

Automated imaging technologies for the diagnosis of glaucoma: a comparative diagnostic study for the evaluation of the diagnostic accuracy, performance as triage tests and cost-effectiveness (GATE study)

Augusto Azuara-Blanco,^{1*} Katie Banister,²
Charles Boachie,³ Peter McMeekin,⁴
Joanne Gray,⁵ Jennifer Burr,⁶ Rupert Bourne,⁷
David Garway-Heath,^{8,9} Mark Batterbury,¹⁰
Rodolfo Hernández,¹¹ Gladys McPherson,²
Craig Ramsay² and Jonathan Cook¹²

¹Centre for Experimental Medicine, Queen's University Belfast, Belfast, UK

²Health Services Research Unit, University of Aberdeen, Aberdeen, UK

³Robertson Centre for Biostatistics, University of Glasgow, Glasgow, UK

⁴Health Economics Group, Institute of Health and Society, Newcastle University, Newcastle upon Tyne, UK

⁵Faculty of Health and Life Sciences, Northumbria University, Newcastle upon Tyne, UK

⁶School of Medicine, University of St Andrews, St Andrews, UK

⁷Vision and Eye Research Unit, Postgraduate Institute, Anglia Ruskin University, Cambridge, UK

⁸National Institute of Health Research Biomedical Research Centre, Moorfields Eye Hospital, London, UK

⁹University College London Institute of Ophthalmology, London, UK

¹⁰St Paul's Eye Unit, Royal Liverpool and Broadgreen University Hospitals NHS Trust, Liverpool, UK

¹¹Health Economics Research Unit, University of Aberdeen, Aberdeen, UK

¹²Nuffield Department of Orthopaedics, Rheumatology and Musculoskeletal Sciences, University of Oxford, Oxford, UK

*Corresponding author

Declared competing interests of authors: David Garway-Heath received grants from the National Institute of Health Research during the conduct of the study, personal fees and non-financial support from Heidelberg Engineering UK, personal fees and non-financial support from Carl Zeiss Meditec, Inc., non-financial support from OptoVue Inc. and non-financial support from Topcon outside the submitted work.

Published January 2016

DOI: 10.3310/hta20080

Scientific summary

Automated imaging technologies for the diagnosis of glaucoma (GATE study)

Health Technology Assessment 2016; Vol. 20: No. 8

DOI: 10.3310/hta20080

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

Scientific summary

Background

Glaucoma describes a group of chronic age-related eye diseases in which there is progressive damage of the optic disc and characteristic visual field loss. Glaucoma is a significant public health problem, as it is the second leading cause of blindness in the UK.

Glaucoma care constitutes a major proportion of the workload of the hospital eye service. In England there are over 1 million glaucoma-related outpatient visits to the acute sector annually. Considerable NHS resources are required to assess referrals to hospital eye services for possible glaucoma, which are typically initiated by community optometrists. However, fewer than one-quarter of referrals are found to have glaucoma, and nearly half of the referred individuals are discharged after their first visit. If referrals could be triaged in a clinically effective and cost-effective manner, resources could be better utilised for other needs.

Glaucoma is diagnosed by clinicians detecting structural changes of the optic nerve head, also known as the optic disc, and corresponding visual field defects. New imaging techniques for assessment of the structural changes have emerged: scanning laser ophthalmoscopy, commercially available as the Heidelberg Retinal Tomograph [HRT; including two diagnostic algorithms, Moorfields regression analysis (HRT-MRA; Heidelberg Engineering, Heidelberg, Germany) and glaucoma probability score (HRT-GPS; Heidelberg Engineering, Heidelberg, Germany)] and scanning laser polarimetry, commercially available as glaucoma diagnostics (GDx; Carl Zeiss Meditec, Dublin, CA, USA) and spectral domain optical coherence tomography (SD-OCT; Heidelberg Engineering, Heidelberg, Germany), with several commercial devices available.

Imaging technologies are being introduced into glaucoma services but their role in the diagnostic pathway is unclear. Imaging tests are user-friendly and safe, provide automated classifications and potentially could reduce the need for an examination by a clinician.

Aim

To assess the relative performance and cost-effectiveness of diagnostic imaging technologies as triage tests in secondary care for identifying people with glaucoma.

Objectives

Primary objective

To compare the diagnostic performance (in terms of sensitivity and specificity) in a cohort of patients referred to hospital eye services with possible glaucoma, of:

- four imaging tests [HRT-MRA, HRT-GPS, GDx and optical coherence tomography (OCT)] for diagnosis of glaucoma
- a composite triage test [combining imaging tests, visual acuity (VA) and intraocular pressure (IOP) measurements] in correctly identifying patients to be discharged from secondary care.

Secondary objectives

- To explore alternative thresholds for determining abnormal tests.
- To evaluate the diagnostic performance of combinations of imaging tests.
- To evaluate the performance of the tests across the spectrum of glaucoma (mild, moderate and severe).
- To evaluate the cost-effectiveness of incorporating imaging in a triage test in hospital eye services compared with current practice of diagnostic examination by a clinician.
- To evaluate patient preferences related to different imaging technologies.

Methods

We designed a pragmatic within-patient comparative diagnostic and triage evaluation of imaging techniques for glaucoma. Participants were adult patients referred from community optometrists or general practitioners with any possible glaucoma-related findings. Five UK NHS centres participated: three academic centres and two district general hospitals.

Participants received all imaging tests: HRT-GPS, HRT-MRA, GDx and OCT. Possible tests results were within normal limits, borderline or outside normal limits.

The HRT uses confocal laser scanning to allow quantitative structural measurement of the optic disc anatomy. There are two main classification tools to relate measurements to normative data: (1) HRT-MRA, which requires user definition of the optic disc boundary, and (2) HRT-GPS, which is fully automated.

The GDx scanning laser polarimeter measures the retinal nerve fibre layer (RNFL) thickness surrounding the optic disc utilising the birefringent properties of the RNFL. The software provides a discriminating classifier termed the nerve fibre indicator, which is fully automated.

Spectral domain OCT is an optical imaging technique providing high-resolution, cross-sectional imaging of the retina analogous to B-scan ultrasonography but using light instead of sound. The Spectralis® optical coherence tomograph (Heidelberg Engineering, Heidelberg, Germany) was used in this study.

The reference standard was a full clinical examination, including visual field testing, by a consultant ophthalmologist with glaucoma expertise to determine (1) a diagnosis of glaucoma (mild, moderate or severe) according to well-defined criteria (diagnosis analysis) and (2) whether or not the patient would be discharged or should be monitored/treated within hospital eye services (triage analysis).

Statistical analysis

Sample size calculations were based on standard diagnostic accuracy study methods. A 5% significance level based on a two-sided test was used, which required a study of 897 individuals to have 90% power to detect an accuracy difference of 9% for the primary outcome of glaucoma diagnosis. Including a 6% indeterminacy rate increased the sample size to 954.

Two diagnostic performance analyses were undertaken: a diagnosis and a triage analysis. For the diagnosis analysis (classification of glaucoma), one eye per patient was used: the eye with more severe disease except for in one sensitivity analysis. The test 'abnormal' definition was an imaging test result of 'outside normal limits', with 'borderline' cases classified as 'normal'. This was compared with a reference standard diagnosis of 'glaucoma'.

For the triage analysis, a composite test (including three components: imaging, IOP measurement and VA) was compared with a reference standard of clinical decision 'do not discharge'. The test categorised a patient as needing evaluation by a clinician if any elements of the composite triage test were themselves 'abnormal' in either eye: imaging classification 'outside normal limits' or IOP > 21 mmHg or VA of 6/12 or poorer.

Primary diagnostic performance outcomes were the sensitivity and specificity of tests. Secondary diagnostic performance outcomes were likelihood ratio and diagnostic odds ratio (DOR). The proportions of indeterminate test results, low-quality imaging and need for pupil dilatation were measured and patient preference for the tests was ranked. The test performance was assessed with respect to the glaucoma spectrum (mild, moderate and severe), when including glaucoma suspects in the reference standard diagnosis, and when including 'borderline' results as abnormal. The diagnostic performance of combinations of tests was also evaluated.

Economic analysis

A current practice pathway model was developed whereby patients referred to hospital eye services were seen by a nurse for VA assessment, a technician for visual field measurement and by a clinician.

In an alternative triage care pathway model, individuals were seen by a nurse for VA examination and IOP measurement and a technician for imaging assessment. The triage test results classified patients as needing referral for clinician diagnosis or as discharged. Those referred were seen by a technician for visual field measurement and examined by a clinician.

The cost-effectiveness of four triage pathways, each using IOP, VA and one of the four imaging technologies (which varied by their diagnostic ability and capital cost), and their subsequent care management pathways was assessed using a multistate Markov model compared with current practice.

The cohort started in one of six health states: normal; at risk of glaucoma; mild glaucoma; moderate glaucoma; severe glaucoma; or sight-impaired. The sensitivity and specificity of each triage strategy determined if diagnosis was correct and, depending on this, the health state that patients would move to and the associated progression of any underlying glaucoma.

Modelled care pathways were developed in consultation with the study team and the independent steering committee and used our previous models in this area and reviewed guidelines, study data and expert opinion.

Consequences were considered in terms of monetary costs (of testing and subsequent management of the patient's condition) to the NHS and in terms of the effects on quality of life (by assigning utility weights). Combining these data with the probabilities of events occurring over time enabled costs, patient outcomes and quality-adjusted life-years (QALYs) to be estimated for a hypothetical cohort of patients undergoing each modelled strategy.

Model results were analysed as incremental cost per QALY and incorporated (1) costs (of testing) and triage diagnostic outcomes, (2) costs (of testing and subsequent management) and (3) QALYs. The base-case analysis used a cohort of 40-year-old males using prevalence data from the Glaucoma Automated Tests Evaluation (GATE) study and for a 50-year time horizon. Cycle length was 1 year. The results were presented in incremental cost-effectiveness ratios (ICERs).

Several deterministic sensitivity analyses were explored, which varied: the annual probability of discharged patients having a sight test; the cost of triage tests; the start age of the cohort; the performance of the diagnosing clinician; the diagnostic performance of imaging technologies; the prevalence of glaucoma in the referred population; and utility weights for those 'at risk of glaucoma'. The possibility of a hypothetical pathway, in which patients diagnosed as 'at risk of glaucoma' were discharged from the service, was explored to investigate the impact in terms of costs and QALYs.

Results

Between April 2011 and July 2013, 2088 participants were identified as potentially eligible: 2013 were invited to take part. Of those invited, 966 (48%) agreed to take part. Following consent, 11 participants were found to be ineligible and did not participate and 12 were excluded as they did not receive all four imaging tests. Therefore, 943 participants were available for the comparisons of tests.

The average age of participants was 60.5 years [standard deviation (SD) 13.8 years] and 51.1% were female. Non-participants had similar age and sex balance. Most participants (89.2%) were of 'white British' ethnicity. The average IOP at referral was 20 mmHg. The most common diagnosis was 'no glaucoma-related findings' (31.7% of participants). Comorbidities were uncommon, except for cataract, which was reported in 8.3% of right eyes and 7.4% of left eyes. Glaucoma was diagnosed in at least one eye in 16.8% of the GATE cohort and 6.5% had glaucoma in both eyes at referral. Overall, 37.9% of GATE participants were discharged after the first visit.

Performance of the imaging tests in diagnosing glaucoma differed. HRT-MRA had the highest sensitivity [87.0%, 95% confidence interval (CI) 80.2% to 92.1%] but the lowest specificity (63.9%, 95% CI 60.2% to 67.4%), GDx had the lowest sensitivity (35.1%, 95% CI 27.0% to 43.8%) but the highest specificity (97.2%, 95% CI 95.6% to 98.3%) and the other two tests provided intermediate results (HRT-GPS sensitivity 81.5%, 95% CI 73.9% to 87.6% and specificity 67.7%, 95% CI 64.2% to 71.2%; OCT sensitivity 76.9%, 95% CI 69.2% to 83.4% and specificity 78.5%, 95% CI 75.4% to 81.4%).

Likelihood ratios showed evidence of being able to both rule in and rule out the presence of glaucoma for all four imaging tests (95% CIs did not contain 1.0). DORs ranged from 9.24 for HRT-GPS to 18.48 for GDx.

When including borderline imaging results as an abnormal test, the sensitivity increased but with a corresponding decrease in specificity. In this sensitivity analysis, HRT-MRA had the highest sensitivity (94.9%, 95% CI 89.8% to 97.9%) but the second lowest specificity (43.9%, 95% CI 40.2% to 47.6%), GDx had the lowest sensitivity (60.4%, 95% CI 51.6% to 68.8%) but the highest specificity (82.8%, 95% CI 79.8% to 85.5%) and the other two tests provided intermediate results.

The impact of combining two imaging tests was improved detection of glaucoma but the effect was marginal and smaller than the loss of specificity.

When considering participants with severe glaucoma, according to our definition of disease stage, OCT had the highest sensitivity (95.2%, 95% CI 76.2% to 99.9%) and the second highest specificity (70.9%, 95% CI 67.7% to 73.9%), GDx had the lowest sensitivity (78.9%, 95% CI 54.4% to 93.9%) but the highest specificity (93.7%, 95% CI 91.8% to 95.2%) and the other two tests provided intermediate results.

The performance of triage tests (a composite assessment comprising imaging test, IOP and VA assessments) in correctly identifying patients to be discharged from secondary care showed that triage including HRT-GPS had the highest sensitivity (86.0%, 95% CI 82.8% to 88.7%) but the lowest specificity (39.1%, 95% CI 34.0% to 44.3%) and GDx had the lowest sensitivity (64.7%, 95% CI 60.7% to 68.7%) but the highest specificity (53.6%, 95% CI 48.2% to 58.9%), the other two tests providing intermediate results [HRT-MRA values were very similar to the HRT-GPS results in sensitivity (86.0%, 95% CI 82.8% to 88.7%) and specificity (53.6%, 95% CI 48.2% to 58.9%) and OCT had lower sensitivity (75.4%, 95% CI 71.6% to 78.9%) but higher specificity (41.0%, 95% CI 35.8% to 46.3%) values than HRT-GPS and HRT-MRA]. Likelihood ratios (and 95% CI) showed evidence of all four triage tests being able to rule in and out the presence of abnormalities for all four triage tests (CIs did not contain 1.0). DORs ranged from 2.1 for GDx to 3.9 for HRT-GPS.

Participant preference for type of imaging test was collected for 890 participants (94%). Almost half of responders (48.2%) had no preference. Of those participants who expressed a preference, OCT was ranked as most preferred (27.6%), followed by GDx (11.9%) and HRT (5.1%). Average time taken to perform the test varied from 5.2 minutes (SD 3.0 minutes) for OCT to 7.6 minutes (SD 5.0 minutes) for HRT.

Economic analysis results

All triage strategies were more cost-effective than current practice but resulted in reduced health because of missing cases (i.e. fewer expected QALYs). The base-case results suggest that, of the triage pathways modelled, a triage including IOP, VA and HRT-MRA is the most cost-effective strategy. Triage including GDx was shown to be the least costly and least effective. Triage including OCT and HRT-GPS were not cost-effective. Compared with GDx, the cost per QALY gained for HRT-MRA was £22,904. The cost per QALY gained with current practice was £156,985 compared with HRT-MRA. Large savings could be made by implementing HRT-MRA but some benefit to patients would be forgone.

These results should be interpreted with some caution, particularly in terms of differences among triage strategies, since the diagnostic accuracy of all tests (except GDx) and their unit costs are very similar. The incremental cost-effectiveness of the triage strategies compared with current practice is very sensitive to the costs included in the model. Indeed, current practice becomes cost-effective when the total cost of a triage test increases to £30 and above. A key assumption used in the model was that clinicians are 100% accurate in their diagnostic ability. Relaxing this assumption increased further the ICER (favouring triage strategies).

Conclusions

Implications for health care

Imaging technologies can be effective to aid the diagnosis of glaucoma. An alternative pathway for patients referred from community to hospital eye services with possible glaucoma, using a triage test that includes imaging, IOP and VA, appears to be cost-effective compared with current practice. Our findings are based on a relatively inexpensive composite triage test (< £30). The most cost-effective strategy would include HRT-MRA imaging. However, triaging would be associated with a loss of health, and the acceptability of this option among users and clinicians has not been evaluated.

Recommendations for research

- Determine the acceptability to patients and health-care providers of implementing an efficient triage glaucoma diagnostic triage system but with reduced health.
- Obtain data on glaucoma disease progression, specifically including patients classified as having glaucoma suspects and ocular hypertension, associated utility, and cost of providing health care.
- Investigate varying the results of the imaging tests beyond the standard options, since the recommended classification may not be the one best suited to the population from which GATE recruited.
- Examine the effectiveness of implementation of a composite triage test.

Funding

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

ISSN 1366-5278 (Print)

ISSN 2046-4924 (Online)

Impact factor: 5.027

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

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

Editorial contact: nhredit@southampton.ac.uk

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

Criteria for inclusion in the *Health Technology Assessment* journal

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

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

HTA programme

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

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

For more information about the HTA programme please visit the website: <http://www.nets.nihr.ac.uk/programmes/hta>

This report

The research reported in this issue of the journal was funded by the HTA programme as project number 09/22/111. The contractual start date was in September 2010. The draft report began editorial review in March 2014 and was accepted for publication in November 2014. The authors have been wholly responsible for all data collection, analysis and interpretation, and for writing up their work. The HTA editors and publisher have tried to ensure the accuracy of the authors' report and would like to thank the reviewers for their constructive comments on the draft document. However, they do not accept liability for damages or losses arising from material published in this report.

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

© Queen's Printer and Controller of HMSO 2016. This work was produced by Azuara-Blanco *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

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

Health Technology Assessment Editor-in-Chief

Professor Hywel Williams Director, HTA Programme, UK and Foundation Professor and Co-Director of the Centre of Evidence-Based Dermatology, University of Nottingham, UK

NIHR Journals Library Editor-in-Chief

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

NIHR Journals Library Editors

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

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

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

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

Professor Aileen Clarke Professor of Public Health and Health Services Research, Warwick Medical School, University of Warwick, UK

Dr Tessa Crilly Director, Crystal Blue Consulting Ltd, UK

Dr Peter Davidson Director of NETSCC, HTA, UK

Ms Tara Lamont Scientific Advisor, NETSCC, UK

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

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

Professor Geoffrey Meads Professor of Health Sciences Research, Health and Wellbeing Research and Development Group, University of Winchester, UK

Professor John Norrie Health Services Research Unit, University of Aberdeen, UK

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

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

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

Professor Helen Roberts Professor of Child Health Research, UCL Institute of Child Health, UK

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

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

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

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

Editorial contact: nihredit@southampton.ac.uk