



City Research Online

City St George's, University of London

Citation: Dabasia, P. L., Fidalgo, B. R., Edgar, D. F., Garway-Heath, D. F. & Lawrenson, J. (2015). Diagnostic accuracy of technologies for glaucoma case-finding in a community setting. *Ophthalmology*, 122(12), pp. 2407-2415. doi: 10.1016/j.ophtha.2015.08.019

This is the accepted version of the paper.

This version of the publication may differ from the final published version. To cite this item please consult the publisher's version.

Permanent repository link: <https://city-ref-test.eprints-hosting.org/id/eprint/12788/>

Link to published version: <https://doi.org/10.1016/j.ophtha.2015.08.019>

Copyright and Reuse: Copyright and Moral Rights remain with the author(s) and/or copyright holders. Copies of full items can be used for personal research or study, educational, or not-for-profit purposes without prior permission or charge, unless otherwise indicated, provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way. For full details of reuse please refer to [City Research Online policy](#).

Title

Diagnostic accuracy of technologies for glaucoma case-finding in a community setting

Authors

Priya L Dabasia (PhD, MCOptom)¹
Bruno R Fidalgo (MSc, MOptom)¹
David F Edgar (FCOptom)¹
David F Garway-Heath (MD, FRCOphth)²
John G Lawrenson (PhD, FCOptom)¹

Institutions

¹ Centre for Public Health Research, School of Health Sciences, City University London, UK

² National Institute for Health Research (NIHR) Biomedical Research Centre at Moorfields Eye Hospital NHS Foundation Trust and UCL Institute of Ophthalmology, London, UK

Corresponding author

John G Lawrenson (PhD, FCOptom)

Meeting Presentation

- European Association for Vision and Eye Research, October 2014
- British Congress of Optometry and Vision Science, September 2014

Financial support

The iVue Optical Coherence Tomographer used in this study was loaned by Optovue Inc., which also provided a part-proportion of funding for the technician to undertake the index tests during data collection. Optovue Inc. made suggestions at the protocol stage of the study but had no input in the data analysis and interpretation. The College of Optometrists, UK provided a PhD studentship to Priya Dabasia to conduct this work, and had no role in the design or conduct of this research. DFG-H is funded in part by the National Institute for Health Research (NIHR) Biomedical Research Centre based at Moorfields Eye Hospital and UCL Institute of Ophthalmology. The views expressed are those of the author(s) and not necessarily those of the NHS, the National Institute for Health Research, or the Department of Health. DFG-H's chair at UCL is supported by funding from the International Glaucoma Association.

Conflict of interest

Dr. Garway-Heath reports personal fees from Heidelberg Engineering, grants from National Institute for Health Research (HTA), outside the submitted work. In addition, Dr. Garway-Heath has a patent ANSWERS pending.

Running head

Accuracy of technologies for glaucoma case-finding

Address for reprints

City University London
Division of Optometry and Visual Science
First floor, Myddleton Street Building
1 Myddleton Street
London
EC1R 1UW

This article contains additional online-only material. The following should appear online-only: Table 1, Table 2, Table 4, Table 6, and Figure 3.

Abstract

Purpose: To assess the case-finding performance of the Frequency Doubling Technology perimeter (FDT), Moorfields Motion Displacement Test (MMDT), iVue Optical Coherence Tomographer (OCT) and Ocular Response Analyser (ORA), used alone or combined, for suspect and definite primary open angle glaucoma (POAG).

Design: Cross-sectional, observational, community-based study.

Participants: 505 subjects aged 60 years and older recruited from a community setting using no pre-defined exclusion criteria.

Methods: Subjects underwent 4 index tests conducted by a technician unaware of subjects' ocular status. FDT and MMDT were used in suprathereshold mode. iVue OCT measured ganglion cell complex and retinal nerve fibre layer (RNFL) thickness. The reference standard was a full ophthalmic examination by an experienced clinician, masked to index test results. Subjects were classified as POAG (open drainage angle, glaucomatous optic neuropathy, and glaucomatous field defect), glaucoma suspect, ocular hypertension (OHT) or non-POAG/non-OHT.

Main Outcome Measures: Test performance evaluated the individual as unit of analysis. Diagnostic accuracy was initially assessed using predefined cut-offs for abnormality to generating sensitivity, specificity, and likelihood ratios.

Continuous data were used to derive estimates of sensitivity at 90% specificity, and partial area under the curve of receiver operating characteristic (AUROC) plots from 90% to 100% specificity.

Results: From the reference standard examination, 26 (5.1%) subjects were POAG and 32 (6.4%) glaucoma suspects. Sensitivity (95% confidence interval) at 90% specificity for detection of glaucoma suspect/POAG combined was 41% (28 to 55) for FDT, 35% (21 to 48) for MMDT, and 57% (44 to 70) for best-performing OCT parameter (inferior quadrant RNFL thickness); for POAG, sensitivity was 62% (39 to 84) for FDT, 58% (37 to 78) for MMDT, and 83% (68 to 98) for inferior quadrant RNFL thickness. The partial AUROC was significantly greater for inferior RNFL thickness than visual-function tests ($p < 0.001$). Post-test probability of glaucoma suspect /POAG combined and definite POAG increased substantially when best-performing criteria were combined for FDT or MMDT, iVue OCT and ORA.

Conclusions: Diagnostic performance of individual tests gave acceptable accuracy for POAG detection. The low specificity of visual-function tests precludes their use in isolation, but case-detection improves by combining RNFL thickness analysis with visual-function tests.

Open angle glaucoma (OAG) is a major cause of visual morbidity, accounting for 10.6% to 13.5% of blindness in high-income countries.¹ However, epidemiological studies in developed countries consistently demonstrated that approximately half of those with OAG remained undetected using current case-finding strategies.²⁻⁸

OAG satisfies Wilson-Jungner criteria for the condition and treatment ideally required to initiate a screening programme.⁹ In 2012, a Comparative Effectiveness Review by the Agency for Healthcare Research and Quality concluded that limited evidence existed on the effectiveness of screening for OAG in adult populations.¹⁰ An earlier UK-based economic modelling study reported that population screening at any age was not cost-effective, but stronger evidence existed in support of targeted screening of high-risk groups.¹¹ A strategy for improving screening cost-effectiveness was proposed, involving initial technology-based assessment, allowing an enriched population to be referred for office-based assessment by an ophthalmologist or optometrist. In the context of case-finding for a low prevalence disease in the general population, an ideal screening test must be simple, fast and combine high specificity (above 90%), with acceptably high sensitivity. However, a 2008 systematic review found no single test, used alone or in combination, provided sufficiently high accuracy for OAG detection.¹² The review highlighted a dearth of high-quality diagnostic accuracy studies for OAG detection. In many cases, reliability and applicability of study findings are limited by methodology, with failure to satisfy the quality assessment of diagnostic accuracy studies (QUADAS) criteria.¹³

129

130 This study aims to determine diagnostic accuracy of modern imaging and
131 visual function testing technologies, used alone and in combination, for
132 detecting OAG in a representative sample of the primary-care population,
133 compared to a reference standard ophthalmic examination including standard
134 automated perimetry (SAP). The study was designed, and findings reported in
135 accordance with Standards for Reporting of Diagnostic Accuracy (STARD)
136 criteria.¹⁴

137 **Methods**

138 This prospective cross-sectional study was conducted in one university-based
139 community eye clinic in London, UK, during 12 months from September 2012.
140 The study was approved by the institutional review board and adhered to the
141 Declaration of Helsinki tenets. All subjects provided written informed consent.
142 Males and females aged 60 years and older were recruited. Study information,
143 together with an invitation to participate, was distributed locally through
144 neighbouring optometry practices and community groups. To ensure a
145 representative sample of the eligible population, no pre-defined exclusion
146 criteria were specified; subjects with known POAG or other ocular morbidities
147 were included.

148
149 All subjects underwent a series of technology-based index tests, followed by a
150 reference standard ophthalmic examination on the same day. Figure 1 shows
151 the study flow diagram. Thresholds of abnormality for the index tests were
152 based on cut-offs commonly reported in previous literature, manufacturers'
153 suggested cut-offs, and comparisons with internal normative databases, and
154 were specified in the protocol prior to data analysis. The technology-based
155 assessment comprised four index tests and was performed by a single,
156 experienced technician with no prior knowledge of subjects' ocular status or
157 findings from the reference standard ophthalmic examination. All equipment
158 used for tests performed during the reference standard ophthalmic
159 examination and technology-based assessment was calibrated daily in
160 accordance with manufacturers' instructions, and examinations were

undertaken in dedicated research rooms based in the community eye clinic to ensure a consistent and reliable testing environment over the 12-month period.

Visual function tests (FDT and MMDT)

The first generation frequency doubling technology (FDT; Carl Zeiss Meditec, Inc., Dublin, CA) perimeter was used in C20-5 suprathreshold mode (software version 4.00.0). Contrast thresholds are evaluated at 17 locations within the central 20° of visual field. A detailed description of measurement principles has been described elsewhere.¹⁵ An abnormal result was defined using two cut-offs: a) one or more location(s) missed at the $p < 5\%$ significance level and b) one or more location(s) missed at the $p < 1\%$ significance level. Further analysis was performed using a scoring system described by Patel et al. which allocates an overall score between 0 and 87 for each FDT result, giving increased importance to more severe defects and locations missed closer to fixation.¹⁶

The Moorfields motion displacement test (MMDT; Moorfields Eye Hospital, London, UK) is a prototype perimeter based on a form of temporal hyperacuity, in which subjects identify oscillation of a vertical bar, the threshold being the smallest displacement seen. Testing was performed using the Enhanced Standard Threshold Algorithm (ESTA) 99.5 suprathreshold program (Pandora, software version v1.7.10) (see <http://www.moorfieldsmdt.co.uk/clinicians.asp> for more details on MMDT technology). The test presents 31 stimuli on a standard laptop LCD display. Displacements seen or not seen are recorded on a pass-fail plot, and this information is used together with the ESTA spatial

filter to generate a probability plot that provides an estimate of the 'probability of true damage' (PTD) between 0 and 100 at each test location. In the present study, an abnormal plot was defined by the developers' recommended threshold of a global PTD ≥ 3.0 .

The testing order between FDT and MMDT was randomized, and these examinations were never performed in immediate succession. Tests were repeated once if one or more locations were missed, or if the result was unreliable (Table 1, available at www.aaojournal.org).

iVue Spectral Domain OCT (SD-OCT)

The iVue optical coherence tomographer (OCT; Optovue Inc., Fremont, CA) is a compact version of the RTVue OCT. Diagnostic data for OAG detection were obtained using the ganglion cell complex (GCC) protocol of the iWellness scan, and glaucoma optic nerve head (ONH) retinal nerve fibre layer (RNFL) scan patterns in software version V3.2.0.42 (details of scan protocols are described elsewhere¹⁷). Scans were initially captured through undilated pupils in dark-room illumination, and repeated following pupil dilation if data quality failed to meet manufacturers' guidelines (8%, 81 of 1009 eyes).

Of the structural parameters for GCC and RNFL thickness, the overall mean, superior hemifield and inferior hemifield thickness were analysed. RNFL thickness was further evaluated by hourglass quadrant: temporal 316 to 45 degrees, superior 46 to 135 degrees, nasal 136 to 225 degrees, and inferior 226 to 315 degrees. GCC thickness data were also represented by two

additional parameters which analyse the pattern of GCC loss using differing levels of focality: Global loss volume (GLV) and Focal loss volume (FLV). Descriptions of procedures deriving these parameters have been reported previously.¹⁸⁻²⁰ The defined cut-off for abnormality was any RNFL or GCC parameter falling outside the 99% normal limit based on manufacturers' integrated normal database.

Ocular response analyzer (ORA)

The ORA (Reichert Ophthalmic Instruments, Depew, NY, USA) is an air-puff tonometer which uses a bi-directional applanation sequence to derive two measures of corneal biomechanical properties: corneal hysteresis (CH) and corneal resistance factor (CRF), and two intraocular pressure (IOP) parameters: IOPg (Goldmann-correlated) and IOPcc (Cornea-compensated).²¹ A minimum of four measurements from each eye was acquired (software version 3.01). The highest waveform score (WS) measurement was used for analysis provided multiple measurements with similar graphical outputs had been attained²² with a WS of 3.5 or greater.^{22, 23} IOPg or IOPcc above 21mmHg was defined as the cut-off for abnormality.

Reference standard ophthalmic examination

All subjects underwent a series of standard tests for glaucoma by an experienced clinician, trained and validated in glaucoma according to UK practice, and masked to results of the preceding index tests. Validation of the reference standard examiner was confirmed by competency-based assessment, with results being compared with classification by a consultant

glaucoma sub-specialist ophthalmologist. Kappa agreement for combined and separate assessment of the optic disc and visual field ranged from 0.70 to 0.89.

Visual field testing was performed with the Humphrey Field Analyzer (HFA; Carl Zeiss Meditec, Inc., Dublin, CA) and the Swedish Interactive Thresholding Algorithm (SITA) 24-2 standard pattern (Model 720i, software version 5.1.2). Where possible, HFA was repeated for unreliable results (false negative responses or fixation losses >33%, false positive responses >15%) and Glaucoma hemifield test (GHT) recordings of 'outside normal limits'. Following full anterior segment assessment by biomicroscope, and measurement of IOP by Goldmann Applanation Tonometer, eyes with a potentially occludable angle identified by the van Herick test²⁴ were evaluated by gonioscopy. Detailed posterior segment examination was performed through dilated pupils using indirect ophthalmoscopy and fundus photography (Topcon TRC-NW8F). Subjects were asked to complete a questionnaire regarding the acceptability of each index test.

The following criteria were used for classification of subjects as definite POAG or as glaucoma suspect based on observations from one or both eyes:

- Definite POAG: open anterior chamber angle, presence of glaucomatous optic neuropathy (either localised absence of neuro-retinal rim, cup/disc ratio (CDR) of ≥ 0.7 or inter-ocular asymmetry in vertical CDR of ≥ 0.2 in similar sized discs) and the presence of a concordant glaucomatous field defect based on criteria amended from Anderson and Patella²⁵ (a cluster of

≥3 points on the pattern deviation plot having $p < 5\%$ with at least one point with $p < 1\%$, none of which can be edge points unless located immediately above or below the nasal horizontal meridian, AND pattern standard deviation (PSD) $p < 5\%$, AND GHT 'outside normal limits').

- Glaucoma suspect: included 'disc suspects' showing features of glaucomatous optic neuropathy but with normal or equivocal fields, and subjects with visual field defects but without concordant disc damage (see 'Definite POAG' above for definitions of glaucomatous optic neuropathy and visual field defects).

The ocular hypertension (OHT) case definition in this study for subjects not taking IOP-lowering medication was based on measurement of IOP above 21mmHg on two separate occasions, with open anterior chamber angles and neither visual field plots nor optic discs meeting the criteria for abnormality.

Sample size calculation

The sample size was based on an anticipated sensitivity of the index tests to detect POAG (based on current case definitions) of 0.75^{12} with a minimal acceptable precision of the sensitivity estimate of ± 0.25 with 0.95 probability. This requires 42 POAG cases. Since prevalence of suspected and definite POAG in the local elderly population would be approximately $10\%^{26}$ it was estimated that at least 420 subjects needed to be recruited.

Statistical analysis

Statistical analysis was performed using SPSS 21.0 software (www.ibm.com/SPSS_Statistics), Medcalc 14.8.1 (www.medcalc.org), and STATA 13.0 (StataCorp. 2013. College Station, TX: StataCorp LP, www.stata.com). Index data were analysed masked to findings from the reference ophthalmic examination. Unreliable results acquired by visual function tests (FDT and MMDT), and data from repeatedly poor quality ORA and OCT acquisitions were removed from analysis. The unit of analysis was the individual, and the comparison was between the most abnormal index test result from either the right or left eye and the overall reference standard classification.

Differences in mean values for demographic characteristics between diagnostic groups were evaluated by ANOVA for normally distributed data, and Kruskal-Wallis test for data with skewed distributions, each together with post-hoc analysis. For all tests, $p < 0.05$ was considered statistically significant. Initial diagnostic accuracy estimates of each index test to detect glaucoma suspect/definite POAG combined and definite POAG were evaluated using the predefined cut-offs for abnormality to generate sensitivity, specificity and likelihood ratios with 95% confidence intervals. To compare index test performance within a clinically relevant range for detection of a low prevalence disease we determined the sensitivity at 90% specificity, and normalized the partial AUROC curves to determine the average sensitivity²⁷ between 90% and 100% specificity. To test for any statistically significant differences between sensitivity at set specificity, and partial AUROC curve estimates the Wald test was used.²⁸ Best performing structural and functional criteria were combined in

311 series to calculate sensitivity and specificity values, and change from pre-test
312 to post-test probability estimates of a given subject having POAG were
313 determined using Bayesian reasoning.

Results

505 subjects entered the study (59% female and 41% male), aged between 60 and 92 years with median (interquartile range) age being 68 (59 to 77) years. Self-reported ethnicities were 88% White, 8% South Asian, 2% Black, 1% Chinese, and 1% 'other'. Based on the reference standard examination, 26 (5.1%) subjects were classified as definite POAG, 32 (6.4%) glaucoma suspect, and 17 (3.4%) OHT. Using Hodapp-Parrish-Anderson criteria,²⁹ 11 (42%) definite POAG cases were classified as early, 6 (23%) as moderate and 9 (35%) as advanced. Demographic and summary clinical data for each group are summarised in Table 2, available at www.aaojournal.org. A high proportion of subjects had ocular co-morbidities, including 9.5% with moderate or advanced AMD and 10.7% with clinically-significant cataract in one or both eyes. Following repeat examination, over 95% of results acquired using each of the four index tests were reliable or of sufficient quality for analysis (Table 1, available at www.aaojournal.org).

Diagnostic performance of visual-function tests

A FDT performance cut-off of 1 or more missed location at $p < 5\%$ level of significance, representing the most common threshold for abnormality in published literature, yielded 72.4% (CI 59.8 to 82.3) sensitivity and 66.7 (CI 62.1 to 71.0) specificity for detection of glaucoma suspect/POAG combined (Table 3). Using the same cut-off, sensitivity to detect POAG alone was 92.3% (CI 75.9 to 97.9) and specificity 65.2% (60.8 to 69.3). Test specificity improved to 79.1% (CI 75.2 to 82.5) using a test failure cut-off of 1 or more location(s) missed at $p < 1\%$ level of significance, while retaining a sensitivity of 88.5% (CI

71.0 to 96.0) for POAG detection (Table 3). The developers' recommended MMDT performance cut-off (global PTD ≥ 3.0) achieved test specificity of over 80% but lower sensitivity of 51.7% (CI 39.2 to 64.1) for glaucoma suspect/POAG combined, and 65.4% (CI 46.2 to 80.6) for POAG detection. Notably, all (100%) cases of moderate and advanced POAG (mean deviation worse than -6dB) were detected by both perimetry index tests. Of the 11 POAG subjects classified with early disease (-6dB or better), only 2 subjects (18%) were test positive using MMDT (global PTD ≥ 3.0), compared with 9 subjects (82%) detected by the less specific FDT criterion (1 or more missed location at $p < 5\%$ level of significance).

Diagnostic performance of the SD-OCT

Best performing parameters based on highest test sensitivity for detection of glaucoma suspect /POAG combined were GCC FLV (46.6%, CI 34.3 to 59.2), and inferior quadrant RNFL thickness (46.6%, CI 34.3 to 59.2). A similar trend followed for detection of POAG (GCC FLV 73.1%, CI 53.9 to 86.3; inferior quadrant thickness 76.9%, CI 57.9 to 89.0) (Table 3). Notably, all 5 GCC and 7 RNFL parameters included for analysis individually provided a test specificity exceeding 90%. In particular, GCC GLV was 97.9% (CI 96.2 to 98.8) specific for discrimination of definite POAG, with the highest positive likelihood ratio of 21.8 (CI 10.4 to 45.8) of all iVue parameters (Table 4, available at www.aaojournal.org). However, a threshold of abnormality defined by any of the 7 RNFL parameters exceeding the 99% normative level provided further diagnostic value by improving sensitivity to 62.1% (CI 49.2 to 73.4) for glaucoma suspect/POAG combined and 88.5% (CI 71.0 to 96.0) for POAG

while achieving specificity above 88%. Using the same cut-off, sensitivity improved to 93.3% (CI 70.2 to 98.8) for distinguishing POAG subjects with moderate and advanced POAG. Moreover, 25 of the 26 (96.1%, CI 81.1 to 99.3) subjects classified as POAG in the reference ophthalmic examination were detected by one or more GCC or RNFL parameter exceeding the 99% normative interval (see Table 3) for a specificity of 81.3% (77.5 to 84.6).

IOP estimates of IOPcc and IOPg generated by the ORA had little diagnostic value for distinguishing glaucoma suspect and POAG subjects from the rest of the sample.

ROC analysis

Sensitivity at 90% specificity, and partial AUROC curve for 90% to 100% specificity are summarized in Table 5 (see Table 6, available at www.aaojournal.org for data on total AUROC curves). Overall, inferior quadrant RNFL thickness measured using the iVue SD-OCT was best performing parameter, providing highest sensitivity (56.9%, CI 44.2 to 69.6 glaucoma suspect/POAG combined; 82.8%, CI 67.6 to 97.9 POAG) and partial AUROC curve estimate (0.46, CI 0.34 to 0.58 glaucoma suspect/POAG combined; 0.70, CI 0.53 to 0.86 POAG) from 90% to 100% specificity. In fact, inferior quadrant RNFL thickness was statistically significantly superior to each of the visual function tests, based on partial AUROC curve estimates (glaucoma suspect/POAG combined FDT and MMDT $p < 0.001$; POAG FDT and MMDT $p < 0.001$) (Figure 2). Of the visual-function tests, FDT Patel et al. score (2000) achieved higher sensitivity (61.5%, CI 39.4 to 83.6) but a lower

partial AUROC curve result (0.35, CI 0.18 to 0.52) compared with MMDT global PTD (57.7%, CI 37.4 to 78.0 sensitivity, 0.44, CI 0.26 to 0.61 partial AUROC curve) for ranges starting from 90% specificity for distinguishing POAG from the rest of the sample, but these observations did not represent a statistically significant difference (sensitivity at set specificity $p=0.598$, partial AUROC curve $p=0.248$) (Figure 2).

Combining index test results

The combination of inferior quadrant RNFL thickness ($p<1\%$) with FDT (1 or more location(s) missed at $p<5\%$ level) in which failure of either test is indicative of abnormality achieves a sensitivity of 79.3% (CI 67.2 to 87.7) for glaucoma suspect/POAG combined and 100.0% (CI 87.1 to 100.0) for POAG detection but with a marked reduction in specificity (glaucoma suspect/ POAG combined 63.3, CI 58.9 to 67.6; POAG 65.2, CI 60.7 to 69.5). On the other hand, stipulating that failure of both tests was indicative of POAG improved specificity to 96.8% (CI 94.8 to 98.1), but this did not represent a statistically significant improvement above test specificity of 95.0% (CI 92.6 to 96.6) achieved by inferior quadrant thickness alone (McNemar, $p=1.0$). Notably, the combination of iVue SD-OCT RNFL inferior quadrant parameter ($p<1\%$) with FDT (1 or more missed location at $p<5\%$ level) detected all 26 subjects classified as POAG (Figure 3, available at www.aaojournal.org).

To further evaluate the diagnostic value of combining index test data using Bayesian probabilistic reasoning, best-performing parameters and cut-offs for abnormality were selected using the highest positive likelihood ratios (Table 4,

415 available at www.aaojournal.org). The probability estimate of a given subject
416 having POAG rose from 5% (pre-test probability) to over 85% (post-test
417 probability) when visual function tests (FDT, 1 or more missed location at
418 $p < 1\%$ level or MMDT, global PTD ≥ 3.0) were combined in series with best
419 performing structural parameters (RNFL inferior quadrant thickness or GCC
420 GLV, $p < 1\%$), and ORA IOPcc (> 21 mmHg). Using these test cut-offs, a post-
421 test probability over 90% was achieved for detection of glaucoma
422 suspect/POAG combined, rising from a pre-test probability of 11.5%.

Discussion

Currently, a national population-based screening programme for OAG has not been implemented in any country. An economic modeling study undertaken in Finland determined that an organized screening programme for glaucoma could be a cost-effective strategy compared to opportunistic case-finding, especially in older age groups.³⁰ A UK-based study using a similar approach to evaluate the clinical and cost-effectiveness of screening for POAG proposed the use of tonometry combined with an initial technology-based assessment, which would allow an enriched population to be referred for an office-based assessment by an ophthalmologist.¹¹ Alternatively, clinical data collected from a technology-based assessment could be transferred digitally and evaluated in a virtual clinic by a glaucoma specialist to improve the positive predictive value of referrals for further ophthalmic investigation.^{31, 32} Cost-effectiveness may be improved by implementing a screening programme that targeted a number of sight-threatening eye diseases.

The current study evaluated the diagnostic performance of structural and visual function tests for the detection of glaucoma in a population of elderly subjects, representative of the target population for screening, in which pathologies other than glaucoma may be present. Data were analyzed using the individual as the unit of analysis. The performance of the FDT using the C20-5 screening program was similar to that reported in previous population screening studies.^{33, 34} However, there has only been one published diagnostic accuracy study evaluating the MMDT.³⁵ This study found sensitivities and specificities of greater than 85%. It is likely that the lower performance of the

MMDT in the current study relates to the high levels of ocular co-morbidity typical of an elderly population, which may have impacted on the overall performance of the vision-function tests. ROC analysis of the FDT and MMDT, based on sensitivities at set specificities and partial AUROC, showed no statistical difference in performance between the two tests for the detection of POAG. However, in view of the MMDTs greater portability, ease of use and relatively lower cost it warrants further evaluation in population studies to further determine its potential as a screening test for glaucoma.

The iVue OCT is a recently developed compact SD-OCT and this is the first study to investigate its diagnostic performance for glaucoma detection using its in-built normative database. The structural parameters selected for the analysis and associated pass-fail criteria (value outside the 99% confidence interval) were established *a priori*. The best performing individual structural parameter (inferior quadrant RNFL thickness) provided a sensitivity of over 75% with a specificity of 95%, which may reflect the vulnerability of the inferior quadrant of the optic disc to glaucomatous damage.^{36, 37} The OCT was particularly effective in identifying subjects with glaucoma, for example using a criterion of any structural parameter at the $p < 1\%$ level the OCT would have identified 25 of 26 glaucoma subjects in our sample. ORA-derived IOP estimates were of limited diagnostic value in our population as half of the 26 glaucoma subjects were already receiving IOP-lowering therapy or had previously undergone surgical or laser interventions.

Early detection and treatment of glaucoma reduces the rate of progression of glaucomatous vision loss and visual field defects,^{38, 39} which is likely to result in a better health-related quality of life for those affected, but concerns have been raised as to the potential overtreatment of individuals who may not be at significant risk of developing advanced glaucoma and visual impairment in their lifetime.¹¹ A retrospective UK study using a large visual field dataset, and modelling projected field loss in the patients' remaining lifetime, determined that only 5.2% of patients were at risk of progressing to statutory blindness in both eyes; more than 90% of these had a visual field mean deviation worse than -6dB in one or both eyes at presentation.⁴⁰ Given that the likelihood of patients suffering significant visual impairment is linked to the level of VF loss at presentation, it is notable that 100% of those in the current study with moderate or advanced glaucoma (mean deviation worse than -6dB) were detected by either the FDT ($p < 5\%$ level), or the MMDT (global PTD ≥ 3.0).

The natural history of glaucoma means that in some people with early disease, structural changes precede functional loss, whilst in others functional abnormalities may be observed before detectable changes in structural parameters.⁴¹ In the current study, thirty-two subjects fell into either category and were classified as 'glaucoma suspects'. Differentiating between suspects and normals presents a significant clinical challenge, as there is a substantial overlap of clinical characteristics between the groups. All four index tests showed poorer discrimination between normal subjects and POAG/glaucoma suspect groups combined than between those with confirmed glaucoma and the rest of the sample. The detection of glaucoma suspects requires a case

definition based on failure on either a structural or functional test. Whilst this strategy is likely to improve sensitivity it is generally at the expense of specificity. An alternative case-finding strategy is to use a Bayesian reasoning approach. In clinical practice, a clinician will intuitively integrate the results of diagnostic tests together with an estimate of the patient's pre-test probability of disease based on age, IOP and family history of glaucoma to estimate an individual's post-test probability. The probability of disease can be formally estimated by calculations using the likelihood ratios of the diagnostic tests. The results of independent tests can be combined in series to revise post-test probability estimates.⁴² However, the lack of true independence between structural and functional criteria may lead to an overestimation of the combined post-test probability. Nevertheless, this Bayesian approach could be used to develop diagnostic algorithms and has great potential for glaucoma case-finding or population screening pathways.⁴³

The present study had a number of strengths: the design, analysis and reporting complied with the principles of the STARD statement¹⁴ and to reduce spectrum bias the target population included consecutive subjects who met the inclusion criteria. Although it is possible that higher numbers of those with previous or family ocular history were more likely to volunteer and agree to participate in the study, the prevalence of OAG in our population (5%) was comparable with that expected for the age demographic. Furthermore, a wide spectrum of disease severity was identified. We therefore feel the population is likely to be broadly representative of those presenting for glaucoma case-finding in the community. The reference standard for OAG corresponded to

that used in a typical hospital glaucoma unit and was based on the results of a standard ophthalmic examination by a validated clinician. At the present time, this examination represents the clinical reference standard for OAG, but as evidence accumulates it is anticipated that OCT may become part of this standard in the future. All index tests and the reference standard examination were undertaken on the same day, and the clinicians performing the reference and index tests were masked to the outcome of either. The study also has some limitations. The sample size of 505 subjects provided only 26 glaucoma subjects. This resulted in wide confidence intervals around our diagnostic sensitivity estimates, which may have masked real differences between index tests. Furthermore, almost 90% of our study population was of White European origin suggesting our findings may not be generalizable to other ethnic groups where glaucoma is more prevalent (e.g. subjects of Black origin). Data collection for this study was undertaken in dedicated research rooms based in a community eye clinic. In a real-world clinic setting, equipment may not be calibrated routinely and it is anticipated that diagnostic performance may be less good. Nevertheless, this study provides useful data to inform the development of further larger multi-center glaucoma screening studies.

Acknowledgements:

The authors wish to thank the UK College of Optometrists who provided a PhD studentship to Priya Dabasia, Optovue Inc. for the loan of the iVue OCT, and Tunde Peto (NIHR BMRC at Moorfields Eye Hospital NHS Foundation

546 Trust and UCL Institute of Ophthalmology, London, UK) for providing the
547 medical retina images and facilitating the core-competency-based
548 assessments for the validation of the reference standard examiner.

Legends for Figures 1 and 2

Figure 1: Study flow diagram. FDT = frequency doubling technology perimeter; MMDT = Moorfields motion displacement test; SD-OCT = spectral domain optical coherence tomographer; ORA = ocular response analyzer; POAG = primary open angle glaucoma; OHT = ocular hypertension.

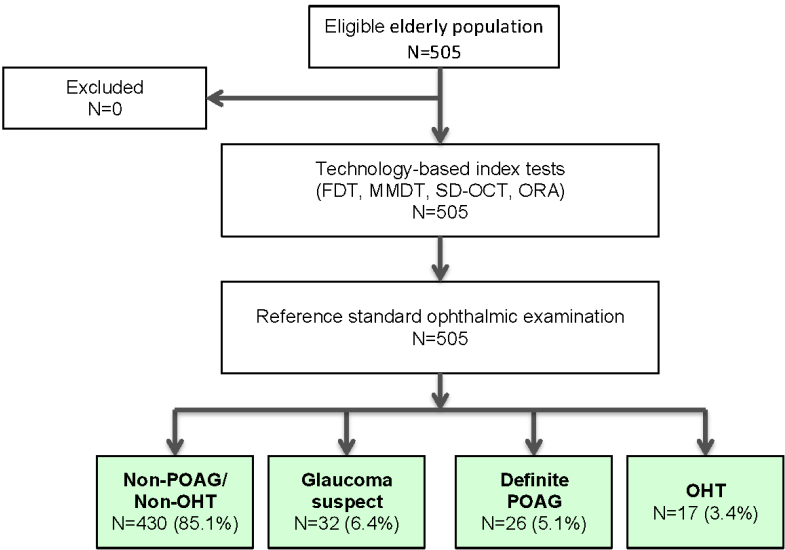
Figure 2: Index test diagnostic effectiveness comparisons using ROC curves with sensitivity at set specificity estimates and associated 95% confidence intervals for detection of glaucoma suspect/POAG (primary open angle glaucoma) combined (a) and POAG (b). FDT = Frequency Doubling Technology Perimeter; MMDT = Moorfields motion displacement threshold test; RNFL = retinal nerve fibre layer thickness.

563 REFERENCES

- 564 1. Bourne RR, Jonas JB, Flaxman SR, et al. Prevalence and causes of
565 vision loss in high-income countries and in Eastern and Central Europe: 1990-
566 2010. *Br J Ophthalmol* 2014;98:629-38.
- 567 2. Tielsch JM, Katz J, Singh K, et al. A population-based evaluation of
568 glaucoma screening: the Baltimore Eye Survey. *Am J Epidemiol*
569 1991;134:1102-10.
- 570 3. Mitchell P, Smith W, Attebo K, et al. Prevalence of open-angle
571 glaucoma in Australia. The Blue Mountains Eye Study. *Ophthalmology*
572 1996;103:1661-9.
- 573 4. Quigley HA & Vitale S. Models of open-angle glaucoma prevalence and
574 incidence in the United States. *Invest Ophthalmol Vis Sci* 1997;38:83-91.
- 575 5. Wensor MD, McCarty CA, Stanislavsky YL, et al. The prevalence of
576 glaucoma in the Melbourne Visual Impairment Project. *Ophthalmology*
577 1998;105:733-9.
- 578 6. Klein BE, Klein R, Sponsel WE, et al. Prevalence of glaucoma. The
579 Beaver Dam Eye Study. *Ophthalmology* 1992;99:1499-504.
- 580 7. Dielemans I, Vingerling JR, Wolfs RC, et al. The prevalence of primary
581 open-angle glaucoma in a population-based study in The Netherlands. The
582 Rotterdam Study. *Ophthalmology* 1994;101:1851-5.
- 583 8. Weih LM, Nanjan M, McCarty CA, et al. Prevalence and predictors of
584 open-angle glaucoma: results from the visual impairment project.
585 *Ophthalmology* 2001;108:1966-72.
- 586 9. Wilson JMG & Jungner G. Principles and practice for screening for
587 disease. *Public Health Pap* 1968;34.
- 588 10. Ervin AM, Boland MV, Myrowitz EH, et al. AHRQ Comparative
589 Effectiveness Reviews. Screening for Glaucoma: Comparative Effectiveness.
590 Rockville (MD): Agency for Healthcare Research and Quality (US); 2012.
- 591 11. Burr JM, Mowatt G, Hernandez R, et al. The clinical effectiveness and
592 cost-effectiveness of screening for open angle glaucoma: a systematic review
593 and economic evaluation. *Health Technol Assess* 2007;11:iii-iv, ix-x, 1-190.
- 594 12. Mowatt G, Burr JM, Cook JA, et al. Screening tests for detecting open-
595 angle glaucoma: systematic review and meta-analysis. *Invest Ophthalmol Vis*
596 *Sci* 2008;49:5373-85.
- 597 13. Whiting P, Rutjes AW, Reitsma JB, et al. The development of QUADAS:
598 a tool for the quality assessment of studies of diagnostic accuracy included in
599 systematic reviews. *BMC Med Res Methodol* 2003;3:25.
- 600 14. Bossuyt PM, Reitsma JB, Bruns DE, et al. The STARD statement for
601 reporting studies of diagnostic accuracy: explanation and elaboration. *Ann*
602 *Intern Med* 2003;138:W1-12.
- 603 15. Johnson CA. FDT perimetry for the detection of glaucomatous visual
604 field loss. *J Glaucoma Today* 2008:26-8.
- 605 16. Patel SC, Friedman DS, Varadkar P, et al. Algorithm for interpreting the
606 results of frequency doubling perimetry. *Am J Ophthalmol* 2000;129:323-7.
- 607 17. Aref AA & Budenz DL. Spectral domain optical coherence tomography
608 in the diagnosis and management of glaucoma. *Ophthalmic Surg Lasers*
609 *Imaging* 2010;41 Suppl:S15-27.
- 610 18. Sinai M. Direct Ganglion Cell Assessment with the RTVue: The
611 Ganglion Cell Complex Analysis. Available from

- 612 http://adaptitda.com.br/conteudo/opto/rtvue_files/1281618322_directganglioncellassessmentwiththertvue_revb_final.pdf (accessed 19 December 2014).
- 613 2008.
- 614 19. Rao HL, Babu JG, Addepalli UK, et al. Retinal nerve fiber layer and
- 615 macular inner retina measurements by spectral domain optical coherence
- 616 tomograph in Indian eyes with early glaucoma. *Eye (London, England)*
- 617 2012;26:133-9.
- 618 20. Tan O, Chopra V, Lu AT, et al. Detection of macular ganglion cell loss in
- 619 glaucoma by Fourier-domain optical coherence tomography. *Ophthalmology*
- 620 2009;116:2305-14.
- 621 21. Terai N, Raiskup F, Haustein M, et al. Identification of biomechanical
- 622 properties of the cornea: the ocular response analyzer. *Curr Eye Res*
- 623 2012;37:553-62.
- 624 22. Vantomme M, Pourjavan S & Detry-Morel M. The range of the
- 625 waveform score of the ocular response analyzer (ora) in healthy subjects. *Bull*
- 626 *Soc Belge Ophtalmol* 2013;91-7.
- 627 23. Lam AK, Chen D & Tse J. The usefulness of waveform score from the
- 628 ocular response analyzer. *Optom Vis Sci* 2010;87:195-9.
- 629 24. Van Herick W, Shaffer RN & Schwartz A. Estimation of width of angle of
- 630 anterior chamber. Incidence and significance of the narrow angle. *Am J*
- 631 *Ophthalmol* 1969;68:626-9.
- 632 25. Anderson DR & Patella VM. *Automated static perimetry* (2nd ed.).
- 633 Mosby, St Louis, MO; 1999:147-159.
- 634 26. Reidy A, Minassian DC, Vafidis G, et al. Prevalence of serious eye
- 635 disease and visual impairment in a north London population: population based,
- 636 cross sectional study. *BMJ* 1998;316:1643-6.
- 637 27. Hillis SL & Metz CE. An analytic expression for the binormal partial area
- 638 under the ROC curve. *Acad Radiol* 2012;19:1491-8.
- 639 28. Pepe M, Longton G & Janes H. Estimation and Comparison of Receiver
- 640 Operating Characteristic Curves. *The Stata journal* 2009;9:1.
- 641 29. Hodapp E, Parrish IRK & Anderson DR. *Clinical decisions in glaucoma*.
- 642 St. Louis, The C.V, Mosby Co.; 1993:52-61.
- 643 30. Vaahtoranta-Lehtonen H, Tuulonen A, Aronen P, et al. Cost
- 644 effectiveness and cost utility of an organized screening programme for
- 645 glaucoma. *Acta Ophthalmol Scand* 2007;85:508-18.
- 646 31. Trikha S, Macgregor C, Jeffery M, et al. The Portsmouth-based
- 647 glaucoma refinement scheme: a role for virtual clinics in the future? *Eye*
- 648 (London, England) 2012;26:1288-94.
- 649 32. Rathod D, Win T, Pickering S, et al. Incorporation of a virtual
- 650 assessment into a care pathway for initial glaucoma management: Feasibility
- 651 study. *Clinical and Experimental Ophthalmology* 2008;36:543-6.
- 652 33. Robin TA, Muller A, Rait J, et al. Performance of community-based
- 653 glaucoma screening using Frequency Doubling Technology and Heidelberg
- 654 Retinal Tomography. *Ophthalmic Epidemiol* 2005;12:167-78.
- 655 34. Detry-Morel M, Zeyen T, Kestelyn P, et al. Screening for glaucoma in a
- 656 general population with the non-mydratic fundus camera and the frequency
- 657 doubling perimeter. *Eur J Ophthalmol* 2004;14:387-93.
- 658 35. Ong EL, Zheng Y, Aung T, et al. Performance of the Moorfields motion
- 659 displacement test for identifying eyes with glaucoma. *Ophthalmology*
- 660 2014;121:88-92.
- 661

36. Jonas JB, Fernandez MC & Sturmer J. Pattern of glaucomatous neuroretinal rim loss. *Ophthalmology* 1993;100:63-8.
37. Hood DC, Raza AS, de Moraes CG, et al. Glaucomatous damage of the macula. *Prog Retin Eye Res* 2013;32:1-21.
38. AGIS. The Advanced Glaucoma Intervention Study (AGIS): 7. The relationship between control of intraocular pressure and visual field deterioration. The AGIS Investigators. *Am J Ophthalmol* 2000;130:429-40.
39. Heijl A, Leske MC, Hyman L, et al. Reduction of intraocular pressure and glaucoma progression: Results from the Early Manifest Glaucoma Trial. *Arch Ophthalmol* 2002;120:1268-79.
40. Saunders LJ, Russell RA, Kirwan JF, et al. Examining visual field loss in patients in glaucoma clinics during their predicted remaining lifetime. *Invest Ophthalmol Vis Sci* 2014;55:102-9.
41. Kass MA, Heuer DK, Higginbotham EJ, et al. The Ocular Hypertension Treatment Study: a randomized trial determines that topical ocular hypotensive medication delays or prevents the onset of primary open-angle glaucoma. *Arch Ophthalmol* 2002;120:701-13; discussion 829-30.
42. Deeks JJ & Altman DG. Diagnostic tests 4: likelihood ratios. *BMJ* 2004;329:168-9.
43. Garway-Heath DF & Friedman DS. How should results from clinical tests be integrated into the diagnostic process? *Ophthalmology* 2006;113:1479-80.



686
687

688

