teleophthalmology. Reducing these referrals could meaningfully relieve pressure on overstretched specialist services.

A recent systematic review of teleophthalmology referral triaging pathways for eye disease demonstrated that such pathways have a beneficial effect on the accuracy and appropriateness of referrals to HES. Specifically, the review focused on general ophthalmology and/or retinal referrals sent by primary care optometrists with photographs

attached. These studies all reported positive impacts on referral accuracy, with 34–48% of participants reviewed virtually identified as not requiring referral for face-to-face hospital review. This value increased to 80.5% in a more recent Danish study,[81–83] perhaps due to the improved quality of ocular imaging available for review. In order to improve the ability of clinicians to efficiently triage participants remotely, more information needs to be uploaded for remote review by hospital-based experts, including advanced ocular imaging such as OCT, which is now more widely available in primary care. Two studies included uploading of OCT images along with fundus photographs.[7,63] The more recent study from the UK[7] assessed referrals, specifically for retinal conditions, and found that 52% of the participants classified into the referral-warranting category did not require specialist hospital in-person attendance.

Previous research has also shown the ability of asynchronous teleophthalmology to successfully reduce the rate of false-negative referrals.[33,116] The HERMES referral platform did demonstrate a lower number of false-negative referrals compared to standard of care, but not significantly. However, in terms of wrong referral urgency, the control arm had > 25% more incorrect referral decisions than the intervention arm, a significant difference at the 5% level. Wrong referral urgency has major impacts on both the NHS and participants. For example, a routine condition being referred urgently causes unnecessary stress and anxiety to participants[7] and also contributes to the existing capacity pressures faced by NHS urgent services. Conversely, when an urgent condition is erroneously ascribed low referral urgency (routine referral), this is associated with a high risk for delayed hospital consultation and initiation of vision-saving treatment for these participants, leading potentially to irreversible, preventable vision loss.

When assessing the impact on nAMD specifically, the asynchronous teleophthalmology arm had approximately a 50% reduction in false-positive referrals compared to CO. The diagnostic accuracy for asynchronous teleophthalmology was also high [area under the curve (AUC) 0.96] compared to CO (AUC 0.73). This is a significant finding, as previous research has found that the vision of nAMD participants is greatly and adversely affected if definitive diagnosis and treatment initiation are delayed. Parfitt *et al.* showed that a large proportion of participants (72%) with any eye condition, including nAMD, had experienced a permanent reduction in VA due to service-related delays. Additionally, their study found that 31% of AMD cases were missed, of which > 50% were by optometrists.[117] Several factors may contribute to the lower observed accuracy of referrals from CO, such as an individual practitioner's years of experience and their ability to correctly identify clinical signs.[19,20,118] A prospective study evaluating the quality of urgent eye referrals, via use of a rapid access form, found that only 37% of participants were correctly diagnosed with nAMD due to difficulties with recognising clinical signs such as drusen and subretinal fluid.[119] In the HERMES study, the overall rate of false-negative referrals for retinal disease was low in both study arms. For urgent, vision-threatening nAMD disease referrals, however, false-negative referral rates were significantly lower for the teleophthalmology arm. This indicates that safety is also an important consideration, specifically for urgent nAMD referrals from community optometry to HES in the UK healthcare setting, likely associated with variability in the level of experience and training of CO for the correct interpretation of OCT scans. Efficiency, however, as indicated by the high rate of false-positive referrals by the community optometry (standard of care) arm in HERMES, is a more pronounced and confirmed challenge for current community optometry referral pathways, and a significant source of compounded capacity pressures for HES, thus justifying our choice of primary outcome for HERMES being the rate of false-positive referrals.

A teleophthalmology triage platform gives an opportunity to help COs strengthen their clinical skills in the form of providing feedback on their referrals. This valuable interaction is not always fulfilled with standard methods of referral. A review had found the referral reply rate to CO had varied from 13% to 16% for different reasons, such as a lack of contact details on a referral form and limited communication with GPs.[34] In addition to the commonly used GOS18 referral forms and letters, EeRS is also available for some COs to use. Its use as a potential teleophthalmology referral triaging model was taken up in certain regions of England; however, it did not fulfil all the characteristics of an optimised efficient teleophthalmology referral pathway, such as the one implemented in the HERMES study. This included both the consistent use of full-volume OCT scans and bi-directional feedback channels between referrers and hospital-based experts as well as tight turnaround times of 2 business days between referral and triaging decisions.

It is important to consider the positive and negative effects that discussed interventions may have on optometrists and/or optometry practices. One positive impact of implementing teleophthalmology referral pathways is the potential for an improved interaction between primary and secondary care. In the current standard-of-care referral pathways, a clinic letter is written by assessing healthcare professionals following hospital consultations, which summarises the

appointment findings, but this is usually addressed to the general medical practitioner only. Early studies found that referral reply rate to optometrists, either through direct reply or by copying-in, varied from 13% to 16%[79,80] due to general medical practitioners not always including the optometrists' contact details on referral letters to hospital-based services and considering CO are transient care providers.[81] Consistent feedback to referring optometrists from hospital-based experts, as a fixed feature of teleophthalmology referral models, such as the one evaluated in the HERMES study, could both keep optometrists up to date with patient outcomes and informed on whether/when they should expect to see referred participants again in their practices, and could also act as a learning aid for future management of similar cases, enabling gradual upskilling and improved referral decisions by CO.

Additionally, a sensitivity analysis was conducted to assess the time from referral to consultation for the intervention arm only. We were able to accurately trace and identify the external factors that resulted in a delay for each patient, such as the Trial Co-ordinator receiving the patient details from the CO late, or limited clinic capacity. As these factors were not due to the failings of the teleophthalmology referral platform, they were excluded, and the effectiveness was reassessed. We were unable to conduct the same analysis for the control arm, as it was not possible to reliably track all external factors which could have led to a delay in the standard care pathway, for example, the patient receiving their referral letter from their CO late, the patient handing the referral letter to their GP late and the GP not forwarding the hard copy or e-mailed referral letter to HES in a timely manner. Apart from the very few referrals which were sent via EeRS, it was not possible to accurately track the referral process for most control arm participants.

A sensitivity analysis was also conducted for the outcome measuring the time from referral to treatment. Although the mean time to consultation was significantly lower in the intervention arm [53 (51 to 55) days] compared to the control arm [89 (87 to 91) days] (comparison using survival analysis: $p = 0.039$) in the post-exclusion analysis, this pathway could have been improved by directly booking participants into the required treatment clinic (e.g. an injection clinic). Unfortunately, this was not possible due to the already established HES policies and protocols. If implemented, similar to virtual clinics, the option of an additional telephone consultation could be arranged to discuss treatment options before providing participants with a face-to-face appointment.

Furthermore, an evaluation of the costs and benefits of the HERMES teleophthalmology referral pathway compared with the standard of care was conducted. In the CEA, the HERMES pathway had a greater effect at a lower cost than the standard pathway, meaning it was the dominant intervention. The results of the DCE showed that the public had a greater preference for a more effective intervention that could be delivered more quickly. Importantly, they had a greater preference for obtaining a correct diagnosis compared to a reduction in the waiting time. A CBA, based on the DCE results, demonstrated a net benefit for the HERMES pathway compared with the standard pathway of £992 for every patient seen. Deterministic and probabilistic sensitivity analyses were carried out to assess the robustness of the conclusions, and, in these, conclusions were found not to change. These results of the economic evaluations provide strong evidence for the efficiency of the HERMES pathway compared with the standard care pathway. These data are significant as previous studies reviewing different referral interventions, including asynchronous teleophthalmology, provide very limited information on this key topic. As summarised in *Figure 9* by Carmichael *et al.*, there are limited data of cost-effectiveness for almost all referral interventions.[34]

Lastly, when analysing the ability of each pathway to safely triage rare diseases, although not all cases were correctly diagnosed, each arm referred every case to HES for further investigation. It was difficult to ascertain if the teleophthalmology referral platform was more efficient than the standard pathway as the analysis was based on a very small number of participants: two from the intervention arm and six from the control arm. We also did not see any serious adverse events related to the study. Therefore, assisting optometrist referrals to the secondary/tertiary care setting through asynchronous teleophthalmology is safe and would help to manage a correct referral as well the referral urgency more appropriately, particularly for common, serious conditions such as nAMD. This potential pathway may also allow participants access to required care within the most appropriate time frame.

Artificial intelligence study

Among the optometry practices recruiting to the HERMES study, 204 out of 396 participants from 17 of the 29 participating CO practices were included, where the majority of participants were from 14 sites, 201 of the 204

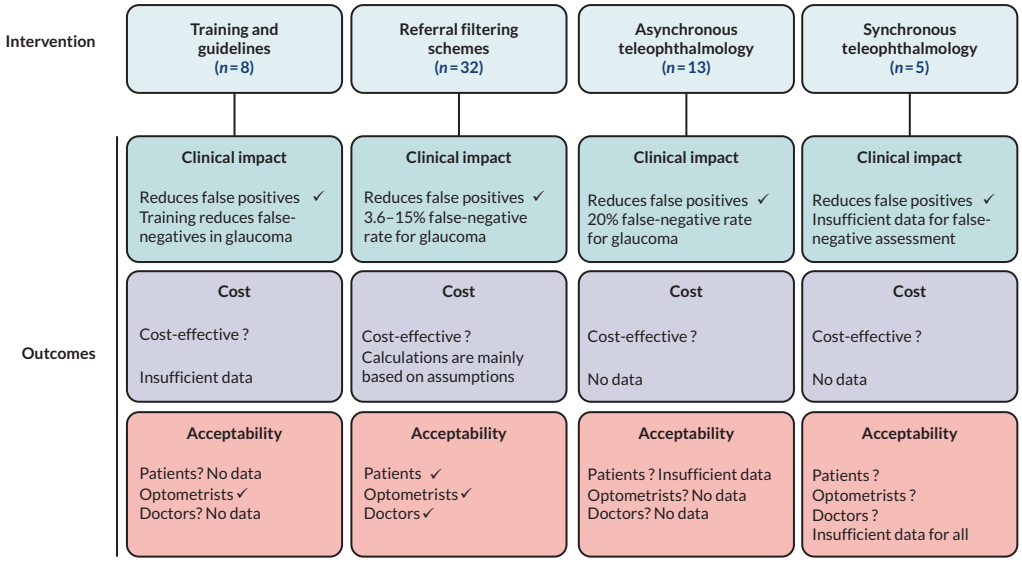

FIGURE 9 A summary of outcome measures in relation to four different types of referral interventions. Outcomes which are supported by evidence are indicated with a '✓'. An outcome with a '?' indicates that it is not fully supported, or evidence is lacking. Reasons on why it is not fully supported are stated.

participants (98.5%); 192 participants were excluded from the AI study. Most (*n* = 179) were excluded because the corresponding OCT images were not suitable for analysis by the Octane AI. Out of the 204 included participants, 152 participants were imaged with OCT devices manufactured by Topcon and 52 were imaged with OCT devices manufactured by Nidek. The Octane model was trained primarily on data from Topcon OCT devices, with a smaller set of data from the Heidelberg Spectralis device. For this reason, Octane is trained and validated to process Topcon and Heidelberg Spectralis OCTs. Additionally, Octane has been trained to process OCT scans of specific dimensions, acquisition protocols and OCT file exporting formats. As a result, Octane did not process approximately 30% of the Topcon OCT scans collected in the HERMES study due to their dimensions (mostly 'wide-field' scans capturing a larger area of the retina) and scan acquisition protocols (mostly 'sparse' OCT volumes with fewer slices/b-scans per volume). The diversity of OCT devices and scan acquisition protocols (scan dimensions and density) encountered in the real-world CO landscape, and reflected in the HERMES data, were significant but reflective of contemporary practice.

The Octane model showed a good referral accuracy compared to the rule-based RS for routine/urgent referral versus no referral and a good diagnostic accuracy for urgent referral versus routine/no referral, with a modest decrease in pathway efficiency for urgent referrals. This suggests that the model shows good performance, when its performance is assessed on the basis of the predetermined referral rules set for the model. The referral accuracy of Octane decreased compared with the clinical RS. This suggests that the model shows inferior performance when compared to real-life referral decisions made by hospital-based human experts and shows comparable performance to community optometrists making real-life clinical decisions. Some of the reasons for this discrepancy can be addressed with simple (e.g. adjusting preset rules for AI referral decisions) or more involved refinements (e.g. incorporating clinical history as input for AI referral decisions) of the AI model. A limitation of the Octane model is its narrow application with respect to OCT devices and the image formats it supports, which is exacerbated in HERMES due to the wide diversity in devices and image acquisition patterns observed in community optometry practices.

It is known that AI clinical support models often report a higher performance accuracy in retrospective validation studies,[120] compared to the ones reported when the same models are exposed to the diversity and noise of real-world prospective validation studies. This issue of generalisability of AI clinical support models is exacerbated in the context of ophthalmic imaging data due to the high prevalence of proprietary file formats specific to the different device manufacturers. This is a particular and well-recognised challenge for OCT data.[121] Device manufacturers have recently taken steps towards addressing this issue by enabling data export in open-source file formats (.dicom, .tiff, .jpeg and.png files). For example, the Heidelberg Spectralis data management and viewing software (HEYEX) (Heidelberg Engineering, Heidelberg, Germany) has evolved from HEYEX1 to HEYEX2, the latter incorporating a .dicom file export functionality.

At the same time, ophthalmic imaging research groups in clinical AI have developed bespoke solutions for open-source imaging data file conversion from the respective proprietary data file formats (such as .fda files for Topcon and .e2e files for Heidelberg), circumventing the above limitations.[122]

The HERMES study contributed significant new insights into this technical limitation for OCT file formats. For example, the HERMES research team (SA and BK) was able to identify and provide guidelines for the .dicom file export process of Nidek OCT scans. Nidek OCT scans were previously not considered amenable to AI-enabled automated processing due to the perceived lack of a .dicom file exporting functionality. It became evident that the Nidek .dicom file format has significant similarities with the Topcon .dicom file format (e.g. the same OCT scan density of 128 slices/b-scans per OCT volume). As a result, the Octane model was able to process and demonstrate comparable diagnostic and referral accuracy for Nidek scans to that reported for Topcon scans in the Octane retrospective validation study. This is likely because images are acquired with a similar protocol and dimensions between the two manufacturers. Despite this, it is clear that there are some differences in the reflectivity of the retinal layers with Nidek scanning, which is different from what Octane was trained on and which occasionally leads to erroneous results. The AI research team concluded that it was reasonable to include the Nidek scans in the HERMES data analysis, with the following caveats: (1) this was not pre-specified in the study protocol; and (2) further fine-tuning of the Octane model would be required to obtain optimal results for Nidek scans. A subanalysis of the HERMES Topcon scans alone, excluding HERMES Nidek scans, showed comparable diagnostic accuracy performance of the model. Of note, Nidek devices are almost ubiquitous at Specsavers practices that represent a significant proportion of the overall number of community optometry practices with OCT equipment.

Heidelberg Spectralis scans have specifications that are quite distinct from the Topcon scans. Heidelberg devices are widely used in CO practices and hospital-based ophthalmology departments globally. Hence, the prospective, real-world evaluation of AI DSS performance of Heidelberg scans should have been an important outcome of the HERMES study. This is particularly relevant for Octane, as it was partially trained and validated on Heidelberg scans in its retrospective validation study. There were 75 participants recruited in the HERMES study via CO practices using the Heidelberg Spectralis device. Unfortunately, the Octane AI was not able to process these scans. The reasons for this were: (1) the technical challenges for the conversion of the proprietary Heidelberg .e2e file format to the open-source .dicom file format were identified by the Octane AI team after completion of scan exporting and transfer from CO practices to Moorfields for AI processing; (2) an open-source file exporting functionality is available in the Heidelberg data management software (HEYEX), but CO practices were not advised to use it by the Octane AI team. Although these OCT scans could not be processed by the AI model, all available OCT scans from participants recruited into HERMES were of sufficient quality and specification was to be reviewed by human experts in the teleophthalmology arm (GG, CD, EM, AM) and for determining the reference standard in both arms (KB, AS).

Accuracy of the Octane model for making referral recommendations and retinal diagnoses needs to, therefore, be considered under the above caveats. The reduced sample of 204 patients may reduce the precision of the diagnostic accuracy estimates, leading to wider CIs for each measure. There should be minimal effects on the estimates of diagnostic accuracy other than the effect on precision.

In the simulated scenario where human assessors were replaced by AI in the HERMES study, it was observed that the AI model would, on occasion, provide a referral recommendation that would differ from the one given by human assessors and, importantly, would also differ from the clinical RS, which was used to produce this secondary outcome. The AI model chose the wrong referral urgency a comparable number of times as assessors in the control arm (referral decisions made by community optometrists). The proportion of wrong referral urgencies made by the AI model was higher than that seen in the intervention arm (referral decisions made by hospital-based experts), with significant overlap of CIs. This result is interesting as it signifies that AI would not reduce the high rate of unnecessary referrals from community optometry seen in this study. On the other hand, it is also likely that AI would not adversely affect the efficiency of hospital expert-based teleophthalmology, which was shown to be significantly higher compared to community optometry in the cRCT reported here.

The referral accuracy of Octane compared with the rule-based RS shows high sensitivity and specificity. This suggests that Octane performs well against its own preset referral recommendation rules. When Octane recommends a referral,

there is a 100% probability that a referral is truly needed (PPV); yet, when it recommends no referral, there is a 69.2% probability that a referral is truly not needed. With absolute numbers of no referrals being small in both pathways; however, these comparisons may not reflect actual differences. Unlike sensitivity and specificity, PPV and NPV are affected by prevalence, that is, PPV increases and NPV decreases due to high prevalence of cases actually requiring a referral. Performance of Octane remains good in identifying the need for an urgent referral, with a reduction in efficiency indicated by the PPV of 78.2% when compared with hospital-based human experts. This suggests that Octane would recommend an unnecessary urgent referral in 21.8% of cases.

Octane performance for referral recommendations decreases when compared to the clinical RS. There are several reasons as to why hospital eye specialists within a teleophthalmology pathway, in particular, may have made different decisions than the AI when determining a referral recommendation and its urgency, yet both recommendations can be clinically acceptable and safe. The AI model was trained using a set of rules, where the referral recommendations were designed to be overcautious to ensure safe practice, for example, all dry AMD cases (drusen and atrophy) led to an AI referral recommendation of routine referral. According to the referral guidance by National Institute for Health and Care Excellence [from 2013 – formerly from 2005 National Institute for Health and Clinical Excellence formerly the National Institute for Clinical Excellence] (NICE),[123] dry AMD referrals do not, in their majority, require a referral to HES. This is a source of discrepancy that does not necessarily indicate underperformance of the AI. The overcautious referral decision rules embedded in the AI classifier have an adverse effect on the apparent efficiency of autonomous AI decision-making. The encoded referral decision rules can be revised, thus eliminating this source of apparent discrepancy.

In other instances of real-life clinical practice, several additional factors would be considered when determining a patient management decision, such as patient preference and level of concern, systemic and ophthalmic comorbidities and only-eye status. Other indicative examples include: a participating CO decided not to refer an asymptomatic patient with mild dry AMD, as per NICE guidance. The same expert opted for routine referral of an advanced case of dry AMD for consideration of sight impairment registration. These additional pieces of clinical information are not reviewed by Octane, where the output is produced solely upon the review of an OCT scan. The reduced efficiency against the clinical RS is mostly associated with decisions for routine referral of non-vision threatening diagnoses, such as dry AMD and VRAs, such as ERM, VMT and partial thickness macular hole (PTMH). Arguably, such diagnoses would not necessarily require a referral to HES, especially in the UK healthcare setting. Presumably, an AI-assisted iteration of the referral pathway, where clinical history and information of patient–clinician interaction can be incorporated in the decision-making process, would improve AI performance in a real-life clinical setting. The modest reduction in pathway efficiency for urgent referral decisions against the rule-based RS, however, may represent a true reduction in efficiency, while pathway safety remains high. Safety is the primary consideration for urgent referrals, which correspond always to a diagnosis of CNV in the HERMES study, the only urgent (< 2 weeks referral-to-treatment time) vision-threatening macular diagnosis considered in HERMES.

The OCT scan quality is another factor of consideration for the role of AI DSS in real-life clinical care. Specific reasons which led to the exclusion of only 4% of OCT scans from the AI analysis in HERMES were poor resolution; missing b-scans; off-centred images; cases where a single b-scan was provided by the community optometrist rather than the full OCT volume; the lesion exceeding the OCT frame and foveal duplication.[124]

The main reasons for the exclusion of a significant proportion of available OCT scans, and of recruited participants into the study, from the AI analysis are the small number of supported OCT devices by Octane, the rigid OCT scan specifications suitable for Octane and technical difficulties, as previously described.

With respect to retinal diagnoses detected by Octane, the false-positive rate for retinal diagnosis, defined as diagnosis of a retinal condition by Octane that is different to the retinal condition diagnosed by human assessors, is comparable between the AI model and human assessors across both study arms. The false-negative rate, defined as finding no diagnosis (normal eye) or unclassified diagnosis by Octane, when human assessors have diagnosed a specific macular condition is higher for the AI model. Overall, the AI model is comparable at detecting the correct retinal diagnosis, when a specific retinal diagnosis is actually present, but it wrongly finds no diagnosis (normal case) or no specific diagnosis, when a specific diagnosis is actually present, more frequently compared to both hospital-based expert clinicians and community optometrists. The diagnostic accuracy results should be interpreted with some caution due to the lack of formally validated probability thresholds for some of the diagnoses detected by Octane (VRA, MO and CSCR). The

Octane AI team recently validated probability thresholds for two important diagnoses; dry AMD (drusen and atrophy) and CNV. CNV is an important diagnosis as it is the only diagnosis that triggers an urgent referral.[123] Diagnostic accuracy of Octane for CNV is good compared to both the clinical- and rule-based RS. Octane is safe for the diagnosis of CNV with a small increase in wrong CNV diagnosis (false positives) compared to hospital-based experts.

The role of clinical AI in disease diagnosis and hospital referral decisions for patients can be considered in two deployment models: AI autonomous and AI-assisted care pathways. A key benefit of AI-assisted systems, compared to AI autonomous systems, is that AI-assisted pathways have in-built quality assurance systems with a human in the loop for the review of ambiguous cases for diagnosis of eye conditions.[125,126] AI-assisted pathways do have additional challenges to AI autonomous pathways, where bias can still be created for the decision-making process using AI DSS.[126] In terms of the design of such systems, further work is needed to optimise decision-making for healthcare staff through appropriate HCI studies.

More recent advances in the rapidly evolving AI landscape and bespoke pre-processing techniques have enabled the development of AI DSS with broad generalisability, that is, the ability to handle diverse OCT file formats and devices. It should be highlighted that the HERMES study was designed to validate AI-enabled referral pathways between community optometry and hospital eye services through a methodology agnostic to the specific AI DSS used. In this sense, the HERMES study generated new insights and produced a robust methodological framework with transferability for the prospective, real-world performance evaluation of other clinical AI models, especially involving medical imaging.

Post implementation substudy

Following on from our pilot study, which demonstrated a potential for teleophthalmology to drastically reduce unnecessary referrals to HES, we aimed to further assess its effectiveness against standard care.[39] In response to the challenges presented by the COVID-19 pandemic, the NHS services underwent rapid and significant adjustments across the board. To minimise unnecessary hospital visits, teleophthalmology pathways (EeRS) were commissioned in some regions of England at speed, where a digital link to facilitate referrals between CO and HES was implemented. Greater Manchester was an early adopter of this approach and provided an opportunity to record the variation in the quality of health care within a local region and assess the real-life effectiveness and efficiency of the teleophthalmology referral pathway.

The post-implementation substudy faced significant challenges that prevented the high-volume patient recruitment that was originally envisaged. Seventeen participants from three community optometry practices were recruited into this exploratory arm of the HERMES study. The small patient sample was not sufficient to inform the safety subanalysis of real-life teleophthalmology referral pathways that were originally envisaged.

The post-implementation substudy was, however, highly informative on the implementation science front. Feedback provided by the local research team highlighted deficiencies in the implementation plan adopted for the EeRS pathways under the extraordinary circumstances of the COVID-19 pandemic, presented in the *Chapter 3, Results*. Pitfalls and key requirements for success of teleophthalmology pathways derived from this pilot implementation can inform planning of future initiatives.

The HERMES teleophthalmology pathway, on the other hand, was designed to enable and facilitate the transfer of full-volume OCT scans, with early and consistent availability of training and support of participating community optometry practices in the relevant tasks, thus enabling the safe and accurate triaging of retinal referrals by hospital-based experts.

Limitations

The study has certain limitations.

Limitations relating to the Octane model are discussed in *Chapter 6, Discussion*.

We did not capture ethnicity data within the study eCRF, as our engagement with community optometry stakeholders during study design indicated potential reluctance of community optometrists to collect these data. This did not allow the evaluation of generalisability of the study to different ethnic groups. Given, however, the wide distribution of participating community optometry practices and the selection of secondary care NHS sites from urban, suburban and rural areas, there is a valid expectation that an ethnically, socioeconomically and culturally diverse patient population was recruited into the study. We collected data through a diverse, multicentre collaboration for both the cRCT and the AI substudy.

Baseline characteristics at the patient level showed a lower number of smokers within the study than expected in a UK population. This may be due to the genuine differences between our sample and real world; however, we feel it may be due to the underdocumentation of smoking status by optometrists involved in the study.

Lastly, there are practical barriers to the implementation of the dual technologies of teleophthalmology and AI DSS in the healthcare setting. We had practical issues of availability of suitable hardware/software infrastructure, such as basic wireless internet connections, during the study process. There were also issues identified with HCI as well as potential for bias created for healthcare staff when using AI DSS.[126] Further work is required to optimise the use of AI DSS, in conjunction with clinical input, for all these scenarios in the clinical setting.

Patient and public involvement

Patient and public members were consulted prior to the trial protocol design. This process was organised and delivered by members of the clinical team piloting the teleophthalmology pathway at the MEH, Croydon. Questionnaires were given to 18 participants, and the perception of teleophthalmology and issues of data privacy, impersonal care, trust in technology and confidence in the quality of care provided through digital means were explored.

Surveyed participants were positive that teleophthalmology could reduce waiting times, curtail unnecessary hospital visits and alleviate anxiety from prolonged uncertainty around diagnosis. When asked what measures would help gain trust in the teleophthalmology clinics, several participants placed an emphasis on the need for clear and detailed information of the pathways, the experience to be expected during their visit, the timescale for obtaining feedback and a point of contact for questions. These comments were considered during the preparation of patient information material, and these helped to ensure it was appropriate and suitably worded to inform on the pathways while alleviating any concerns. Patient input also helped reinforce the importance of including the comprehensive qualitative element to the study through the HCI analysis to capture patient perceptions around digital models of care.

A patient and public involvement (PPI) group based at the MEH was also consulted during the pre-application phase to inform on the development of patient-facing material, including patient information and consent forms for both the main study and the dedicated qualitative study (HCI).

During the development phase of the study, the COVID-19 pandemic presented difficulties to the PPI activity, prompting more extensive use of digital means of engagement and communication. The patient and public representative on the TSC, Geraldine Hoad (from the Macular Society), who also offered the perspective of eye charities, reviewed the PIS and the informed consent form as part of the TSC meeting preparatory material. In addition, the PPI contributors provided advice on barriers to recruitment, issues surrounding geographical spread of optometry sites and obtaining informed consent. Their input was critical in finalising the study protocol.

Following the commencement of the study, PPI activity mainly continued via the equality, diversity and inclusion (EDI) advisors' group. Recommendations and actions have been detailed below. Additionally, the PPI representative on the TSC, representing the Macular Society, the largest UK-based eye charity, provided valuable input into matters of barriers to recruitment, and amendments to the study protocol, ensuring the core patient-centric objectives of the HERMES study are continuously pursued.

Patient and public involvement contributors approached via the dedicated Moorfields and University College London (UCL)/Institute of Ophthalmology National Institute for Health Research (NIHR) Biomedical Research Centre (BRC) patient and public involvement and engagement (PPIE) team were invited to review the topic guide for the HCI patient and practitioner interviews. We also sought advice from PPI contributors with respect to proposed adjustments to study design and particularly the proposed increase in the number of NHS sites and community optometry practices participating in the study and their geographical distribution. PPI input has been critical for navigating the study through the challenging period of the COVID-19 pandemic and its aftermath and for making appropriate adjustments to study design that ensured the viability of the study and enhanced the value of expected output from the study.

An end-of-study debrief is scheduled with all PPI contributors and the PPI Representative on the TSC, which will include discussion of the prioritisation and dissemination of study results both to the public and relevant healthcare professionals. We will leverage the extensive dissemination and communications network of the UK's leading eye charity, the Macular Society, to ensure widespread and effective dissemination of our findings to participants and carers, the public, NHS stakeholders and policy decision-makers.

Equality, diversity and inclusion

The EDI advisors group at the NIHR BRC at MEH and UCL Institute of Ophthalmology provided useful suggestions on how EDI could be considered in the study. The EDI group highlighted several at-risk patient groups and causative factors for exclusion and provided several well-informed recommendations. The research team, within the constraints of its resources, partially implemented the recommendations whenever possible.

Several actions which met the EDI criteria were conducted. We thoroughly considered factors, such as indices of social deprivation in areas, where new candidate practices were located as well as ethnic and cultural diversity of local populations. We had also considered various geographical locations of the primary care sites, which enabled participants from different ethnic backgrounds to participate in the study. One of the participating NHS sites was deliberately selected outside of large urban centres in North West Anglia Trust, and an associated community optometry practice referring participants to the NHS was selected in a rural location within the wider catchment area of the corresponding NHS Trust. Additionally, participants were not selected on the grounds of their understanding of any digital equipment, thus eliminating the factor of digital exclusion.

During the selection of optometry sites, the EDI had recommended that all candidate community optometry practices should be assessed on factors, such as potential communication barriers with local populations, their awareness of the ethnic, linguistic and cultural diversity of the people visiting their practice as well as the availability of essential language skills within the practice, to ensure effective communication of the study. Although a great effort was made to implement this, relevant specialist training to undertake risk assessments for EDI under diverse real-life circumstances would have helped to ensure that productive and structured discussions had taken place.

Furthermore, engagement to encourage all participating community optometry sites to take initiatives for identifying underserved groups in their local areas was suggested. This was an excellent recommendation; however, the business model of community optometry practices makes it very challenging for CO to engage in additional activities and dedicate time away from direct patient care, especially without reimbursement. There is a pressing need for specialist training of research and all clinical staff in an array of skills, domain knowledge and communication techniques to successfully design and deliver initiatives involving underserved groups, such as the homeless, persons with mental health issues, learning disabilities as well as different cultural norms and communication codes. This should be accompanied by the development of dedicated guidance documents for the diverse care needs of all people in a comprehensive and effective EDI policy for clinical care and research.

The research team were also recommended to conduct targeted engagement with support groups, charities and networks specialising in improving the quality of life of vulnerable, at-risk groups. Also, the development of communication material in different languages and formats as well as efficient strategies for distribution and dissemination was recommended. However, both suggestions could not be executed due to prioritisation of

recruitment, limited time and resources. An establishment of a dedicated EDI manager role and the investment in greater resources in terms of administrative support would have helped to achieve this.

Impact and learning

The impact and dissemination programme leverages NHS digital innovation initiatives in the area of eye care. Also, in response to the requirement to address the pressing priority for NHS to innovate, in particular through introduction of digitally enabled care pathways, in order to address the currently experienced capacity pressures, with large appointment backlogs in ophthalmology, Moorfields Eye Hospital (under the leadership of the Trust's Director of Digital Innovation) is piloting a Centralised Telemedicine Services programme for London Integrated Care Systems, supported by NHS England. Within this programme, referral refinement is the primary use-case and priority objective, as proof of concept studies have demonstrated the potential for teleophthalmology to reduce observed delays and barriers in access to eye care. In this context, evidence from HERMES will be influential for NHS policy decisions on whether to expand such teleophthalmology pathways nationally or not. An additional area of significant impact of the HERMES study is that it provided the possibility, for the first time, to develop and sustain a Community Optometry Research Networks. Acting as a starting point and catalyst, the research-active cohort of community optometry practices that participated in the HERMES study motivated and promoted synergistic initiatives, like the Ulster University project for the development of research active community networks led by Padraig Mulholland (Moorfields Optometry Education Lead) and assisted by our exceptionally capable HERMES Assistant PI Anitta Sharma (Moorfields Specialist Optometrist). These Community Optometry Research Networks will be essential for the propagation and successful implementation of teleophthalmology (and potentially AI-assisted) referral refinement models in real-life clinical practice between the NHS (Hospital Eye Services) and non-NHS sectors (community optometry practices).

With the invaluable contribution of our Charity Partner, the Macular Society, as well as the Moorfields Eye Charity and PPIE Networks, the visibility and dissemination of HERMES study findings have been further enhanced.

An extensive network of experienced academics, innovating clinical leaders, and NHS policy stakeholders has been established. This highly engaged group is actively working on the design of follow-up research work to complement the evidence base on AI-enabled clinical care. Evidence gaps relate to the scalability and safety of AI-enabled care pathways on implementation science evidence on human factors, human–AI Interaction, organisational and cultural adaptations and barriers and enablers to real-life deployment of AI-enabled eye care.

Research recommendations

(1) The cRCT: While the teleophthalmology arm demonstrated a clinically significant reduction in the time to treatment, research on further streamlining the process, whereby patients are directly assessed in a treatment clinic following confirmation of activity via the teleophthalmology pathway, could be conducted.

(2) AI: 'Secondly' RCTs are the gold standard for safety and efficacy assessment of medical technology. The fast pace of AI technological evolution challenges the regulatory landscape. Post hoc external validations often meet the evidence standard for regulatory approval of medical AI. The prospective, observational, real-world validation framework applied in the HERMES study generates highest quality evidence, sort of an RCT. Similar studies have greater feasibility. HERMES demonstrated their necessity for regulatory approval of medical AI. Further studies in the same and other high-volume ophthalmic pathologies and AI models can benefit from the HERMES methodological template.

(3) AI: Collaborative research initiatives for the development and validation of medical AI for eye care, leveraging the transition from discriminatory AI to generative AI; research into novel clinical trial and validation methodologies ('smart trials') so that technological evolution keeps pace with evidence generation of its safety and efficacy.

(4) AI: Health services research would allow evaluating themes on medical AI implementation:

human–AI interaction
effect of AI on clinical decision-making by human experts (automation bias)
role of AI in upskilling or deskilling eye care specialists
AI-enabled clinical pathway redesigning, AI-assisted versus AI autonomous clinical pathways
Organisational-, cultural- and human factor-related enablers and barriers to the clinical deployment of AI and modelling of service efficiencies and cost-effectiveness of the deployment of clinical AI in the eye care referral process.

(5) AI: Prospective validation studies (or RCTs) focusing on the two other major aspects of the nAMD patient care journey; the period of active treatment with intraocular injections and the period of post-treatment monitoring of participants for disease reactivation. The latter prospective study (community optometry AI-assisted monitoring of participants with stable/post-treatment nAMD to detect potential reactivation) should be informed by the study reported here and the, soon to report, NIHR Health Technology Assessment Programme FENETRE Study (addressing the specific use-case of post-treatment monitoring with regular visits to community optometry for repeat OCT scans).

Conclusions

The HERMES study is the first interventional, real-life, RCT to evaluate teleophthalmology referral pathways, in a robust, statistically powered study design. Asynchronous teleophthalmology was shown definitively to reduce the number of unnecessary referrals/hospital visits, and robust evidence was provided for the significant superiority of teleophthalmology in reducing the number of erroneous urgent referrals for serious, vision-threatening disease, like nAMD. Teleophthalmology referral triaging pathways were also shown to provide more appropriate timescales for referral across the board and some positive downstream effects on nAMD treatment pathways by reducing participants' waiting time from referral to NHS hospital consultation and treatment, compared to standard of care. These are powerful findings which suggest teleophthalmology is an effective enabler of digital eye services transformation, enabling the delivery of ophthalmology triaging services more efficiently, particularly for erroneous urgent referrals. The teleophthalmology pathway was shown to be highly cost-effective, with the intervention (teleophthalmology referral triaging pathway) characterised as 'dominant' in the economic evaluation, a rare finding in health economics, indicating significantly greater cost-effectiveness of teleophthalmology in all the economic evaluation models used to evaluate this referral pathway in the HERMES study.

The HERMES study additionally implemented a robust methodological framework for the performance evaluation of an AI-enabled referral triaging pathway for retinal referrals in a prospective, observational, real-life study design. It generated new insights into the challenges of real-life deployment of clinical AI and highlighted the important issue of AI generalisability, especially in real-world clinical settings.

Teleophthalmology was shown to be a safe, efficient, acceptable, technically and logistically feasible referral pathway for filtering the unnecessary urgent referrals, which disproportionately disrupt hospital eye services. These findings indicate the positive impact of adoption of teleophthalmology referral pathways. Clinical AI integration of the Octane model in eye referral pathways shows promise in terms of safety, but our study highlighted certain technical and logistical challenges as well as potential routes for addressing these.

Additional information

CRediT contribution statement

Anitta Sharma (https://orcid.org/0009-0001-5133-2844) (Associate Principal Investigator, Optometrist): Data curation (equal), project administration (equal), Supervision (equal), Validation (equal), Visualisation (equal), Writing - original draft (lead), Writing - review and editing (equal).

Rima Hussain (https://orcid.org/0009-0005-9796-5486) (Trial Project Manager): Project administration (lead), Resources (equal), Supervision (equal).

Annastazia E Learoyd (https://orcid.org/0000-0001-6964-9041) (Senior Clinical Trials Statistician): Formal analysis (equal), Methodology (equal), Software (equal).

Angela Aristidou (https://orcid.org/0000-0002-0290-4244) (Associate Professor): Investigation (equal), Supervision (equal), Writing - review and editing (equal).

Taha Soomro (https://orcid.org/0000-0003-2756-3674) (Ophthalmologist): Data curation (equal), Writing - original draft (equal).

Ann Blandford (https://orcid.org/0000-0002-3198-7122) (Professor of Human Computer Interaction): Conceptualisation (equal), Formal analysis (equal), Methodology (equal), Supervision (equal), Visualisation (equal), Writing - review and editing (equal).

John G Lawrenson (https://orcid.org/0000-0002-2031-6390) (Professor of Clinical Visual Science, Optometrist): Conceptualisation (equual), Methodology (equal), Writing - review and editing (equal).

Gabriela Grimaldi (https://orcid.org/0000-0003-2435-7867) (Ophthalmologist): Investigation (equal), Supervision (equal).

Abdel Douiri (https://orcid.org/0000-0002-4354-4433) (Professor of Medical Statistics): Methodology (equal).

Ashleigh Kernohan (https://orcid.org/0000-0002-5514-3186) (Senior Research Associate, Health Economics): Data curation (equal), Methodology (equal), Visualisation (equal), Writing - review and editing (equal).

Tomos Robinson (https://orcid.org/0000-0001-8695-9738) (Senior Research Associate, Health Economics): Data curation (equal), Formal analysis (equal), Methodology (equal).

Najmeh Moradi (https://orcid.org/0000-0003-4172-4411) (Research Associate, Health Economics): Data curation (equal), Formal analysis (equal).

Christiana Dinah (https://orcid.org/0000-0002-0815-4771) (Principal Investigator and Consultant Ophthalmologist): Investigation (equal), Supervision (equal), Writing - review and editing (equal).

Evangelos Minos (https://orcid.org/0000-0002-0033-134X) (Principal Investigator and Consultant Ophthalmologist): Investigation (equal), Supervision (equal), Writing - review and editing (equal).

Dawn Sim (https://orcid.org/0000-0002-6363-7805) (Consultant Ophthalmologist): Methodology (equal).

Tariq Aslam (https://orcid.org/0000-0002-9739-7280) (Principal Investigator and Consultant Ophthalmologist): Methodology (equal), Writing - review and editing (equal).

Avinash Manna (https://orcid.org/0009-0005-2227-8550) (Principal Investigator and Consultant Ophthalmologist): Investigation (equal), Supervision (equal).

Alastair K Denniston (https://orcid.org/0000-0001-7849-0087) (Honorary Professor and Consultant Ophthalmologist): Formal analysis (equal), Methodology (equal), Supervision (equal), Writing - review and editing (equal).

Praveen J Patel (https://orcid.org/0000-0001-8682-4067) (Consultant Ophthalmologist): Data curation (equal), Supervision (equal), Writing - review and editing (equal).

Pearse A Keane (https://orcid.org/0000-0002-9239-745X) (Professor of Artificial Medical Intelligence and Consultant Ophthalmologist): Data curation (equal), Validation (equal), Writing - review and editing (equal).

Catey Bunce (https://orcid.org/0000-0002-0935-3713) (Principal Statistician): Conceptualisation (equal), Data curation (equal), Methodology (equal), Supervision (equal), Writing - review and editing (equal).

Luke Vale (https://orcid.org/0000-0001-8574-8429) (Professor of Health Economics): Conceptualisation (equal), Formal analysis (equal), Methodology (equal), Supervision (equal), Visualisation (equal), Writing - review and editing (equal).

Konstantinos Balaskas (https://orcid.org/0000-0003-2034-8920) (Professor of Artificial Medical Intelligence, Chief Investigator of the HERMES trial and Consultant Ophthalmologist): Conceptualisation (lead), Data curation (lead), Formal analysis (lead), Funding acquisition (lead), Methodology (lead), Project administration (equal), Resources (equal), Software (equal), Supervision (equal), Validation (equal), Visualisation (equal), Writing - original draft (equal), Writing - review and editing (lead).

Acknowledgements

We thank the NIHR Health Technology Assessment programme for providing funding (NIHR127773) and MEH for sponsoring the trial. We thank the NIHR Biomedical Research Centre based at MEH, the NIHR North Thames Clinical Research Network and UCL Institute of Ophthalmology for their support.

The HERMES Research Team members also wish to extend their gratitude to the following members of the HERMES Trial Advisory Committees, PPI and EDI advisors' group for their expert advice and guidance to study planning and delivery:

TSC: Chair: Professor Richard Gale, York and Scarborough Teaching Hospitals NHS foundation Trust. Members: Mrs Irene Stratton, University of Southampton. Dr Bushra Mushtaq, Sandwell and West Birmingham NHS Foundation Trust. Dr Gabriella De Salvo, University of Southampton. Geraldine Hoad, the Macular Society.

DMC**:** Chair: Dr Louise Downey, Hull University Teaching Hospitals NHS Foundation Trust. Members: Dr Gabriela Czanner, Liverpool John Moores University. Dr Markus Groppe, Buckinghamshire Healthcare NHS Trust.

Moorfields PPI and BRC EDI advisors' group: Ms Jocelyn Cammack, Ms Helen Baker, Ms Subhojit Chakraborty, Mrs Geraldine Hoad (the Macular Society).

The authors would like to specially thank all HERMES trial participants. We also wish to express sincere gratitude to all recruiting community optometry sites and their corresponding NHS sites for their valuable contribution:

Moorfields Eye Hospital NHS Foundation Trust

Research Team: Konstantinos Balaskas, Rima Hussain, Anitta Sharma, Taha Soomro, Praveen Patel, Gabriela Grimaldi, Bhavna Ramchurn, Matilda Fenn, Taniya Chowdhury, Richard Seeberan, Wen Xing, Daniela Florea, Ana Paula Ribeiro Reis, Alan Sousa, Gunjan Naik, Ismail Moghul, Nikolas Pontikos

Community optometry sites: A View Opticians, Auerbach and Steele, Boots Opticians Camden, Boots Opticians Wimbledon, Day Lewis Opticians, Eyevalue Opticians, Gavzey Opticians, Hing Opticians, Hynes Optometrists, Parker and Hammond (Northwood & South Harrow), Rose Opticians, Specsavers Camden, Specsavers Wimbledon

University Hospitals Birmingham NHS Foundation Trust

Research Team: Avinash Manna, Ramesh Sivaraj, Sobha Joesph, Andrej Kidess, Nick Capewell, Donna Holding, Ray Rayhan

Community optometry sites: Albright Opticians, Alan Tyler Opticians, Charles Morris Optometrists, Eye Opticians

Central Middlesex Hospital at London North West University Healthcare NHS Trust

Research Team: Christiana Dinah, Arevik Ghulakhszian, Laasya Mandalapu, Jennifer Maynard

Community optometry sites: DF Optometrists, Insight Opticians, Specsavers Wembley, The Eyecare Studio, Visioncare Medical Eye Centre, Visual Eyes Opticians

North West Anglia NHS Foundation Trust

Research Team: Evangelos Minos, Hani Hasan, Athanasios Georgas, Catherine Wilshire, Sarah Tester, Ionela Sinanovic

Community optometry sites: Harrison Eye Clinic, Krog and Whitehead Optometrists, Werrington Opticians

Manchester University NHS Foundation Trust

Research Team: Tariq Aslam, Laura Berg, Fiona Crawford

Community optometry sites: Leeming Opticians, Specsavers Glossop, Wallwork Opticians

Patient data statement

This work uses data provided by participants and collected by the NHS as part of their care and support. Using patient data is vital to improve health and care for everyone. There is huge potential to make better use of information from people's patient records, to understand more about disease, develop new treatments, monitor safety and plan NHS services. Patient data should be kept safe and secure, to protect everyone's privacy, and it's important that there are safeguards to make sure that they are stored and used responsibly. Everyone should be able to find out about how patient data are used. #datasaveslives You can find out more about the background to this citation here: https://understandingpatientdata.org.uk/data-citation

Data-sharing statement

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

Ethics statement

Ethical approval was obtained on 26 January 2021 from the London-Bromley Research Ethics Committee (REC 20/LO/1299). Health Research Authority approval was also received on the same date.

Information governance statement

Moorfields Eye Hospital NHS Trust 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, Moorfields Eye Hospital NHS Trust 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 (www.moorfields.nhs.uk/content/privacy-notice-nhs-participants).

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/QNDF3325.

Primary conflicts of interest: Dr Pearse Keane has acted as Consultant for Google DeepMind.

Disclaimer

Every effort has been made to obtain the necessary permissions for reproduction, to credit original sources appropriately and to respect copyright requirements. However, despite our diligence, we acknowledge the possibility of unintentional omissions or errors and we welcome notifications of any concerns regarding copyright or permissions.

Publications

Han JED, Liu X, Bunce C, Douiri A, Vale L, Blandford A, *et al.* Teleophthalmology-enabled and artificial intelligence-ready referral pathway for community optometry referrals of retinal disease (HERMES): a Cluster Randomised Superiority Trial with a linked Diagnostic Accuracy Study – HERMES study report 1 – study protocol. *BMJ Open* 2022;**12**:e055845. https://doi.org/10.1136/bmjopen-2021-055845

A second paper on the HERMES study detailed qualitative Human Computer Interaction analysis Protocol and methodology has been published in *BMJ Open*.

Blandford A, Abdi S, Aristidou A, Carmichael J, Cappellaro G, Hussain R, Balaskas K. Protocol for a qualitative study to explore acceptability, barriers and facilitators of the implementation of new teleophthalmology technologies between community optometry practices and hospital eye services. *BMJ Open* 2022;**12**:e060810. https://doi.org/10.1136/bmjopen-2022-060810

Sharma A, Hussain R, Learoyd AE, Aristidou A, Soomro T, Blandford A, *et al*. Teleophthalmology versus standard of care for community optometry referrals of retinal disease (HERMES): a cluster randomised controlled trial with linked prospective diagnostic accuracy assessment of artificial intelligence support. Lancet Primary Care 2025 [published online ahead of print 21 October 2025]. https://doi.org/10.1016/j.lanprc.2025.100041

A scientific presentation of the main outcomes of the HERMES study has been accepted to take place in May 2024, Seattle, USA, during the Association for Research in Vision and Ophthalmology (ARVO) conference, the major annual international scientific conference in Ophthalmology.

References

1. Wong WL, Su X, Li X, Cheung CMG, Klein R, Cheng CY, Wong TY. Global prevalence of age-related macular degeneration and disease burden projection for 2020 and 2040: a systematic review and meta-analysis. *Lancet Glob Health* 2014;**2**:e106–16.
2. Hendrick AM, Gibson MV, Kulshreshtha A. Diabetic retinopathy. *Prim Care* 2015;**42**:451–64.
3. Wang Y, Zhong Y, Zhang L, Wu Q, Tham Y, Rim TH, *et al*. Global incidence, progression, and risk factors of age-related macular degeneration and projection of disease statistics in 30 years: a modeling study. *Gerontology* 2022;**68**:721–35.
4. NHS Digital. *Hospital Outpatient Activity 2022–23*. NHS Digital. URL: https://digital.nhs.uk/data-and-information/publications/statistical/hospital-outpatient-activity/2022-23 (accessed 18 January 2024).
5. Royal College of Ophthalmologists. *Facing Workforce Shortages and Backlogs in the Aftermath of COVID-19: The 2022 Census of the Ophthalmology Consultant, Trainee and SAS Workforce*. The Royal College of Ophthalmologists; 2023. URL: www.rcophth.ac.uk/wp-content/uploads/2023/03/2022-Ophthalmology-census-Facing-workforce-shortages-and-backlogs-in-the-aftermath-of-COVID-19.pdf (accessed 18 January 2024).
6. Resnikoff S, Lansingh VC, Washburn L, Felch W, Gauthier TM, Taylor HR, *et al*. Estimated number of ophthalmologists worldwide (International Council of Ophthalmology update): will we meet the needs? *Br J Ophthalmol* 2020;**104**:588–92.
7. Resnikoff S, Felch W, Gauthier TM, Spivey B. The number of ophthalmologists in practice and training worldwide: a growing gap despite more than 200,000 practitioners. *Br J Ophthalmol* 2012;**96**:783–7.
8. Foot B, MacEwen C. Surveillance of sight loss due to delay in ophthalmic treatment or review: frequency, cause and outcome. *Eye* 2017;**31**:771–5.
9. Lim JH, Wickremasinghe SS, Xie J, Chauhan DS, Baird PN, Robman LD, *et al*. Delay to treatment and visual outcomes in participants treated with anti-vascular endothelial growth factor for age-related macular degeneration. *Am J Ophthalmol* 2012;**153**:678–86, 686.e1.
10. Stone LG, Grinton ME, Talks JS. Delayed follow-up of medical retina participants due to COVID-19: impact on disease activity and visual acuity. *Arbeitsphysiologie* 2021;**259**:1773–80.
11. Thomas D, Weegen L, Walendzik A, Wasem J, Jahn R. *Comparative Analysis of Delivery of Primary Eye Care in Three European Countries*. IBES Diskussionsbeitrag; 2011. URL: www.econstor.eu/bitstream/10419/52720/1/671799908.pdf (accessed 20 January 2024).
12. Soroka M. The New York State optometry workforce study. *J Community Health* 2012;**37**:448–57.
13. Davey CJ, Green C, Elliott DB. Assessment of referrals to the hospital eye service by optometrists and GPs in Bradford and Airedale: quality of referrals to the hospital eye service. *Ophthalmic Physiol Opt* 2011;**31**:23–8.
14. Evans BJW, Edgar DF, Jessa Z, Yammouni R, Campbell P, Soteri K, *et al*. Referrals from community optometrists to the hospital eye service in England. *Ophthalmic Physiol Opt* 2021;**41**:365–77.
15. Looker HC, Nyangoma SO, Cromie DT, Olson JA, Leese GP, Black MW, *et al*.; Scottish Diabetes Research Network Epidemiology Group. Rates of referable eye disease in the Scottish National Diabetic Retinopathy Screening Programme. *Br J Ophthalmol* 2014;**98**:790–5.
16. Konstantakopoulou E, Harper RA, Edgar DF, Larkin G, Janikoun S, Lawrenson JG. Clinical safety of a minor eye conditions scheme in England delivered by community optometrists. *BMJ Open Ophthalmol* 2018;**3**:e000125.
17. Han JED, Liu X, Bunce C, Douiri A, Vale L, Blandford A, *et al*. Teleophthalmology-enabled and artificial intelligence-ready referral pathway for community optometry referrals of retinal disease (HERMES): a Cluster Randomised Superiority Trial with a linked Diagnostic Accuracy Study-HERMES study report 1-study protocol. *BMJ Open* 2022;**12**:e055845.

18. Cameron JR, Ahmed S, Curry P, Forrest G, Sanders R. Impact of direct electronic optometric referral with ocular imaging to a hospital eye service. *Eye (London, England)* 2009;**23**:1134–40.
19. Davey CJ, Scally AJ, Green C, Mitchell ES, Elliott DB. Factors influencing accuracy of referral and the likelihood of false positive referral by optometrists in Bradford, United Kingdom. *J Optom* 2016;**9**:158–65.
20. Wagner SK, Chopra R, Ledsam JR, *et al*. Diagnostic accuracy and interobserver variability of macular disease evaluation using optical coherence tomography. *Invest Ophthalmol Vis Sci* 2019;**60**:1849–1849.
21. Sandholdt CT, Jønsson ABR, Reventlow S, Bach-Holm D, Line K, Kolko M, *et al*. DETECT: DEveloping and testing a model to identify preventive vision loss among older participants in gEneral praCTice – protocol for a complex intervention in Denmark. *BMJ Open* 2023;**13**:e069974.
22. Muttuvelu DV, Buchholt H, Nygaard M, Rasmussen MLR, Sim D. Danish teleophthalmology platform reduces optometry referrals into the national eye care system. *BMJ Open Ophthalmol* 2021;**6**:e000671.
23. Tan O, Chen A, Li Y, Bailey S, Hwang TS, Lauer AK, *et al*. Prospective evaluation of optical coherence tomography for disease detection in the Casey mobile eye clinic. *Exp Biol Med (Maywood)* 2021;**246**:2214–21.
24. Sapru S, Berktold J, Crews JE, Katz LJ, Hark L, Girkin CA, *et al*. Applying RE-AIM to evaluate two community-based programs designed to improve access to eye care for those at high-risk for glaucoma. *Eval Program Plann* 2017;**65**:40–6.
25. Young K. *Specsavers Reaches OCT Installation Milestone of 1000 Devices*. URL: www.aop.org.uk/ot/industry/high-street/2022/09/29/specsavers-reaches-oct-installation-milestone-of-1000-devices (accessed 18 January 2024).
26. Dabasia PL, Edgar DF, Garway-Heath DF, Lawrenson JG. A survey of current and anticipated use of standard and specialist equipment by UK optometrists. *Ophthalmic Physiol Opt* 2014;**34**:592–613.
27. Toth CA. Optical coherence tomography and eye care. *N Engl J Med* 2023;**389**:1526–9.
28. Kern C, Fu DJ, Kortuem K, Huemer J, Barker D, Davis A, *et al*. Implementation of a cloud-based referral platform in ophthalmology: making telemedicine services a reality in eye care. *Br J Ophthalmol* 2020;**104**:312–7.
29. Nivison-Smith L, Milston R, Madigan M, Kalloniatis M. Age-related macular degeneration: linking clinical presentation to pathology. *Optom Vis Sci* 2014;**91**:832–48.
30. Ly A, Nivison-Smith L, Zangerl B, Assaad N, Kalloniatis M. Advanced imaging for the diagnosis of age-related macular degeneration: a case vignettes study. *Clin Exp Optom* 2018;**101**:243–54.
31. Kilduff CL, Thomas AA, Dugdill J, Casswell EJ, Dabrowski M, Lovegrove C, *et al*. Creating the Moorfields' virtual eye casualty: video consultations to provide emergency teleophthalmology care during and beyond the COVID-19 pandemic. *BMJ Health Care Inform* 2020;**27**:e100179.
32. Faes L, Bachmann LM, Sim DA. Home monitoring as a useful extension of modern tele-ophthalmology. *Eye* 2020;**34**:1950–3.
33. Korot E, Pontikos N, Drawnel FM, Jaber A, Fu DJ, Zhang G, *et al*. Enablers and barriers to deployment of smartphone-based home vision monitoring in clinical practice settings. *JAMA Ophthalmol* 2022;**140**:153–60.
34. Carmichael J, Abdi S, Balaskas K, Costanza E, Blandford A. The effectiveness of interventions for optometric referrals into the hospital eye service: a review. *Ophthalmic Physiol Opt* 2023;**43**:1510–23.
35. Blandford A, Abdi S, Aristidou A, Carmichael J, Cappellaro G, Hussain R, Balaskas K. Protocol for a qualitative study to explore acceptability, barriers and facilitators of the implementation of new teleophthalmology technologies between community optometry practices and hospital eye services. *BMJ Open* 2022;**12**:e060810.
36. Patel D, Abdi S, Carmichael J, Balaskas K, Blandford A. What are the experiences of teleophthalmology in optometric referral pathways? A qualitative interview study with participants and clinicians. *BMJ Open* 2024;**14**:e078161.
37. Esteva A, Chou K, Yeung S, Naik N, Madani A, Mottaghi A, *et al*. Deep learning-enabled medical computer vision. *NPJ Digit Med* 2021;**4**:5.

38. De Fauw J, Ledsam JR, Romera-Paredes B, Nikolov S, Tomasev N, Blackwell S, *et al*. Clinically applicable deep learning for diagnosis and referral in retinal disease. *Nat Med* 2018;**24**:1342–50.
39. Kortuem K, Fasler K, Charnley A, Khambati H, Fasolo S, Katz M, *et al*. Implementation of medical retina virtual clinics in a tertiary eye care referral centre. *Br J Ophthalmol* 2018;**102**:1391–5.
40. Agresti A, Coull BA. Approximate is better than 'exact' for interval estimation of binomial proportions. *Am Stat* 1998;**52**:119.
41. Drummond MF, Sculpher MJ, Claxton K, Stoddart GL, Torrance GW. *Methods for the Economic Evaluation of Health Care Programmes*. Oxford: Oxford University Press; 2015.
42. NHS England. *NHS England » 2021/22 National Cost Collection Data Publication*. URL: www.england.nhs.uk/publication/2021-22-national-cost-collection-data-publication/ (accessed 1 February 2024).
43. GOV.UK. *Letter Setting Out General Ophthalmic Services Fees from 1 April 2023*. GOV.UK. URL: www.gov.uk/government/publications/nhs-general-ophthalmic-service-fees-and-optical-voucher-values-from-april-2023/letter-setting-out-general-ophthalmic-services-fees-from-1-april-2023 (accessed 1 February 2024).
44. Miller L. *Optometry Wales Confirms Future Fees for Clinical Services*. URL: www.aop.org.uk/ot/professional-support/optical-organisations/2022/09/28/optometry-wales-confirms-future-fees-for-clinical-services (accessed 1 February 2024).
45. Ferris T. *General Ophthalmic Services (GOS): Directorate of Primary Care Dentistry, Optometry and Audiology Division, Scottish Government*. 2022. URL: www.sehd.scot.nhs.uk/pca/PCA2022(O)05.pdf (accessed 7 February 2024).
46. NICE. *Economic Evaluation | NICE Health Technology Evaluations: The Manual | Guidance*. NICE. URL: www.nice.org.uk/process/pmg36/chapter/economic-evaluation (accessed 1 February 2024).
47. Jones KC, Weatherly H, Birch S, Castelli A, Chalkley M, Dargan A, *et al*. *Unit Costs of Health and Social Care 2022 Manual*. London: PSSRU; 2023.
48. Pezzullo L, Streatfeild J, Simkiss P, Shickle D. The economic impact of sight loss and blindness in the UK adult population. *BMC Health Serv Res* 2018;**18**:63.
49. Ryan M, Gerard K, Amaya-Amaya M. *Using Discrete Choice Experiments to Value Health and Health Care*. 2008th edn. New York: Springer; 2006.
50. ChoiceMetrics. *Ngene 1.3 User Manual & Reference Guide: The Cutting Edge in Experimental Design*. ChoiceMetrics. URL: www.choice-metrics.com/features (accessed 7 February 2024).
51. Mariel P, Hoyos D, Meyerhoff J, Czajkowski M, Dekker T, Glenk K, *et al*. *Environmental Valuation with Discrete Choice Experiments: Guidance on Design, Implementation and Data Analysis*. 1st edn. Cham: Springer Nature; 2020.
52. Kuhfeld WF, Tobias RD, Garratt M. Efficient experimental design with marketing research applications. *J Marketing Res* 1994;**31**:545.
53. Dynata. *World's Largest First Party Data Platform*. Dynata; 2021. URL: www.dynata.com/ (accessed 1 February 2024).
54. Bridges JFP, Mohamed AF, Finnern HW, Woehl A, Hauber AB. participants' preferences for treatment outcomes for advanced non-small cell lung cancer: a conjoint analysis. *Lung Cancer* 2012;**77**:224–31.
55. Oluboyede Y, Ternent L, Vale L, Allen J. Using a discrete-choice experiment to estimate the preferences of clinical practitioners for a novel non-invasive device for diagnosis of peripheral arterial disease in primary care. *PharmacoEcon Open* 2019;**3**:571–81.
56. McFadden D, Train K. Mixed MNL models for discrete response. *J Appl Econometrics* 2000;**15**:447–70.
57. McIntosh E, Clarke P, Frew E, Louviere J. *Applied Methods of Cost-Benefit Analysis in Health Care*. Oxford: Oxford University Press; 2010.

58. Sharafeldin N, Kawaguchi A, Sundaram A, Campbell S, Rudnisky C, Weis E, *et al*. Review of economic evaluations of teleophthalmology as a screening strategy for chronic eye disease in adults. *Br J Ophthalmol* 2018;**102**:1485–91.

59. Mason T, Jones C, Sutton M, Konstantakopoulou E, Edgar DF, Harper RA, *et al*. Retrospective economic analysis of the transfer of services from hospitals to the community: an application to an enhanced eye care service. *BMJ Open* 2017;**7**:e014089.

60. Bajre M, Lim E, Pennington N, Rose J. Artificial intelligence (AI) based telemedicine for cost-effective cataract surgery follow-up at NHS: an economic evaluation. *Value Health* 2023;**26**:S71.

61. Abdi S, Patel D, Carmichael J, Balaskas K, Blandford A (under review). Implementing a teleophthalmology referral platform in routine practice: a study of understanding digital health intervention implementation using normalisation process theory. *Digital Health* 2025;**11**. https://doi.org/10.1177/20552076241303812

62. Bowling B, Chen SDM, Salmon JF. Outcomes of referrals by community optometrists to a hospital glaucoma service. *Br J Ophthalmol* 2005;**89**:1102–4.

63. Pierscionek TJ, Moore JE, Pierscionek BK. Referrals to ophthalmology: optometric and general practice comparison. *Ophthalmic Physiol Opt* 2009;**29**:32–40.

64. Tham YC, Husain R, Teo KYC, Tan ACS, Chew ACY, Ting DS, *et al*. New digital models of care in ophthalmology, during and beyond the COVID-19 pandemic. *Br J Ophthalmol* 2022;**106**:452–7.

65. Caffery LJ, Farjian M, Smith AC. Telehealth interventions for reducing waiting lists and waiting times for specialist outpatient services: a scoping review. *J Telemed Telecare* 2016;**22**:504–12.

66. Fatehi F, Jahedi F, Tay-Kearney ML, Kanagasingam Y. Teleophthalmology for the elderly population: a review of the literature. *Int J Med Inform* 2020;**136**:104089.

67. Sreelatha OK, Ramesh SV. Teleophthalmology: improving patient outcomes? *Clin Ophthalmol* 2016;**10**:285–95.

68. Walsh L, Hong SC, Chalakkal RJ, Ogbuehi KC. A systematic review of current teleophthalmology services in New Zealand compared to the four comparable countries of the United Kingdom, Australia, United States of America (USA) and Canada. *Clin Ophthalmol* 2021;**15**:4015–27.

69. Korot E, Wood E, Weiner A, Sim DA, Trese M. A renaissance of teleophthalmology through artificial intelligence. *Eye* 2019;**33**:861–3.

70. Wong TY, Bressler NM. Artificial intelligence with deep learning technology looks into diabetic retinopathy screening. *JAMA* 2016;**316**:2366–7.

71. Topol EJ. High-performance medicine: the convergence of human and artificial intelligence. *Nat Med* 2019;**25**:44–56.

72. Gunasekeran DV, Wong TY. Artificial intelligence in ophthalmology in 2020: a technology on the cusp for translation and implementation. *Asia Pac J Ophthalmol (Phila)* 2020;**9**:61–6.

73. Conway MP, Forristal MT, Treacy MP, Duignan ES. Investigating the role of optometrists in teleophthalmology and the implications of increasing access to advanced imaging techniques and digital referral: a systematic search and review. *Telemed J E Health* 2021;**27**:974–81.

74. Lee JX, Manjunath V, Talks SJ. Expanding the role of medical retina virtual clinics using multimodal ultra-widefield and optical coherence tomography imaging. *Clin Ophthalmol* 2018;**12**:2337–45.

75. Blandford A, Wesson J, Amalberti R, AlHazme R, Allwihan R. Opportunities and challenges for telehealth within, and beyond, a pandemic. *Lancet Glob Health* 2020;**8**:e1364–5.

76. Nagra M, Allen PM, Norgett Y, Beukes E, Bowen M, Vianya-Estopa M. The effect of the COVID-19 pandemic on working practices of UK primary care optometrists. *Ophthalmic Physiol Opt* 2021;**41**:378–92.

77. Greenhalgh T, Wherton J, Papoutsi C, Lynch J, Hughes G, A'Court C, *et al*. Beyond adoption: a new framework for theorizing and evaluating nonadoption, abandonment, and challenges to the scale-up, spread, and sustainability of health and care technologies. *J Med Internet Res* 2017;**19**:e367.

78. Ong BN, Hodgson D, Small N, Nahar P, Sanders C. Implementing a digital patient feedback system: an analysis using normalisation process theory. *BMC Health Serv Res* 2020;**20**:387.

79. Chambers D, Cantrell A, Booth A. *Rapid Evidence Review: Challenges to Implementing Digital and Data-Driven Technologies in Health and Social Care*. 2021. URL: https://eprints.whiterose.ac.uk/175758/ (accessed 7 February 2024).

80. Murray E, Hekler EB, Andersson G, Collins LM, Doherty A, Hollis C, *et al*. Evaluating digital health interventions: key questions and approaches. *Am J Prev Med* 2016;**51**:843–51.

81. Liu Y, Zupan NJ, Swearingen R, Jacobson N, Carlson JN, Mahoney JE, *et al*. Identification of barriers, facilitators and system-based implementation strategies to increase teleophthalmology use for diabetic eye screening in a rural US primary care clinic: a qualitative study. *BMJ Open* 2019;**9**:e022594. URL: https://bmjopen.bmj.com/content/9/2/e022594.abstract (accessed 10 February 2024).

82. Piyasena MMPN, Murthy GVS, Yip JLY, Gilbert C, Zuurmond M, Peto T, *et al*. Systematic review on barriers and enablers for access to diabetic retinopathy screening services in different income settings. *PLOS ONE* 2019;**14**:e0198979.

83. Bastos de Carvalho A, Lee Ware S, Belcher T, Mehmeti F, Higgins EB, Sprang R, *et al*. Evaluation of multi-level barriers and facilitators in a large diabetic retinopathy screening program in federally qualified health centers: a qualitative study. *Implement Sci Commun* 2021;**2**:54.

84. Denscombe M. Critical Incidents and Learning About Risks: The Case of Young People and Their Health. In Hammersley M, editor. *Educational Research and Evidence-based Practice*. London: Sage and The Open University. pp. 204–19.

85. Holtzblatt K, Beyer HR. *8.1 Motivations and Key Principles*. URL: https://monet.en.kku.ac.th/courses/EN814707/Contextual%20Design3.html (accessed 10 February 2024).

86. Blandford A, Furniss D, Makri S. *Qualitative HCI Research: Going Behind the Scenes*. San Rafael, CA: Morgan & Claypool Publishers; 2016.

87. Blandford A. *Semi-Structured Qualitative Studies*. The Interaction Design Foundation. URL: www.interaction-design.org/literature/book/the-encyclopedia-of-human-computer-interaction-2nd-ed/semi-structured-qualitative-studies (accessed 7 February 2024).

88. Hoffman RR, Crandall B, Shadbolt N. Use of the Critical Decision Method to elicit expert knowledge: a case study in the methodology of cognitive task analysis. *Hum Factors* 1998;**40**:254–76.

89. Blandford A, William Wong BL. Situation awareness in emergency medical dispatch. *Int J Hum-Comput Stud* 2004;**61**:421–52.

90. Braun V, Clarke V. Using thematic analysis in psychology. *Qual Res Psychol* 2006;**3**:77–101.

91. May CR, Albers B, Bracher M, Finch TL, Gilbert A, Girling M, *et al*. *Translational Framework for Implementation Evaluation and Research: A Normalisation Process Theory Coding Manual for Qualitative Research and Instrument Development*. 2022. URL: https://hdl.handle.net/10779/uos.23309108.v1 (accessed 7 February 2024).

92. Clarke V, Braun V. Successful qualitative research: a practical guide for beginners. *Qmip Bull* 2013;**2013**:39. https://doi.org/10.53841/bpsqmip.2014.1.18.39

93. May C, Finch T. Implementing, embedding, and integrating practices: an outline of normalization process theory. *Sociology* 2009;**43**:535–54.

94. Murray E, Treweek S, Pope C, MacFarlane A, Ballini L, Dowrick C, *et al*. Normalisation process theory: a framework for developing, evaluating and implementing complex interventions. *BMC Med* 2010;**8**:63.

95. Kelly SP, Wallwork I, Haider D, Qureshi K. Teleophthalmology with optical coherence tomography imaging in community optometry. Evaluation of a quality improvement for macular participants. *Clin Ophthalmol* 2011;**5**:1673–8.

96. Davey CJ, Harley C, Elliott DB. Levels of state and trait anxiety in participants referred to ophthalmology by primary care clinicians: a cross sectional study. *PLOS ONE* 2013;**8**:e65708.

97. NHS England. *PRN00242-General-Ophthalmic-Mandatory-Services-Model-Contract-September-2023*. URL: www.england.nhs.uk/publication/2018-19-addendum-to-the-gp-it-operating-model-securing-excellence-in-gp-it-services-2016-18-revisions/ (accessed 9 December 2023).

98. Harvey K, Edgar DF, Agarwal R, Benwell MJ, Evans BJ. Referrals from community optometrists in England and their replies: a mixed methods study. *Ophthalmic Physiol Opt* 2022;**42**:454–70.

99. Baldwin DM, Quintela J, Duclos C, Staton EW, Pace WD. Patient preferences for notification of normal laboratory test results: a report from the ASIPS Collaborative. *BMC Fam Pract* 2005;**6**:11.

100. O'Malley AS, Reschovsky JD. Referral and consultation communication between primary care and specialist physicians: finding common ground. *Arch Intern Med* 2011;**171**:56–65.

101. Townsend D, Reeves BC, Taylor J, Chakravarthy U, O'Reilly D, Hogg RE, Mills N. Health professionals' and service users' perspectives of shared care for monitoring wet age-related macular degeneration: a qualitative study alongside the ECHoES trial. *BMJ Open* 2015;**5**:e007400.

102. Gagnon MP, Desmartis M, Labrecque M, Car J, Pagliari C, Pluye P, *et al*. Systematic review of factors influencing the adoption of information and communication technologies by healthcare professionals. *J Med Syst* 2012;**36**:241–77.

103. Ross J, Stevenson F, Lau R, Murray E. Factors that influence the implementation of e-health: a systematic review of systematic reviews (an update). *Implement Sci* 2016;**11**:1–12.

104. McCrorie C, Benn J, Johnson OA, Scantlebury A. Staff expectations for the implementation of an electronic health record system: a qualitative study using normalisation process theory. *BMC Med Inform Decis Mak* 2019;**19**:222.

105. Ramchandran RS, Yousefi-Nooraie R, Dadgostar P, Yilmaz S, Basant J, Dozier AM. Implementation of teleophthalmology to improve diabetic retinopathy surveillance: Qualitative interview study of clinical staff informed by implementation science frameworks. *JMIR Diabetes* 2022;**7**:e32162.

106. Ross J, Stevenson F, Dack C, Pal K, May C, Michie S, *et al*. Developing an implementation strategy for a digital health intervention: an example in routine healthcare. *BMC Health Serv Res* 2018;**18**:794.

107. Skivington K, Matthews L, Simpson SA, Craig P, Baird J, Blazeby JM, *et al*. A new framework for developing and evaluating complex interventions: update of Medical Research Council guidance. *BMJ* 2021;**374**:n2061.

108. Blandford A. Interaction Design for Healthcare Technologies. In Vanderdonckt J, Palanque P, Winckler M, *editors. Handbook of Human Computer Interaction*. Cham: Springer, 2023. pp. 1–24.

109. de Grood C, Raissi A, Kwon Y, Santana MJ. Adoption of e-health technology by physicians: a scoping review. *J Multidiscip Healthc* 2016;**9**:335–44.

110. Bokolo AJ. Exploring the adoption of telemedicine and virtual software for care of outpatients during and after COVID-19 pandemic. *Ir J Med Sci* 2021;**190**:1–10.

111. Solomon DH, Rudin RS. Digital health technologies: opportunities and challenges in rheumatology. *Nat Rev Rheumatol* 2020;**16**:525–35.

112. Gunasekeran DV, Tham YC, Ting DSW, Tan GSW, Wong TY. Digital health during COVID-19: lessons from operationalising new models of care in ophthalmology. *Lancet Digit Health* 2021;**3**:e124–34.

113. Awad A, Trenfield SJ, Pollard TD, Ong JJ, Elbadawi M, McCoubrey LE, *et al*. Connected healthcare: improving patient care using digital health technologies. *Adv Drug Deliv Rev* 2021;**178**:113958.

114. Perski O, Short CE. Acceptability of digital health interventions: embracing the complexity. *Transl Behav Med* 2021;**11**:1473–80.

115. Sharma U, Reed J, Doyle C, Bell D. Challenges in evaluating telehealth through RCT-the problem of randomization. *Stud Health Technol Inform* 2012;**180**:323–7.

116. Kotecha A, Brookes J, Foster PJ. A technician-delivered 'virtual clinic' for triaging low-risk glaucoma referrals. *Eye* 2017;**31**:899–905.

117. Parfitt A, Boxell E, Amoaku WM, Bradley C. Patient-reported reasons for delay in diagnosis of age-related macular degeneration: a national survey. *BMJ Open Ophthalmol* 2019;**4**:e000276.

118. Carmichael J, Abdi S, Balaskas K, Costanza E, Blandford A. Assessment of optometrists' referral accuracy and contributing factors: a review. *Ophthalmic Physiol Opt* 2023;**43**:1255–77.

119. Muen WJ, Hewick SA. Quality of optometry referrals to neovascular age-related macular degeneration clinic: a prospective study. *JRSM Short Rep* 2011;**2**:64.

120. Lehman CD, Wellman RD, Buist DSM, Kerlikowske K, Tosteson ANA, Miglioretti DL; Breast Cancer Surveillance Consortium. Diagnostic accuracy of digital screening mammography with and without computer-aided detection. *JAMA Intern Med* 2015;**175**:1828–37. https://doi.org/10.1001/jamainternmed.2015.5231

121. Lee CS, Brandt JD, Lee AY. Big data and artificial intelligence in ophthalmology: where are we now? *Ophthalmol Sci* 2021;**1**:100036.

122. Zhang G, Fu DJ, Liefers B, Faes L, Glinton S, Wagner S, *et al*. Clinically relevant deep learning for detection and quantification of geographic atrophy from optical coherence tomography: a model development and external validation study. *Lancet Digit Health* 2021;**3**:e665–75.

123. NICE. *Recommendations | Age-Related Macular Degeneration | Guidance*. NICE. URL: www.nice.org.uk/guidance/ng82/chapter/recommendations (accessed 7 February 2024).

124. Bazvand F, Ghassemi F. Artefacts in macular optical coherence tomography. *J Curr Ophthalmol* 2020;**32**:123–31.

125. Saenz AD, Harned Z, Banerjee O, Abràmoff MD, Rajpurkar P. Autonomous AI systems in the face of liability, regulations and costs. *NPJ Digit Med* 2023;**6**:185. https://doi.org/10.1038/s41746-023-00929-1

126. Carmichael J, Costanza E, Blandford A, Struyven R, Keane P, Balaskas K. Diagnostic decisions of specialist optometrists exposed to ambiguous deep-learning outputs. *Research Square* 2023. https://doi.org/10.21203/rs.3.rs-3291751/v1

Appendix 1 Case report forms for the HERMES study

Moorfields Eye Hospital, Case Report Form
BALK1006

Moorfields Eye Hospital NHS Foundation Trust

Study No: ☐☐ - ☐☐☐

HERMES Study

PATIENT DEMOGRAPHICS (Community Optometry) **Page 1/2**

ALL FIELDS ARE MANDATORY

Patient Details		
Study site	MEH ☐ Manchester ☐ Birmingham ☐ Central Middlesex ☐ North West Anglia ☐	
Optometrist site		
Randomisation Cluster number		
Randomisation Arm	Control ☐	Intervention ☐
Age	*(≥18)* ☐☐☐	
Sex	Male ☐	Female ☐
Medical history	Yes No Heart attack ☐ ☐ COPD ☐ ☐ Diabetes ☐ ☐ Hypertension ☐ ☐	Yes No TIA/Stroke ☐ ☐ Impaired Mobility ☐ ☐ Asthma ☐ ☐ Other ☐ ☐
If other please specify		
Medication for eye condition	Yes ☐	No ☐
If yes specify glaucoma drops?	Prostaglandins ☐ CA inhibitors ☐	B-blockers ☐
If yes specify AREDS	Yes ☐	No ☐
Smoker?	Ex-smoker ☐ Smoker ☐	Non-smoker ☐

	Right Eye	**Left eye**
Ocular history	Yes No Wet AMD ☐ ☐ Dry AMD ☐ ☐ Central Serous Chorioretinopathy ☐ ☐ Central Retinal Vein Occlusion ☐ ☐ Branch Retinal Vein Occlusion ☐ ☐ Diabetic Macular Oedema ☐ ☐ Inherited Eye Disease ☐ ☐ Other ☐ ☐	Yes No Wet AMD ☐ ☐ Dry AMD ☐ ☐ Central Serous Chorioretinopathy ☐ ☐ Central Retinal Vein Occlusion ☐ ☐ Branch Retinal Vein Occlusion ☐ ☐ Diabetic Macular Oedema ☐ ☐ Inherited Eye Disease ☐ ☐ Other ☐ ☐
If other, specify		

BALK1006_ CRF_PACK _v2.0.pdf
Rev 2.0 – 26/05/2021

CONFIDENTIAL
/2021

Page 2 of 14

Moorfields Eye Hospital, Case Report Form
BALK1006

Moorfields Eye Hospital NHS Foundation Trust

Study No: ☐☐-☐☐☐

HERMES Study

PATIENT DEMOGRAPHICS (Community Optometry) **Page 2/2**

ALL FIELDS ARE MANDATORY

Patient Details

		Yes	No		Yes	No
Previous eye procedures	Cataract surgery	☐	☐	Cataract surgery	☐	☐
	Glaucoma surgery	☐	☐	Glaucoma surgery	☐	☐
	Eyelid surgery	☐	☐	Eyelid surgery	☐	☐
	Other	☐	☐	Other	☐	☐
If other specify						

Comments

I can confirm that the patient meets all eligibility criteria for the study and I have completed this form in full and take full responsibility for any missing data.

Signed:	**Print:**	**Date:**
		☐☐/☐☐/☐☐☐☐
Office use only, data entry completed by:		
Signed:	**Print:**	**Date:**
		☐☐/☐☐/☐☐☐☐

Moorfields Eye Hospital, Case Report Form
BALK1006

Moorfields Eye Hospital NHS Foundation Trust

Study No: ☐☐-☐☐☐

HERMES Study

BASELINE VISIT (Community Optometry) **Page 1/2**

ALL FIELDS ARE MANDATORY

Baseline Exam		
Date of visit	*(dd mm yyyy)* ☐☐/☐☐/☐☐☐☐	
Years since diagnosis of any retinal disease?	☐☐	
OCT Device	Topcon 3D OCT-2000 ☐ Heidelberg OCT1 ☐ Other ☐	
If other specify		
	Right Eye	**Left eye**
Visual acuity (ETDRS)	*(ETDRS letters, range 0-95)* ☐☐	*(ETDRS letters, range 0-95)* ☐☐
Visual acuity (Snellen)	☐☐/☐☐	☐☐/☐☐
Diagnosis by the Optometrist	Yes / No Wet AMD ☐ ☐ Dry AMD ☐ ☐ Central Serous Chorioretinopathy ☐ ☐ Central Retinal Vein Occlusion ☐ ☐ Branch Retinal Vein Occlusion ☐ ☐ Diabetic Macular Oedema ☐ ☐ Inherited Eye Disease ☐ ☐ Other ☐ ☐	Yes / No Wet AMD ☐ ☐ Dry AMD ☐ ☐ Central Serous Chorioretinopathy ☐ ☐ Central Retinal Vein Occlusion ☐ ☐ Branch Retinal Vein Occlusion ☐ ☐ Diabetic Macular Oedema ☐ ☐ Inherited Eye Disease ☐ ☐ Other ☐ ☐
If other specify		
Clinical findings	Yes / No Macular Haemorrhage ☐ ☐ Other RetinalHaemorrhage ☐ ☐ Exudates ☐ ☐ Disc Swelling ☐ ☐ Macular Atrophy ☐ ☐ Cotton Wool Spot ☐ ☐ Other ☐ ☐	Yes / No Macular Haemorrhage ☐ ☐ Other RetinalHaemorrhage ☐ ☐ Exudates ☐ ☐ Disc Swelling ☐ ☐ Macular Atrophy ☐ ☐ Cotton Wool Spot ☐ ☐ Other ☐ ☐
If other specify		

Moorfields Eye Hospital, Case Report Form
BALK1006

Moorfields Eye Hospital NHS Foundation Trust

Study No: ☐☐-☐☐☐

HERMES Study

BASELINE VISIT (Community Optometry) **Page 2/2**

ALL FIELDS ARE MANDATORY

Baseline Exam	**Right Eye**	**Left eye**
Intraocular pressure	*(0-60)* ☐☐	*(0-60)* ☐☐
OCT taken?	Yes ☐ No ☐	Yes ☐ No ☐
OCT qualitative	SRF ☐ IRF ☐ PED ☐ SHRM ☐	SRF ☐ IRF ☐ PED ☐ SHRM ☐
Referral Recommendation by Optometrist (both arms)	Urgent ☐ Routine ☐ No referral ☐	
If not referred, specify reason		

Comments:

I have completed this form in full and take full responsibility for any missing data.

Signed:	**Print:**	**Date:**
		☐☐/☐☐/☐☐☐☐
Office use only, data entry completed by:		
Signed:	**Print:**	**Date:**
		☐☐/☐☐/☐☐☐☐

Moorfields Eye Hospital NHS Foundation Trust NHS

Study No: ☐☐ - ☐☐☐

HERMES Study

REFERRALS (tele-HES) **Page 2/3**

ALL FIELDS ARE MANDATORY

Tele - HES Review		
OCT from referral reviewed by tele-HES?	Yes ☐	No ☐
	Right Eye	**Left eye**
OCT from referral *qualitative by tele-HES*	SRF ☐ IRF ☐ PED ☐ SHRM ☐	SRF ☐ IRF ☐ PED ☐ SHRM ☐
Clinical findings	Yes No Macular Haemorrhage ☐ ☐ Other Retinal Haemorrhage ☐ ☐ Exudates ☐ ☐ Disc Swelling ☐ ☐ Macular Atrophy ☐ ☐ Cotton Wool Spot ☐ ☐ Other ☐ ☐	Yes No Macular Haemorrhage ☐ ☐ Other Retinal Haemorrhage ☐ ☐ Exudates ☐ ☐ Disc Swelling ☐ ☐ Macular Atrophy ☐ ☐ Cotton Wool Spot ☐ ☐ Other ☐ ☐
If other specify		
Tele-HES Review	**Right Eye**	**Left eye**
Diagnosis from referring optometrist confirmed by tele-HES	Yes ☐ No ☐	
Diagnosis by tele-HES	Yes No Wet AMD ☐ ☐ Dry AMD ☐ ☐ Central Serous Chorioretinopathy ☐ ☐ Central Retinal Vein Occlusion ☐ ☐ Branch Retinal Vein Occlusion ☐ ☐ Diabetic Macular Oedema ☐ ☐ Inherited Eye Disease ☐ ☐ Other ☐ ☐	Yes No Wet AMD ☐ ☐ Dry AMD ☐ ☐ Central Serous Chorioretinopathy ☐ ☐ Central Retinal Vein Occlusion ☐ ☐ Branch Retinal Vein Occlusion ☐ ☐ Diabetic Macular Oedema ☐ ☐ Inherited Eye Disease ☐ ☐ Other ☐ ☐
If other specify		

Moorfields Eye Hospital NHS Foundation Trust NHS

Study No: ☐☐-☐☐☐

HERMES Study

REFERRALS (tele-HES) **Page 3/3**

ALL FIELDS ARE MANDATORY			
Tele - HES Review			
Referral recommendation by referring optometrist confirmed by tele-HES	Yes ☐	No ☐	
If yes, eye on which referral decision was based	Right ☐	Left ☐	Both ☐
Referral Decision by tele-HES	Urgent ☐	Routine ☐	No referral ☐

Comments:

I have completed this form in full and take full responsibility for any missing data.

Signed:	**Print:**	**Date:**
		☐☐/☐☐/☐☐☐☐
Office use only, data entry completed by:		
Signed:	**Print:**	**Date:**
		☐☐/☐☐/☐☐☐☐

Moorfields Eye Hospital, Case Report Form
BALK1006

Moorfields Eye Hospital NHS Foundation Trust

Study No: ☐☐-☐☐☐

HERMES Study

HES First Visit

Page 1/2

ALL FIELDS ARE MANDATORY		
HES Review		
Follow up status	Attended ☐ Cancelled ☐ DNA ☐	
Date of consultation	*(dd mm yyyy)* ☐☐/☐☐/☐☐☐☐	
Change in eye medication	Yes ☐ No ☐	
If yes, specify		
Date of first treatment (if applicable)	*(dd mm yyyy)* ☐☐/☐☐/☐☐☐☐	
	Right Eye	**Left eye**
Intraocular pressure	*(0-60)* ☐☐	*(0-60)* ☐☐
OCT taken?	Yes ☐ No ☐	Yes ☐ No ☐
OCT qualitative	SRF ☐ IRF ☐ PED ☐ SHRM ☐	SRF ☐ IRF ☐ PED ☐ SHRM ☐
Visual acuity (ETDRS)	*(ETDRS letters, range 0-95)* ☐☐	*(ETDRS letters, range 0-95)* ☐☐
Visual acuity (Snellen)	☐☐/☐☐	☐☐/☐☐
Clinical findings by HES	Yes No Macular Haemorrhage ☐ ☐ Other Retinal Haemorrhage ☐ ☐ Exudates ☐ ☐ Disc Swelling ☐ ☐ Macular Atrophy ☐ ☐ Cotton Wool Spot ☐ ☐ Other ☐ ☐	Yes No Macular Haemorrhage ☐ ☐ Other Retinal Haemorrhage ☐ ☐ Exudates ☐ ☐ Disc Swelling ☐ ☐ Macular Atrophy ☐ ☐ Cotton Wool Spot ☐ ☐ Other ☐ ☐
If other specify		
Diagnosis by HES	Yes No Wet AMD ☐ ☐ Dry AMD ☐ ☐ Central Serous Chorioretinopathy ☐ ☐ Central Retinal Vein Occlusion ☐ ☐ Branch Retinal Vein Occlusion ☐ ☐ Diabetic Macular Oedema ☐ ☐ Inherited Eye Disease ☐ ☐ Other ☐ ☐	Yes No Wet AMD ☐ ☐ Dry AMD ☐ ☐ Central Serous Chorioretinopathy ☐ ☐ Central Retinal Vein Occlusion ☐ ☐ Branch Retinal Vein Occlusion ☐ ☐ Diabetic Macular Oedema ☐ ☐ Inherited Eye Disease ☐ ☐ Other ☐ ☐
If other specify		

BALK1006_ CRF_PACK _v2.0.pdf
Rev 2.0 – 26/05/2021

CONFIDENTIAL
/2021

Page 9 of 14

Moorfields Eye Hospital, Case Report Form
BALK1006

Moorfields Eye Hospital NHS Foundation Trust

Study No: ☐☐-☐☐☐

HERMES Study

HES First Visit

Page 2/2

ALL FIELDS ARE MANDATORY		
HES Review	**Right Eye**	**Left eye**
Additional Diagnostic Procedures	Yes No OCT - Angio ☐ ☐ IGCA ☐ ☐ FA ☐ ☐ Optos ☐ ☐ Ultrasound B Scan ☐ ☐ Visual Field Test ☐ ☐ Other ☐ ☐	Yes No OCT - Angio ☐ ☐ IGCA ☐ ☐ FA ☐ ☐ Optos ☐ ☐ Ultrasound B Scan ☐ ☐ Visual Field Test ☐ ☐ Other ☐ ☐
If other specify		

Comments:

I have completed this form in full and take full responsibility for any missing data.

Signed:	**Print:**	**Date:**
		☐☐/☐☐/☐☐☐☐
Office use only, data entry completed by:		
Signed:	**Print:**	**Date:**
		☐☐/☐☐/☐☐☐☐

Appendix 2 Clinical and rule-based reference standard diagnoses

TABLE 41 The corresponding clinical and rule-based RS diagnoses

Clinical RS diagnosis	Rule-based RS diagnosis
Wet AMD (nAMD)	CNV
Dry AMD	Dry AMD
CSCR	CSCR
DMO	MO
BRVO with MO	MO
CRVO with MO	MO
IED	Other
Other (myopic CNV, peripapillary CNV)	CNV
Other (ERM/VMT/VMA/FTMH/PTMH)	VRA
Other (ERM/VMT/VMA/FTMH/PTMH) with cysts/fluid	MO
Other (Mac-Tel, poppers, vitelliform lesion, etc.)	Other
No pathology	Normal

Appendix 3 Cluster RCT post-hoc analysis excluding control arm site

TABLE 42 Proportion of false positive referrals in each study arm following exclusion of control arm site

Definition of false positive referral[a]	Control arm, *N* = 120	Intervention arm, *N* = 158	Difference in proportions[b]	Unadjusted odds ratio[c]	Unadjusted *p*-value[c]
Referral when not needed					
Proportion of all patients	3/120 (3%)	2/158 (1%)	1% (−2% to 5%)	1.99 (0.22, 24)	0.745
ICC	*< 0.001*	*< 0.001*			
Proportion of referrals	3/109 (3%)	2/124 (1%)	1% (−3% to 5%)	1.72 (0.19, 21)	0.879
ICC	*< 0.001*	*< 0.001*			
Urgent referral when not needed					
Proportion of all patients	19/120 (16%)	1/158 (1%)	15% (9% to 22%)	29.27 (4.51, 1235)	< 0.001
Proportion of urgent referrals	19/32 (59%)	1/27 (4%)	56% (36% to 75%)	35.69 (4.70, 1638)	< 0.001
ICC	*0.055*	*< 0.001*			

a Two possible definitions were provided for false positive referral. Both are compared against the clinical reference standard.
b Difference in proportions (control arm − intervention arm) reported with 95% confidence interval adjusting for cluster using arm-specific calculated Intraclass correlation coefficient (ICC).
c Odds ratio and unadjusted *p*-value are obtained from an exact logistic regression model not accounting for clustering. Odds ratio is presented as control arm/intervention arm.

EME
HSDR
HTA
PGfAR
PHR

Part of the NIHR Journals Library
www.journalslibrary.nihr.ac.uk

Published by the NIHR Journals Library