Cochrane Library

Cochrane Database of Systematic Reviews

Optic nerve head and fibre layer imaging for diagnosing glaucoma (Review)

Michelessi M, Lucenteforte E, Oddone F, Brazzelli M, Parravano M, Franchi S, Ng SM, Virgili G

Michelessi M, Lucenteforte E, Oddone F, Brazzelli M, Parravano M, Franchi S, Ng SM, Virgili G.
Optic nerve head and fibre layer imaging for diagnosing glaucoma.
Cochrane Database of Systematic Reviews 2015, Issue 11. Art. No.: CD008803.
DOI: 10.1002/14651858.CD008803.pub2.

www.cochranelibrary.com

WILEY

TABLE OF CONTENTS

HEADER 1
ABSTRACT 1
PLAIN LANGUAGE SUMMARY 2
SUMMARY OF FINDINGS 4
BACKGROUND 6
OBJECTIVES 7
METHODS 8
RESULTS 10
Figure 1. 11
Figure 2. 12
Figure 3. 14
Figure 4. 15
DISCUSSION 16
AUTHORS' CONCLUSIONS 18
ACKNOWLEDGEMENTS 18
REFERENCES 19
CHARACTERISTICS OF STUDIES 27
DATA 267
Test 1. GDx: Inferior average 270
Test 2. GDx: NFI 271
Test 3. GDx: Superior average 272
Test 4. GDx: TSNIT average 273
Test 5. OCT: RNFL average 274
Test 6. OCT: RNFL inferior quadrant 276
Test 7. OCT: RNFL nasal quadrant 277
Test 8. OCT: RNFL superior quadrant 278
Test 9. OCT: RNFL temporal quadrant 279
Test 10. HRT: Bathija function 280
Test 11. HRT: Cup area 280
Test 12. HRT: C/D area ratio 280
Test 13. HRT: vertical C/D ratio 281
Test 14. HRT: Cup shape measure 281
Test 15. HRT: Cup volume 281
Test 16. HRT: FSM discriminant function o Mikelberg function 282
Test 17. HRT: MRA 282
Test 18. HRT: Rim area 282
Test 19. HRT: RB discriminant function 283
Test 20. HRT: Rim Volume 283
Test 21. OCT: GCC RTVue average thickness 284
Test 22. OCT: GCC RTVue superior thickness 285
Test 23. OCT: GCC RTVue inferior thickness 286
Test 24. OCT: GCC RTVue FLV 287
Test 25. OCT: GCC RTVue GLV 287
Test 26. OCT: GCC 3DTopcon average thickness 288
Test 27. OCT: GCC 3DTopcon superior thickness 288
Test 28. OCT: GCC 3DTopcon inferior thickness 288
Test 29. OCT: GCIPL Cirrus average thickness 289
Test 30. OCT: GCIPL Cirrus minimum thickness 289
Test 31. OCT: GCIPL Cirrus superior thickness 289
Test 32. OCT: GCIPL Cirrus inferior thickness 290
Test 33. OCT: ONH Disc area 290

Test 34. OCT: ONH Cup area 291
Test 35. OCT: ONH Rim area 291
Test 36. OCT: ONH Rim volume 292
Test 37. OCT: ONH Nerve head volume 292
Test 38. OCT: ONH Cup volume 293
Test 39. OCT: ONH C/D area ratio 293
Test 40. OCT: ONH horizontal C/D ratio 294
Test 41. OCT: ONH vertical C/D ratio 294
Test 42. OCT: GCIPL Cirrus Inferonasal quadrant 295
Test 43. OCT: GCIPL Cirrus Inferotemporal quadrant 295
Test 44. OCT: GCIPL Cirrus Superonasal quadrant 295
Test 45. OCT: GCIPL Cirrus Superotemporal quadrant 296
Test 46. OCT: GCC Spectralis average thickness 296
Test 47. Direct comparison: GDx NFI 296
Test 48. Direct comparison: OCT RNFL average 296
ADDITIONAL TABLES 297
APPENDICES 303
WHAT'S NEW 311
HISTORY 312
CONTRIBUTIONS OF AUTHORS 312
DECLARATIONS OF INTEREST 312
SOURCES OF SUPPORT 312
DIFFERENCES BETWEEN PROTOCOL AND REVIEW 313
INDEX TERMS 313

[Diagnostic Test Accuracy Review]

Optic nerve head and fibre layer imaging for diagnosing glaucoma

Manuele Michelessi[1], Ersilia Lucenteforte[2], Francesco Oddone[1], Miriam Brazzelli[3], Mariacristina Parravano[1], Sara Franchi[4], Sueko M Ng[5], Gianni Virgili[6]

[1]Ophthalmology, Fondazione G.B. Bietti per lo studio e la ricerca in Oftalmolologia-IRCCS, Rome, Italy. [2]Department of Clinical and Experimental Medicine, University of Pisa, Pisa, Italy. [3]Health Services Research Unit, University of Aberdeen, Aberdeen, UK. [4]Azienda Ospedaliero Universitaria Careggi, Florence, Italy. [5]Department of Ophthalmology, School of Medicine, University of Colorado, Aurora, CO, USA. [6]Department of Neurosciences, Psychology, Drug Research and Child Health (NEUROFARBA), University of Florence, Florence, Italy

Contact address: Gianni Virgili, gianni.virgili@unifi.it.

Editorial group: Cochrane Eyes and Vision Group.
Publication status and date: Edited (no change to conclusions), published in Issue 8, 2020.

Citation: Michelessi M, Lucenteforte E, Oddone F, Brazzelli M, Parravano M, Franchi S, Ng SM, Virgili G. Optic nerve head and fibre layer imaging for diagnosing glaucoma. *Cochrane Database of Systematic Reviews* 2015, Issue 11. Art. No.: CD008803. DOI: 10.1002/14651858.CD008803.pub2.

ABSTRACT

Background

The diagnosis of glaucoma is traditionally based on the finding of optic nerve head (ONH) damage assessed subjectively by ophthalmoscopy or photography or by corresponding damage to the visual field assessed by automated perimetry, or both. Diagnostic assessments are usually required when ophthalmologists or primary eye care professionals find elevated intraocular pressure (IOP) or a suspect appearance of the ONH. Imaging tests such as confocal scanning laser ophthalmoscopy (HRT), optical coherence tomography (OCT) and scanning laser polarimetry (SLP, as used by the GDx instrument), provide an objective measure of the structural changes of retinal nerve fibre layer (RNFL) thickness and ONH parameters occurring in glaucoma.

Objectives

To determine the diagnostic accuracy of HRT, OCT and GDx for diagnosing manifest glaucoma by detecting ONH and RNFL damage.

Search methods

We searched several databases for this review. The most recent searches were on 19 February 2015.

Selection criteria

We included prospective and retrospective cohort studies and case-control studies that evaluated the accuracy of OCT, HRT or the GDx for diagnosing glaucoma. We excluded population-based screening studies, since we planned to consider studies on self-referred people or participants in whom a risk factor for glaucoma had already been identified in primary care, such as elevated IOP or a family history of glaucoma. We only considered recent commercial versions of the tests: spectral domain OCT, HRT III and GDx VCC or ECC.

Data collection and analysis

We adopted standard Cochrane methods. We fitted a hierarchical summary ROC (HSROC) model using the *METADAS* macro in SAS software. After studies were selected, we decided to use 2 x 2 data at 0.95 specificity or closer in meta-analyses, since this was the most commonly-reported level.

Main results

We included 106 studies in this review, which analysed 16,260 eyes (8353 cases, 7907 controls) in total. Forty studies (5574 participants) assessed GDx, 18 studies (3550 participants) HRT, and 63 (9390 participants) OCT, with 12 of these studies comparing two or three tests. Regarding study quality, a case-control design in 103 studies raised concerns as it can overestimate accuracy and reduce the applicability of the results to daily practice. Twenty-four studies were sponsored by the manufacturer, and in 15 the potential conflict of interest was unclear.

Comparisons made within each test were more reliable than those between tests, as they were mostly based on direct comparisons within each study.The Nerve Fibre Indicator yielded the highest accuracy (estimate, 95% confidence interval (CI)) among GDx parameters (sensitivity: 0.67, 0.55 to 0.77; specificity: 0.94, 0.92 to 0.95). For HRT measures, the Vertical Cup/Disc (C/D) ratio (sensitivity: 0.72, 0.60 to 0.68; specificity: 0.94, 0.92 to 0.95) was no different from other parameters. With OCT, the accuracy of average RNFL retinal thickness was similar to the inferior sector (0.72, 0.65 to 0.77; specificity: 0.93, 0.92 to 0.95) and, in different studies, to the vertical C/D ratio.

Comparing the parameters with the highest diagnostic odds ratio (DOR) for each device in a single HSROC model, the performance of GDx, HRT and OCT was remarkably similar. At a sensitivity of 0.70 and a high specificity close to 0.95 as in most of these studies, in 1000 people referred by primary eye care, of whom 200 have manifest glaucoma, such as in those who have already undergone some functional or anatomic testing by optometrists, the best measures of GDx, HRT and OCT would miss about 60 cases out of the 200 patients with glaucoma, and would incorrectly refer 50 out of 800 patients without glaucoma. If prevalence were 5%, e.g. such as in people referred only because of family history of glaucoma, the corresponding figures would be 15 patients missed out of 50 with manifest glaucoma, avoiding referral of about 890 out of 950 non-glaucomatous people.

Heterogeneity investigations found that sensitivity estimate was higher for studies with more severe glaucoma, expressed as worse average mean deviation (MD): 0.79 (0.74 to 0.83) for MD < -6 db versus 0.64 (0.60 to 0.69) for MD ≥ -6 db, at a similar summary specificity (0.93, 95% CI 0.92 to 0.94 and, respectively, 0.94; 95% CI 0.93 to 0.95; P < 0.0001 for the difference in relative DOR).

Authors' conclusions

The accuracy of imaging tests for detecting manifest glaucoma was variable across studies, but overall similar for different devices. Accuracy may have been overestimated due to the case-control design, which is a serious limitation of the current evidence base.

We recommend that further diagnostic accuracy studies are carried out on patients selected consecutively at a defined step of the clinical pathway, providing a description of risk factors leading to referral and bearing in mind the consequences of false positives and false negatives in the setting in which the diagnostic question is made. Future research should report accuracy for each threshold of these continuous measures, or publish raw data.

PLAIN LANGUAGE SUMMARY

Tests for imaging the optic nerve and its fibres for diagnosing glaucoma

Review question

We reviewed the evidence about the accuracy of confocal scanning laser ophthalmoscopy (commercially available as the Heidelberg Retinal Tomogram (HRT)), optical coherence tomography (OCT) and scanning laser polarimetry (as used by the GDx device) for diagnosing glaucoma in people who are at risk. These tests can measure the structure of the optic nerve head or measure the thickness of the nerve's fibres, or both.

Background

Glaucoma is a progressive neurodegenerative disease that affects the optic nerve, with corresponding damage to the visual field. The course of the disease can be slowed or halted by reducing intraocular pressure with eye drops or surgery.

Study characteristics

We found 106 studies, mostly assessing a single device, which analysed 16,260 eyes (8353 cases, 7907 controls). Forty studies (5574 participants) assessed GDx, 18 studies (3550 participants) HRT, and 63 (9390 patients) OCT. Twenty-four studies were sponsored by the manufacturer, and in 15 the study funding was unclear. The final diagnosis of glaucoma had to be confirmed by clinical examination, including visual field testing or clinical optic nerve examination or both. However, we could not find studies comparing two tests, the most robust way to test these instruments, and including a series of consecutive patients at risk as seen in routine care, as we had hoped. Rather, we found studies assessing the performance of a single test in people without glaucoma as opposed to its performance in people with a previous diagnosis of glaucoma. The study search is current to 19 February 2015.

Key results

The performance of all devices was very variable across studies, but overall similar. In 1000 people referred by primary eye care, of whom 200 (20%) have manifest glaucoma, such as in those who have already undergone some functional or anatomic testing by optometrists, the best measures of GDx, HRT and OCT would miss about 60 cases out of the 200 patients with glaucoma (sensitivity 70%), and would incorrectly refer 50 out of 800 patients without glaucoma (at specificity 95%). If prevalence were 5%, for example, in people referred only

because of family history of glaucoma, the corresponding figures would be 15 patients missed out of 50 with manifest glaucoma, avoiding referral of about 890 out of 950 non-glaucomatous people.

The tests were better at detecting more severe glaucoma compared to early glaucoma.

Quality of the evidence
The selection of two well-defined groups of healthy and glaucoma eyes in nearly all studies, rather than the use of these imaging tests in a series of patients at risk of glaucoma as in the real world, may overestimate the accuracy of these devices compared to what could be achieved in daily practice.

Summary of findings 1. Summary of diagnostic accuracy of the best measure of all tests

What is the accuracy of GDx, HRT and OCT for diagnosing manifest glaucoma?									
Patients/population	Patients with manifest glaucoma compared to healthy controls								
Prior testing	Unclear (case-control design and insufficient reporting for nearly all studies)								
Settings	Studies carried out at glaucoma clinics								
Index test	Scanning Laser Polarimetry (GDx), Heidelberg Retina Tomograph II (HRT), Optical Coherence Tomography (OCT)								
Importance	Objective and reproducible test								
Reference standard	Clinical assessment of visual field or optic nerve head or both								
Studies	Case-control design for all studies								
Quality and Comments	Case-control design overestimates accuracy and makes inference difficult								
Test parameter	**N. studies (participants)**	**Sensitivity (95% CI)**	**Specificity (95% CI)**	**Implications in 1000 patients referred from primary care for clinician's assessment**					
				Manifest glaucoma prevalence 5% 50 cases out of 1000 referrals			**Manifest glaucoma prevalence 20% 200 cases out of 1000 referrals**		
				Glaucoma detected	**Missed**	**Referred, but no glaucoma**	**Glaucoma detected**	**Missed**	**Referred, but no glaucoma**
GDx NFI	35 (4958)	0.76 (0.70 to 0.81)	0.92 (0.90 to 0.94)	38	12	76	152	64	48
HRT vertical C/D ratio	8 (1849)	0.67 (0.55 to 0.77)	0.94 (0.92 to 0.95)	34	16	57	134	66	48
OCT C/D vertical ratio	15 (2389)	0.72 (0.60 to 0.81)	0.94 (0.92 to 0.95)	36	14	57	144	56	48

OCT RNFL inferior sector	57 (8223)	0.72 (0.65 to 0.77)	0.93 (0.92 to 0.95)	36	14	67	140	56	56
Heterogeneity investigation: sensitivity was better for detecting more severe glaucoma cases (MD ≥ 6 db: 0.79, 95% CI 0.74 to 0.83) as compared to milder cases (MD < 6 db: 0.64, 95% CI 0.60 to 0.69, P = < 0.0001) at a specificity of 0.04 (95% CI 0.93 to 0.95) and 0.93 (95% CI 0.92 to 0.94) respectively.									
CAUTION: The results on this table should not be interpreted in isolation from the results of the individual included studies contributing to each summary test accuracy measure. These are reported in the main body of the text of the review									

BACKGROUND

Target condition being diagnosed

Glaucoma is a group of progressive optic neuropathies that have in common a slow progressive degeneration of retinal ganglion cells and their axons, resulting in a distinct appearance of the optic disc and retinal nerve fibre layer (RNFL) and a concomitant pattern of visual loss (Weinreb 2004).

Without adequate treatment, glaucoma can progress to visual disability and eventual blindness (Quigley 2006). Vision loss caused by glaucoma is irreversible, and glaucoma is the second leading cause of blindness in the world. It is estimated that glaucoma affects more than 66 million individuals worldwide with at least eight million bilaterally blind.

The overall risk of developing glaucoma increases with the number and strength of risk factors. It increases substantially with the level of intraocular pressure (IOP) elevation and with increasing age (OHTS 2002). Other strong risk factors include some visual field (VF) abnormalities seen in otherwise usual baseline visual field examinations, high myopia and family history of glaucoma. Recently, a thin cornea and a vertical or horizontal cup-to-disc ratio of greater than 0.4 (as determined from stereoscopic disc photographs) have been added to the list of risk factors for developing glaucoma (Coleman 2008; OHTS 2002).

Disease progression rates in primary open angle glaucoma, the most common form of glaucoma in Europe, differ strongly between patients from rapid to very slow. Many patients show no or only small deterioration, even after years of follow-up (EMGT 1999; Wilson 2002). Most cases of glaucoma are not discovered until vision has already been permanently lost because clinical signs of early glaucoma are subtle, even to an eye specialist (Weinreb 2004). In most cases, the loss of vision caused by glaucoma can be limited or prevented by currently available therapies if the disease is identified in its early stages (AGIS 1994; CIGTS 1999; EMGT 1999).

The goal of glaucoma treatment is to maintain the visual function and related quality of life at a sustainable cost (EGS 2008 Guidelines). Currently, the only approach proven to be efficient in preserving visual function is lowering the IOP (AGIS 1994; CIGTS 1999; EMGT 1999; OHTS 2002). It has been estimated that each single mmHg of pressure reduction obtained with treatment accounts for a 10% to 19% reduction of risk of progression (Chauhan 2008; EMGT 1999).

The diagnosis of glaucoma is traditionally based on the finding of visual field damage with automated perimetry, glaucomatous damage to the optic nerve head (ONH), or both (EGS 2008 Guidelines). Diagnostic assessments are usually required when ophthalmologists or primary care physicians find an elevated IOP or a suspected anomaly of the optic nerve head such as a large cup/disc ratio or a focal rim notch.

Visual field damage is commonly assessed with automated perimetry. A variety of visual field scoring systems or algorithms have been adopted in cohort studies to diagnose the presence of glaucoma (AGIS 1994; Brusini 2006b; CIGTS 1999; EMGT 1999; Mills 2006; Spaeth 2006). However, no scoring system has yet been accepted as a reference standard. Furthermore, visual field examination is not completely reliable and repeated testing may be needed to diagnose cases with modest damage (Katz 1995; Spry 2003). Moreover, ONH deterioration is thought to precede visual field damage; there is evidence that about 40% of nerve fibres may be lost before impairment of visual function (Sommer 1991). The main pathological ONH changes are progressive neuroretinal rim thinning and enlargement of the cup/disc ratio, or a definite disc cupping in more severe cases (Spaeth 2006). Optic disc assessment is usually based on fundus biomicroscopy or photography. A disadvantage of direct optic disc evaluation with biomicroscopy or photography is that these methods, especially biomicroscopy, rely on the ability and experience of the physician who is performing the assessment, and therefore lead to considerable variation amongst assessments (Abrams 1994). Imaging methods provide more reliable and quantitative results. In clinical practice, imaging investigations might contribute to standardising the diagnosis of glaucoma and improvement of follow-up.

Even though ONH and RNFL imaging is already a well-established alternative to biomicroscopy or photography for the evaluation of ONH appearance, no method has yet been recognised as optimal.

Index test(s)

Clinical ONH and RNFL assessment is limited by poor reproducibility and by a wide variation in the normal anatomy of these structures between individuals (Lichter 1976). Confocal scanning laser ophthalmoscopy, commercially available as the Heidelberg Retinal Tomogram (HRT), optical coherence tomography (OCT) and scanning laser polarimetry (SLP), commercially available as GDx, are relatively new techniques for the measurement of the structural changes of the optic nerve and RNFL (Mai 2007; Medeiros 2004; Oddone 2008; Strouthidis 2008).

These devices allow measurement of RNFL thickness as well as various morphological optic disc parameters.

HRT: HRT uses a diode laser (670 nm) to scan the retinal surface at multiple consecutive parallel focal planes. The pixel with the highest reflectivity on the z-axis across the focal planes for each x, y location is used to identify the retinal surface and to construct a topographic image of the ONH. Relative topographic heights are then calculated from a reference ring placed on the retinal surface at the periphery of the scanned area.

After image acquisition, the operator using HRT needs to set an optic disc contour line manually, after which the instrument calculates ONH stereometric parameters. Besides stereometric parameters, the HRT 3 provides two different classification algorithms of the ONH morphology: the Moorfields Regression Analysis (MRA), which requires the placement of the contour line; and the more recent, contour-line independent, Glaucoma Probability Score (GPS).

GDx: The GDx is a scanning laser polarimeter that measures RNFL thickness using polarised, near-infrared (780 nm) light. The GDx measures the RNFL birefringence, which is correlated to the RNFL thickness. The cornea and lens are also birefringent structures which affect the total retardation measured, thus the GDx measures and individually compensates for the anterior segment (cornea and lens), isolating the signal from the RNFL. Individual anterior segment compensation late-generation models result in more accurate RNFL measures.

OCT: Optical coherence tomography (OCT) is a high-resolution imaging device that uses low coherent light from a broadband light source produced from a super-luminescent diode to acquire in vivo images of the retina. Optical coherence tomography applies the principle of interferometry to interpret reflectance data from a series of multiple side-by-side A-scans combined to form a cross-sectional image.

Classification algorithms are implemented in HRT, GDx and OCT, based on normative databases to discriminate between normal and diseased eyes. It has been estimated that the availability of imaging devices for the diagnosis and management of glaucoma ranges from 12.5% for the GDx to 43.9% for the HRT and 45.2% for the OCT in hospital practice in the UK (Gordon-Bennet 2008). It is likely that these figures are lower in primary care services and in low- and middle-income countries. As technology advances, different versions of glaucoma imaging devices have been released in the market in the last 10 years with improvements in terms of resolution, accuracy, reproducibility and availability of normative databases. In this review, we consider only versions equipped with normative databases, thus providing classifications, and versions with latest, mature technology (Spectral Domain OCT devices, HRT 3 and GDx VCC or ECC).

A health technology assessment (HTA) conducted in 2005 found poor performance of both HRT and GDx in cross-sectional and longitudinal groups of patients suspected of glaucomatous visual field loss (Kwartz 2005). However, the assessment was based on the results of a single clinical study and did not include a systematic review of the literature. Moreover, the GDx and HRT versions considered are no longer available. More recently, Burr 2007 assessed the HRT II, an older model not included in our review, and yielded meta-analytic estimates of sensitivity and specificity of 86% and 89% in three studies using a common cut-off.

Clinical pathway

We expect that ONH and RNFL imaging is used in people who have already been tested by means of clinical examination at primary care level, including ONH clinical assessment, IOP measurement and even visual field testing. Thus, these devices will generally be used as an add-on test. Patients may be screened for or suspected of having glaucoma for several reasons. Apart from population-based screening programmes, which are still uncommon (Heijl 2013), people may refer themselves to optometrists, orthoptists or ophthalmologists, depending on the setting, for refractive error or routine eye check. In the USA and Canada, referrals to glaucoma specialists are made both by ophthalmologists and optometrists (Cheng 2014). Those with a family history of glaucoma may know that they are particularly at risk and seek periodic consultation. An eye care professional will prescribe further tests for glaucoma in the presence of ocular hypertension (above 21 mmHg) or ONH changes at fundus examination. Visual field testing is needed to confirm manifest or perimetric glaucoma, but it has to be interpreted by an experienced professional in the context of a full eye examination. After visual field testing, an examination by an ophthalmologist is the gold standard for manifest glaucoma, whereas suspected glaucoma may require longitudinal follow-up demonstrating either changes to the visual field or ONH or both. Furthermore, there are glaucoma specialists or ophthalmologists with greater experience in glaucoma, to whom other ophthalmologists may refer difficult cases.

Prior test(s)

Ratnarajan 2013 has recently reported on suspected glaucoma referral patterns by optometrists with or without special interest in glaucoma in the UK. They concluded that a referral for suspected glaucoma is based characteristically on finding an elevated IOP, an abnormal optic disc appearance, an abnormal visual field, or a combination of these. The frequency of manifest glaucoma was about 5% to 15% when elevated IOP was the main reason for referral, and rose to 20% to 30%, the higher figure being detected by optometrists with special interest in glaucoma, when optic disc anomalies were also considered.

Role of index test(s)

How ONH and RNFL imaging could affect glaucoma referrals and diagnosis in real-world clinical settings is unclear, according to the studies we retrieved to prepare this review. Even among general ophthalmologists, the value of ONH and RNFL imaging may be enhanced by the large variability in diagnostic accuracy among clinicians, and the often moderate intra-observer agreement between clinicians in a large study of 243 ophthalmologists in 11 European countries (Reus 2010), which makes an objective and reproducible measure attractive. Reus 2010 also found that common imaging devices outperform most clinicians in classifying optic discs. An objective test providing continuous anatomical measures may therefore considerably improve clinical performance, as also found by Andersson 2011.

Alternative test(s)

A previous systematic review has examined a range of tests that can be used for the screening of glaucoma, as well as in diagnostic settings (Burr 2007; Mowatt 2008). However, our review focuses on studies of patients referred from primary care or self-referred patients, or studies of patients already followed in secondary- or tertiary-care glaucoma clinics. We considered the three tests (GDx, HRT, OCT) as equally relevant and no further test as a comparator.

Rationale

Imaging of the ONH and of the RNFL is increasingly used as an objective tool to diagnose glaucomatous disc and RNFL changes.

Each imaging device provides several continuous parameters and classification algorithms characterised by a broad spectrum of sensitivity and specificity. We therefore deemed a systematic assessment of the diagnostic accuracy of new imaging methods for the diagnosis of glaucoma to be useful.

OBJECTIVES

- To determine the diagnostic accuracy of HRT, OCT and GDx for diagnosing manifest glaucoma by detecting ONH and RNFL damage.

Secondary objectives

- To determine which morphometric measure or diagnostic algorithm yields the highest diagnostic accuracy within each device.
- To compare the relative diagnostic accuracy of the three devices.
- To explore potential causes of heterogeneity of diagnostic performance across studies.

We planned to investigate the following sources of clinical heterogeneity:

A. Heterogeneity related to the choice of reference standard: type of reference standard (optic disc assessment, visual field, or both); definitions of visual field damage.

B. Heterogeneity related to characteristics of the study population: severity of glaucoma.

C. Heterogeneity related to issues of methodological quality.

As we expected a large number of included studies to be case-control, we considered a particular type of bias resembling incorporation bias for these studies. Usually the investigator assessing the presence of glaucoma does not rely exclusively on valid perimetric criteria to allocate patients to the glaucoma group, but also on optic disc appearance such as cupping. Diseased patients may have larger cups than expected, thus enhancing the ability of imaging methods to detect disease based on disc morphology algorithms. For this reason, we investigated heterogeneity between case-control studies using visual field only versus case-control studies using visual field plus optic disc as a reference standard. We considered visual field alone the preferred, unbiased reference standard method (Garway-Heath 1998).

We originally planned to investigate heterogeneity based on specific methodological issues of included studies (Appendix 1): inclusion of a representative spectrum of patients; reporting of uninterpretable results; choice of unit of analysis. However, we then adopted QUADAS 2 and used its domains for heterogeneity investigation.

Finally, we planned an exploratory subgroup analysis based on the overall level of missing data, regardless of their cause (including withdrawals and any patients who may have been excluded because of uninterpretable index test results), using the median level of missing data across studies to define better versus worse quality, as well as a level of 10% missing data for the same purpose. We planned further subgroup analyses to investigate the contribution of studies that did not report any missing data but did not explicitly state that there were no missing data.

METHODS

Criteria for considering studies for this review

Types of studies

We planned to include all prospective and retrospective cohort studies and case-control studies that evaluate the accuracy of OCT, HRT or the GDx for diagnosing glaucoma. We included both single studies assessing each imaging method and comparative studies assessing more than one imaging method in the same patient population. We included only studies that provide data to allow calculation of sensitivity and specificity estimates.

A first draft of this review was submitted based on a literature search conducted until 15 June 2013, which identified a large number of case-control studies. During the revision of the final version of this review, we updated the search to 15 February 2015 and found some additional case-control studies. We decided not to include these additional case-control studies, as they are known to be prone to methodological biases and unlikely to change the current evidence base. Future updates of this review should only focus on studies where the patient population is enrolled consecutively, with the same set of inclusion criteria, such as referable patients identified in primary care.

We applied no language restriction to the inclusion criteria of the studies.

Participants

The tests on which this review focuses have not been extensively studied in population-based screening studies, which should be the subject of a future Cochrane review on screening tests for glaucoma. The published protocol for this review stated we would include glaucoma suspects, but did not fully specify the professional and clinical pathway stage at which such a question is made. Framing the question in a well-defined pathay is also difficult due to variation of eye care pattens in different health care settings. In retrospect, the findings of this review could be used in an add-on setting which could be a primary care, or a triage setting when somebody has already been referred from primary care to secondary care as suspect glaucoma and needs triage by a non glaucoma specialist

Index tests

We assessed the following imaging devices: confocal scanning laser ophthalmoscopy (HRT); optical coherence tomography (OCT); and scanning laser polarimetry (GDx). For each test we extracted and analysed all parameters which can be obtained with standard commercial software and are measuring RNFL or ONH morphology.

During the review process, we decided to extract OCT measures that are not related to RNFL and ONH morphology, but to macular cell layers affected by glaucoma, such as ganglion cell complex (GCC) and ganglion cell inner plexiform layer (GCIPL), as these parameters have gained popularity in recent years.

Target conditions

The target condition of interest was manifest glaucoma.

Reference standards

There is no universally-accepted reference standard for the diagnosis of manifest glaucoma. Both optic disc and visual field damage are used to diagnose the presence of glaucoma. Several systems have been proposed to score visual field and optic disc damage and have been tested in multicentre randomised controlled trials (RCTs) (AGIS 1994; CIGTS 1999; EMGT 1999). While we accepted any diagnosis of glaucoma given by the study investigators, we conducted subgroup analyses to assess whether differences between studies could be explained by the choice of the reference standard.

Search methods for identification of studies

Electronic searches

We searched the Database of Abstracts of Reviews of Effects (DARE), the Health Technology Assessment Database (HTA) and the NHS Economic Evaluation Database (NHSEED) (Cochrane Library 2015, Issue 1), Ovid MEDLINE, Ovid MEDLINE In-Process and Other Non-Indexed Citations, Ovid MEDLINE Daily, Ovid OLDMEDLINE (January 1946 to February 2015), EMBASE (January 1950 to February 2015), MEDION (www.mediondatabase.nl/) (2002 to 2012, database

archived in 2012) and the Aggressive Research Intelligence Facility database (ARIF) (147.188.28.230/rmwp) We did not use any date or language restrictions in the electronic searches for studies. We last searched the electronic databases on 19 February 2015.

See: Appendices for details of search strategies for the Cochrane Library (Appendix 2), MEDLINE (Appendix 3), EMBASE (Appendix 3), MEDION (Appendix 4) and ARIF (Appendix 5).

Searching other resources

We handsearched the reference lists of the included studies for further relevant studies.

Data collection and analysis

Selection of studies

Pairs of review authors (MM, EL, GV, SF) independently examined the titles and abstracts of all citations identified by the electronic searches. We classified the abstracts as (a) definitely included, (b) unsure or (c) definitely excluded. We obtained and re-assessed full-text copies of those classified as (a) definitely included and (b) unsure. We subsequently classified the studies as (1) included, (2) awaiting assessment or (3) excluded. Because of the huge volume of identified evidence, we did not contact the authors of studies classified as awaiting assessment for further clarification, but we planned to re-assess the studies if further information should become available. Due to the large number of retrieved and assessed full-text papers, we chose not to list all studies classified by the two review authors as (3) excluded in the 'Characteristics of excluded studies ' table. We are happy to provide a list of these studies upon request. We assessed all studies identified as (1) included for methodological quality and data extraction. The review authors were not masked to the names of study authors and institutions. We resolved any disagreement between the two review authors by discussion or by referral to a third review author (GV).

Data extraction and management

Pairs of review authors (SF, EL, MM, SN) independently extracted the following information from each included study: the number of true positives (TP), i.e. patients categorised as diseased by both the reference and index test; the number of false negatives (FN), i.e. patients categorised as diseased by the reference test, but as non-diseased by the index test; the number of true negatives (TN), i.e. patients categorised as non-diseased by both the reference and index tests; the number of false positives (FP), i.e. patients categorised as non-diseased by the reference test, but as diseased by the index test; the number of patients with uninterpretable index test results; the number of patients for whom the assessment of both eyes was included in the statistical analyses; the number with missing data (patients who were not included in the analyses).

We summarised the Characteristics of included studies using the items shown in Appendix 6.

Assessment of methodological quality

Pairs of review authors (SF, EL, MM, SN) independently assessed the methodological quality of included studies using the QUADAS 2 checklist (Appendix 7), which has recently replaced the original QUADAS checklist (Whiting 2003) (Appendix 1). We also followed the recommendations provided in Chapter 9 of the *Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy* (Reitsma 2009). We resolved any disagreement by discussion or by referral to a third author (GV).

Statistical analysis and data synthesis

For each imaging test we extracted indices of diagnostic performance or derived them from the data reported in each primary study. Where possible we recorded the number of true positive cases, false positive cases, false negative cases and true negative cases by 2 x 2 contingency tables, where the columns reveal the true status (diseased or not diseased) of the condition under investigation and the rows show the dichotomised index test results. From the 2 x 2 tables we calculated: sensitivity (the proportion of diseased people correctly diagnosed) and specificity (the proportion of non-diseased people incorrectly diagnosed) with 95% confidence intervals. Initially, we explored heterogeneity by visual inspection of the forest plots of pairs of sensitivity and specificity, and of plotted data on a receiver operating characteristic (ROC) plot (sensitivity on the vertical axis and (1 - specificity) on the horizontal axis).

We had planned to conduct meta-analyses of correlated pairs of sensitivity and specificity using the hierarchical summary ROC (HSROC) model (Rutter 1995; Rutter 2001). However, when we had completed the data extraction, we noticed that studies compared several measures of each device and presented data at fixed levels of specificity (such as 0.80, 0.90 and 0.95), without reporting any cut-off used, sometimes presenting sensitivity at more than one specificity level. We extracted all data and presented them in forest plots regardless of the specificity level chosen by the study authors. Thereafter, we decided to use 2 x 2 data at 0.95 specificity or closer in meta-analyses, since this was the most commonly reported level and because ONH and RNFL imaging tests might have a role as a triage test when the target condition is manifest glaucoma, especially in primary care settings, which is then confirmed by an ophthalmologist by means of clinical and visual field examination.

Because of the data structure, we expected and found little variation in specificity. Thus, we deviated from the protocol and fitted a bivariate model using the *METADAS* macro in SAS (Takwoingi 2008), focusing on summary sensitivity when reporting data, despite the fact that thresholds were not reported. Harbord 2007 has shown that the bivariate (Reitsma 2005) and the HSROC models are mathematically equivalent and, as a result, *METADAS* simultaneously derives pooled sensitivity and specificity.

Because of the large number of test parameters, we faced the issue of conducting a huge number of comparisons and decided to limit multiple testing by adopting the following strategy: first, we considered that direct comparisons are more reliable than indirect comparisons in diagnostic accuracy studies (Takwoingi 2013). Nearly all studies included a single device, but compared several parameters within the same imaging device, making within-test comparisons more robust than between-test comparisons. We used a covariate coding for each test parameter in the bivariate model and, given limited variation of specificity, we reported the significance of testing for the sensitivity of each parameter versus that with the highest sensitivity. We conducted such comparisons including two parameters at a time, to avoid problems with missing data for other parameters. In order to conduct indirect comparisons between tests, but still reducing the amount of significance testing, we included in the analysis the parameters with the two highest

levels of sensitivity within each test and again compared them to that with the best sensitivity among all.

Investigations of heterogeneity

We had planned to use forest plots to look for evidence of heterogeneity within sensitivity and within specificity, and ROC plots to look for evidence of a threshold effect and heterogeneity due to differences in accuracy.

Although we planned to incorporate covariates in the hierarchical model to examine the effect of potential sources of heterogeneity on threshold parameters, accuracy parameters or both, we adopted a bivariate regression model and focused on reporting sensitivity, as there was minimal variation in specificity as explained above.

Sensitivity analyses

We planned to undertake sensitivity analyses for individual quality items, in particular for 'Type of study design' by omitting case-control studies. However, as nearly all the included studies were case-control we did not perform this analysis.

Assessment of reporting bias

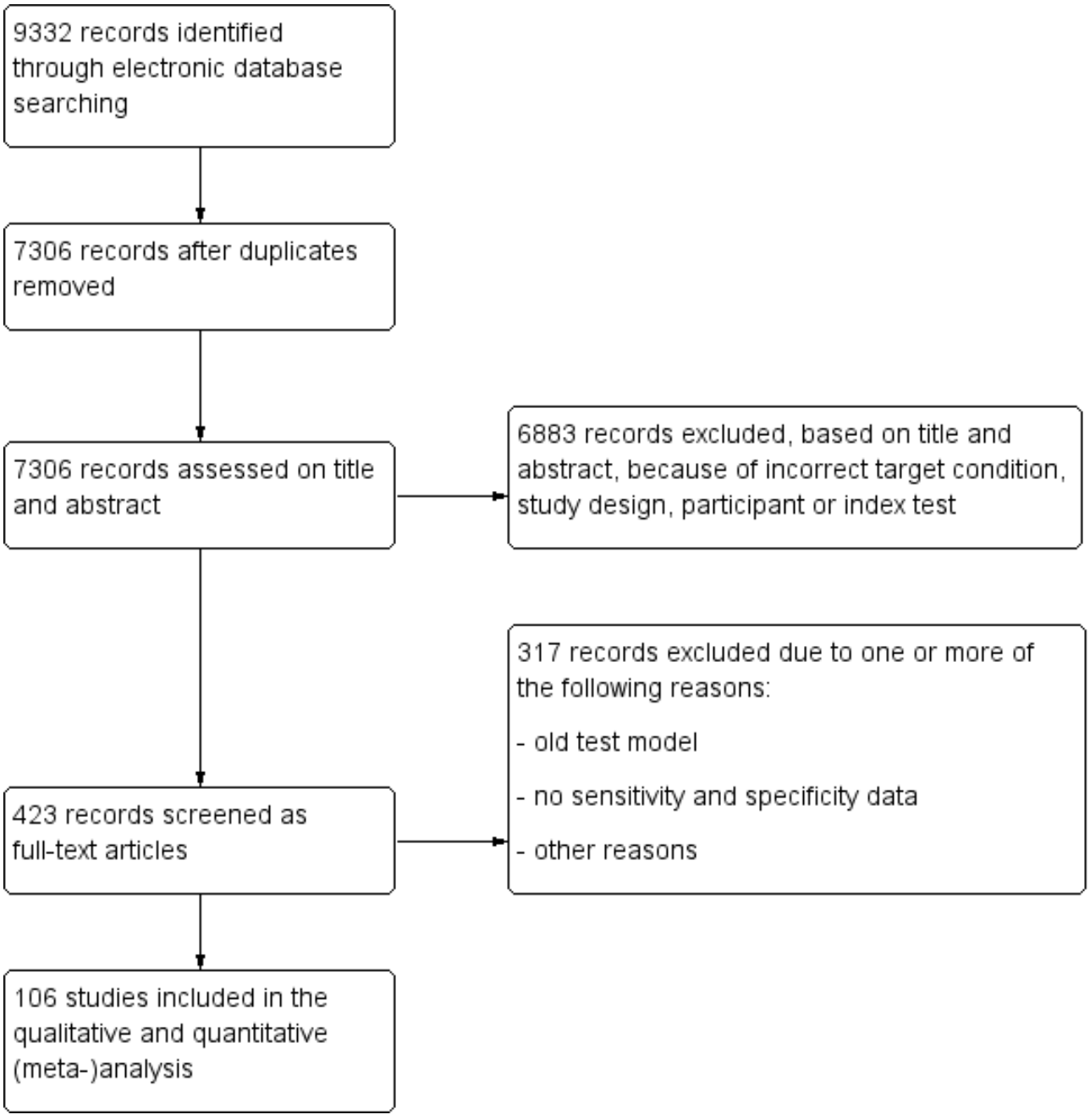
We had planned to assess publication bias using funnel plots displaying lnDOR on the x-axis and $1/ESS^{½}$ (where ESS is the effective sample size) on the y-axis, as recommended by Deeks 2005, provided that 10 or more studies are included in the analyses. We decided not to conduct these analyses in the review phase.

RESULTS

Results of the search

We updated the searches used for this review in February 2015.The electronic searches yielded a total of 9332 records (Figure 1). After deduplication we screened 7306 reports, of which we considered 6883 records not to be relevant, based on title and abstract, because of incorrect target condition, index test, participants, or study design. In total we screened 423 full-text reports of studies, of which we excluded 317 for one or more of the following reasons, mainly because they evaluated an old test version or did not provide suitable data (references available upon request). Finally, we identified 106 relevant studies with a total of 16,260 eyes. One hundred-and-three studies were case-control studies, one study was a consecutive cohort study and the study design was unclear for the remaining two studies. The sample size ranged from 61 to 435 patients (median 143). Most studies were conducted in Asia (44), followed by Europe (31), North America (24), South America (2) and Oceania (1). Four studies did not report sufficient information to determine study setting. Almost all studies enrolled one eye per person (90 studies, 85%).

Figure 1. Flow diagram.

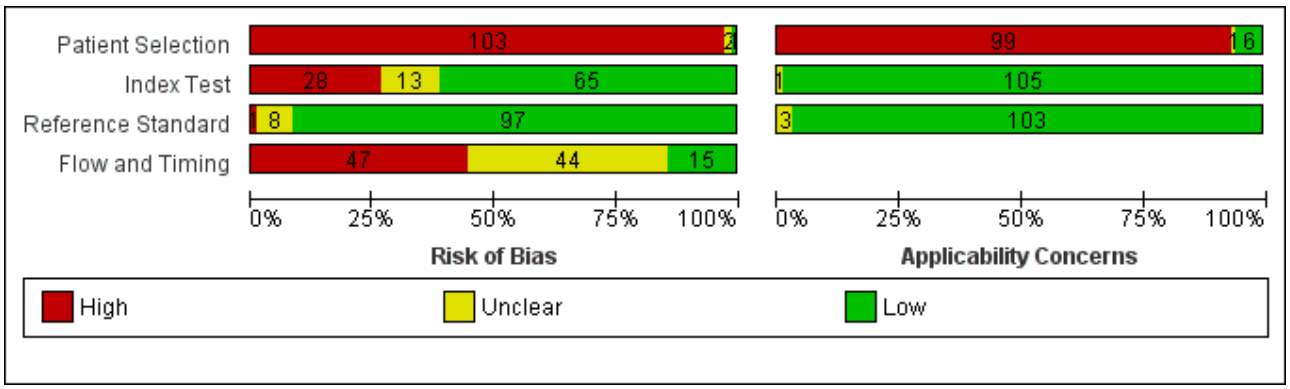


Forty studies (5574 patients) assessed GDx, 18 studies (3550 patients) HRT, and 63 (9390 patients) OCT. Twelve of these studies compared two or three tests. Sixty-seven studies used VF damage plus ONH glaucomatous optic neuropathy as the reference standard; the remaining 37 studies relied on either VF damage only (29 studies) or ONH/RNFL damage only (10 studies) as definition criteria for confirming glaucoma. There was limited opportunity to explore the variability of controls regarding risk factors for glaucoma, as well as to investigate subgroups of severity of glaucoma based on studies' inclusion criteria of cases. We therefore used the study average mean deviation (MD) for this purpose, with values ranging from -0.16 db to -11.4 db.

Methodological quality of included studies

We present a summary of methodological quality assessment in Figure 2. The main quality issue was the case-control design (103 studies) or unclear design (two studies) of all included studies except one. This led to a high risk of bias for the Patient Selection domain in QUADAS 2, and raised concerns about the applicability of our findings to clinical practice, particularly when the purpose is to triage patients to be referred to glaucoma centres.

Figure 2. Risk of bias and applicability concerns graph: review authors' judgements about each domain presented as percentages across included studies



There were some concerns about the conduct of the index test. In fact, we assumed that the use of fixed specificities equalled threshold prespecification in all but one study (Chen 2008).

Quality of images, which we chose as an additional signalling question because it is known to affect the accuracy of RNFL thickness (Rao 2013), was assessed and used in 99 out of 106 studies. Conflict of interest was of high concern in 24 studies, of unclear concern in 15 studies, and of no concern in 67 studies.

Reference standard was rated as good when visual field only was used to detect the presence of glaucoma (27 studies). As reported below, confirmation of glaucoma using visual field testing means that the patient's function is affected, which is more relevant, and also explores a different dimension compared to that assessed by ONH/RNFL imaging tests. Masking of reference test to index test results was unclear (75 studies) or not adopted (one study), with only 30 studies reporting its masked interpretation with respect to index test results.

With regard to the Flow and Timing domain, 101 out of 106 studies used the same reference standard for all patients and 59 studies excluded fewer than 10% of the patients from the analyses; we judged the remaining studies to be at unclear or high risk of bias. However, exclusions were often due to poor-quality images, which we considered a good quality criterion for the assessment of the Index test domain. For this reason, we decided not to carry out sensitivity analyses on this issue, as its interpretation would have been difficult. Finally, adopting a strict criterion of less than one month between index and reference tests, we classified 28 studies at high risk of bias, and most of the remaining at unclear risk of bias.

Findings

One hundred-and-six studies reported sensitivity values of several parameters at given specificity values, mainly at approximately 0.80, 0.90 and 0.95. Our revised analysis plan was to present the accuracy of all reported parameters for each test (Table 1), and then compare parameters to that with the best diagnostic odds ratio (DOR) (Table 2). Because ONH parameters obtained with OCT were reported in a substantially smaller set of studies compared with RNFL parameters, we present them separately to maintain the validity of within-test comparisons.

Finally, macular/GCC and GCIPL parameters have increasingly been investigated as OCT-based parameters for detecting glaucoma, but were not among the structural dimensions we originally planned to investigate in this review (i.e. ONH and RNFL). Nonetheless, 32 studies assessed these new measures, and we report on them separately without carrying out any statistical testing on the differences versus other parameters (Table 3).

Accuracy of test parameters and within-test comparisons

Table 1 presents the accuracy of all parameters of each test. Sensitivities were very heterogenous, as seen in forest plots, while specificities were above 0.80 by design. Statistical modelling of relative DOR within each instrument is shown in Table 2, where sensitivity and specificity may slightly differ from Table 1 due the introduction of covariates and the assumption of parallel HSROC curves in the model to assist interpretability.

GDx

Forty studies (5574 participants) investigated GDx, with each parameter assessed in 30 to 35 studies, indicating that most of them carried out direct comparisons (Table 1). Point estimates of summary sensitivity varied between 0.61 (for superior and inferior RNFL thickness; temporal superior nasal inferior temporal (TSNIT) average) and 0.76 nerve fibre indicator (NFI). There was minimal variation in specificity (0.92 to 0.93) across these parameters, as expected, due to the design of the included studies and our data extraction strategy.

The DOR of the NFI was significantly better than that of other parameters (Table 2).

HRT

Eighteen studies (3550 participants) investigated HRT (Table 1). Eight studies obtained MRA, but only two of these reported other measures. Comparing MRA to other HRT parameters was therefore based mostly on indirect comparisons. The MRA had the highest sensitivity (0.69), with the Vertical C/D ratio as the second best (0.67). However, the specificity was better for the Vertical C/D ratio (0.94 versus 0.89), suggesting threshold effects. For other parameters, sensitivity varied between 0.32 (Cup volume) and 0.58 (Frederick S. Mikelberg (FSM) discriminant function) and specificity was 0.94 to 0.95 for all parameters.

When we compared overall accuracy using DOR, we found no significant differences between the Vertical C/D ratio and the best four parameters, including MRA (Table 2).

OCT

Sixty-three studies (9390 participants) assessed OCT (Table 1). Of these, 57 assessed mean RNFL thickness, 45 and 43 assessed the inferior and superior sectors respectively, which are believed to be clinically more informative than temporal and nasal sectors (assessed in 30 studies each). Point estimates of sensitivity varied between 0.29 (nasal) and 0.72 (inferior) with modest variation in specificity (0.93 to 0.94).

The DOR of the average RNFL thickness was not significantly better than the inferior sector, whereas it was better than the superior, nasal and temporal parameters (Table 2).

Other ONH parameters were evaluated in four to 17 studies, yielding sensitivities between 0.16 (Disc area) and 0.72 (Vertical C/D area ratio) and specificities between 0.92 and 0.95. The Vertical C/D ratio was no better than the C/D Area Ratio, but was superior to all other parameters (Table 2).

Alternative data extraction at the lowest reported specificity

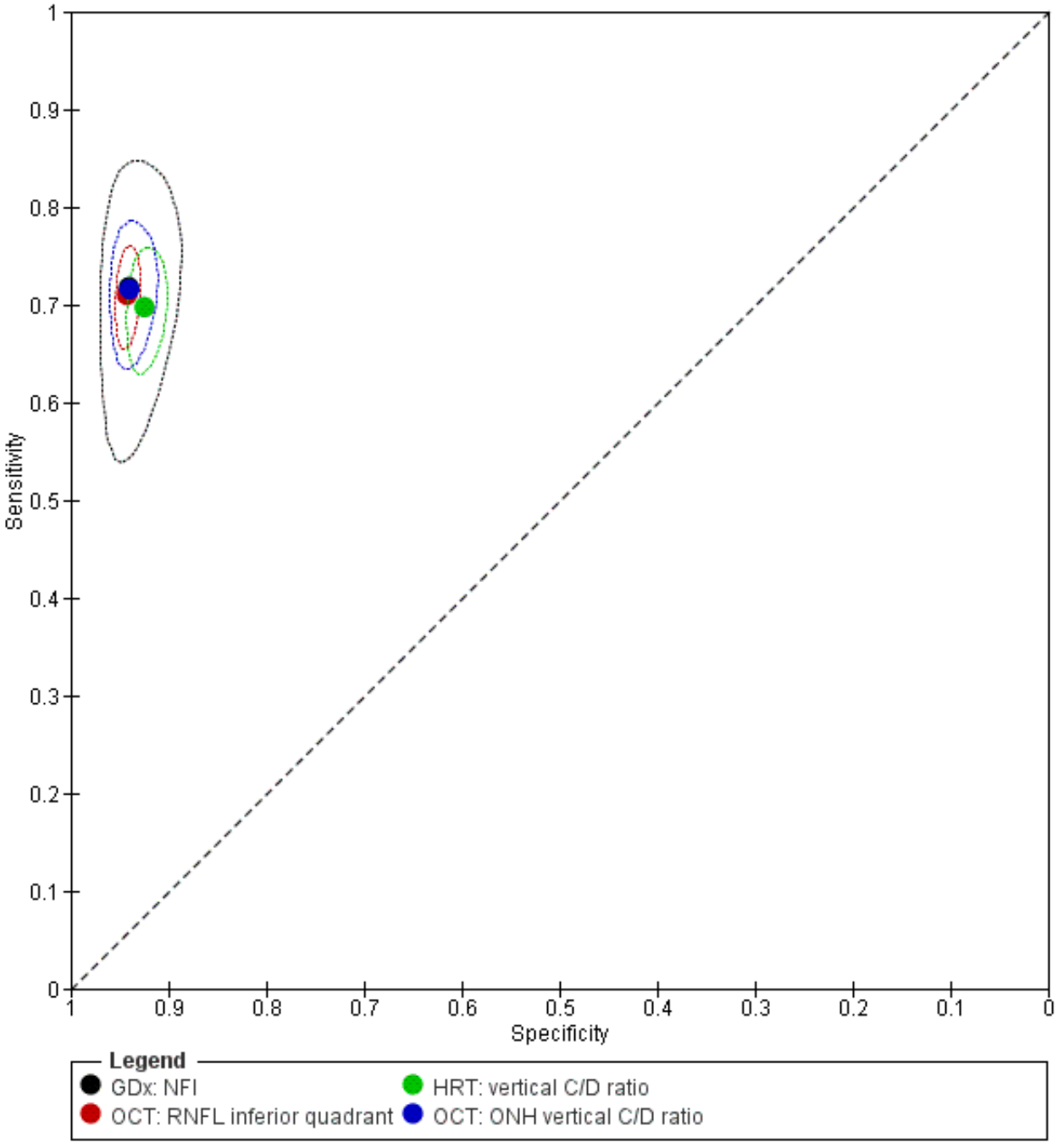
Table 4 presents diagnostic accuracy obtained by extraction data at the lowest rather than the highest reported specificity. The pooled specificity of the best-performing parameters of GDx and OCT decreased to 0.86 to 0.87, and sensitivity increased to about 0.80.

Comparisons of parameters between tests

Overall comparisons

We focused on the parameter with the highest DOR for GDx, HRT, and separately for RNFL and ONH measures of OCT, in single parameter analyses as estimated in Table 1. These were compared including a covariate in the HSROC model: pooled estimates of sensitivity/specificity and DOR were almost identical, (Figure 3; Table 5).

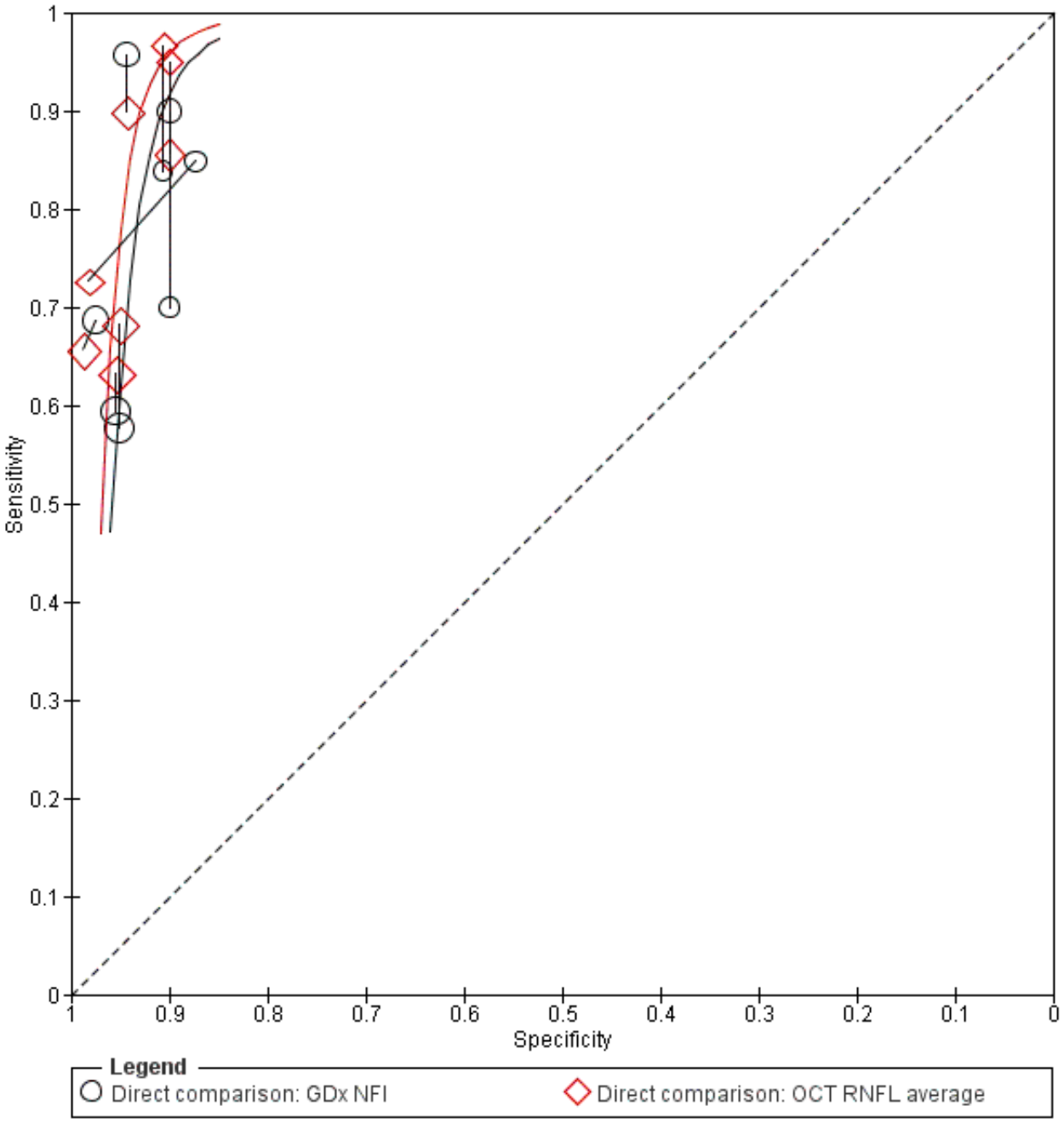
Figure 3. Summary ROC Plot of tests with data extracted at the highest specificity in case of multiple study measures for the same parameter: 2 GDx: NFI, 4 GDx: TSNIT average, 5 OCT: mean RNFL thickness, 6 OCT: RNFL at inferior quadrant, 13 HRT: vertical C-D ratio, 17 HRT: MRA, 39 OCT: ONH C/D area ratio, 41 OCT: ONH C/D vertical ratio.



Direct comparisons

We compared the best parameter for each test by restricting the analysis to direct comparisons. However, direct comparisons of the best-performing parameters were sufficient for meta-analysis only for GDx NFI versus OCT RNFL average (eight studies, Figure 4). The DOR of OCT RNFL average (75.92; 95% CI 44.25 to 130.28) was non-significantly superior to that of GDx NFI (relative DOR: 0.68; 95% CI 0.38 to 1.21; P = 0.190).

Figure 4. Summary ROC Plot of tests: 47 Direct comparison: GDx NFI, 48 Direct comparison: OCT RNFL average.



Accuracy of GCC/GCIPL OCT parameters

Table 3 shows the summary sensitivity and specificity for all GCC/GCIPL parameters with any of three different OCT tests in up to 35 studies for each parameter. Sensitivities and specificities were in the range of those observed for ONH and RNFL parameters. However, we did not compare these parameters formally, since this was not an aim of our review.

Heterogeneity investigation and effect of methodological quality

We restricted these analyses to the best parameter identified in indirect comparisons (NFI for GDx, vertical C/D ratio for HRT, and mean RNFL thickness for OCT) using all available studies, given the similar accuracy of performance. We present the results of these analyses in Table 6.

The main finding was the lower sensitivity estimated for detecting milder glaucoma cases (MD better than -6 Db, 65 studies, 9720

patients: 0.64; 95% CI 0.60 to 0.69), as compared to more severe glaucoma cases (MD -6 Db or worse: 49 studies, 7,598 patients: 0.79; 95% CI 0.74 to 0.83) at about the same specificity (0.93, 95%CI 0.92 to 0.94 and, respectively, 0.94; 95% CI 0.93 to 0.95; P <0.0001 for the difference in relative DOR).

We found no significant difference in sensitivity when adopting a functional reference standard, such as the visual field, as compared to a combination of anatomic and functional reference standards.

All studies were at high risk of bias for the Patient Selection domain, which could not be used as a covariate. We found no difference in accuracy for the domains Index Test, Reference Test or Flow and Timing, as seen in Table 6.

Interpretation of findings

Because the performance of GDx, HRT and OCT was remarkably similar comparing the parameters with the highest DOR in a single HSROC model, we applied our accuracy estimates to the following scenarios (Summary of findings 1). Based on Ratnarajan 2013, who recently investigated glaucoma referral patterns by optometrists with or without special interest in glaucoma in UK, referrals by optometrists with no special interest in glaucoma are diagnosed manifest glaucoma in 3.5% when elevated IOP is the reason for referral, up to about 20% when anomalies of disc and IOP or disc and visual field are reasons for referral. The corresponding figures for optometrists with an interest in glaucoma are about 15% and 30%. Though people finally diagnosed with suspect glaucoma would be more than twice as many as those with manifest glaucoma among primary care referrals, investigating the accuracy of imaging devices for diagnosis of suspect glaucoma is outside the scope of our review. Therefore, we present two referral scenarios, one with a low prevalence of manifest glaucoma (5%) and another with a high prevalence (20%), In both scenarios we also assume a sensitivity of 0.70 and a high specificity close to 0.95 as in most of these studies.

If 50 out of 1000 referrals have manifest glaucoma, for example for people who are found elevated IOP or a family history of glaucoma in a non-specialised primary care setting, these tests would correctly identify about 35 glaucomatous patients and miss 15 out of the 50 patients, while avoiding referral of about 890 out of 950 non-glaucomatous people.

Assuming 200 of 1000 referrals are finally found manifest glaucoma, e.g. on the basis of prior testing such as combined disc and visual field assessment in specialised primary care, these tests would correctly identify about 140 glaucomatous patients and miss 60 out of the 200, while avoiding referral of about 750 out of 800 non-glaucomatous patients.

DISCUSSION

Summary of main results

This review evaluates the accuracy of GDx, HRT and OCT used for imaging the ONH and RNFL for the diagnosis of manifest glaucoma. Considering the use of these devices as stand-alone tests to inform decision making, the findings of this review could be used in an add-on setting which could be a primary care, or a triage setting when somebody has already been referred from primary care to secondary care as suspect glaucoma and needs triage by a non glaucoma specialist.

All 106 included studies used several types of parameters for a single test, with the large majority reporting sensitivities at approximate fixed and high specificity levels, mostly at 0.95. Hence, comparisons between different types of parameters within each test were based largely on direct comparisons. We found that NFI was the most accurate parameter for GDx, whereas for OCT the sensitivity of mean RNFL thickness was not significantly different from that of the inferior sector, but was better than the other sectors. With regard to HRT, we did not observe differences among vertical C/D ratio, C/D area ratio, MRA and FSM or Reinhard O.W. Burk (RB) discriminant functions, but the vertical C/D ratio was superior to all other cup and rim morphological parameters.

The heterogeneity of sensitivity estimates between studies, assessed in forest plots, was large for most devices and parameters at all specificity levels, potentially making indirect comparisons between tests unreliable (Takwoingi 2013). Nonetheless, the performance of the best parameter of each test was remarkably similar.

The main limitation of this assessment, despite the large number of studies on the use of GDx, HRT and OCT for detecting manifest glaucoma, was the case-control design of nearly all included studies. Case-control studies are likely to overestimate diagnostic accuracy due to the sharp separation of the measurements between cases and controls, unless a nested design is used. Furthermore, the applicability of the findings to patients referred to glaucoma specialists by primary eye-care professionals may be limited.

Strengths and weaknesses of the review

The strength of this review is in the systematic assessment of a considerable number of studies, including double data extraction and quality assessment according to recommended standards (QUADAS 2).

A weakness of this review is that we did not provide an explicit description of the potential clinical pathways in the original protocol. However, for the management of glaucoma, the mapping of clinical pathways is a complex and difficult process and is likely to be setting-specific at least at a country/local level. Consequently, the unclear applicability of our findings can also be the result of the differences in the care pathway of patients with glaucoma among different countries, unless such pathways are actively monitored (Ratnarajan 2013). Overall, we find the methodology for such reviews has evolved during the process, particularly the importance of specifying the clinical context in which the review is set.

Comparison with other reviews

We found other relevant reviews of diagnostic accuracy studies.

Recent narrative reviews have supported the use of ONH and RNFL imaging for detecting glaucoma. Two reviews (Bussel 2013; Sung 2011) focused on the role of spectral-domain OCT for the diagnosis and management of glaucoma. They observed that RNFL measurement is the most accurate parameter for the detection of glaucoma, but ONH and segmented macular analyses have shown in many studies a diagnostic capability overlapping and comparable to that of RNFL peripapillary analysis. Bussel 2013 also highlighted a number of limitations of the available evidence, which influence applicability of findings, and concluded that

OCT is a valuable tool for glaucoma diagnosis and detection of progression, but that it lacks the necessary diagnostic performance for general population glaucoma screening. These reviews did not include a systematic search of evidence, nor did they carry out a meta-analysis.

Burr 2007 and Mowatt 2008 published different version of a systematic review of tests for screening and diagnosing glaucoma. Burr 2007 also assessed the cost effectiveness of screening programmes and considered three test categories:

- tests for intraocular pressure measurement: contact and non-contact tonometry;
- tests for structural optic nerve damage: optic disc assessment by means of ophthalmoscopy or photography, RNFL photography, and tests for quantitative analysis of the optic nerve head and RNFL also included by us, such as HRT, GDx and OCT;
- visual function tests: frequency doubling technology (FDT), motion detection technology, oculokinetic perimetry, short-wavelength automated perimetry, standard automated perimetry.

Among imaging tests, HRT II, an older model not included in our review, yielded meta-analytic estimates of sensitivity and specificity of 0.86 and 0.89 in three studies using a common cut-off. It is difficult to compare these results with those of our review, because we included different test models and far more studies.

In a systematic comparative effectiveness review searching for studies up to June 2011, Ervin 2012 investigated the diagnostic performance of a similar set of optic nerve structure and function tests for screening of glaucoma, including 17 studies on HRT II, 11 studies on HRT III, 47 studies on different OCT models and 27 studies on different GDx devices. They found sensitivity estimates of 0.68 and 0.72 at a fixed specificity of 92%, for the best HRT III parameters GPS and MRA, respectively. Sensitivity and specificity estimates for OCT average RNFL thickness ranged from 0.24 to 0.96 and from 0.66 to 1.00, respectively. For the NFI of GDx-VCC, sensitivity estimates ranged from 0.28 to 0.99 at specificity levels between 0.53 and 0.95. The authors concluded that "the ability of these devices to identify glaucoma in a screening setting is not well understood [...] due to the lack of a single diagnostic standard for glaucoma and the high degree of variability in the design and conduct of largely cross-sectional studies of diagnostic accuracy".

Ervin 2012 also included studies assessing older imaging test models and studies conducted in population-based or screening settings. For population-based studies, Ervin 2012 retrieved two HRT II studies, and no OCT or GDx studies up to June 2011. In addition, we found two population-based studies using more recent imaging tests. However, the estimates in these studies were imprecise, since Kamdeu 2011 identified four cases of manifest glaucoma in 197 screened patients, and Bengtsson 2012 identified five cases in 170 screened patients.

Bussel 2013 conducted a narrative review of spectral-domain OCT studies and reported seven selected studies on glaucoma detection, and six studies on glaucoma progression. They concluded that RNFL remains the dominant parameter for glaucoma diagnosis and detection of progression, but that OCT still currently lacks the diagnostic performance for glaucoma screening. Burr 2014 published a modelling study that found that a randomised glaucoma screening trial would not be cost-effective in the UK scenario, but they used conventional tests such as tonometry, visual field, and photography, and not OCT. Meier 2014 remarked that to date the US Food and Drug Administration has not cleared or approved an OCT device for glaucoma diagnosis and screening.

We did not include screening studies in our review. Interestingly, Li 2013 reported on the use of GDx-VCC in a community-based study on volunteer participants with risk factors for glaucoma. They found that the best-performing parameter was the GDx NFI using a cut-off of 35 with a sensitivity of 75% (95% CI 19.4 to 99.4) at a specificity of 95% (95% CI 91.3 to 97.3), and concluded that the GDx-VCC has inadequate sensitivity for screening of definitive glaucoma. Springelkamp 2014 published the results of the population-based Rotterdam study, which detected 41 glaucoma cases with no known glaucoma risk factor and 1081 controls after excluding 96 patients with risk factors. Mean RGCL thickness in the inferior half of the macular region showed the highest sensitivity (53.7%; 95% CI 38.7 to 68.0%) at 97.5% specificity. The mean thickness of the peripapillary RNFL had a sensitivity of 24.4% (95% CI 13.7 to 39.5%).

Our review focused on RNFL and ONH parameters, but there has been an increasing interest in GCC/GCIPL parameters using OCT,in recent years. We did not formally compare such parameters to RNFL and ONH parameters, but overall found similar ranges of sensitivity when they were reported. Lee 2014 observed that GCC may be less sensitive than RNFL parameters to optic disc torsion.

Finally, newer OCT with better tissue penetration, such as the swept-source OCT, are being used to select new imaging parameters by detecting the posterior border of the sclera and lamina cribrosa, which we have not included in our review.

Applicability of findings to the review question

When we planned this review, we were aware of potential variability in care pathways across settings and healthcare systems. We intended to support decisions about patients referred by optometrists and primary eye care professionals (Cheng 2014; Ratnarajan 2013). Studies considered in this review should have included consecutive participants at risk of glaucoma identified by primary eye care professionals, using these devices in an add-on setting, which could be optometrists in primary care, or a triage setting when somebody has already been referred from primary care to secondary care as suspect glaucoma and needs triage by a non glaucoma specialist. However, we ended in including almost only case-control studies including healthy participants and glaucoma patients identified a priori, which not only overestimate accuracy, but also makes it difficult to translate study results to a specific setting.

How ONH and RNFL imaging could affect glaucoma referrals and diagnosis in different real-world clinical settings is still unclear. Even among general ophthalmologists, the value of ONH and RNFL imaging may be enhanced by the large variability in diagnostic accuracy among clinicians. In fact, a large study including 243 ophthalmologists in 11 European countries (Reus 2010) found only moderate intra-observer agreement between clinicians, which makes the use of imaging tests attractive, since they provide an objective and reproducible anatomic measure.

Another applicability issue of the included studies relates to their estimate of sensitivity at fixed specificity (e.g. 95%). Although

this makes the comparison of several measures easier, the lack of a definite measurement cut-off makes inference more difficult for users. Morevoer, overall accuracy at high sensitivity, rather than high specificity, was not available in studies. However, since the standard of care is referral of all patients with glaucoma risk factors in primary eye care, achieving a high sensitivity to avoid missing patients with glaucoma may be a better strategy, provided that the burden of referrals is reduced. As an example, OCT has been used to limit referrals in a UK screening programme of people with diabetes who were screen-positive for diabetic maculopathy on fundus photographs, ruling out diabetic macular oedema when OCT macular retinal thickness is normal (Olson 2013). Although assessing accuracy is a useful step of diagnostic test investigation, mapping patient flow during the whole clinical pathway is necessary to implement screening programmes in public health.

AUTHORS' CONCLUSIONS

Implications for practice

Despite the large number of studies exploring the use of imaging tests for detecting manifest glaucoma, their accuracy has been studied only partially. The accuracy of these tests varied across studies and was suboptimal in many, despite the fact that it may have been overestimated due to the case-control design. As a consequence of these limitations, the studies included in this review should be considered exploratory, and our results would only indirectly inform clinical decisions on referrals in primary eye care settings.

The findings of this review indicate that the best parameters for diagnosing glaucoma in a triage setting are NFI for GDx, average or inferior sector RNFL thickness for OCT, and the vertical C/D ratio or some others for HRT. Although the studies had various methodological shortcomings, we consider these findings useful and reliable because they are mostly based on direct comparisons.

On the other hand, comparisons among tests were hampered by the presence of heterogeneity and the lack of direct comparisons. Overall, the accuracy of the best parameters of GDx, HRT and OCT was remarkably similar.

The implications of using our estimates for clinical decision making is highly dependent on the care pathway and the diagnostic alternatives available, which goes beyond the scope of this review.

Implications for research

Further case-control studies are not useful in this research field. Given the limitations we found, we suggest the following improvements for studies assessing the accuracy of imaging devices for the diagnosis of manifest glaucoma, which should:

- include consecutive patients based on a single set of inclusion/exclusion criteria;
- be conducted in a specific clinical setting;
- clearly specify the clinical decision problem (in order to render the care pathway explicit);
- report relevant information both on patients' prior clinical assessments and on reasons for referral;
- present sensitivity/specificity estimates and counts in 2 x 2 tables at relevant cut-off values of each test parameter which is obtained as a continuous measure;
- discuss the potential consequences for false positives (over-referrals) and false negatives (under-referrals), adopting the new test as compared to existing practice.

Combination of imaging test results with clinical information, such as IOP, age, family history, etc., should also be considered in future research. We need reviews of studies on the ability of longitudinal ONH changes, detected by means of imaging tests (Mansoori 2011), to detect perimetric glaucoma progression.

ACKNOWLEDGEMENTS

The Cochrane Eyes and Vision (CEV) editorial base created the search strategies for the electronic databases. We thank Jennifer Burr and the Cochrane Diagnostic Test Accuracy (DTA) Review Unit for their comments on this review. We thanks Anupa Shah, Managing Editor for CEV for her assistance through the review process.

Takehiro Yamashita helped with providing details on studies in Japanese.

REFERENCES

References to studies included in this review

Akashi 2013 *{published data only}*

Akashi A, Kanamori A, Nakamura M, Fujihara M, Yamada Y, Negi A. Comparative assessment for the ability of Cirrus, RTVue, and 3D-OCT to diagnose glaucoma. *Investigative Ophthalmology and Visual Science* 2013;**54**(7):4478-84.

Aptel 2010 *{published data only}*

Aptel F, Sayous R, Fortoul V, Beccat S, Denis P. Structure-function relationships using spectral-domain optical coherence tomography: comparison with scanning laser polarimetry. *American Journal of Ophthalmology* 2010;**150**(6):825-33.

Arintawati 2013 *{published data only}*

Arintawati P, Sone T, Akita T, Tanaka J, Kiuchi Y. The applicability of ganglion cell complex parameters determined from SD-OCT images to detect glaucomatous eyes. *Journal of Glaucoma* 2013;**22**(9):713-8.

Badala 2007 *{published data only}*

Badalà F, Nouri-Mahdavi K, Raoof DA, Leeprechanon N, Law SK, Caprioli J. Optic disk and nerve fiber layer imaging to detect glaucoma. *American Journal of Ophthalmology* 2007;**144**(5):724-32.

Barella 2013 *{published data only}*

Barella KA, Costa VP, Goncalves Vidotti V, Silva FR, Dias M, Gomi ES. Glaucoma diagnostic accuracy of machine learning classifiers using retinal nerve fiber layer and optic nerve data from SD-OCT. *Journal of Ophthalmology.* 2013;**2013**:Article ID 789129.

Begum 2014a *{published data only}*

Begum VU, Addepalli UK, Yadav RK, Shankar K, Senthil S, Garudadri CS, et al. Ganglion cell-inner plexiform layer thickness of high definition optical coherence tomography in perimetric and preperimetric glaucoma. *Investigative Ophthalmology and Visual Science* 2014;**55**(8):4768-75.

Begum 2014b *{published data only}*

Begum VU, Jonnadula GB, Yadav RK, Addepalli UK, Senthil S, Choudhari NS, et al. Scanning the macula for detecting glaucoma. *Indian Journal of Ophthalmology* 2014;**62**(1):82-7.

Benitez-del-Castillo 2011 *{published data only}*

Benítez-del-Castillo J, Martinez A, Regi T. Diagnostic capability of scanning laser polarimetry with and without enhanced corneal compensation and optical coherence tomography. *European Journal of Ophthalmology* 2011;**21**(3):228-36.

Bertuzzi 2014 *{published data only}*

Bertuzzi F, Benatti E, Esempio G, Rulli E, Miglior S. Evaluation of retinal nerve fiber layer thickness measurements for glaucoma detection: GDx ECC versus spectral-domain OCT. *Journal of Glaucoma* 2014;**23**(4):232-9.

Borque 2008 *{published data only}*

Borque E, Ferreras A, Polo V, Larrosa JM, Pablo LE, Honrubia FM. Diagnostic ability of GDx VCC for glaucoma diagnosis. *Archivos de la Sociedad Espanola de Oftalmologia* 2008;**83**(6):357-64.

Bowd 2005 *{published data only}*

Bowd C, Medeiros FA, Zhang Z, Zangwill LM, Hao J, Lee TW, et al. Relevance vector machine and support vector machine classifier analysis of scanning laser polarimetry retinal nerve fiber layer measurements. *Investigative Ophthalmology and Visual Science* 2005;**46**(4):1322-9.

Bozkurt 2010 *{published data only}*

Bozkurt B, Irkec M, Arslan U. Diagnostic accuracy of Heidelberg Retina Tomograph III classifications in a Turkish primary open-angle glaucoma population. *Acta Ophthalmologica* 2010;**88**(1):125-30.

Brusini 2005 *{published data only}*

Brusini P, Salvetat ML, Parisi L, Zeppieri M, Tosoni C. Discrimination between normal and early glaucomatous eyes with scanning laser polarimeter with fixed and variable corneal compensator settings. *European Journal of Ophthalmology* 2005;**15**(4):468-76.

Brusini 2006a *{published data only}*

Brusini P, Salvetat ML, Zeppieri M, Tosoni C, Parisi L, Felletti M. Comparison between GDx VCC scanning laser polarimetry and Stratus OCT optical coherence tomography in the diagnosis of chronic glaucoma. *Acta Ophthalmologica Scandinavica* 2006;**84**(5):650-5.

Calvo 2014 *{published data only}*

Calvo P, Ferreras A, Abadia B, Ara M, Figus M, Pablo LE, Frezzotti P. Assessment of the optic disc morphology using spectral-domain optical coherence tomography and scanning laser ophthalmoscopy. *BioMed Research International* 2014;**2014**:Article ID 275654.

Chen 2007 *{published data only}*

Chen HY, Huang ML, Tsai YY, Hung PT. Diagnostic value of GDx polarimetry in a Taiwan Chinese population. *Optometry and Vision Science* 2007;**84**(7):640-6.

Chen 2008 *{published data only}*

Chen HY, Huang ML, Tsai YY, Hung PT, Lin EJ. Comparing glaucomatous optic neuropathy in primary open angle and primary angle closure glaucoma eyes by scanning laser polarimetry-variable corneal compensation. *Journal of Glaucoma* 2008;**17**(2):105-10.

Chen 2013 *{published data only}*

Chen HY, Chang YC, Wang IJ, Chen WC. Comparison of glaucoma diagnoses using Stratus and Cirrus optical coherence tomography in different glaucoma types in a Chinese population. *Journal of Glaucoma* 2013;**22**(8):638-46.

Cho 2011 *{published data only}*

Cho JW, Sung KR, Hong JT, Um TW, Kang SY, Kook MS. Detection of glaucoma by spectral domain-scanning laser ophthalmoscopy/optical coherence tomography (SD-SLO/OCT) and time domain optical coherence tomography. *Journal of Glaucoma* 2011;**20**(1):15-20.

Choi 2013 *{published data only}*

Choi YJ, Jeoung JW, Park KH, Kim DM. Glaucoma detection ability of ganglion cell-inner plexiform layer thickness by spectral-domain optical coherence tomography in high myopia. *Investigative Ophthalmology and Visual Science* 2013;**54**(3):2296-304.

Da Pozzo 2005 *{published data only}*

Da Pozzo S, Iacono P, Marchesan R, Fantin A, Ravalico G. Scanning laser polarimetry with variable corneal compensation and detection of glaucomatous optic neuropathy. *Graefe's Archive for Clinical and Experimental Ophthalmology* 2005;**243**(8):774-9.

Da Pozzo 2006 *{published data only}*

Da Pozzo S, Fuser M, Vattovani O, Di Stefano G, Ravalico G. GDx-VCC performance in discriminating normal from glaucomatous eyes with early visual field loss. *Graefe's Archive for Clinical and Experimental Ophthalmology* 2006;**244**(6):689-95.

De Leon-Ortega 2006 *{published data only}*

Deleón-Ortega JE, Arthur SN, McGwin G Jr, Xie A, Monheit BE, Girkin CA. Discrimination between glaucomatous and nonglaucomatous eyes using quantitative imaging devices and subjective optic nerve head assessment. *Investigative Ophthalmology and Visual Science* 2006;**47**(8):3374-80.

De Leon-Ortega 2007 *{published data only}*

De León-Ortega JE, Sakata LM, Monheit BE, McGwin G Jr, Arthur SN, Girkin CA. Comparison of diagnostic accuracy of Heidelberg Retina Tomograph II and Heidelberg Retina Tomograph 3 to discriminate glaucomatous and nonglaucomatous eyes. *American Journal of Ophthalmology* 2007;**144**(4):525-32.

Essock 2005 *{published data only}*

Essock EA, Zheng Y, Gunvant P. Analysis of GDx-VCC polarimetry data by Wavelet-Fourier analysis across glaucoma stages. *Investigative Ophthalmology and Visual Science* 2005;**46**(8):2838-47.

Fang 2010 *{published data only}*

Fang Y, Pan YZ, Li M, Qiao RH, Cai Y. Diagnostic capability of Fourier-Domain optical coherence tomography in early primary open angle glaucoma. *Chinese Medical Journal* 2010;**123**(15):2045-50.

Ferreras 2007 *{published data only}*

Ferreras A, Pajarín AB, Polo V, Larrosa JM, Pablo LE, Honrubia FM. Diagnostic ability of Heidelberg Retina Tomograph 3 classifications: glaucoma probability score versus Moorfields regression analysis. *Ophthalmology* 2007;**114**(11):1981-7.

Ferreras 2008a *{published data only}*

Ferreras A, Pablo LE, Pajarín AB, Larrosa JM, Polo V, Pueyo V. Diagnostic ability of the Heidelberg Retina Tomograph 3 for glaucoma. *American Journal of Ophthalmology* 2008;**145**(2):354-9.

Ferreras 2008b *{published data only}*

Ferreras A, Pablo LE, Larrosa JM, Polo V, Pajarín AB, Honrubia FM. Discriminating between normal and glaucoma-damaged eyes with the Heidelberg Retina Tomograph 3. *Ophthalmology* 2008;**115**(5):775-81.e2.

Garas 2011 *{published data only}*

Garas A, Vargha P, Holló G. Diagnostic accuracy of nerve fibre layer, macular thickness and optic disc measurements made with the RTVue-100 optical coherence tomograph to detect glaucoma. *Eye* 2011;**25**(1):57-65.

Garas 2012 *{published data only}*

Garas A, Vargha P, Holló G. Comparison of diagnostic accuracy of the RTVue Fourier-domain OCT and the GDx-VCC/ECC polarimeter to detect glaucoma. *European Journal of Ophthalmology* 2012;**22**(1):45-54.

Garudadri 2012 *{published data only}*

Garudadri CS, Rao HL, Parikh RS, Jonnadula GB, Selvaraj P, Nutheti R, et al. Effect of optic disc size and disease severity on the diagnostic capability of glaucoma imaging technologies in an Indian population. *Journal of Glaucoma* 2012;**21**(7):475-80.

Gonzales de la Rosa 2013 *{published data only}*

Gonzalez de la Rosa M, Gonzalez-Hernandez M, Sanchez-Garcia M, Rodriguez de la Vega R, Diaz-Aleman T, Pareja Rios A. Oculus-Spark perimetry compared with 3 procedures of glaucoma morphologic analysis (GDx, HRT, and OCT). *European Journal of Ophthalmology* 2013;**23**(3):316-23.

Harizman 2006 *{published data only}*

Harizman N, Zelefsky JR, Ilitchev E, Tello C, Ritch R, Liebmann JM. Detection of glaucoma using operator-dependent versus operator-independent classification in the Heidelberg retinal tomograph-III. *British Journal of Ophthalmology* 2006;**90**(11):1390-2.

Hoesl 2013 *{published data only}*

Hoesl LM, Tornow RP, Schrems WA, Horn FK, Mardin CY, Kruse FE, et al. Glaucoma diagnostic performance of GDxVCC and spectralis OCT on eyes with atypical retardation pattern. *Journal of Glaucoma* 2013;**22**(4):317-24.

Hong 2007 *{published data only}*

Hong S, Ahn H, Ha SJ, Yeom HY, Seong GJ, Hong YJ. Early glaucoma detection using the Humphrey Matrix Perimeter, GDx VCC, Stratus OCT, and retinal nerve fiber layer photography. *Ophthalmology* 2007;**114**(2):210-5.

Huang 2010 *{published data only}*

Huang ML, Chen HY, Huang WC, Tsai YY. Linear discriminant analysis and artificial neural network for glaucoma diagnosis using scanning laser polarimetry-variable cornea compensation measurements in Taiwan Chinese population. *Graefe's*

Archive for Clinical and Experimental Ophthalmology 2010;**248**(3):435-41.

Huang 2011 *{published data only}*

Huang JY, Pekmezci M, Mesiwala N, Kao A, Lin S. Diagnostic power of optic disc morphology, peripapillary retinal nerve fiber layer thickness, and macular inner retinal layer thickness in glaucoma diagnosis with fourier-domain optical coherence tomography. *Journal of Glaucoma* 2011;**20**(2):87-94.

Hwang 2012 *{published data only}*

Hwang YH, Kim YY. Glaucoma diagnostic ability of quadrant and clock-hour neuroretinal rim assessment using cirrus HD optical coherence tomography. *Investigative Ophthalmology and Visual Science* 2012;**53**(4):2226-34.

Iester 2008 *{published data only}*

Iester M, Perdicchi A, Capris E, Siniscalco A, Calabria G, Recupero SM. Comparison between discriminant analysis models and "glaucoma probability score" for the detection of glaucomatous optic nerve head changes. *Journal of Glaucoma* 2008;**17**(7):535-40.

Jeoung 2010 *{published data only}*

Jeoung JW, Park KH. Comparison of Cirrus OCT and Stratus OCT on the ability to detect localized retinal nerve fiber layer defects in preperimetric glaucoma. *Investigative Ophthalmology and Visual Science* 2010;**51**(2):938-45.

Jeoung 2013 *{published data only}*

Jeoung JW, Choi YJ, Park KH, Kim DM. Macular ganglion cell imaging study: glaucoma diagnostic accuracy of spectral-domain optical coherence tomography. *Investigative Ophthalmology and Visual Science* 2013;**54**(7):4422-9.

Jindal 2010 *{published data only}*

Jindal S, Dada T, Sreenivas V, Gupta V, Sihota R, Panda A. Comparison of the diagnostic ability of Moorfield's regression analysis and glaucoma probability score using Heidelberg retinal tomograph III in eyes with primary open angle glaucoma. *Indian Journal Ophthalmology* 2010;**58**(6):487-92.

Kanamori 2006 *{published data only}*

Kanamori A, Nagai-Kusuhara A, Escaño MF, Maeda H, Nakamura M, Negi A. Comparison of confocal scanning laser ophthalmoscopy, scanning laser polarimetry and optical coherence tomography to discriminate ocular hypertension and glaucoma at an early stage. *Graefe's Archive for Clinical and Experimental Ophthalmology* 2006;**244**(1):58-68.

Kang 2012 *{published data only}*

Kang SY, Sung KR, Na JH, Choi EH, Cho JW, Cheon MH, et al. Comparison between deviation map algorithm and peripapillary retinal nerve fiber layer measurements using Cirrus HD-OCT in the detection of localized glaucomatous visual field defects. *Journal of Glaucoma* 2012;**21**(6):372-8.

Kim 2011 *{published data only}*

Kim NR, Lee ES, Seong GJ, Kang SY, Kim JH, Hong S, et al. Comparing the ganglion cell complex and retinal nerve fibre layer measurements by Fourier domain OCT to detect glaucoma in high myopia. *British Journal of Ophthalmology* 2011;**95**(8):1115-21.

Kim 2013a *{published data only}*

Kim NR, Hong S, Kim JH, Rho SS, Seong GJ, Kim CJ. Comparison of macular ganglion cell complex thickness by Fourier-domain OCT in normal tension glaucoma and primary open-angle glaucoma. *Journal of Glaucoma* 2013;**22**(2):133-9.

Kim 2013b *{published data only}*

Kim KE, Ahn SJ, Kim DM. Comparison of two different spectral domain optical coherence tomography devices in the detection of localized retinal nerve fiber layer defects. *Japanese Journal of Ophthalmology* 2013;**57**(4):347-58.

Kim 2014a *{published data only}*

Kim MJ, Jeoung JW, Park KH, Choi YJ, Kim DM. Topographic profiles of retinal nerve fiber layer defects affect the diagnostic performance of macular scans in preperimetric glaucoma. *Investigative Ophthalmology & Visual Science* 2014;**55**(4):2079-87.

Kim 2014b *{published data only}*

Kim YJ, Kang MH, Cho HY, Lim HW, Seong M. Comparative study of macular ganglion cell complex thickness measured by spectral-domain optical coherence tomography in healthy eyes, eyes with preperimetric glaucoma, and eyes with early glaucoma. *Japanese Journal of Ophthalmology* 2014;**58**(3):244-51.

Kita 2013 *{published data only}*

Kita Y, Kita R, Takeyama A, Takagi S, Nishimura C, Tomita G. Ability of optical coherence tomography-determined ganglion cell complex thickness to total retinal thickness ratio to diagnose glaucoma. *Journal of Glaucoma* 2013;**22**(9):757-62.

Koh 2014 *{published data only}*

Koh KM, Jin S, Hwang YH. Cirrus high-definition optical coherence tomography versus spectral optical coherence tomography/scanning laser ophthalmoscopy in the diagnosis of glaucoma. *Current Eye Research* 2014;**39**(1):62-8.

Kook 2005 *{published data only}*

Kook MS, Cho HS, Seong M, Choi J. Scanning laser polarimetry using variable corneal compensation in the detection of glaucoma with localized visual field defects. *Ophthalmology* 2005;**112**(11):1970-8.

Kotowski 2012 *{published data only}*

Kotowski J, Folio LS, Wollstein G, Ishikawa H, Ling Y, Bilonick RA, et al. Glaucoma discrimination of segmented cirrus spectral domain optical coherence tomography (SD-OCT) macular scans. *British Journal of Ophthalmology* 2012;**96**(11):1420-5.

Kratz 2014 *{published data only}*

Kratz A, Lim R, Goldberg I. Optic nerve head assessment: Comparison of Cirrus optic coherence tomography and Heidelberg Retinal Tomograph 3. *Clinical and Experimental Ophthalmology* 2014;**42**(8):734-44.

Lee 2010 *{published data only}*

Lee S, Sung KR, Cho JW, Cheon MH, Kang SY, Kook MS. Spectral-domain optical coherence tomography and scanning laser polarimetry in glaucoma diagnosis. *Japanese Journal of Ophthalmology* 2010;**54**(6):544-9.

Leite 2011 *{published data only}*

Leite MT, Rao HL, Zangwill LM, Weinreb RN, Medeiros FA. Comparison of the diagnostic accuracies of the Spectralis, Cirrus, and RTVue optical coherence tomography devices in glaucoma. *Ophthalmology* 2011;**118**(7):1334-9.

Leung 2010 *{published data only}*

Leung CK, Lam S, Weinreb RN, Liu S, Ye C, Liu L, et al. Retinal nerve fiber layer imaging with spectral-domain optical coherence tomography: analysis of the retinal nerve fiber layer map for glaucoma detection. *Ophthalmology* 2010;**117**(9):1684-91.

Lisboa 2013 *{published data only}*

Lisboa R, Paranhos A Jr, Weinreb RN, Zangwill LM, Leite MT, Medeiros FA. Comparison of different spectral domain OCT scanning protocols for diagnosing preperimetric glaucoma. *Investigative Ophthalmology and Visual Science* 2013;**54**(5):3417-25.

Mai 2007 *{published data only}*

Mai TA, Reus NJ, Lemij HG. Diagnostic accuracy of scanning laser polarimetry with enhanced versus variable corneal compensation. *Ophthalmology* 2007;**114**(11):1988-93.

Mansoori 2011 *{published data only}*

Mansoori T, Viswanath K, Balakrishna N. Ability of spectral domain optical coherence tomography peripapillary retinal nerve fiber layer thickness measurements to identify early glaucoma. *Indian Journal Ophthalmology* 2011;**59**(6):455-9.

Medeiros 2004a *{published data only}*

Medeiros FA, Zangwill LM, Bowd C, Mohammadi K, Weinreb RN. Comparison of scanning laser polarimetry using variable corneal compensation and retinal nerve fiber layer photography for detection of glaucoma. *Archives of Ophthalmology* 2004;**122**(5):698-704.

Medeiros 2004b *{published data only}*

Medeiros FA, Zangwill LM, Bowd C, Weinreb RN. Comparison of the GDx VCC scanning laser polarimeter, HRT II confocal scanning laser ophthalmoscope, and stratus OCT optical coherence tomograph for the detection of glaucoma. *Archives of Ophthalmology* 2004;**122**(6):827-37.

Medeiros 2005 *{published data only}*

Medeiros FA, Zangwill LM, Bowd C, Sample PA, Weinreb RN. Use of progressive glaucomatous optic disk change as the reference standard for evaluation of diagnostic tests in glaucoma. *American Journal of Ophthalmology* 2005;**139**(6):1010-8.

Moreno 2011 *{published data only}*

Moreno PA, Konno B, Lima VC, Castro DP, Castro LC, Leite MT, et al. Spectral-domain optical coherence tomography for early glaucoma assessment: analysis of macular ganglion cell complex versus peripapillary retinal nerve fiber layer. *Canadian Journal of Ophthalmology* 2011;**46**(6):543-7.

Moreno-Montañés 2008 *{published data only}*

Moreno-Montañés J, Antón A, García N, Mendiluce L, Ayala E, Sebastián A. Glaucoma probability score vs Moorfields classification in normal, ocular hypertensive, and glaucomatous eyes. *American Journal of Ophthalmology* 2008;**145**(2):360-8.

Moreno-Montañés 2010 *{published data only}*

Moreno-Montañés J, Olmo N, Alvarez A, García N, Zarranz-Ventura J. Cirrus high-definition optical coherence tomography compared with Stratus optical coherence tomography in glaucoma diagnosis. *Investigative Ophthalmology and Visual Science* 2010;**51**(1):335-43.

Mwanza 2012 *{published data only}*

Mwanza JC, Durbin MK, Budenz DL, Sayyad FE, Chang RT, Neelakantan A, et al. Glaucoma diagnostic accuracy of ganglion cell-inner plexiform layer thickness: comparison with nerve fiber layer and optic nerve head. *Ophthalmology* 2012;**119**(6):1151-8.

Mwanza 2013 *{published data only}*

Mwanza JC, Warren JL, Budenz DL, Ganglion Cell Analysis Study Group. Combining spectral domain optical coherence tomography structural parameters for the diagnosis of glaucoma with early visual field loss. *Investigative Ophthalmology and Visual Science* 2013;**54**(13):8393-400.

Mwanza 2014 *{published data only}*

Mwanza JC, Budenz DL, Godfrey DG, Neelakantan A, Sayyad FE, Chang RT, et al. Diagnostic performance of optical coherence tomography ganglion cell--inner plexiform layer thickness measurements in early glaucoma. *Ophthalmology* 2014;**121**(4):849-54.

Na 2013a *{published data only}*

Na JH, Lee K, Lee JR, Baek S, Yoo SJ, Kook MS. Detection of macular ganglion cell loss in preperimetric glaucoma patients with localized retinal nerve fibre defects by spectral-domain optical coherence tomography. *Clinical and Experimental Ophthalmology* 2013;**41**(9):870-80.

Na 2013b *{published data only}*

Na JH, Lee KS, Lee JR, Lee Y, Kook MS. The glaucoma detection capability of spectral-domain OCT and GDx-VCC deviation maps in early glaucoma patients with localized visual field defects. *Graefe's Archive for Clinical and Experimental Ophthalmology* 2013;**251**(10):2371-82.

Nakatani 2011 *{published data only}*

Nakatani Y, Higashide T, Ohkubo S, Takeda H, Sugiyama K. Evaluation of macular thickness and peripapillary retinal nerve fiber layer thickness for detection of early glaucoma using spectral domain optical coherence tomography. *Journal of Glaucoma* 2011;**20**(4):252-9.

Nouri-Mahdavi 2013 *{published data only}*

Nouri-Mahdavi K, Nowroozizadeh S, Nassiri N, Cirineo N, Knipping S, Giaconi J, et al. Macular ganglion cell/inner

plexiform layer measurements by spectral domain optical coherence tomography for detection of early glaucoma and comparison to retinal nerve fiber layer measurements. *American Journal of Ophthalmology* 2013;**156**(6):1297-307.

Oddone 2008 *{published data only}*

Oddone F, Centofanti M, Rossetti L, Iester M, Fogagnolo P, Capris E, et al. Exploring the Heidelberg Retinal Tomograph 3 diagnostic accuracy across disc sizes and glaucoma stages: a multicenter study. *Ophthalmology* 2008;**115**(8):1358-65.

Oddone 2011 *{published data only}*

Oddone F, Centofanti M, Tanga L, Parravano M, Michelessi M, Schiavone M, et al. Influence of disc size on optic nerve head versus retinal nerve fiber layer assessment for diagnosing glaucoma. *Ophthalmology* 2011;**118**(7):1340-7.

Pablo 2010 *{published data only}*

Pablo LE, Ferreras A, Fogagnolo P, Figus M, Pajarin AB. Optic nerve head changes in early glaucoma: a comparison between stereophotography and Heidelberg retina tomography. *Eye* 2010;**24**(1):123-30.

Pueyo 2006 *{published data only}*

Pueyo V, Polo V, Larrosa JM, Ferreras A, Martínez-de-la-Casa JM, Honrubia FM. Diagnostic usefulness of optical coherence tomography (OCT), scanning laser tomography (HRT-II) and laser polarimetry (GDx) in open-angle glaucoma. *Archivos de la Sociedad Espanola de Oftalmologia* 2006;**81**(12):693-700.

Rao 2010a *{published data only}*

Rao HL, Babu GJ, Sekhar GC. Comparison of the diagnostic capability of the Heidelberg Retina Tomograph 2 and 3 for glaucoma in the Indian population. *Ophthalmology* 2010;**117**(2):275-81.

Rao 2010b *{published data only}*

Rao HL, Zangwill LM, Weinreb RN, Sample PA, Alencar LM, Medeiros FA. Comparison of different spectral domain optical coherence tomography scanning areas for glaucoma diagnosis. *Ophthalmology* 2010;**117**(9):1692-9.

Rao 2012a *{published data only}*

Rao HL, Kumbar T, Addepalli UK, Bharti N, Senthil S, Choudhari NS, et al. Effect of spectrum bias on the diagnostic accuracy of spectral-domain optical coherence tomography in glaucoma. *Investigative Ophthalmology and Visual Science* 2012;**53**(2):1058-65.

Rao 2012b *{published data only}*

Rao HL, Babu JG, Addepalli UK, Senthil S, Garudadri CS. Retinal nerve fiber layer and macular inner retina measurements by spectral domain optical coherence tomograph in Indian eyes with early glaucoma. *Eye* 2012;**26**(1):133-9.

Rao 2013 *{published data only}*

Rao HL, Addepalli UK, Chaudhary S, Kumbar T, Senthil S, Choudhari NS, et al. Ability of different scanning protocols of spectral domain optical coherence tomography to diagnose preperimetric glaucoma. *Investigative Ophthalmology and Visual Science* 2013;**54**(12):7252-7.

Rao 2014 *{published data only}*

Rao HL, Yadav RK, Addepalli UK, Chaudhary S, Senthil S, Choudhari NS, et al. Retinal nerve fiber layer evaluation of spectral domain optical coherence tomograph and scanning laser polarimeter to diagnose glaucoma. *Eye* 2014;**28**(6):654-61.

Reus 2004 *{published data only}*

Reus NJ, Lemij HG. Diagnostic accuracy of the GDx VCC for glaucoma. *Ophthalmology* 2004;**111**(10):1860-5.

Reus 2007 *{published data only}*

Reus NJ, De Graaf M, Lemij HG. Accuracy of GDx VCC, HRT I, and clinical assessment of stereoscopic optic nerve head photographs for diagnosing glaucoma. *British Journal of Ophthalmology* 2007;**91**(3):313-8.

Rho 2014 *{published data only}*

Rho S, Sung Y, Kang T, Kim NR, Kim CY. Improvement of diagnostic performance regarding retinal nerve fiber layer defect using shifting of the normative database according to vessel position. *Investigative Ophthalmology and Visual Science* 2014;**55**(8):5116-24.

Roberti 2014 *{published data only}*

Roberti G, Centofanti M, Oddone F, Tanga L, Michelessi M, Manni G. Comparing optic nerve head analysis between confocal scanning laser ophthalmoscopy and spectral domain optical coherence tomography. *Current Eye Research* 2014;**39**(10):1026-32.

Rolle 2011 *{published data only}*

Rolle T, Briamonte C, Curto D, Grignolo FM. Ganglion cell complex and retinal nerve fiber layer measured by fourier-domain optical coherence tomography for early detection of structural damage in patients with preperimetric glaucoma. *Clinical Ophthalmology* 2011;**5**:961-9.

Schrems 2010 *{published data only}*

Schrems WA, Mardin CY, Horn FK, Juenemann AG, Laemmer R. Comparison of scanning laser polarimetry and optical coherence tomography in quantitative retinal nerve fiber assessment. *Journal of Glaucoma* 2010;**19**(2):83-94.

Sehi 2007 *{published data only}*

Sehi M, Ume S, Greenfield DS. Scanning laser polarimetry with enhanced corneal compensation and optical coherence tomography in normal and glaucomatous eyes. *Investigative Ophthalmology and Visual Science* 2007;**48**(5):2099-104.

Seong 2010 *{published data only}*

Seong M, Sung KR, Choi EH, Kang SY, Cho JW, Um TW, et al. Macular and peripapillary retinal nerve fiber layer measurements by spectral domain optical coherence tomography in normal-tension glaucoma. *Investigative Ophthalmology and Visual Science* 2010;**51**(3):1446-52.

Shah 2006 *{published data only}*

Shah NN, Bowd C, Medeiros FA, Weinreb RN, Sample PA, Hoffmann EM, et al. Combining structural and functional testing for detection of glaucoma. *Ophthalmology* 2006;**113**(9):1593-602.

Shin 2013 *{published data only}*

Shin JW, Uhm KB, Lee WJ, Kim YJ. Diagnostic ability of retinal nerve fiber layer maps to detect localized retinal nerve fiber layer defects. *Eye* 2013;**27**(9):1022-31.

Sullivan-Mee 2013 *{published data only}*

Sullivan-Mee M, Ruegg CC, Pensyl D, Halverson K, Qualls C. Diagnostic precision of retinal nerve fiber layer and macular thickness asymmetry parameters for identifying early primary open-angle glaucoma. *American Journal of Ophthalmology* 2013;**156**(3):567-77.

Sung 2013 *{published data only}*

Sung MS, Yoon JH, Park SW. Diagnostic validity of macular ganglion cell-inner plexiform layer thickness deviation map algorithm using cirrus HD-OCT in preperimetric and early glaucoma. *Journal of Glaucoma* 2014;**23**(8):e144-51.

Takahashi 2008 *{published data only}*

Takahashi H, Chihara E. Impact of diabetic retinopathy on quantitative retinal nerve fiber layer measurement and glaucoma screening. *Investigative Ophthalmology and Visual Science* 2008;**49**(2):687-92.

Weinreb 2003 *{published data only}*

Weinreb RN, Bowd C, Zangwill LM. Glaucoma detection using scanning laser polarimetry with variable corneal polarization compensation. *Archives of Ophthalmology* 2003;**121**(2):218-24.

Wu 2012 *{published data only}*

Wu H, De Boer JF, Chen TC. Diagnostic capability of spectral-domain optical coherence tomography for glaucoma. *American Journal of Ophthalmology* 2012;**153**(5):815-26.

Yamada 2014 *{published data only}*

Yamada H, Hangai M, Nakano N, Takayama K, Kimura Y, Miyake M, et al. Asymmetry analysis of macular inner retinal layers for glaucoma diagnosis. *American Journal of Ophthalmology* 2014;**158**(6):1318-29.

Yang 2014 *{published data only}*

Yang Z, Tatham AJ, Zangwill LM, Weinreb RN, Zhang C, Medeiros FA. Diagnostic ability of retinal nerve fiber layer imaging by swept-source optical coherence tomography in glaucoma. *American Journal of Ophthalmology* 2014;**159**(1):193-201.

Yoshida 2014 *{published data only}*

Yoshida T, Iwase A, Hirasawa H, Murata H, Mayama C, Araie M, et al. Discriminating between glaucoma and normal eyes using optical coherence tomography and the 'random forests' classifier. *PLoS ONE* 2014;**9**:e106117.

Zelefsky 2006 *{published data only}*

Zelefsky JR, Harizman N, Mora R, Ilitchev E, Tello C, Ritch R, et al. Assessment of a race-specific normative HRT-III database to differentiate glaucomatous from normal eyes. *Journal of Glaucoma* 2006;**15**(6):548-51.

Zeppieri 2010 *{published data only}*

Zeppieri M, Brusini P, Parisi L, Johnson CA, Sampaolesi R, Salvetat ML. Pulsar perimetry in the diagnosis of early glaucoma. *American Journal of Ophthalmology* 2010;**149**(1):102-12.

Zhang 2014 *{published data only}*

Zhang C, Tatham AJ, Weinreb RN, Zangwill LM, Yang Z, Zhang JZ, et al. Relationship between ganglion cell layer thickness and estimated retinal ganglion cell counts in the glaucomatous macula. *Ophthalmology* 2014;**121**(12):2371-9.

Zheng 2008 *{published data only}*

Zheng W, Baohua C, Qun C, Zhi Q, Hong D. Retinal nerve fiber layer images captured by GDx-VCC in early diagnosis of glaucoma. *Ophthalmologica* 2008;**222**(1):17-20.

Additional references

Abrams 1994

Abrams LS, Scott IU, Spaeth GL, Quigley HA, Varma R. Agreement among optometrists, ophthalmologists, and residents in evaluating the optic disc for glaucoma. *Ophthalmology* 1994;**101**(10):1662-7.

AGIS 1994

Anonymous. The Advanced Glaucoma Intervention Study (AGIS): 1. Study design and methods and baseline characteristics of study patients. *Controlled Clinical Trials* 1994;**15**(4):299-325.

Andersson 2011

Andersson S, Heijl A, Bengtsson B. Optic disc classification by the Heidelberg Retina Tomograph and by physicians with varying experience of glaucoma. *Eye* 2011;**25**(11):1401-7.

Bengtsson 2012

Bengtsson B, Andersson S, Heijl A. Performance of time-domain and spectral-domain Optical Coherence Tomography for glaucoma screening. *Acta Ophthalmologica* 2012;**90**(4):310-5.

Brusini 2006b

Brusini P, Filacorda S. Enhanced Glaucoma Staging System (GSS 2) for classifying functional damage in glaucoma. *Journal of Glaucoma* 2006;**15**(1):40-6.

Burr 2007

Burr JM, Mowatt G, Hernandez R, Siddiqui MA, Cook J, Lourenco T, et al. The clinical effectiveness and cost-effectiveness of screening for open angle glaucoma: a systematic review and economic evaluation. *Health Technology Assessment* 2007;**11**(41):iii-iv, ix-x, 1-190.

Burr 2014

Burr J, Hernández R, Ramsay C, Prior M, Campbell S, Azuara-Blanco A, et al. Is it worthwhile to conduct a randomized controlled trial of glaucoma screening in the United Kingdom? *Journal of Health Services Research and Policy* 2014;**19**(1):42-51.

Bussel 2013

Bussel II, Wollstein G, Schuman JS. OCT for glaucoma diagnosis, screening and detection of glaucoma progression. *British Journal of Ophthalmology* 2013;**98**(Suppl 2):15-9.

Chauhan 2008

Chauhan BC, Mikelberg FS, Balaszi AG, LeBlanc RP, Lesk MR, Trope GE, et al. Canadian Glaucoma Study: 2. risk factors for the progression of open-angle glaucoma. *Archives of Ophthalmology* 2008;**126**(8):1030-6.

Cheng 2014

Cheng J, Beltran-Agullo L, Trope GE, Buys YM. Assessment of the quality of glaucoma referral letters based on a survey of glaucoma specialists and a glaucoma guideline. *Ophthalmology* 2014;**121**(1):126-33.

CIGTS 1999

Musch DC, Lichter PR, Guire KE, Standardi CL. The Collaborative Initial Glaucoma Treatment Study: study design, methods, and baseline characteristics of enrolled patient. *Ophthalmology* 1999;**106**(4):653-62.

Coleman 2008

Coleman AL, Miglior S. Risk factors for glaucoma onset and progression. *Survey of Ophthalmology* 2008;**53**(Suppl 1):S3-10.

Deeks 2005

Deeks JJ, Macaskill P, Irwig L. The performance of tests of publication bias and other sample size effects in systematic reviews of diagnostic test accuracy was assessed. *Journal of Clinical Epidemiology* 2005;**58**(9):882-93.

EGS 2008 Guidelines

European Glaucoma Society. Terminology and Guidelines for Glaucoma. www.icoph.org/dynamic/attachments/resources/egs_guidelines_4_english.pdf (accessed 12 February 2015).

EMGT 1999

Leske MC, Heijl A, Hyman L, Bengtsson B. Early Manifest Glaucoma Trial: design and baseline data. *Ophthalmology* 1999;**106**(11):2144-53.

Ervin 2012

Ervin AM, Boland MV, Myrowitz EH, Prince J, Hawkins B, Vollenweider D, et al. Screening for Glaucoma: Comparative Effectiveness. Comparative Effectiveness Review No. 59. www.ncbi.nlm.nih.gov/books/NBK95371/?report=reader (accessed 16 February 2014).

Garway-Heath 1998

Garway-Heath DF, Hitchings RA. Sources of bias in studies of optic disc and retinal nerve fibre layer morphology. *British Journal of Ophthalmology* 1998;**82**(9):986.

Gordon-Bennet 2008

Gordon-Bennett PS, Ioannidis AS, Papageorgiou K, Andreou PS. A survey of investigations used for the management of glaucoma in hospital service in the United Kingdom. *Eye* 2008;**22**(11):1410-8.

Harbord 2007

Harbord RM, Deeks JJ, Egger M, Whiting P, Sterne JA. A unification of models for meta-analysis of diagnostic accuracy studies. *Biostatistics* 2007;**8**(2):239-51.

Heijl 2013

Heijl A, Bengtsson B, Oskarsdottir SE. Prevalence and severity of undetected manifest glaucoma: results from the early manifest glaucoma trial screening. *Ophthalmology* 2013;**120**(8):1541-5.

Kamdeu 2011

Kamdeu Fansi AA, Agoumi Y, Harasymowycz PJ. Screening for glaucoma with Moorfields regression analysis and glaucoma probability score in confocal scanning laser ophthalmoscopy. *Canadian Journal of Ophthalmology* 2011;**46**(3):254-60.

Katz 1995

Katz J, Quigley HA, Sommer A. Repeatability of the Glaucoma Hemifield Test in automated perimetry. *Investigative Ophthalmology and Visual Science* 1995;**36**(8):1658-64.

Kwartz 2005

Kwartz AJ, Henson DB, Harper RA, Spencer AF, McLeod D. The effectiveness of the Heidelberg Retina Tomograph and laser diagnostic glaucoma scanning system (GDx) in detecting and monitoring glaucoma. *Health Technology Assessment* 2005;**9**(46):1-132, iii.

Lee 2014

Lee KH, Kim CY, Kim NR. Variations of retinal nerve fiber layer thickness and ganglion cell-inner plexiform layer thickness according to the torsion direction of optic disc. *Investigative Ophthalmology and Visual Science* 2014;**55**(2):1048-55.

Li 2013

Li G, Fansi AK, Harasymowycz P. Screening for glaucoma using GDx-VCC in a population with >1 risk factors. *Canadian Journal of Ophthalmology* 2013;**48**(4):279-85.

Lichter 1976

Lichter PR. Variability of expert observers in evaluating the optic disc. *Transactions of the American Ophthalmological Society* 1976;**74**:532-72.

Medeiros 2004

Medeiros FA, Zangwill LM, Bowd C, Weinreb RN. Comparison of the GDx VCC scanning laser polarimeter, HRT II confocal scanning laser ophthalmoscope, and stratus OCT optical coherence tomograph for the detection of glaucoma. *Archives of Ophthalmology* 2004;**122**(6):827-37.

Meier 2014

Meier KL, Greenfield DS, Hilmantel G, Kahook MY, Lin C, Rorer EM, et al. Special commentary: Food and drug administration and American glaucoma society co-sponsored workshop: The validity, reliability, and usability of glaucoma imaging devices. *Ophthalmology* 2014;**121**(11):2116-23.

Mills 2006

Mills RP, Budenz DL, Lee PP, Noecker RJ, Walt JG, Siegartel LR, et al. Categorizing the stage of glaucoma from pre-diagnosis

to end-stage disease. *American Journal of Ophthalmology* 2006;**141**(1):24-30.

Mowatt 2008

Mowatt G, Burr JM, Cook JA, Siddiqui MA, Ramsay C, Fraser C, et al. Screening tests for detecting open-angle glaucoma: systematic review and meta-analysis. *Investigative Ophthalmology and Visual Science* 2008;**49**(12):5373-85.

OHTS 2002

Gordon MO, Beiser JA, Brandt JD, Heuer DK, Higginbotham EJ, Johnson CA, et al. The Ocular Hypertension Treatment Study: baseline factors that predict the onset of primary open-angle glaucoma. *Archives of Ophthalmology* 2002;**120**(6):714-20.

Olson 2013

Olson J, Sharp P, Goatman K, Prescott G, Scotland G, Fleming A, et al. Improving the economic value of photographic screening for optical coherence tomography-detectable macular oedema:a prospective, multicentre, UK study. *Health Technology Assessment* 2013;**17**(51):1-142.

QUADAS 2

Whiting PF, Rutjes AW, Westwood ME, Mallett S, Deeks JJ, Reitsma JB, et al. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Annals of Internal Medicine* 2011;**155**(8):529-36.

Quigley 2006

Quigley HA, Broman AT. The number of people with glaucoma worldwide in 2010 and 2020. *British Journal of Ophthalmology* 2006;**90**(3):262-7.

Ratnarajan 2013

Ratnarajan G, Newsom W, Vernon SA, Fenerty C, Henson D, Spencer F, et al. The effectiveness of schemes that refine referrals between primary and secondary care--the UK experience with glaucoma referrals: the Health Innovation & Education Cluster (HIEC) Glaucoma Pathways Project. *BMJ Open* 2013;**3**(7):pii:e002715.

Reitsma 2005

Reitsma JB, Glas AS, Rutjes AW, Scholten RJ, Bossuyt PM, Zwinderman AH. Bivariate analysis of sensitivity and specificity produces informative summary measures in diagnostic reviews. *Journal of Clinical Epidemiology* 2005;**58**(10):982-90.

Reitsma 2009

Reitsma JB, Rutjes AWS, Whiting P, Vlassov VV, Leeflang MMG, Deeks JJ. Chapter 9: Assessing methodological quality. In: Deeks JJ, Bossuyt PM, Gatsonis C (editors), Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy Version 1.0.0. The Cochrane Collaboration, 2009. Available from: srdta.cochrane.org.

Reus 2010

Reus NJ, Lemij HG, Garway-Heath DF, Airaksinen PJ, Anton A, Bron AM, et al. Clinical assessment of stereoscopic optic disc photographs for glaucoma: the European Optic Disc Assessment Trial. *Ophthalmology* 2010;**117**(4):717-23.

Rutter 1995

Rutter CM, Gatsonis CA. Regression methods for meta-analysis of diagnostic test data. *Academic Radiology* 1995;**2**(Suppl 1):S48-56.

Rutter 2001

Rutter CM, Gatsonis CA. A hierarchical regression approach to meta-analysis of diagnostic test accuracy evaluations. *Statistics in Medicine* 2001;**20**(19):2865-84.

Sommer 1991

Sommer A, Katz J, Quigley HA, Miller NR, Robin AL, Richter RC, et al. Clinically detectable nerve fiber atrophy precedes the onset of glaucomatous field loss. *Archives of Ophthalmology* 1991;**109**(1):77-83.

Spaeth 2006

Spaeth GL, Lopes JF, Junk AK, Grigorian AP, Henderer J. Systems for staging the amount of optic nerve damage in glaucoma: a critical review and new material. *Survey of Ophthalmology* 2006;**51**(4):293-315.

Springelkamp 2014

Springelkamp H, Lee K, Wolfs RC, Buitendijk GH, Ramdas WD, Hofman A, et al. Population-based evaluation of retinal nerve fiber layer, retinal ganglion cell layer, and inner plexiform layer as a diagnostic tool for glaucoma. *Investigative Ophthalmology and Visual Science* 2014;**55**(12):8428-38.

Spry 2003

Spry PG, Johnson CA, McKendrick AM, Turpin A. Measurement error of visual field tests in glaucoma. *British Journal of Ophthalmology* 2003;**87**(1):107-12.

Strouthidis 2008

Strouthidis NG, Garway-Heath DF. New developments in Heidelberg retina tomograph for glaucoma. *Current Opinion in Ophthalmology* 2008;**19**(2):141-8.

Sung 2011

Sung KR, Kim JS, Wollstein G, Folio L, Kook MS, Schuman JS. Imaging of the retinal nerve fibre layer with spectral domain optical coherence tomography for glaucoma diagnosis. *British Journal of Ophthalmology* 2011;**95**(7):909-14.

Takwoingi 2008 [Computer program]

Available at srdta.cochrane.org/software-development. The Cochrane Collaboration METADAS: A SAS macro for meta-analysis of diagnostic accuracy studies. Takwoingi Y, Deeks JJ. Available at srdta.cochrane.org/software-development. The Cochrane Collaboration, 2008.

Takwoingi 2013

Takwoingi Y, Leeflang MM, Deeks JJ. Empirical evidence of the importance of comparative studies of diagnostic test accuracy. *Annals of Internal Medicine* 2013;**158**(7):544-54.

Weinreb 2004

Weinreb RN, Khaw PT. Primary open-angle glaucoma. *Lancet* 2004;**363**(9422):1711-20.

Whiting 2003

Whiting P, Rutjes AW, Reitsma JB, Bossuyt PM, Kleijnen J. The development of QUADAS: a tool for the quality assessment of studies of diagnostic accuracy included in systematic reviews. *BMC Medical Research Methodology* 2003;**3**:25.

Wilson 2002

Wilson MR, Kosoko O, Cowan CL Jr, Sample PA, Johnson CA, Haynatzki G, et al. Progression of visual field loss in untreated glaucoma patients and glaucoma suspects in St. Lucia, West Indies. *American Journal of Ophthalmology* 2002;**134**(3):399-405.

References to other published versions of this review

Oddone 2010

Oddone F, Virgili G, Parravano M, Brazzelli M, Novielli N, Michelessi M. Optic nerve head and fibre layer imaging for diagnosing glaucoma. *Cochrane Database of Systematic Reviews* 2010, Issue 11. Art. No: CD008803. [DOI: 10.1002/14651858.CD008803]

CHARACTERISTICS OF STUDIES

Characteristics of included studies *[ordered by study ID]*

Akashi 2013

Study characteristics	
Patient Sampling	Observational cross-sectional study in which Japanese glaucomatous and normal people were enrolled. If eligible, both eyes of the same patients were included in the study.
Patient characteristics and setting	**Sample size**: 232 participant enrolled, 145 glaucoma (75 of whom considered as early glaucoma) and 87 controls. **Age**: all glaucoma patients mean ± SD, 47.6 ± 9.4 years; early glaucoma patients mean ± SD, 48.3 ± 10.6 years; controls 43.5 ± 12.8 years. **Sex**: 102 men (68 glaucoma, 34 controls) and 130 women (77 glaucoma, 53 controls). **Ethnicity**: Japanese. **Country**: Japan. **Setting**: Kobe University Hospital. **Ocular comorbidities**: Patient with BCVA worse than 20/40, spherical refraction < -6 D, a cylinder correction > ±3 D were not included. Patients with any previous ocular surgery, VF loss due to vitreoretinal diseases, and optic nerve or RNFL abnormality unrelated to glaucomatous optic neuropathy, were excluded. **Spectrum of glaucoma severity**: the mean ± SD MD on the VF test were -7.12 ± 6.62 dB for glaucoma. According to Anderson and Patella's classification, patient with MD > -6 were considered as early glaucoma. **Control participants: IOP** ≤ 21 mmHg and reliable VF test result with no abnormal finding suggestive of glaucoma.
Index tests	**Optical coherence tomography**: Cirrus HD-OCT (software version 6.1.0.96; Carl Zeiss Meditec). The optic disc cube protocol 200 x 200 and macular cube 200 x 200 protocols were used. Images with signal strength < 6 were excluded. **Optical coherence tomography:** RTVue-100 (software version 4.0.5.39; Optovue, Inc., Fremont, CA, USA). The ONH map and GCC protocols were used. Only images with a signal strength index > 30 were accepted. **Optical coherence tomography**: 3D OCT-2000 (software version 8.00; Topcon, Inc., Tokyo, Japan). The 3D 7 x 7 mm scan disc and 3D macular protocols were used. Images with a quality factor < 60 were excluded. No authors had conflict of interest.

Akashi 2013 *(Continued)*

Target condition and reference standard(s)	**Manifest glaucoma:** eyes with glaucomatous optic nerve appearance (defined as neuroretinal rim damage, an increased cup-to-disc ratio, rim thinning, and notches with or without RNFL defects) and glaucomatous VF defects (defined as 2+ contiguous points with a PSD sensitivity loss of $P < 0.01$, 3+ contiguous points with sensitivity loss of $P < 0.05$ not crossing the horizontal meridian line, or a 10-dB difference across the nasal horizontal midline at 2+ adjacent locations, and GHT outside normal limit). **Visual field testing:** Humphrey Field Analyzer, 30-2 SITA standard programme (Carl Zeiss Meditec). **Optic disc evaluation**: no details were reported.
Flow and timing	Index tests and reference standard were performed within 6 months. No patients were reported by the authors as excluded from the analysis.
Comparative	
Notes	None.

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Was a case-control design avoided?	No		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (All tests)			
If a threshold was used, was it prespecified?	Yes		
Were imaging test's quality assessed?	Yes		
Were any conflict of interest avoided	Yes		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern

Akashi 2013 *(Continued)*

DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	No		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
Did all patients receive a reference standard	Yes		
Could the patient flow have introduced bias?		Unclear risk	

Aptel 2010

Study characteristics	
Patient Sampling	Prospective investigation conducted in a French university-affiliated glaucoma centre. 166 patients were initially screened. One eye from each of 120 patients were finally included in the analysis: 40 with glaucoma, 40 with suspected glaucoma, and 40 healthy participants.
Patient characteristics and setting	**Sample size**: 166 patients initially screened, 120 eyes of 120 patients included in the analysis (40 glaucoma, 40 suspected glaucoma, 40 controls). **Age**: glaucoma patients mean ± SD, 63.4 ± 11.2 years; suspected glaucoma 61.7 ± 12.7 years; controls 60.9 ± 13.1. **Sex**: 46 men (14 glaucoma, 15 suspected glaucoma, 17 controls) and 74 women (26 glaucoma, 25 suspected glaucoma, 23 controls). **Ocular comorbidities**: no retinal disease, BCVA < 20/40, SE < -6 or > +3 D, non-glaucomatous optic neuropathy or intraocular surgery except for uncomplicated cataract surgery.

Aptel 2010 *(Continued)*

	Setting: French university-affiliated glaucoma centre. **Country**: France. **Spectrum of glaucoma severity**: mean ± SD MD and PSD on the VF test were -9.88 ± 6.93 dB and 4.42 ± 4.85 dB for glaucoma, -1.73 ± 2.16 dB and 2.06 ± 0.54 dB for suspected glaucoma. **Control participants**: IOP < 22 mmHg, normal optic disc appearance and no repeatable abnormal SAP results.
Index tests	**Scanning laser polarimetry**: GDx VCC, software version 5.5.0 (Carl Zeiss Meditec, Inc.). Only well-focused, well-centred images with a quality scan score >8 were used. **Optical coherence tomography**: Cirrus OCT, software version 3.0 (Carl Zeiss Meditec, Inc, Dublin, California, USA). Imaging was performed using the 200 x 200 protocol optic disc cube. Only well-focused, well-centred images, without eye movement and with a signal strength of 7/10 or more, were used. The authors indicate no financial conflict of interest.
Target condition and reference standard(s)	**Manifest glaucoma:** abnormal SAP result (GHT and PSD outside 95% of normal limits), and optic nerve damage (asymmetric cup-to-disc ratio > 0.2, rim thinning, notching, excavation, or RNFL defect). **Suspected glaucoma:** glaucomatous optic nerve damage (asymmetric cup-to-disc ratio > 0.2, rim thinning, notching, excavation, or retinal nerve fibre layer defect) without repeatable abnormal SAP results. **Visual field testing:** Humphrey Field Analyzer, 24-2 SITA–standard strategy (Zeiss-Humphrey Systems, Dublin, CA, USA). **Optic disc evaluation:** dilated fundus was examined by indirect ophthalmoscopy.
Flow and timing	Time interval between reference standard and index tests was not reported. 14 patients were excluded from the analysis for poor OCT quality criteria, 23 for poor GDx VCC quality criteria, and 28 for poor VF quality or reliability criteria.
Comparative	
Notes	None.

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Was a case-control design avoided?	No		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		High risk	

Aptel 2010 *(Continued)*

Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (All tests)			
If a threshold was used, was it pre-specified?	Yes		
Were imaging test's quality assessed?	Yes		
Were any conflict of interest avoided	Yes		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Unclear		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	No		
Did all patients receive a reference standard	Yes		
Could the patient flow have introduced bias?		High risk	

Arintawati 2013

Study characteristics	
Patient Sampling	Retrospective study in which new glaucoma and glaucoma-suspect patients, referred to the Department of Ophthalmology, between March 2008 and April 2011, were recruited. 164 patients were studied. 261 eyes were included in the analysis.
Patient characteristics and setting	**Sample size**: 261 eyes included in the analysis (80 advanced glaucoma, 81 early glaucoma, 32 preperimetric glaucoma and 68 controls). **Age**: glaucoma mean ± SD, 61.49 ± 14.21 years (advanced glaucoma 64.56 ± 10.89; early glaucoma 60.16 ± 16.77; preperimetric glaucoma 58.94 ± 12.15 years); controls 59.65 ± 16.88 years. **Sex**: 113 men and 150 women **Ethnicity**: not specified. **Setting:** Department of Ophthalmology, Hiroshima University Hospital. **Country**: Japan. **Ocular comorbidities**:patient with refractive errors (spherical equivalent) > +3.00 D or < 7.00 D, and those with retinal disease that could cause VF defects or optic disc abnormalities were excluded. **Spectrum of glaucoma severity**: The mean ± SD mean deviation and PSD on the VF test were –6.05 ± 6.22 and 6.57 ± 4.88 for glaucoma group overall (-0.11 ± 1.55 and 1.58 ± 0.31 respectively for the preperimetric eyes, -2.68 ± 1.79 and 4.03 ± 2.57 respectively for the early glaucoma, -11.99 ± 5.29 and 11.26 ± 3.47 respectively for advanced glaucoma). **Control participants**: IOP < 22 mmHg, normal optic disc appearance, and normal ophthalmological findings.
Index tests	RTVue Fourier-domain OCT system (OptovueInc., Fremont, CA, USA); software version 4.0.5.100). Imaging was performed using GCC and RNFL 3.45 mode analysis. Images with misalignment of the surface detection algorithm, or decentration of the measurement circle and the signal strength index < 40, were excluded. The authors indicate no financial conflict of interest.
Target condition and reference standard(s)	**Manifest glaucoma:** VF defects (defined as the pattern deviation plot with more than 3 contiguous points with P < 0.05 and at least 1 with P < 0.01 level on the same side of the horizontal meridian and GHt outside the normal limit) and glaucomatous optic disc appearance (neuroretinal rim loss, notching, focal thinning of the nerve fibre layer, disc haemorrhages, or vertical elongation of the optic cup). Preperimetric Glaucoma: glaucomatous optic disc appearance but normal VF results. **Visual field test:** Humphrey Field Analyzer (Carl Zeiss Meditec, Dublin, CA, USA); 24-2 SITA–standard strategy. **Optic nerve evaluation**: Dilated fundus biomicroscopy.
Flow and timing	164 patients were originally studied. Patients with SD-OCT not good were excluded from this study. 261 eyes were included in the analysis, but details about number of exclusions were not reported. Time interval between reference standard and index tests was not reported.
Comparative	

Arintawati 2013 *(Continued)*

Notes	None.

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Was a case-control design avoided?	No		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (All tests)			
If a threshold was used, was it pre-specified?	Yes		
Were imaging test's quality assessed?	Yes		
Were any conflict of interest avoided	Yes		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			

Arintawati 2013 *(Continued)*

Was there an appropriate interval between index test and reference standard?	Unclear	
Did all patients receive the same reference standard?	Yes	
Were all patients included in the analysis?	Unclear	
Did all patients receive a reference standard	Yes	
Could the patient flow have introduced bias?		Unclear risk

Badala 2007

Study characteristics	
Patient Sampling	Cases were extracted from the clinical database of the Glaucoma Division at Jules Stein Eye Institute (University of California, LA) choosing from patients who underwent VF testing and optic disk imaging with OCT, CSLO, SLP and stereoscopic optic disk photographs at the same visit between April 1 2003 and April 1 2006. Normal patients were recruited among staff, patients' spouses, and volunteers.
Patient characteristics and setting	**Sample size**: 92 eyes of 92 patients (46 glaucoma, 46 healthy controls). **Age:** glaucoma patients mean ± SD, 61.8 ± 9.7 years; controls 58.9 ± 6.8. **Sex:** 37 men (20 glaucoma, 17 controls) and 55 women (26 glaucoma, 29 controls). **Ethnicity:** glaucoma: 31 white, 5 black, 4 Hispanic and 6 Asian. Controls: 25 white, 1 black, 9 Hispanic and 11 Asian. **Country:** USA. **Ocular comorbidities**: no ocular disease other than glaucoma, BCVA < 20/40, SE > ±5 D, and no history of ocular surgery/trauma. **Setting:** Glaucoma Division, Jules Stein Eye Institute, University of California, Los Angeles. **Spectrum of glaucoma severity:** mean ± SD MD and PSD on the VF test were −4.0 ± 2.5 dB and 5.5 ± 2.5 dB. No patients had MD < -8 dB. **Control participants:** normal optic disc, IOP ≤ 21 mmHg and a normal SAP (GHT within normal limits and a PSD with a P > 0.05 on 2 consecutive examinations).
Index tests	**Scanning laser polarimetry**: GDx-VCC, software version 5.2.3 (Laser Diagnostic Technologies, San Diego, CA, USA). The image quality scores were averaged and reported. The authors indicate no financial conflict of interest.
Target condition and reference standard(s)	**Manifest Glaucoma:** early defect on SAP (defined as GHT results outside normal limits, a PSD with P < 0.05 and a MD of more than -8 dB) and open angle by gonioscopy. **Visual field testing:** Humphrey Field Analyzer, model 750, 24-2 SITA-Standard strategy (Allergan Humphrey, San Leandro, CA., USA). Only patients with reliable

Badala 2007 *(Continued)*

	fields (fixation loss rate < 33%; false-positive and false-negative rates < 20%) were included. Optic disc appearance was not part of the reference standard.
Flow and timing	Reference standard and imaging tests were performed during the same day. All patients enrolled were included in the analysis.
Comparative	
Notes	None.

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Was a case-control design avoided?	No		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (All tests)			
If a threshold was used, was it pre-specified?	Yes		
Were imaging test's quality assessed?	Yes		
Were any conflict of interest avoided	Yes		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		

Badala 2007 *(Continued)*

Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
Did all patients receive a reference standard	Yes		
Could the patient flow have introduced bias?		Low risk	

Barella 2013

Study characteristics	
Patient Sampling	Observational, case-control study, enrolling 103 eyes of 103 participants (46 control patients and 57 glaucoma). One eye per person was randomly selected.
Patient characteristics and setting	**Sample size**: 103 eyes of 103 patients (57 glaucoma and 46 controls). **Age**: glaucoma mean ± SD, 59.9 ± 9.0 years; controls, 56.5 ± 8.9 years. **Sex**: 51 men (28 glaucoma, 23 controls) and 52 women (29 glaucoma, 23 controls) **Ethnicity**: 78 white (43 glaucoma and 35 controls); 25 African-American (14 glaucoma and 11 controls). **Clinical setting**: Glaucoma Service of the University of Campinas (UNICAMP). **Country**: Brazil. **Ocular comorbidities**: patient with retinal diseases, uveitis, pseudophakia or aphakia, non-glaucomatous optic neuropathy, and significant cataract were excluded. **Spectrum of glaucoma severity**: the mean ± SD mean deviation and PSD on the VF test were -4.0 ± 2.4 and 4.3 ± 2.4 respectively, for glaucomatous eyes. 86% had early VF damage, 14%, moderate VF damage. **Control participants**: IOP ≤ 21 mmHg with no history of elevated IOP or glaucoma cases in the family and 2 consecutive and reliable normal VFs.
Index tests	**Optical coherence tomography**: Cirrus SD-OCT (version 5.1.1.6, Carl Zeiss Meditec Inc., Dublin, CA, USA). ONH modes scan was used to measure RNFL thickness and ONH topography measurement. Poor-quality images with incorrect identification of the vitreoretinal surface, horizontal eye motion within the measurement circle, and misidentification of Bruch's membrane, or a signal strength < 6 were excluded. All

Barella 2013 *(Continued)*

	images were acquired with undilated pupils by a single, well-trained ophthalmologist, masked for the diagnosis. No conflicts of interest were reported
Target condition and reference standard(s)	**Manifest glaucoma:** IOP measurements > 21 mmHg and a glaucomatous VF defect confirmed in 2 recent and reliable examinations. VF defects were defined as 2 of the following criteria: cluster of 3 points with P < 5% on a pattern deviation map in a single hemifield, including at least 1 point with P < 1% or GHT outside normal limits, or PSD outside normal limits. **Visual field test:** Humphrey Field Analyzer II (Carl Zeiss Meditec, Dublin, CA, USA); 24-2 SITA–standard strategy. **Optic nerve appearance**: dilated slit lamp fundus examination.
Flow and timing	No details reported. Time interval between reference standard and index tests was not reported.
Comparative	
Notes	None.

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Was a case-control design avoided?	No		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 2: Index Test (All tests)			
If a threshold was used, was it pre-specified?	Yes		
Were imaging test's quality assessed?	Yes		
Were any conflict of interest avoided	Yes		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern

Barella 2013 *(Continued)*

DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Unclear		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Unclear		
Did all patients receive a reference standard	Yes		
Could the patient flow have introduced bias?		Unclear risk	

Begum 2014a

Study characteristics	
Patient Sampling	Cross-sectional, case-control study of the baseline examinations of participants included in a prospective longitudinal study (LOGES), enrolling glaucoma, glaucoma suspects and normal controls.
Patient characteristics and setting	**Sample size**: 304 eyes of 174 patients enrolled. 136 eyes of 112 patients included in the analysis (62 eyes of 46 perimetric glaucoma; 21 eyes of 18 preperimetric glaucoma and 53 eyes of 38 control patients **Age**: perimetric glaucoma median (IQR), 53 (45, 58) years; preperimetric glaucoma median (IQR), 47 (36, 60) years; controls, 42 (33, 53) years. **Sex**: 67 men (34 perimetric glaucoma, 12 preperimetric glaucoma, 21 controls) and 35 women (12 perimetric glaucoma, 6 preperimetric glaucoma, 17 controls) **Ethnicity**: Indian **Clinical setting**:L V Prasad Eye Institute, Hyderabad, **Country**: India

Begum 2014a *(Continued)*

	Ocular comorbidities: patient with any media opacities that prevented good quality optic disc photographs and other imaging tests, and any retinal (including macular) or neurologic disease other than glaucoma, were excluded. **Spectrum of glaucoma severity**: the median (IQR) mean deviation and PSD on the VF test were -1.9 (-2.9, -0.8) and 1.7 (1.3, 1.9) respectively, for preperimetric glaucomatous; -11.4 (-17.5, -4.9) and 7.6 (4.9, 10.0) respectively, for perimetric glaucoma. **Control participants**: non-glaucomatous optic discs appearance and normal VF result.
Index tests	**Optical coherence tomography**: Cirrus HD-OCT (software version 6.0). Macular cube 200 x 200 and optic disc cube 200 x 200, were the scanning protocol used. Only good-quality scans with signal strength > 6, absence of motion and blinking artefacts, and segmentation failure were used for the analysis. No conflict of interest with manufacturer were reported.
Target condition and reference standard(s)	**Manifest perimetric glaucoma:** glaucomatous optic discs (defined as the presence of focal or diffuse neuroretinal rim thinning, localised notching, or nerve fibre layer defects and glaucomatous) and glaucomatous VF result (defined as the PSD < 5% and GHT outside normal limits). **Visual field test:** Humphrey Field Analyzer, model 750i (Zeiss Humphrey Systems, Dublin, CA, USA), with the SITA-standard programme. The VFs were considered reliable if the fixation losses, false-positive and false-negative response rates were < 20%. A single observer masked to the optic disc classification, SD-OCT findings and the other eye status, graded all VFs. **Optic disc evaluation**: stereoscopic optic disc photographs using digital fundus camera (FF 450 plus with VISUPAC 4.2.2; Carl Zeiss Meditec Systems GmbH, Pirmasens, Germany). Optic disc photograph was evaluated independently by 2 experts masked to the clinical details of the patients.
Flow and timing	21 eyes were excluded due to poor disc photographs. 28 eyes due to unreliable VFs were excluded and 57 eyes were excluded due to poor quality HD-OCT scans. 106 out of 242 eyes (> 10%) originally considered, were not included in the analysis. VF data of the same imaging day were reported but time interval between all the reference standard and imaging session are unclearly reported.
Comparative	
Notes	None.

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Was a case-control design avoided?	No		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		High risk	

Begum 2014a *(Continued)*

Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 2: Index Test (All tests)			
If a threshold was used, was it prespecified?	Yes		
Were imaging test's quality assessed?	Yes		
Were any conflict of interest avoided	Yes		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Unclear		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	No		
Did all patients receive a reference standard	Yes		
Could the patient flow have introduced bias?		High risk	

Begum 2014b

Study characteristics	
Patient Sampling	Retrospective, cross-sectional study. 295 eyes were randomly selected (after the exclusion of eyes with poor index or reference-test quality results) from 678 eyes of 382 patients referred for glaucoma evaluation to a tertiary care clinic.
Patient characteristics and setting	**Sample size**: 295 eyes (68 with perimetric glaucoma, 62 with preperimetric glaucoma and 165 normal control eyes). **Age**: perimetric glaucoma median (IQR), 56 (48, 61) years; preperimetric glaucoma median (IQR), 54 (41, 62) years; controls, 54 (41, 63) years. **Sex**: not reported. **Ethnicity**: not reported. **Setting:** L V Prasad Eye Institute, Banjara Hills, Hyderabad, Andhra Pradesh. **Country**: India. **Ocular comorbidities**: patient with any media opacities that prevented good-quality optic disc photographs and other imaging tests, and any retinal (including macular) or neurologic disease other than glaucoma, were excluded. **Spectrum of glaucoma severity**: the median (IQR) mean deviation and PSD on the VF test were -9.1 (-14.8, -4.8) and 8.2 (3.7, 10.5) respectively, for perimetric glaucoma, -2.3 (-3.9, -0.9) and 1.8 (1.5, 2.2) respectively, for preperimetric glaucoma. **Control participants**: non-glaucomatous optic discs appearance and normal VF result.
Index tests	**Optical coherence tomography**: RTVue (Optovue Inc, Fremont, CA, USA), software version 5.1.0.90. GCC scanning protocol was used for imaging the macula. Only well-centred images with a signal strength index of ≥ 30 were used for analysis. The authors declared no conflict of interest.
Target condition and reference standard(s)	**Manifest perimetric glaucoma:** glaucomatous optic disc (defined as the presence of focal or diffuse neuroretinal rim thinning, localised notching, or nerve fibre layer defects and glaucomatous) and glaucomatous VF result (defined as the PSD < 5% and GHT outside normal limits). **Visual field test:** Humphrey Field Analyzer, model 750i (Zeiss Humphrey Systems, Dublin, CA, USA), with the SITA-standard programme. The VFs were considered reliable if the fixation losses, false-positive and false-negative response rates were < 20%. A single observer masked to the optic disc classification, SD-OCT findings and the other eye status, graded all VFs. **Optic disc evaluation**: stereoscopic optic disc photographs using digital fundus camera (FF 450 plus with VISUPAC 4.2.2; Carl Zeiss Meditec Systems GmbH, Pirmasens, Germany). Optic disc photograph was evaluated independently by 2 experts masked to the clinical details of the patients.
Flow and timing	42 eyes with unreliable VFs, 7 eyes with poor quality disc photographs and 18 eyes with poor OCT images quality, were excluded from the analysis. So, fewer than 10% of the patients enrolled were excluded. Index test and reference standard were performed on the same day.
Comparative	

Begum 2014b *(Continued)*

Notes	None.

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Unclear		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		Unclear risk	
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 2: Index Test (All tests)			
If a threshold was used, was it pre-specified?	Yes		
Were imaging test's quality assessed?	Yes		
Were any conflict of interest avoided	Yes		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern

Begum 2014b *(Continued)*

DOMAIN 4: Flow and Timing		
Was there an appropriate interval between index test and reference standard?	Yes	
Did all patients receive the same reference standard?	Yes	
Were all patients included in the analysis?	Yes	
Did all patients receive a reference standard	Yes	
Could the patient flow have introduced bias?		Low risk

Benitez-del-Castillo 2011

Study characteristics	
Patient Sampling	Healthy volunteers and patients with glaucoma who met the eligibility criteria were consecutively enrolled in this prospective, observational case-control study. Normal participants consisted of volunteers such as office employees and friends or family members of patients with glaucoma. Only one eye per person, selected randomly, was enrolled.
Patient characteristics and setting	**Sample size**: 117 patients enrolled, 88 eyes of 88 patients included in the analysis (33 glaucoma, 55 controls). **Age:** glaucoma patients mean ± SD, 63.8 ± 13.3 years; controls 59.1 ± 7.5. **Sex:** 45 men (23 glaucoma, 22 controls) and 43 women (10 glaucoma, 33 controls). **Country:** Spain. **Ocular comorbidities**: No ocular disease other than glaucoma or cataract, BCVA < 20/40, SE < -7 or > +3 D, neurologic disorders, retinal disease, or intraocular surgery except for uncomplicated cataract extraction. **Setting:** Glaucoma Unit, Hospital General del S.A.S. de Jerez. **Spectrum of glaucoma severity:** mean (95% CI) MD and PSD on the VF test were -6.69 (-8.07 to -5.31) dB and 6.22 (4.8 to 7.65) dB respectively. According to Hodapp et al. grading scale, 18 eyes had early disease and 15 eyes moderate. **Control participants:** IOP ≤ 21 mmHg, normal optic disc appearance and 2 normal SAP results (define as GHT within normal limits, MD and PSD with P > 5%).
Index tests	**Scanning laser polarimetry**: GDx-VCC and GDx-ECC, software version 5.5.0 (Carl Zeiss Meditec, Inc.). 3 consecutive scans were obtained with VCC and ECC on the same day by the same examiner, through undilated pupils. An average of the 3 measurements was used for the analysis. Images that were obtained during eye movement were excluded, as well as unfocused, poorly centred images or images with a quality scan score of < 8. **Optical coherence tomography**: Cirrus OCT, software version 3.0 (Carl Zeiss Meditec, Inc.). Test was performed through undilated pupils using a fast RNFL thickness acquisition proto-

Benitez-del-Castillo 2011 *(Continued)*

	col on the same day by the same examiner. The average of 2 measurements was used for the analysis. Images that were obtained during eye movement or were unfocused, were poorly centred, or had signal strength of < 7 were excluded. The authors stated no conflict of interest.
Target condition and reference standard(s)	**Manifest glaucoma:** glaucomatous optic nerve damage (defined as cup-to-disc asymmetry between fellow eyes of greater than 0.2, rim thinning, notching, excavation, and/or RNFL defect) and corresponding abnormal SAP result (GHT and PSD outside 95% of normal limits). **Visual field testing:** Humphrey Field Analyzer, 24-2 SITA-standard strategy (Carl-Zeiss Meditec, Inc.). VF with rate of fixation losses, false positives, and false negatives > 33% were considered unreliable. **Optic disc evaluation:** dilated fundus stereoscopic examination and photography.
Flow and timing	Index tests were performed on the same day, but no detail reported about reference standard's execution time. A total of 117 eyes were enrolled. 9 participants were not included in the control group: 4 for quality SLP-VCC scan < 8, 3 OCT signal strength < 7, and 2 for unreliable VF. 20 glaucoma patients were not included: 9 for quality SLP-VCC scan < 8, 6 OCT signal strength < 7 and 5 for unreliable VF.
Comparative	
Notes	None.

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	No		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (All tests)			
If a threshold was used, was it prespecified?	Yes		
Were imaging test's quality assessed?	Yes		
Were any conflict of interest avoided	Yes		

Benitez-del-Castillo 2011 *(Continued)*

Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Unclear		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	No		
Did all patients receive a reference standard	Yes		
Could the patient flow have introduced bias?		High risk	

Bertuzzi 2014

Study characteristics	
Patient Sampling	Case-control study in which patients attending the glaucoma clinic and healthy volunteers were enrolled between September 2009 and October 2010. One eye per person (randomly selected if both eligible) was considered.

Bertuzzi 2014 *(Continued)*

Patient characteristics and setting	**Sample size**: 205 eyes of 205 participants (70 glaucoma, 65 ocular hypertension, 70 normal controls). **Age**: perimetric glaucoma mean ± SD, 65.87 ± 11.90 years; controls, 56.80 ± 11.16 years. **Sex**: 69 men (38 glaucoma, 31 controls) and 71 women (32 glaucoma, 39 controls). **Ethnicity**: not reported. **Setting:** Glaucoma Service of Policlinico di Monza Hospital (University of Milan-Bicocca). **Country**: Italy. **Ocular comorbidities**: eyes with significant lens opacity, systemic diseases with ophthalmic involvement, co-existing retinal disease, uveitis, or non-glaucomatous optic neuropathy were excluded. **Spectrum of glaucoma severity**: the mean ± SD mean deviation and PSD on the VF test were -6.49 ± 6.46 and 6.39 ± 3.97 respectively, for glaucoma. **Control participants**: IOP of < 21 mmHg, no history of high IOP, and 2 reliable normal VFs (PSD and GHT within normal limits).
Index tests	**Optical coherence tomography**: RTVue (Optovue Inc.), software version 4.0.5.39. ONH and GCC scanning protocol were used for the analysis. Only good-quality images, defined as a signal strength index of Z50 without motion artefacts, were used for the analysis. The authors declare no conflict of interest.
Target condition and reference standard(s)	**Manifest perimetric glaucoma:** glaucomatous VF damage defined as PSD outside the 95% normal confidence limits or a GHT result outside the 99% normal confidence limits, in at least 2 consecutive and reliable VF examinations. **Visual field test:** automated perimetry model 750i (Carl Zeiss Meditec Inc.), with 24-2 SITA-algorithm. Tests were considered reliable only with fixation loss of < 30%, and false-positive and false-negative response rates of < 20%.
Flow and timing	No details were reported about patients exclusion or time interval between index and reference test.
Comparative	
Notes	None.

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Was a case-control design avoided?	No		
Did the study avoid inappropriate exclusions?	Yes		

Bertuzzi 2014 *(Continued)*

Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (All tests)			
If a threshold was used, was it pre-specified?	Yes		
Were imaging test's quality assessed?	Yes		
Were any conflict of interest avoided	Yes		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Unclear		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Unclear		
Did all patients receive a reference standard	Yes		
Could the patient flow have introduced bias?		Unclear risk	

Borque 2008

Study characteristics			

Borque 2008 *(Continued)*

Patient Sampling	Patients were chosen prospectively and consecutively from the outpatient clinics from January 2006 to December 2006. One eye per person was randomly selected.
Patient characteristics and setting	**Sample size**: 440 patients were assessed, 417 eyes of 417 patients were included in the analysis (71 perimetric glaucoma, 68 preperimetric glaucoma, 218 OHT, 60 healthy controls). **Age:** perimetric glaucoma patients mean ± SD, 64.83 ± 9.23 years; preperimetric glaucoma patients 59.57 ± 10.18 years; OHT patients 53.21 ± 12.01 years; controls 59.85 ± 10.78 years. **Ethnicity:** all participants were white. **Setting:** "Miguel Servet" University Hospital in Zaragoza **Country:** Spain. **Ocular comorbidities**: no history of eye surgery or serious trauma, systemic diseases with ophthalmic repercussions; BCVA ≥ 20/30, spherical refraction > ±5 D, cylinder refraction > ±3 D, transparent optic media. **Spectrum of glaucoma severity:** mean ± SD MD on the VF test were -6.10 ± 5.43 dB, for perimetric glaucoma eyes, -0.43 ± 1.30 dB, for preperimetric glaucoma; -0.26 ± 1.06 dB for OHT. **Control participants:** normal eye exam, IOP < 21 mmHg, normal morphology of the optic nerve and normal VF result.
Index tests	**Scanning laser polarimetry:** GDx-VCC, (version 5.4.1.35, Laser Diagnostic Technologies, Inc., San Diego, CA, USA). Images were taken under midriasis by experienced technicians. Tests were accepted only if of high quality (> 7), centred on the optic nerve, with images perfectly and uniformly focused and lighted with no movement artefacts. No details about author's conflict of interest were reported.
Target condition and reference standard(s)	**Manifest perimetric glaucoma:** glaucomatous optic nerve appearance (defined as neuroretinal rim thinning, focal or diffuse with an increase of the cup, the presence of notches, or both) and glaucomatous VF defects (defined as the presence of a group of at least 3 altered points with a P < 5% or a group(not near the blind spot) with at least 2 altered points with a P < 1% and/or SD from the mean with a P < 5% and/or GHT outside normal limits. **Preperimetric glaucoma:** IOP ≥ 21 mmHg, papillary morphology compatible with glaucoma and normal VF result. **Ocular hypertensive:** IOP ≥ 21 mmHg, normal papillary morphology and normal VF result. **Visual field testing**: Humphrey Field Analyzer, model 750, 24-2 SITA standard programme (Zeiss-Humphrey, Dublin, CA, USA). VF reliability criteria (false positives, false negatives and loss of focus) were considered to accept each test but the cut-off values considered were not specified. **Optic disc evaluation:** papillary stereophotographs by 2 glaucoma specialists unaware of the patient's medical history.
Flow and timing	11 patients were excluded due to poor-quality images, 5 did not sign the informed consent form and 7 did not attend all the appointments to complete the examination protocol.Therefore 23 patients (< 10%) were not included in the final analysis. The time interval between reference standard and index test was not reported.
Comparative	
Notes	None.

Borque 2008 *(Continued)*

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	No		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (All tests)			
If a threshold was used, was it prespecified?	Yes		
Were imaging test's quality assessed?	Yes		
Were any conflict of interest avoided	Unclear		
Could the conduct or interpretation of the index test have introduced bias?		Unclear risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			

Borque 2008 *(Continued)*

Was there an appropriate interval between index test and reference standard?	Unclear	
Did all patients receive the same reference standard?	Yes	
Were all patients included in the analysis?	Yes	
Did all patients receive a reference standard	Yes	
Could the patient flow have introduced bias?		Unclear risk

Bowd 2005

Study characteristics	
Patient Sampling	Participants were enrolled in the University of California, San Diego, Diagnostic Innovations in Glaucoma Study (DIGS). One randomly-selected eye from each patient was included in this observational cross-sectional study.
Patient characteristics and setting	**Sample size**: 164 eyes of 164 patients (92 glaucoma and 72 healthy controls). **Age:** glaucoma patients mean ± SD, 66.9 ± 8.9 years; controls 64.3 ± 8.8 years. **Country:** USA. **Ocular comorbidities**: no co-existing retinal disease, BCVA < 20/40, spherical refraction > ±5 D, cylinder refraction > ± 3D, uveitis, or non-glaucomatous optic neuropathy. **Setting:** University of California, San Diego. **Spectrum of glaucoma severity:** mean ± SD MD on the VF test was -5.32 ± 4.0 dB (range, -20.14 dB to -0.26 dB). According to Hodapp et al. grading scale, 54 patient had early, 24 had moderate and 14 had severe glaucoma. **Control participants:** healthy-appearing ONH on clinical examination, SAP results (MD, PSD, GHT) within normal limits, and no history of IOP > 22 mmHg.
Index tests	**Scanning laser polarimetry**: GDx VCC, software version 5.01 (Laser Diagnostic Technologies, San Diego, CA., USA). 2 machine learning classifiers were tested: the support vector machine and the relevance vector machine. Only well-focused, evenly illuminated, and centred scans with residual anterior segment retardation < 15.0 nm and atypical scan scores < 25, determined by GDx VCC software, were included. One author had financial disclosure.
Target condition and reference standard(s)	**Manifest glaucoma**: repeatable (2 consecutive) SAP results outside normal limits by PSD (P < 5%) or GHT. **Visual field testing:** Humphrey Field Analyzer, model II, 24-2 SITA-standard strategy (Carl Zeiss Meditec, Dublin, CA, USA).

Bowd 2005 *(Continued)*

	Optic disc appearance was not part of the reference standard.
Flow and timing	The first abnormal SAP was on or before the imaging date but no other information about time delay between tests was reported. No patients were excluded from the analysis.
Comparative	
Notes	None.

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Was a case-control design avoided?	No		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (All tests)			
If a threshold was used, was it pre-specified?	Yes		
Were imaging test's quality assessed?	Yes		
Were any conflict of interest avoided	No		
Could the conduct or interpretation of the index test have introduced bias?		High risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	

Bowd 2005 *(Continued)*

Are there concerns that the target condition as defined by the reference standard does not match the question?		Low concern
DOMAIN 4: Flow and Timing		
Was there an appropriate interval between index test and reference standard?	Unclear	
Did all patients receive the same reference standard?	Yes	
Were all patients included in the analysis?	Yes	
Did all patients receive a reference standard	Yes	
Could the patient flow have introduced bias?		Unclear risk

Bozkurt 2010

Study characteristics	
Patient Sampling	Healthy and glaucoma patients were enrolled prospectively. Normal eyes were consecutively recruited from patients referred for refraction who underwent routine examination or from hospital staff. No further details about glaucoma patients enrolment. One eye per person was randomly selected.
Patient characteristics and setting	**Sample size**: 342 participants were enrolled (158 glaucoma and 184 healthy controls). **Age:** glaucoma patients mean ± SD, 63.0 ± 10.7 years, controls 59.6 ± 9.7 years. **Sex:** 121 men (60 glaucoma, 61 controls) and 221 women (98 glaucoma, 123 controls) **Ethnicity:** Turkish. **Country:** Turkey. **Ocular comorbidities:** BCVA ≥ 20/40, refractive error of < 5 spherical dioptres and 2 D of cylinder and transparent ocular media. No parapapillary atrophy, tilted discs or indistinct disc borders. **Setting:** Hacettepe University School of Medicine. **Spectrum of glaucoma severity:** mean ± SD MD and PSD on the VF test were -6.97 ± 4.98 dB and 4.28 ± 3.33 dB, respectively. **Control participants:** IOP < 20 mmHg, ONH appearance no suspicious for glaucoma and normal SAP.
Index tests	**Confocal scanning laser ophthalmoscopy:** Heidelberg Retina Tomograph (HRT; Heidelberg Engineering GmbH, Heidelberg, Germany). ONH topography (through undilated pupils) and contour line drawing were performed by the same experienced operator using HRT II, with HRT III software version 3.0. Good image quality was defined as follows: acquisition sensitivity < 90%; topography SD < 35 mm; > 75% of the disc within the target circle; minimal movement during the acquisition movie; no floaters over the disc. No details about authors' conflict of interest were reported.

Bozkurt 2010 *(Continued)*

Target condition and reference standard(s)	**Manifest perimetric glaucoma:** ONH or RNFL structural abnormalities (diffuse thinning, focal narrowing or notching of the optic disc rim; documented progression of cupping of the optic disc; diffuse or localised abnormalities of the peripapillary RNFL; disc rim or peripapillary RNFL haemorrhages; neural rim asymmetry between the 2 eyes consistent with loss of neural tissue) and/or VF result abnormalities (defined as a cluster of 3 points with P < 5%, a cluster of 2 points with P < 1% on pattern deviation probability plots, or a PSD with P < 5% or GHT outside normal limits. **Visual field testing**: Humphrey Field Analyzer II, 30-2 SITA standard programme (Carl Zeiss-Humphrey, Dublin, CA, USA). VF reliability criteria included fixation losses rates, false-positive and false-negative rates < 25%. **Optic disc evaluation:** no details reported.
Flow and timing	No patients were reported by the authors as excluded from the analysis. All tests and imaging were carried out within a 2-week period.
Comparative	
Notes	None.

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Was a case-control design avoided?	No		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (All tests)			
If a threshold was used, was it pre-specified?	Yes		
Were imaging test's quality assessed?	Yes		
Were any conflict of interest avoided	Unclear		
Could the conduct or interpretation of the index test have introduced bias?		Unclear risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern

Bozkurt 2010 *(Continued)*

DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
Did all patients receive a reference standard	Yes		
Could the patient flow have introduced bias?		Low risk	

Brusini 2005

Study characteristics	
Patient Sampling	Consecutive patients with early-to-moderate primary open-angle glaucoma and controls were considered. One eye per person was selected.
Patient characteristics and setting	**Sample size**: 80 eyes of 80 patients (40 glaucoma and 40 healthy controls). **Age:** glaucoma patients mean ± SD, 65.8.9 ± 8.5 years; controls 57 ± 7.8 years. **Country**: not specified. **Ocular comorbidities**: no ocular pathologies other than glaucoma, BCVA < 32/40, SE > ±5 D, mild nuclear sclerosis, drusen, large peripapillary atrophy, previous intraocular surgery, diabetes mellitus, or neurologic disorders. **Setting**: not specified. **Spectrum of glaucoma severity:** mean ± SD MD and PSD on the VF test were -3.1 ± 1.6 dB and 3.1 ± 0.9 dB. Patient with SAP test result having a MD > -9 dB and a PSD < 8 dB were included. **Control participants:** normal IOP and normal SAP results.

Brusini 2005 *(Continued)*

Index tests	**Scanning laser polarimetry**: GDx-FCC (Nerve Fibre Analyzer, version 2.0.09, Laser Diagnostic Technologies, Inc., San Diego, CA, USA) and GDx-VCC (software version 5.1.0, Laser Diagnostic Technologies, Inc. San Diego, CA, USA). According to the GDx-normative database, values labelled as outside normal limits and the Number > 70 were considered abnormal. A new cut-off point was determined for each GDx parameter. The inclusion criteria included a good SLP image quality. No details about authors' conflict of interest were reported.
Target condition and reference standard(s)	**Manifest glaucoma:** IOP > 21 mmHg before treatment and reproducible SAP glaucomatous defects (defined as at least 1 of the following: a cluster of > 3 points in the pattern deviation probability plot, located in areas that are typical of glaucoma, having a probability level of < 5%, with at least 1 point having a probability level of < 1%; PSD probability level of < 5%; GHT outside normal limits). **Visual field testing:** Humphrey Field Analyzer, model II 750, 30-2 SITA-standard strategy (Carl Zeiss Meditec Inc., Dublin, CA, USA). Reliable criteria for VF tests included false-positive and false-negative responses of < 33% and fixation losses of < 20%. Optic disc appearance was not part of the reference standard.
Flow and timing	Reference standard and index test were performed within 3 months. All patients were included in the analysis.
Comparative	
Notes	None.

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	No		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (All tests)			
If a threshold was used, was it pre-specified?	Yes		
Were imaging test's quality assessed?	Yes		
Were any conflict of interest avoided	Unclear		

Brusini 2005 *(Continued)*

Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	No		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
Did all patients receive a reference standard	Yes		
Could the patient flow have introduced bias?		High risk	

Brusini 2006a

Study characteristics	
Patient Sampling	Glaucoma patients were recruited from those under the care of the Glaucoma Service of the Department of Ophthalmology. Normal participants were recruited from staff members and volunteers. One eye per person was randomly selected.
Patient characteristics and setting	**Sample size**: 157 eyes of 157 participants (95 glaucoma and 62 healthy controls). **Age:** glaucoma patients mean ± SD, 71 ± 10 years; controls 66 ± 9.9. **Country:** Italy. **Ocular comorbidities**: no ocular pathologies other than glaucoma, BCVA < 0.7, SE > ±5 D, papillary anomalies, large peripapillary atrophy, previous intraocular surgery, diabetes, or neurological disorders. **Setting:** Glaucoma Service, Department of Ophthalmology at the Santa Maria della Misericordia Hospital, Udine.

Brusini 2006a *(Continued)*

	Spectrum of glaucoma severity: mean ± SD MD and PSD on the VF test were -3.7 ±3 dB and 4.5 ± 2.7 dB. According to the GSS, 45 eyes as stage 1 (Md > -5.0 dB) and 41 as stage 2 (MD range -5.0 to -9.0). **Control participants:** normal IOP, normal ONH/RNFL appearance (no diffuse or focal rim thinning, cupping, optic disc haemorrhage or RNFL defects), and normal SAP results (MD and PSD within 95% CI, and a GHT within normal limits).
Index tests	**Scanning laser polarimetry**: GDx-VCC, software version 5.1.0 (Laser Diagnostic Technologies, Inc. San Diego, CA, USA).The mean of 3 measurements was used. All images with quality score gradings < 8 were excluded. No details about authors' conflict of interest were reported.
Target condition and reference standard(s)	**Manifest glaucoma:** IOP > 21 mmHg before treatment and reproducible SAP glaucomatous defects (defined as at least 1 of the following: a cluster of > 3 points in the pattern deviation probability plot, located in areas that are typical of glaucoma, having a probability level of < 5%, with at least 1 point having a probability level of < 1%; PSD probability level of < 5%; GHT outside normal limits). **Visual field testing:** Humphrey Field Analyzer, model II 750, 30-2 SITA-standard strategy (Carl Zeiss Meditec Inc., Dublin, CA, USA). Reliable criteria for VF tests included false-positive and false-negative responses of < 33% and fixation losses of < 20%. Optic disc appearance was not part of the reference standard.
Flow and timing	Reference standard and index test were conducted within a period of 3 months. All patients were included in the analysis.
Comparative	
Notes	None.

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Was a case-control design avoided?	No		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (All tests)			
If a threshold was used, was it pre-specified?	Yes		

Brusini 2006a *(Continued)*

Were imaging test's quality assessed?	Yes		
Were any conflict of interest avoided	Unclear		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	No		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
Did all patients receive a reference standard	Yes		
Could the patient flow have introduced bias?		High risk	

Calvo 2014

Study characteristics	
Patient Sampling	Case-control study, in which patients with glaucoma were recruited consecutively from an ongoing longitudinal follow-up study at the Miguel Servet University Hospital, and normal eyes were consecutively recruited from patients referred for refraction that underwent routine examination, hospital staff, and relatives of patients. One eye per person was randomly selected.
Patient characteristics and setting	**ample size**: 338 eyes of 338 participants (156 glaucoma and 182 controls). **Age**: glaucoma mean ± SD, 61.05 ± 9.4 years; controls, 59.55 ± 9.7 years.

Calvo 2014 *(Continued)*

	Sex: 125 men (68 glaucoma, 57 controls) and 213 women (88 glaucoma, 125 controls) **Ethnicity**: white. **Clinical Setting:** Miguel Servet University Hospital, Zaragoza. **Country**: Spain. **Ocular comorbidities**: patient with previous intraocular surgery, diabetes or other systemic diseases, history of ocular or neurologic disease, or current use of a medication that could affect VF sensitivity were excluded. **Spectrum of glaucoma severity**: the mean ± SD MD and PSD on the VF test were -6.64 ± 6.0 and 6.03 ± 3.8 respectively, for glaucomatous eyes. **Control participants**: no specific details reported.
Index tests	**Optical coherence tomography**: Cirrus SD-OCT (Carl Zeiss Meditec Inc., Dublin, CA, USA), software version 6.2. Optic disc cube 200 x 200 scan protocol was used for the analysis. All images had to have a quality > 6. **Confocal scanning laser ophthalmoscopy**: HRT III (Heidelberg Engineering, Heidelberg, Germany). The margin of the optic disc was manually traced by the same glaucoma specialist, masked to the patients' identity and clinical history. All scans had to have an interscan SD < 30 µm. No conflict of interest were reported
Target condition and reference standard(s)	**Manifest glaucoma:** IOP measurements > 21 mmHg and a glaucomatous VF defect, defined as a PSD with P < 0.5% and GHt outside normal limits. No details about ophthalmic characteristics of controls. **Visual field test:** Humphrey Field Analyzer model 750i (Carl Zeiss Meditec, Dublin, CA, USA); 24-2 SITA–standard strategy.
Flow and timing	12 patients (< 10%) were excluded from the analysis: 4 with no reliable standard automated perimetry after 3 attempts and 8 which did not complete the visits included in the study protocol. All exams were performed within 6 weeks of the person's date of enrolment into the study.
Comparative	
Notes	None.

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	No		
Did the study avoid inappropriate exclusions?	Yes		

Calvo 2014 *(Continued)*

Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (All tests)			
If a threshold was used, was it pre-specified?	Yes		
Were imaging test's quality assessed?	Yes		
Were any conflict of interest avoided	Yes		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Unclear		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Unclear risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	No		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
Did all patients receive a reference standard	Yes		
Could the patient flow have introduced bias?		High risk	

Chen 2007

Study characteristics			

Chen 2007 *(Continued)*

Patient Sampling	Healthy controls, early glaucoma patients and glaucoma suspects were prospectively enrolled. Control participants were volunteers from the staff or their family members at the China Medical University Hospital. No details to assess the number of eyes for each person.
Patient characteristics and setting	**Sample size**: 210 eyes were enrolled, 189 actually included in the analysis (82 early glaucoma, 45 glaucoma suspects and 62 controls). **Age:** early glaucoma patients mean ± SD, 48.55 ± 15.36 years, glaucoma suspects 44.2 ± 15.97, controls 44.7 ± 12.55 years. **Sex**: 89 men (41 glaucoma, 19 suspects, 29 controls) and 100 women (41 glaucoma, 26 suspects, 33 controls). **Ethnicity:** Taiwan Chinese population. **Country:** Taiwan. **Ocular comorbidities:** BCVA < 20/40, a spherical equivalent outside ±5.0 D, and a cylinder correction > 3.0 D were excluded. **Setting:** China Medical University Hospital (Taiwan). **Spectrum of glaucoma severity:** mean ± SD MD and PSD on the VF test were -3.32 ± 2.2 dB and 3.09 ± 2.2 dB, respectively for early glaucoma patients; -2.43 ± 2.16 dB and 2.45 ± 1.6 dB, respectively for glaucoma suspects. **Control participants:** IOP < 21 mmHg, open angle on gonioscopy, normal optic disc appearance and normal VF result (GHT and CPSD within normal limits).
Index tests	**Scanning Laser polarimetry:** GDx-VCC, software 5.5.0 (Carl Zeiss Meditec inc.) The exams were performed by the same experienced technician, through undilated pupils. All images had to be well focused, with centred optic disc, without any motion artefact and a minimum score of 8. No details about authors' conflict of interest were reported.
Target condition and reference standard(s)	**Manifest perimetric glaucoma:** glaucomatous optic nerve appearance (defined as notching or thinning of the neuroretinal rim) and glaucomatous corresponding VF defects (defined by 2 or more contiguous points with a pattern deviation sensitivity loss of P < 0.01, or 3 or more contiguous points with sensitivity loss of P < 0.05 in the superior or inferior arcuate areas, or a 10-dB difference across the nasal horizontal midline at 2 or more adjacent locations and an abnormal result on the GHT), and open angle by gonioscopy. All patients had VF MD > -6 dB. **Glaucoma suspects**: abnormal disc consistent with glaucoma with a normal VF test. **Visual field testing**: Humphrey Field Analyzer, model 750 II, full-threshold automated perimetry, 30–2 mode (Carl Zeiss-Humphrey, Dublin, CA, USA). VF reliability criteria included fixation losses rates, false-positive and false-negative rates < 20%. **Optic disc evaluation:** stereoscopic fundus examination.
Flow and timing	21 eyes (< 10%) enrolled were excluded from the analysis because good images could not be obtained. All tests and imaging were carried out within 4 weeks.
Comparative	
Notes	None.
Methodological quality	

Chen 2007 *(Continued)*

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Was a case-control design avoided?	No		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (All tests)			
If a threshold was used, was it pre-specified?	Yes		
Were imaging test's quality assessed?	Yes		
Were any conflict of interest avoided	Unclear		
Could the conduct or interpretation of the index test have introduced bias?		Unclear risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			

Chen 2007 *(Continued)*

Was there an appropriate interval between index test and reference standard?	Yes	
Did all patients receive the same reference standard?	Yes	
Were all patients included in the analysis?	Yes	
Did all patients receive a reference standard	Yes	
Could the patient flow have introduced bias?		Low risk

Chen 2008

Study characteristics	
Patient Sampling	Prospective cross-sectional study including early-to-moderate glaucomatous eyes (high-tension primary open angle glaucoma and primary angle closure glaucoma) and age-matched participants. The glaucoma patients were followed for at least 6 months between December 2004 and August 2005. Participantsts with normal eyes were volunteers from the staff or family members at the China Medical University Hospital. One eye per person was randomly selected.
Patient characteristics and setting	**Sample size**: 88 eyes of 88 glaucoma patients (47 POAG and 41 PACG); 45 eyes from 45 normal participants. **Age:** glaucoma patients mean ± SD, 61.7 ± 9.9 years for POAG and 61.8 ± 8.5 years for PACG; controls 57.9 ± 9.0 years. **Sex:** 60 men (22 controls, 31 POAG and 7 PACG), and 71 women (21 controls, 16 POAG and 34 PACG). **Ethnicity:** Taiwan Chinese. **Country:** China. **Ocular comorbidities**: no peripapillary atrophy, BCVA < 20/40, SE > ±5 D or secondary angle closure, such as lens-induced glaucoma, neovascular glaucoma, or uveitis. **Setting:** Glaucoma Service, China Medical University Hospital. **Spectrum of glaucoma severity:** mean ± SD MD on the VF test was -4.54 ± 5.43 dB for POAG eyes and -4.62 ± 3.99 dB for PACG eyes. Patients with VF results < -15 dB were excluded. **Control participants:** IOP < 21 mmHg, open angle on gonioscopy, normal optic disc appearance and normal VF result (GHT and CPSD within normal limits).
Index tests	**Scanning Laser polarimetry:** GDx VCC (Carl Zeiss Meditec, Inc, Dublin, CA, USA; version 5.5.0). Measurements were obtained by the same trained and experienced technician. All images had to be of high quality, with a score > 7, a centred optic disc, well-focused, even and just illuminated through the images, and without any motion artefact. Each patient could undergo multiple GDx VCC scans. Only 1 successful scan was saved into the hard disc and was printed out. All of the print-outs were evaluated by the same doctor. None of the authors had conflict of interest.

Chen 2008 *(Continued)*

Target condition and reference standard(s)	**Manifest primary open angle glaucoma:** glaucomatous optic neuropathy (defined as either cup/disc asymmetry between fellow eyes of > 0.2, rim thinning, notching, excavation, or RNFL defect), VF defects (defined as 2 or more contiguous points with a pattern deviation sensitivity loss of P < 0.01, or 3 or more contiguous points with P < 0.05 in the superior or inferior arcuate areas, or a 10-dB difference across the nasal horizontal midline at 2 or more adjacent locations and an abnormal GHT result), open angle on gonioscopy, and initial IOP > 21 mmHg. **Primary angle closure glaucoma:** glaucomatous optic neuropathy with corresponding VF loss associated with gonioscopic finding of at least 180° of peripheral anterior synechiae, and IOP > 21 mmHg on 2 separate occasions. **Visual field testing:** Humphrey Field Analyzer, model 750, 30-2 central full threshold strategy (Carl Zeiss Meditec, Inc). VF reliability criteria included fixation losses and false-positive and false-negative rates of < 20%. **Optic disc evaluation:** stereoscopic fundus examination.
Flow and timing	Time interval between reference standard and index test was < 4 weeks. Some patients without good GDx VCC imaging data were excluded.
Comparative	
Notes	None.

Methodological quality

Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Was a case-control design avoided?	No		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (All tests)			
If a threshold was used, was it prespecified?	No		
Were imaging test's quality assessed?	Yes		
Were any conflict of interest avoided	Yes		

Chen 2008 *(Continued)*

Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Unclear		
Did all patients receive a reference standard	Yes		
Could the patient flow have introduced bias?		Unclear risk	

Chen 2013

Study characteristics	
Patient Sampling	Prospective, case-control study. Glaucoma patients had received regular treatment or follow-up care at the Glaucoma department whereas the normal controls were volunteers recruited from the staff and their families. 1 eye per person was randomly chosen.

Chen 2013 *(Continued)*

Patient characteristics and setting	**Sample size**: 161 eyes of 161 participants (35 POAG, 26 PACG, 27 glaucoma suspects, 21 ocular hypertension and 52 controls). **Age**: glaucoma mean ± SD, 44.71 ± 13.69 years; PACG, 64.81 ± 6.81 years; glaucoma suspects, 34.56 ± 16.46 years; ocular hypertension, 30.0 ± 13.8 years controls, 35.27 ± 15.29 years. **Sex**: no details reported **Ethnicity**: Chinese **Clinical Setting:** Glaucoma Service of the Department of Ophthalmology at China Medical University Hospital. **Country**: China **Ocular comorbidities**: patients with a BCVA < 20/40, a spherical equivalent > ±5.0 D, or a cylinder correction > 3.0 D, or with co-existing retinal disease, uveitis, or non-glaucomatous optic neuropathy were excluded **Spectrum of glaucoma severity**: the mean ± SD mean deviation and PSD on the VF test were -5.47 ± 7.99 and 4.82 ± 7.31 respectively, for POAG eyes; -4.87 ± 5.65 and 5.21 ± 3.92 respectively, for PACG eyes; -1.85 ± 1.44 and 2.12 ± 1.18 respectively, for glaucomatous-suspected eyes. **Control participants**: IOP < 21 mmHg, no history of increased IOP, normal-looking optic disc heads, and normal VF results (MD and PSD with P > 5% and GHT within normal limits).
Index tests	**Optical coherence tomography**: Cirrus SD-OCT (software version 3.0; Carl Zeiss Meditec Inc.). Optic disc cube 200 x 200 scan protocol was used for the analysis. All images had to have focused ocular fundus images, a centred circular ring around the optic disc and a signal strength > 5. The authors declare no conflicts of interest.
Target condition and reference standard(s)	**Manifest primary open angle glaucoma:** IOP > 21 mmHg, open angle on gonioscopy, glaucomatous optic disc appearance (defined as > 0.2 cup/disc asymmetry between the eyes, rim thinning, notching, excavation, or RNFL defect) and a reproducible glaucomatous VF defect (defined as ≥ 2 contiguous points with a pattern deviation with P < 0.01, ≥ 3 contiguous points with a sensitivity loss of P < 0.05 in the superior or inferior arcuate areas, or a 10-dB difference across the nasal horizontal midline at ≥ 2 adjacent locations and GHT outside normal limits). **Manifest primary angle closure glaucoma**: a gonioscopic finding with at least 180° of peripheral anterior synechiae, IOP > 21 mmHg and glaucomatous optic disc appearance. **Glaucoma suspects**: abnormal disc appearance consistent with glaucoma along with a normal VF result. **Visual field test:** Humphrey Field Analyzer model 750 (Carl Zeiss Meditec, Dublin, CA, USA); 30-2 SITA–standard strategy. All exams had fixation losses and false-positive and false-negative rates of < 20%.
Flow and timing	No details about exclusion were reported. The time interval between index and reference test was not reported.
Comparative	
Notes	None.
Methodological quality	

Chen 2013 *(Continued)*

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Was a case-control design avoided?	No		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (All tests)			
If a threshold was used, was it prespecified?	Yes		
Were imaging test's quality assessed?	Yes		
Were any conflict of interest avoided	Yes		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	No		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			

Chen 2013 *(Continued)*

Was there an appropriate interval between index test and reference standard?	Unclear
Did all patients receive the same reference standard?	Yes
Were all patients included in the analysis?	Yes
Did all patients receive a reference standard	Yes
Could the patient flow have introduced bias?	Unclear risk

Cho 2011

Study characteristics	
Patient Sampling	Glaucoma patients were recruited prospectively, in a consecutive manner between August 2008 and February 2009. Age-matched healthy eyes formed the control group. One eye per person was randomly selected.
Patient characteristics and setting	**Sample size**: 108 eyes initially enrolled, 92 actually included in the analysis (49 glaucoma, 43 healthy controls). **Age:** glaucoma patients mean ± SD, 51.8 ± 14.2 years, controls 46.6 ± 16.3 years. **Ethnicity:** Asian. **Country:** South Korea. **Ocular comorbidities**: no ophthalmic disease that could affect VF result, no history of diabetes mellitus; BCVA ≥ 20/30, with a spherical equivalent within ± 5 D and a cylinder correction within +3 D. **Setting:** Asan Medical Center (Seoul, Korea). **Spectrum of glaucoma severity:** mean ± SD MD and PSD on the VF test were -6.39 ± 6.03 dB and 6.38 ± 4.69 dB, respectively. **Control participants:** IOP < 22 mmHg, no history of IOP elevation, and normal based on VF examination.
Index tests	**Optical coherence tomography:** SD-SLO/OCT (OTI, Opkos. Toronto, Canada). No details about author's conflict of interest were reported.
Target condition and reference standard(s)	**Manifest glaucoma:** glaucomatous optic disc appearance (defined as vertical cup/disc ratio of > 0.6, a difference in vertical cup-disc ratio of more than 0.2 between the eyes, diffuse or focal neural rim thinning, haemorrhage, or nerve fibre layer defects) and a glaucomatous VF defect (defined as a cluster of 3 points with P < 5% on the pattern deviation map in at least 1 hemifield, including at least 1 point with a P < 1%; or a cluster of 2 points with a < 1% and a GHT result outside normal limits; or a PSD outside 95% of the normal limits).

Cho 2011 *(Continued)*

	Visual field testing: Humphrey Field Analyzer, SITA standard, 24-2 programme (Carl Zeiss Meditec, Inc., Dublin, CA, USA). VF reliability criteria included false-positive and false-negative rates < 15%, and a fixation loss < 20%. **Optic disc evaluation:** stereoscopic optic nerve photography**.**
Flow and timing	16 subjects (> 10%) were excluded from the analysis due to poor image quality. The time interval between index and reference standard was not specified.
Comparative	
Notes	None.

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	No		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (All tests)			
If a threshold was used, was it pre-specified?	Yes		
Were imaging test's quality assessed?	Yes		
Were any conflict of interest avoided	Unclear		
Could the conduct or interpretation of the index test have introduced bias?		Unclear risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		

Cho 2011 *(Continued)*

Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Unclear		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	No		
Did all patients receive a reference standard	Yes		
Could the patient flow have introduced bias?		High risk	

Choi 2013

Study characteristics	
Patient Sampling	Participants were consecutively enrolled from October 2011 to April 2012. Healthy controls were enrolled among people undergoing routine eye examination. One eye per person was randomly selected.
Patient characteristics and setting	**Sample size**: 207 patients examined, 181 eventually included in the analysis. The patients were divided into 2 groups: a highly myopic group (spherical equivalent > -6.00 D and > -20.00 D) and a non-highly myopic group (spherical equivalent > -6.00 D and < -0.25 D): 71 highly myopic patients (49 glaucoma, 22 controls) and 110 non-highly myopic (54 glaucoma, 56 controls). **Age**: glaucoma highly myopic eyes mean ± SD, 46.57 ± 11.37 years; highly myopic controls 44.05 ± 15.14 years; glaucoma non-highly myopic eyes mean ± SD, 53.85 ± 12.52 years; non-highly myopic controls 49.27 ± 13.42 years. **Sex**: 97 men (61 glaucoma, 36 controls) and 84 women (42 glaucoma, 42 controls). **Ethnicity**: Korean. **Country**: Korea. **Setting**: Glaucoma Clinic of Seoul National University Hospital, Seoul. **Ocular comorbidities**: eyes with retinal pathology, diabetes, BCVA < 20/40 or non-glaucomatous optic nerve diseases, and eyes with previous laser therapy or ocular surgery, were excluded. **Spectrum of glaucoma severity**: the mean ± SD MD and PSD on the VF test were -7.44 ± 4.85 dB and 8.90 ± 4.73 dB respectively for glaucoma highly myopic eyes; were -7.31 ± 6.64 dB and 9.00 ± 4.36 dB respectively for glaucoma non-highly myopic eyes. **Control participants**: IOP < 22 mmHg, normal appearance of ONH and normal VF test.

Choi 2013 *(Continued)*

Index tests	**Optical coherence tomography:** Cirrus HD-OCT (Carl Zeiss Meditec, Dublin, CA, USA). The Macular cube 200 x 200 and 1 optic disc cube 200 x 200 scans were acquired through dilated pupil. Images with a signal strength < 6, visible eye motion, blinking artefacts, or algorithm segmentation failure were excluded. No authors had conflict of interest.
Target condition and reference standard(s)	**Manifest glaucoma:** glaucomatous optic disc change (defined as a large cupping (> 0.7 vertical cup/disc ratio), cup/disc asymmetry between the glaucomatous and normal eyes greater than 0.2, neuroretinal rim thinning, notching, or excavation) and glaucomatous VF defects (defined as GHT outside normal limits; a PSD with P < 0.05; a cluster of 3+ non-edge contiguous points in the pattern deviation plot in the same hemifield with P < 0.05, including 1+ with P < 0.01). **Visual field testing:** Humphrey Field Analyzer, 30-2 SITA standard programme (Carl Zeiss Meditec). VF reliability criteria were fixation losses < 20%, and false-positive and false-negative rates < 15%. **Optic disc evaluation**: stereoscopic colour disc photography, assessed by 2 glaucoma specialists in a masked fashion without knowing clinical data or OCT results.
Flow and timing	The time interval between index tests and reference standard was not reported. 26/207 (> 10%) eyes were excluded: 10 eyes were excluded from the study because of retinal disease, as well as 4 eyes due to optic nerve disease. 11 eyes were excluded owing to unreliable VF, and 1 eye due to poor OCT signal strength.
Comparative	
Notes	None.

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	No		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (All tests)			
If a threshold was used, was it prespecified?	Yes		
Were imaging test's quality assessed?	Yes		

Choi 2013 *(Continued)*

Were any conflict of interest avoided	Yes		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Unclear		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	No		
Did all patients receive a reference standard	Yes		
Could the patient flow have introduced bias?		High risk	

Da Pozzo 2005

Study characteristics	
Patient Sampling	Patients were selected among those referred to the Glaucoma Unit at Trieste University Eye Clinic between January and July 2004 for periodical scheduled visits. Healthy participants were recruited among staff members, friends or spouses of patients, or normal volunteers. One eye per person was randomly selected for inclusion.

Da Pozzo 2005 *(Continued)*

Patient characteristics and setting	**Sample size**: 141 eyes initially enrolled, 124 eyes of 124 participants included in the analysis (59 glaucoma and 65 healthy controls). **Age:** glaucoma patients mean ± SD, 67.1 ± 9.1 years; controls 64.6 ± 7.5. **Country:** Italy. **Ocular comorbidities**: no corneal or lens opacity, BCVA < 20/40, SE > ± 4 D, peripapillary atrophy falling under ellipse measurement, tilted disc, uveitis, significant vitreous floaters, or diffuse/localised retinal or macular disease. **Setting:** Glaucoma Unit, Trieste University Eye Clinic. **Spectrum of glaucoma severity:** mean ± SD MD and PSD on the VF test were −7.66 ± 6.19 dB and 7.46 ± 4.18 dB respectively. **Control participants:** normal VF result (MD and PSD within 95% confidence limits, GHT within normal limit), IOP < 21 mmHg, and healthy optic disc with intact neuroretinal rim.
Index tests	**Scanning Laser polarimetry:** GDx-VCC, software 5.3.4 (Laser Diagnostic Technologies, San Diego, California, USA). Scans with evidence of atypical pattern on the thickness map or a quality score < 8 as automatically provided by device software, were excluded from the study. None of the authors had conflict of interest.
Target condition and reference standard(s)	**Manifest glaucoma:** glaucomatous optic disc appearance(cupping, rim notching, or diffuse thinning) and reproducible VF defects (defined as GHT outside normal limits or PSD with P < 5%). **Visual field testing:** Humphrey Field Analyzer, 24-2 SITA-standard strategy (Humphrey Systems, Dublin, CA, USA). VF reliability criteria included fixation losses and false-positive and false-negative rates of < 20%. **Optic disc evaluation:** stereo biomicroscopy with the aid of a +90 D lens after pupil dilation.
Flow and timing	Time interval between reference standard and index test was within 2 months. 17 patients were excluded for poor imaging quality: 11 presented atypical patterns on the retardation map, 4 did not pass the 4-scan quality check or saw their RNFL readings flagged as “incompatible with normative database,” and 2 had poor fixation.
Comparative	
Notes	None.

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Was a case-control design avoided?	No		
Did the study avoid inappropriate exclusions?	Yes		

Da Pozzo 2005 *(Continued)*

Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (All tests)			
If a threshold was used, was it pre-specified?	Yes		
Were imaging test's quality assessed?	Yes		
Were any conflict of interest avoided	Yes		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	No		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	No		
Did all patients receive a reference standard	Yes		
Could the patient flow have introduced bias?		High risk	

Da Pozzo 2006

Study characteristics	
Patient Sampling	Patients were selected among those referred to the Glaucoma Unit at Trieste University Eye Clinic between January and October 2004 for periodic scheduled visit. Healthy participants were recruited among staff members, friends or spouses of patients, or normal volunteers. One eye per person was randomly selected.
Patient characteristics and setting	**Sample size**: 110 eyes of 110 participants (48 glaucoma and 62 healthy controls). **Age:** glaucoma patients mean ± SD, 66.8 ± 8.8 years; controls 64.7 ± 6.5 years. **Country:** Italy. **Ocular comorbidities**: no corneal or lens opacity, BCVA < 20/40, SE > ± 4 D, peripapillary atrophy falling under ellipse measurement, tilted disc, uveitis, significant vitreous floaters, or diffuse/localised retinal or macular disease. **Setting:** Glaucoma Unit, Trieste University Eye Clinic. **Spectrum of glaucoma severity:** mean ± SD MD and PSD on the VF test were -1.74 ± 1.69 dB and 3.56 ± 1.5 dB. **Control participants:** normal VF result (MD and PSD within 95% confidence limits, GHT within normal limit), IOP < 21 mmHg, and healthy optic disc with intact neuroretinal rim.
Index tests	**Scanning Laser polarimetry:** GDx VCC (software 5.3.4; Carl Zeiss Meditec, CA, USA). The correct positioning of ellipse on inner margin of peripapillary scleral ring was rechecked on all eyes by a trained technician. Scans with evidence of atypical pattern on the printout retardation map or a score < 7 on the 4-scan quality checks performed by software (alignment, fixation, refraction and illumination) were excluded. No details about authors' conflict of interest were reported.
Target condition and reference standard(s)	**Manifest glaucoma:** glaucomatous optic disc appearance(cupping, rim notching, or diffuse thinning) and reproducible VF defects (defined as GHT outside normal limits or PSD with P < 5%). **Visual field testing:** Humphrey Field Analyzer, 24-2 SITA-standard strategy (Humphrey Systems, Dublin, CA, USA). VF reliability criteria included fixation losses and false-positive and false-negative rates of < 20%. **Optic disc evaluation:** stereo biomicroscopy with the aid of a +90 D lens after pupil dilation.
Flow and timing	Time interval between reference standard and index test was within 2 months. 14 patients were excluded for poor imaging quality: 6 presented atypical patterns on the retardation map, 2 did not pass the 4-scan quality check, 3 saw their RNFL readings flagged as “incompatible with normative database” and 3 had poor fixation.
Comparative	
Notes	None.

Methodological quality			
Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			

Da Pozzo 2006 *(Continued)*

Was a consecutive or random sample of patients enrolled?	Unclear		
Was a case-control design avoided?	No		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (All tests)			
If a threshold was used, was it pre-specified?	Yes		
Were imaging test's quality assessed?	Yes		
Were any conflict of interest avoided	Unclear		
Could the conduct or interpretation of the index test have introduced bias?		Unclear risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	No		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	No		

Da Pozzo 2006 *(Continued)*

Did all patients receive a reference standard	Yes
Could the patient flow have introduced bias?	High risk

De Leon-Ortega 2006

Study characteristics	
Patient Sampling	Data were obtained from patients who had undergone optic disc imaging and visual functional testing between January 2003 and February 2005 as part of ongoing longitudinal glaucoma studies. Controls were obtained primarily from referrals and University of Alabama employees. One eye per person was randomly selected.
Patient characteristics and setting	**Sample size**: 228 eyes of 228 participants (79 glaucoma and 149 healthy controls). **Age:** glaucoma patients mean ± SD, 56.0 ± 13.9 years; controls 40.3 ± 11.3 years. **Sex:** 63 men (25 glaucoma and 38 controls) and 165 women (54 glaucoma and 111 controls). **Ethnicity:** 42 of 79 in the glaucoma group and 82 of 149 in the controls were African-American. **Country:** USA. **Ocular comorbidities**: no BCVA < 20/40, SE > ± 5 D, comorbid ophthalmic, or neurologic surgery/disease. **Setting:** University of Alabama at Birmingham. **Spectrum of glaucoma severity:** mean ± SD MD on the VF test was -3.8 ± 3.6 dB. According to Hodapp et al. grading scale, 44 eyes had an early glaucoma, 31 moderate, and 4 severe. **Control participants:** IOP < 22 mmHg, bilateral normal eye examination findings and bilateral normal VF results (defined as PSD within the 95% normal limits and a GHT result within 99% limits).
Index tests	**Scanning laser polarimetry**: GDx VCC (Carl Zeiss Meditec, Inc., Dublin, CA, USA). The mean of 3 images was calculated. Images were considered of good quality if there was good fixation, minimal eye movement, and good illumination on the reflectance image, with no artefacts on the retardance image. No author had conflict of interest.
Target condition and reference standard(s)	**Manifest glaucoma:** glaucomatous VF loss (defined as PSD outside 95% normal limits or GHT outside 99% normal limits) confirmed with a second VF test. **Visual field testing:** no details about how it was conducted and which instrument was used. VF reliability criteria included fixation losses and false-positive and false-negative rates of < 30%. Optic disc appearance was not part of the reference standard.
Flow and timing	Reference standard and index tests were completed within 1 to 8 weeks.

De Leon-Ortega 2006 *(Continued)*

	45 glaucoma patients (> 10%) were excluded due to poor-quality images.
Comparative	
Notes	None.

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Was a case-control design avoided?	No		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (All tests)			
If a threshold was used, was it pre-specified?	Yes		
Were imaging test's quality assessed?	Yes		
Were any conflict of interest avoided	Yes		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Unclear risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern

De Leon-Ortega 2006 *(Continued)*

DOMAIN 4: Flow and Timing		
Was there an appropriate interval between index test and reference standard?	No	
Did all patients receive the same reference standard?	Yes	
Were all patients included in the analysis?	No	
Did all patients receive a reference standard	Yes	
Could the patient flow have introduced bias?		High risk

De Leon-Ortega 2007

Study characteristics	
Patient Sampling	Data were obtained from the University of Alabama at Birmingham Optic Nerve Imaging Center database, which consists of functional and imaging data from glaucoma patients and controls enrolled in clinical studies from January 2000 to December 2004. Glaucoma patients were recruited by chart review and referrals, while controls were university employees, or were recruited from the general population. One eye per person was randomly selected.
Patient characteristics and setting	**Sample size**: 374 participants were initially enrolled, 78 glaucoma (44 African-American, 34 European), 89 healthy controls (51 African-American, 38 European) actually included in the analysis. **Age:** glaucoma African-American patients mean ± SD, 49.5 ± 9.8 years, glaucoma European ancestry 49.4 ± 17.2 years, controls African-American 47.3 ± 9.5 years, controls European ancestry 47.5 ± 8.8 years. **Ethnicity:** African-American and European ancestry. **Country:** USA. **Ocular comorbidities**: no history of intraocular surgery (except uncomplicated cataract surgery), cataracts, problems affecting colour vision other than glaucoma, use of medication or any comorbid condition affecting visual function. BCVA ≥ 20/40, spherical refraction within ±5 D, and cylinder correction within ± 3D. **Setting:** University of Alabama at Birmingham. **Spectrum of glaucoma severity:** mean ± SD MD and PSD on the VF test were -3.6 ± 3.6 dB and 4.3 ± 3.1 dB, for glaucoma African-American; -3.3 ± 3.2 dB and 4.1 ± 3.1 dB, for glaucoma European ancestry, respectively. **Control participants:** IOP < 22 mmHg, no past history of increased IOP, no family history of glaucoma, normal VF test results, and normal optic nerve appearance.
Index tests	**Confocal scanning laser ophthalmoscopy:** Heidelberg Retina Tomography (Heidelberg Engineering, Heidelberg, Germany). An experienced operator evaluated the image quality and outlined the disc margin, masked to the patient diagnosis. After obtaining the HRT 2 results, all scans with their respective contour lines were exported to a personal computer with the HRT 3 software. Images were excluded if they had: acquisition sensitivity > 89%, SD > 39,

De Leon-Ortega 2007 *(Continued)*

	results, ONH not centred, excessive eye movement occurred during the acquisition movie, floaters over or adjacent to the disc. One author was a consultant for Carl Zeiss Meditec.
Target condition and reference standard(s)	**Manifest glaucoma:** glaucomatous VF result, defined as either GHT outside the 99% normal limits or a PSD outside the 95% normal limits, and at least 1 cluster of 3+ test points outside 95% confidence interval in the pattern deviation probability plot, without crossing the horizontal hemifield. **Visual field testing:** Humphrey Field Analyzer II, SITA standard, 24-2 programme (Carl Zeiss Meditec, Inc., Dublin, CA, USA). VF reliability criteria included a fixation loss, false-positive and false-negative rates < 33%. **Optic disc evaluation:** dilated fundus examination, simultaneous stereoscopic optic disc photography.
Flow and timing	Of 374 patients initially enrolled, 167 were actually included in the analysis. 31 (> 10%) were excluded due to poor image quality, 5 patients were excluded due to poor quality in the stereophotograph.
Comparative	
Notes	None.

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Was a case-control design avoided?	No		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (All tests)			
If a threshold was used, was it prespecified?	Yes		
Were imaging test's quality assessed?	Yes		
Were any conflict of interest avoided	Yes		

De Leon-Ortega 2007 *(Continued)*

Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	No		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	No		
Did all patients receive a reference standard	Yes		
Could the patient flow have introduced bias?		High risk	

Essock 2005

Study characteristics	
Patient Sampling	Patients were enrolled prospectively from the outpatient clinics of glaucoma specialists. Both eyes were selected and enrolled for some patients.
Patient characteristics and setting	**Sample size**: 134 eyes of 134 participants (67 glaucoma and 67 control subjects). **Age:** glaucoma patients mean age, 67.22 years; controls 64.61 years.

Essock 2005 *(Continued)*

	Country: USA. **Ocular comorbidities**: no significant ocular media opacity. **Setting:** The Institute of Ophthalmology and Visual Science, University of Medicine and Dentistry of New Jersey; New York Eye and Ear Infirmary (New York, NY); Eye Care Center (San Diego, CA, USA). **Spectrum of glaucoma severity:** mean ± SD MD and PSD on the VF test were -6.82 ± 6.2 dB and 6.25 ± 4.2 dB respectively. **Control participants:** normal IOP and normal appearance of ONH. VFs were measured in most, but not all, cases and were normal.
Index tests	**Scanning laser polarimetry:** GDx VCC (Laser Diagnostic Technologies, Inc., San Diego, CA, USA). The measurements were obtained in 3 different clinics and were performed by experienced technicians. No details about scan's quality assessment were reported. Some authors had conflict of interest.
Target condition and reference standard(s)	**Manifest glaucoma:** patients with VF defects of GSS stage 1 or greater. **Visual field testing:** Humphrey Field Analyzer, model II 30-2 or 24-2 threshold, standard or full SITA strategy (Humphrey-Zeiss Instruments, Dublin, CA, USA). All VFs had good reliability, no further specified. Optic disc appearance was not part of the reference standard.
Flow and timing	Time interval between reference standard and index test was not specified. No patients were reported as excluded from the analysis.
Comparative	
Notes	All healthy participants had normal IOP and had normal appearance of optic discs. VFs were measured in most, but not all, cases.

Methodological quality

Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Was a case-control design avoided?	No		
Did the study avoid inappropriate exclusions?	Unclear		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (All tests)			

Essock 2005 *(Continued)*

If a threshold was used, was it pre-specified?	Yes		
Were imaging test's quality assessed?	Unclear		
Were any conflict of interest avoided	No		
Could the conduct or interpretation of the index test have introduced bias?		High risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Unclear		
Did all patients receive the same reference standard?	No		
Were all patients included in the analysis?	Yes		
Did all patients receive a reference standard	Yes		
Could the patient flow have introduced bias?		High risk	

Fang 2010

Study characteristics	
Patient Sampling	Consecutive outpatients were enrolled from July 2008 to March 2009. One eye per person was selected.
Patient characteristics and setting	**Sample size**: 90 eyes of 90 participants were enrolled. 76 eyes were actually included in the analysis (34 glaucoma, 42 healthy controls) **Age:** glaucoma patients mean ± SD, 58.4 ± 11.0 years; controls 56.3 ± 13.7 years. **Sex:** 27 men (15 glaucoma, 12 controls) and 49 women (19 glaucoma, 30 controls). **Country:** China.

Fang 2010 *(Continued)*

	Ocular comorbidities: no co-existing retinal disease, BCVA < 20/30, SE < -6 D or > +4 D, optic neuropathy, uveitis, trauma and past intraocular surgery, diabetes, hypertension. **Setting:** Department of Ophthalmology, Peking University First Hospital, Beijing. **Spectrum of glaucoma severity:** mean ± SD MD and PSD on the VF test were -2.28 ± 1.8 dB and 3.68 ± 2.14 dB. **Control participants:** IOP < 21 mmHg, healthy ONH/RNFL appearance and normal VF test result.
Index tests	**Optical coherence tomography:** RTvue FD-OCT, version 3.0. (Optovue Inc., Fremont, Ca, USA). Each patient was scanned using 3 patterns, including RNFL 3.45 scan, NHM4 scan, and MM7 scan. Quality FD-OCT scans were defined as those with a signal-strength index > 30. No details about author's conflict of interest were reported.
Target condition and reference standard(s)	**Manifest glaucoma:** glaucomatous optic disc or RNFL appearance (rim thinning, notching, excavation, or haemorrhage), open angle by gonioscopy, and glaucomatous VF defects (defined as GHT outside normal limits, PSD with P < 5%, or a cluster of ≥ 3 points in the pattern deviation plot in a single hemifield (superior or inferior) with P < 0.05, one of which should have a P < 0.01). **Visual field testing:** Humphrey Field Analyzer, model 750, 24-2 SITA fast strategy (Carl Zeiss Meditec, Dublin, CA, USA). VF reliability criteria included fixation losses and false-positive and false-negative rates of < 30%. **Optic disc evaluation:** dilated fundus examination.
Flow and timing	Time interval between reference standard and index test was not reported. 12 patients with early glaucoma were excluded owing to poor image quality.
Comparative	
Notes	None.

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	No		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (All tests)			

Fang 2010 *(Continued)*

If a threshold was used, was it pre-specified?	Yes		
Were imaging test's quality assessed?	Yes		
Were any conflict of interest avoided	Unclear		
Could the conduct or interpretation of the index test have introduced bias?		Unclear risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Unclear		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	No		
Did all patients receive a reference standard	Yes		
Could the patient flow have introduced bias?		High risk	

Ferreras 2007

Study characteristics	
Patient Sampling	Participants were prospectively pre-enrolled from January 2006 to June 2006. Glaucomatous eyes were recruited consecutively from an ongoing longitudinal follow-up study. Normal eyes were consecutively recruited from patients referred for refraction who underwent routine examination without abnormal ocular findings, hospital staff, and relatives of patients in our hospital. One eye per person was selected.
Patient characteristics and setting	**Sample size**: 201 eyes of 201 participants enrolled, 186 eyes of 186 participants included in the analysis (115 glaucoma, 71 healthy controls).

Ferreras 2007 *(Continued)*

	Age: glaucoma patients mean ± SD, 61.9±7.29 years; controls 59.0 ± 9.8. **Ethnicity:** white. **Country:** Spain. **Ocular comorbidities**: no previous intraocular surgery, BCVA < 20/40, refractive spherical error < -5 D/cylinder error > +2 D, diabetes or other systemic diseases, history of ocular or neurologic disease. **Setting:** Miguel Servet University Hospital, Department of Ophthalmology, Zaragoza. **Spectrum of glaucoma severity:** mean ± SD of MD and PSD on the VF test were −6.49 ± 6.08 dB and 5.08 ± 3.63 dB respectively. According to Hodapp et al. grading scale, 62 eyes had early glaucoma, 32 moderate and 21 severe. **Control participants:** IOP < 20 mmHg, no optic disc morphology suspicious for glaucoma, and a normal SAP.
Index tests	**Confocal scanning laser tomography:** HRT3 (Heidelberg Engineering, Heidelberg, Germany) with a diode laser (670 nm wavelength). Topographic images were obtained through dilated pupils and analysed using the Advanced Glaucoma Analysis 3.0 software. Only scans with “acceptable,” “good,” or “very good” image quality scores were included. The margin of the optic discs was manually traced by the same glaucoma specialist with at least a 4-point contour line. No author had conflict of interest.
Target condition and reference standard(s)	**Manifest glaucoma:** IOP of > 21 mmHg (on > 3 readings on different days), open angle by gonioscopy and SAP defects (defined as the presence of a cluster of 3 points lower than P < 5%, a cluster of 2 points lower than P < 1% on pattern deviation probability plots, or a PSD with P < 2% or GHT outside the normal limits). **Visual field testing:** Humphrey Field Analyzer, model 745, 24-2 SITA-standard strategy (Zeiss Humphrey Systems, Dublin, CA, USA). VF reliability criteria included fixation losses, false-positive and false-negative rates of < 20%. The participants completed the perimetry tests before undergoing any clinical examination or structural test. Optic disc appearance was not part of the reference standard.
Flow and timing	Reference standard and index test were performed within 1 month. 15 participants (< 10%) were excluded from the analysis: 2 participants did not provide informed consent, 2 participants did not complete all of the required tests, 3 participants were unable to perform at least 1 of the tests expected; in 8 participants, GPS analyses produce only a global result or no results.
Comparative	
Notes	None.

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	No		

Ferreras 2007 *(Continued)*

Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (All tests)			
If a threshold was used, was it prespecified?	Yes		
Were imaging test's quality assessed?	Yes		
Were any conflict of interest avoided	Yes		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		

Ferreras 2007 *(Continued)*

Did all patients receive a reference standard	Yes
Could the patient flow have introduced bias?	Low risk

Ferreras 2008a

Study characteristics	
Patient Sampling	Participants with normal eyes were recruited from among patients referred for refraction who underwent routine examination without abnormal ocular findings, hospital staff, and relatives of patients in the hospital. Patients with glaucoma were recruited from an ongoing longitudinal follow-up study, including those who underwent imaging of the optic disk with the HRT2 from September 1, 2005 through April 30, 2007. One eye per person was selected.
Patient characteristics and setting	**Sample size**: 183 eyes of 183 participants (90 glaucoma and 93 controls). **Age**: glaucoma patients mean ± SD, 60.45 ± 9.08 years; controls 56.43 ± 9.87. **Sex**: 79 men (41 glaucoma, 38 controls) and 104 women (49 glaucoma, 55 controls). **Ethnicity**: white. **Country**: Spain. **Ocular comorbidities**: no previous intraocular surgery, BCVA < 20/40, SE > ± 5 D, lens opacity, diabetes, or other ocular or neurologic disease. **Setting**: Department of Ophthalmology of Miguel Servet University Hospital. **Spectrum of glaucoma severity**: mean ± SD MD and PSD on the VF test for were -6.03 ± 6.33 dB and 4.01 ± 3.61 dB respectively. **Control subjects**: IOP < 21 mmHg (on at least 3 readings on different days) and a normal SAP test result.
Index tests	**Confocal scanning laser tomography**: HRT 2 (Heidelberg Engineering, Heidelberg, Germany). All scans had to have an interscan SD of < 30 µm. The margin of the optic disks was traced manually by the same glaucoma specialist, who was masked to participant identity and clinical history. Scans were analysed using first the HRT2 software and, afterward, the Advanced Glaucoma Analysis 3.0 software. No author had conflict of interest.
Target condition and reference standard(s)	**Manifest glaucoma:** IOP > 21 mmHg, open angle by gonioscopy and typical glaucomatous SAP defects (defined as the presence of a cluster of 3 points with a P < 0.05 or a cluster of 2 points with a P < 0.01 on the pattern deviation plot, a PSD with P < 5%, a GHT outside normal limits, or a combination thereof). **Visual field testing:** Humphrey Field Analyzer, model 750, 24-2 SITA-standard strategy (Zeiss Humphrey Systems, Dublin, CA, USA). VF reliability criteria included fixation losses, false-positive and false-negative rates of < 20%. The participants completed the perimetry tests before undergoing any clinical examination or structural test. Optic disc appearance was not part of the reference standard.
Flow and timing	Reference standard and index test were performed within 2 months.

Ferreras 2008a *(Continued)*

	Patients were enrolled consecutively. No details about participants excluded from the analysis were reported.
Comparative	
Notes	None.

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	No		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (All tests)			
If a threshold was used, was it pre-specified?	Yes		
Were imaging test's quality assessed?	Yes		
Were any conflict of interest avoided	Yes		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	

Ferreras 2008a *(Continued)*

Are there concerns that the target condition as defined by the reference standard does not match the question?		Low concern
DOMAIN 4: Flow and Timing		
Was there an appropriate interval between index test and reference standard?	No	
Did all patients receive the same reference standard?	Yes	
Were all patients included in the analysis?	Unclear	
Did all patients receive a reference standard	Yes	
Could the patient flow have introduced bias?		High risk

Ferreras 2008b

Study characteristics	
Patient Sampling	From April, 2006, through December, 2006, 2 samples (one population for obtaining the LDF and a second independent population for testing the LDF) of consecutive healthy control participants and glaucoma patients were pre-enrolled prospectively. Normal eyes were recruited from among patients referred for refraction who underwent routine examination without abnormal ocular findings, from among hospital staff, and from among relatives of patients in the hospital. Patients with glaucoma were recruited from an ongoing longitudinal follow-up study. One eye per person was randomly selected.
Patient characteristics and setting	**Sample size**: 2 samples were enrolled. A first sample of 166 eyes (85 glaucoma/ 81 controls) to calculate a discriminant analysis. A second sample of 435 eyes: 225 controls and 210 glaucomatous eyes (163 POAG, 34 PEX and 13 pigmentary glaucoma). **Age**: glaucoma mean ± SD, 61.10 ± 10.07 years; controls 57.46 ± 9.84 years, for the first sample. Glaucoma mean ± SD, 61.37 ± 10.4 years; controls 57.67 ± 10.19 years, for the second sample. **Ethnicity**: white. **Country**: Spain. **Ocular comorbidities**: BCVA < 20/40, SE > ± 5 D, no previous intraocular surgery, lens opacity, diabetes, or other ocular or neurologic disease. **Setting**: Department of Ophthalmology of Miguel Servet University Hospital. **Spectrum of glaucoma severity**: mean ± SD MD and PSD on the VF test were -5.79 ± 5.74 dB and 4.93 ± 3.78 dB for the first sample, -5.34 ± 4.87 dB and 4.87 ± 3.95 dB for the second sample. **Control** participan**ts**: IOP < 21 mmHg (on at least 3 readings on different days), and a normal SAP test result.
Index tests	**Confocal scanning laser tomography**: HRT 3 (Heidelberg Engineering, Heidelberg, Germany). Topographic images were obtained through dilated pupils and were analysed us-

Ferreras 2008b *(Continued)*

	ing the Advanced Glaucoma Analysis 3.0 software. All scans had to have an interscan SD of < 30 µm. The margin of the optic disc was traced manually by the same glaucoma specialist who was masked to the patients' identity and clinical history. No author had conflict of interest.
Target condition and reference standard(s)	**Manifest glaucoma:** IOP > 21 mmHg and typical SAP defects (defined as a PSD with a P < 5% and/or a GHT outside normal limits). **Visual field testing:** Humphrey Field Analyzer, model 745, 24-2 SITA-standard strategy (Zeiss Humphrey Systems, Dublin, CA, USA). VF reliability criteria included fixation losses and false-positive and false-negative rates of < 20%. The participants completed the perimetry tests before undergoing any clinical examination or structural test. Optic disc appearance was not part of the reference standard.
Flow and timing	Reference standard and index test were performed within 1 month. 21 participants (< 10%) were excluded from the analysis: 3 participants did not provide informed consent, 11 participants did not complete all of the required tests, and 7 participants were unable to perform at least 1 of the tests expected.
Comparative	
Notes	None.

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	No		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (All tests)			
If a threshold was used, was it prespecified?	Yes		
Were imaging test's quality assessed?	Yes		
Were any conflict of interest avoided	Yes		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	

Ferreras 2008b *(Continued)*

Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
Did all patients receive a reference standard	Yes		
Could the patient flow have introduced bias?		Low risk	

Garas 2011

Study characteristics	
Patient Sampling	White individuals referred for detection or exclusion of glaucoma, who underwent RNFLT, GCC, and ONH measurements made with the RTVue-100 Fourier-domain OCT between 1 January and 30 November 2009, were enrolled in the study. One eye per person was randomly selected.
Patient characteristics and setting	**Sample size**: 286 eyes of 286 participants (111 with perimetric glaucoma, 46 with preperimetric glaucoma, 36 with ocular hypertension and 93 healthy control participants). **Age:** perimetric glaucoma patients mean ± SD, 62.2 ± 14.7 years; preperimetric glaucoma patients 57.6 ± 11.8 years; OHT patients 51.5 ± 16.5 years; controls 54.9 ± 15.9 years.

Garas 2011 *(Continued)*

	Sex: 126 male, 160 women. **Ethnicity:** white. **Country:** Hungary. **Ocular comorbidities**: no macular pathology, diabetic retinopathy, cornea degeneration, or non-glaucomatous optic neuropathies. **Setting:** Glaucoma Centre of Semmelweis University in Budapest. **Spectrum of glaucoma severity:** mean ± SD MD on the VF test were -0.1 ± 1.2 dB for ocular hypertension, 0.1 ± 1.8 dB for preperimetric group and 9.8 ± 7.8 dB for perimetric group. According to the modified Bascom Palmer staging system, the perimetric glaucoma group consists of 26 stage 1 patients, 34 at stage 2, 21 at stage 3, 24 at stage 4 and 6 at stage 5. **Control participants:** no ONH damages, normal VF tests (MD < 2 dB), and IOP < 21 mmHg.
Index tests	**Optical coherence tomography:** RTVue-100 Fourier-domain OCT, software version 4.0 (Optovue Inc., Froemont, CA, USA). For RNFLT, GCC and ONH measurements the standard glaucoma protocol was used. Scans were acquired through undilated pupils. To be included in the analysis, images had to have a signal strength index > 40. One author is an unpaid consultant of Optovue, Inc and Carl Zeiss Meditec, Inc.
Target condition and reference standard(s)	**Perimetric manifest glaucoma:** glaucomatous neuroretinal rim loss and VF defect typical for glaucoma (inferior and/or superior paracentral or arcuate scotomas, nasal step, hemifield defect or generalised depression with MD > 2 dB). **Preperimetric manifest glaucoma:** glaucomatous neuroretinal rim loss (diffuse/localised neuroretinal rim thinning) and normal visual field with MD < 2 dB. **Ocular Hypertension:** normal ONH, normal visual field with MD < 2 dB and untreated IOP consistently > 21 mmHg. **Visual field testing:** Octopus field analyser, normal or dynamic G2 threshold visual field testing. No details about reliability criteria were reported. **Optic disc evaluation:** stereoscopic ONH photography by a glaucoma specialist.
Flow and timing	Reference standard and index test were performed within 2 months. Of the 316 referred patients 30 (< 10%) did not meet the inclusion criteria and were not enrolled in the study.
Comparative	
Notes	None.

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Was a case-control design avoided?	No		

Garas 2011 *(Continued)*

Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (All tests)			
If a threshold was used, was it pre-specified?	Yes		
Were imaging test's quality assessed?	Yes		
Were any conflict of interest avoided	No		
Could the conduct or interpretation of the index test have introduced bias?		High risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	No		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
Did all patients receive a reference standard	Yes		

Garas 2011 *(Continued)*

Could the patient flow have introduced bias?	High risk

Garas 2012

Study characteristics	
Patient Sampling	Consecutive white individuals referred for detection of glaucoma by their family doctors, optometrists, or local ophthalmologists in the Glaucoma Centre who underwent OCT and GDx imaging session between January 1 and October 31, 2009, and fitting eligibility criteria, were enrolled in the study. One eye per person was randomly selected.
Patient characteristics and setting	**Sample size**: 177 eyes of 177 participants enrolled (66 perimetric glaucoma, 33 preperimetric glaucoma, 28 hypertensive, 50 healthy eyes). **Age**: perimetric glaucoma patients 64.3 ± 12.9 years; preperimetric glaucoma patients 56.2 ± 12.1 years; OHT patients mean ± SD, 50.8 ± 15.6 years; controls 50.2 ± 17.3 years. **Sex**: 75 men (24 perimetric glaucoma, 16 preperimetric glaucoma, 13 OHT, 22 control) and 102 women (42 perimetric glaucoma, 17 preperimetric glaucoma, 15 OHT, 28 control). **Ethnicity**: white. **Country**: Hungary. **Setting**: Glaucoma Centre of Semmelweis University in Budapest. **Ocular comorbidities**: participants with refractive error ≤ ± 10 D, no sufficient central vision for optimal fixation and clinically significant cataract, were not included. **Spectrum of glaucoma severity**: the mean ± SD MD and PSD on the VF test were 0.3 ± 1.7 dB and 9.6 ± 6.8 dB for preperimetric and perimetric glaucoma respectively. **Control participants**: eyes with no structural or functional damage including healthy eyes with normal optic nerve appearance, normal VF result and IOP consistently < 21 mmHg, and hypertensive participants with normal optic nerve appearance, normal VF result and IOP untreated > 21 mmHg.
Index tests	**Optical coherence tomography**: RTVue-100 Fourier-domain OCT (Optovue Inc., Fremont, CA, USA). The ONH scan protocol was used. All images were taken by the same operator and only images with signal strength index > 40 were used. Images with insufficient quality or with any artefact were rejected and reacquired. **Scanning laser polarimetry:** GDx VCC instrument (software version 5.5.1; Carl Zeiss Meditec Inc., Dublin, CA, USA). Both variable corneal compensation or enhanced corneal compensation or both were used. All images were acquired by the same operator and quality score > 8 was required to be accepted. One author is an unpaid consultant of Carl Zeiss, inc. and Optovue, inc.
Target condition and reference standard(s)	Glaucoma group comprised: **Preperimetric glaucoma**: glaucomatous neuroretinal rim loss (diffuse or localised neuroretinal rim thinning, notching with bared circumlinear vessels and corresponding angulation of the vessels at the disc margin) and normal visual field with MD < 2 dB. **Perimetric glaucoma:** glaucomatous neuroretinal rim loss and VF defect typical for glaucoma (inferior and/or superior paracentral or arcuate scotomas, nasal step, hemifield defect) or generalised depression with MD > 2 dB. The glaucoma groups comprised both open-angle and angle-closure glaucoma cases.

Garas 2012 *(Continued)*

	Visual field testing: Octopus Normal or Dynamic G2 threshold. **Optic disc evaluation**: detailed slit-lamp evaluation and stereoscopic ONH photography evaluated by a glaucoma specialist.
Flow and timing	The time interval between index tests and reference standard was not reported. No patients were reported as excluded from the analysis by the authors.
Comparative	
Notes	None.

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	No		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (All tests)			
If a threshold was used, was it prespecified?	Yes		
Were imaging test's quality assessed?	Yes		
Were any conflict of interest avoided	No		
Could the conduct or interpretation of the index test have introduced bias?		High risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			

Garas 2012 *(Continued)*

Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Unclear		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
Did all patients receive a reference standard	Yes		
Could the patient flow have introduced bias?		Unclear risk	

Garudadri 2012

Study characteristics	
Patient Sampling	Prospective cross-sectional study including normal participants and glaucoma patients evaluated between July 2003 and March 2005 at a tertiary eye care centre. One eye per person was randomly selected.
Patient characteristics and setting	**Sample size**: 220 eyes of 220 participants enrolled (125 glaucoma, 95 controls). **Age**: glaucoma eyes mean ± SD, 57.46 ± 9.65 years; controls 50.39 ± 10.76 years. **Sex**: 145 men (86 glaucoma, 59 controls) and 75 women (39 glaucoma, 36 controls). **Ethnicity**: Indian. **Country**: India. **Setting**: LV Prasad Eye Institute, Hyderabad. **Ocular comorbidities**: all eyes had to have BCVA ≥ 20/40, refractive error within ±5 D sphere and ±3 D cylinder of plano. Patients with intraocular surgery or laser within

Garudadri 2012 *(Continued)*

	past 6 months, history or evidence of retinal or macular pathology, evidence of any systemic diseases or neurological disorders, which could produce a field defect were excluded. **Spectrum of glaucoma severity**: the mean ± SD MD on the VF test were -9.55 ± 8.61 dB for glaucoma. **Control participants**: IOP ≤ 22 mmHg, normal posterior segment evaluation and normal VF result.
Index tests	**Scanning laser polarimetry:** GDx VCC (software version 5.5.1; Carl Zeiss Meditec). Only properly-focused and well-centred images of the ONH with an image score ≥ 8 in both eyes were included in the study. Imaging was performed by 1 of 2 trained optometrists masked to the hypothesis and diagnosis. No conflict of interest with the device's manufacturer were reported by the authors.
Target condition and reference standard(s)	**Manifest Glaucoma:** glaucomatous ONH appearance (defined as focal or diffuse neuroretinal rim thinning, localised notching, or nerve fibre layer defects) and corresponding VF defects, defined as 2 of the following 3: the presence of a cluster of 3 points on pattern deviation probability plot with a P < 5%, one of which had a P < 1%, or a PSD with a P < 5%, or a GHT result outside normal limits. **Visual field testing:** Humphrey Field Analyzer (Carl Zeiss Meditec, Inc, Dublin, CA, USA) using the 30-2 or 24-2 SITa standard programme. **Optic disc evaluation**: indirect fundus ophthalmoscopy using a 78D or 90D lens.
Flow and timing	Index tests and reference standard were performed within 3 months. No patients were reported as excluded from the analysis by the authors.
Comparative	
Notes	Garudadri CS was supported by Allergan and Merck, Parikh RS was supported by Merck, and Thomas R was supported by Allergan.

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Was a case-control design avoided?	No		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (All tests)			

Garudadri 2012 *(Continued)*

If a threshold was used, was it pre-specified?	Yes		
Were imaging test's quality assessed?	Yes		
Were any conflict of interest avoided	Yes		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	No		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
Did all patients receive a reference standard	Yes		
Could the patient flow have introduced bias?		High risk	

Gonzales de la Rosa 2013

Study characteristics	
Patient Sampling	Case-control study including eyes with ocular hypertension considered to be at risk, with suspected (IOP > 25 mmHg, or IOP > 21 with CCT < 500 µm or with family history of glaucoma) or confirmed open-angle glaucoma and control eyes. One eye per person was selected.

Gonzales de la Rosa 2013 *(Continued)*

Patient characteristics and setting	**Sample size**: 206 eyes of 206 participants (104 eyes with suspected or confirmed open-angle glaucoma and 102 controls). **Age**: not reported. **Sex**: not reported. **Ethnicity**: not reported. **Clinical Setting:** not reported. **Country**: Spain. **Ocular comorbidities**: not reported **Manifest glaucoma:** focal (localised notching) or diffuse neuroretinal rim narrowing with concentric enlargement of the optic cup, or both, or reproducible glaucomatous VF defects (no further details reported) or both, regardless of the IOP values. **Visual field test:** not reported.
Index tests	**Optical coherence tomography**: Cirrus OCT (Carl Zeiss Meditec, Jena, Germany). **Confocal scanning laser ophthalmoscopy**: Heidelberg Retinal Tomograph HRT III (Heidelberg Engineering, Heidelberg, Germany). No further details reported. Two authors had proprietary interest in one of the index test analysed.
Target condition and reference standard(s)	**Manifest glaucoma:** focal (localised notching) or diffuse neuroretinal rim narrowing with concentric enlargement of the optic cup, or both, or reproducible glaucomatous VF defects or both (no further details reported), regardless of the IOP values. **Visual field test:** no details reported.
Flow and timing	No details about exclusion were reported. The time interval between index and reference test was not reported
Comparative	
Notes	None.

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Was a case-control design avoided?	No		
Did the study avoid inappropriate exclusions?	Unclear		

Gonzales de la Rosa 2013 *(Continued)*

Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (All tests)			
If a threshold was used, was it pre-specified?	Unclear		
Were imaging test's quality assessed?	No		
Were any conflict of interest avoided	Unclear		
Could the conduct or interpretation of the index test have introduced bias?		High risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Unclear
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Unclear		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Unclear risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Unclear
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Unclear		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
Did all patients receive a reference standard	Yes		
Could the patient flow have introduced bias?		Unclear risk	

Harizman 2006

Study characteristics	
Patient Sampling	Normal participants, those suspected of having glaucoma and patients with glaucoma were enrolled. One eye per person was randomly selected.

Harizman 2006 *(Continued)*

Patient characteristics and setting	**Sample size**: 220 eyes of 220 participants enrolled, 217 eyes included in the analysis (83 glaucoma and 134 healthy controls). **Age:** glaucoma patients mean ± SD, 58.5 ± 11.8 years; controls 45.5 ± 13.6. **Country**: not specified. **Ethnicity:** 93 white (62 control, 31 glaucoma), 124 black (72 control, 52 glaucoma). **Ocular comorbidities**: no narrow angle, BCVA < 20/40, SE > ±5 D, retinal disease, ocular surface disease, non-glaucomatous optic neuropathy or previous intraocular surgery other than uncomplicated cataract surgery. **Spectrum of glaucoma severity:** mean ± SD of MD and PSD on the VF test were −7.31 ± 6.66 dB and 6.58 ± 3.85 dB, respectively. Control participants**:** VFs in both eyes unremarkable (PSD with P < 5% and GHT within 97% normal limits) and the clinical examination normal.
Index tests	**Confocal scanning laser tomography:** HRT 2, software version 1.1.1 (Heidelberg Engineering, Germany). A mean topographic image was automatically obtained from 3 scans using HRT2 software V.1.4.1. Good image quality was assessed (acquisition sensitivity < 90%, topography SD < 40 micron, more than ¾ of the disc within the target circle, minimal movement during the acquisition movie, no floaters over the disc, and good imaging clarity and exposure). A trained technician outlined the optic disc margin on the mean topographic image. HRT2 data results were exported to the HRT3 software (V.3.0) and the appropriate racial database was selected before analysis. No author had conflict of interest.
Target condition and reference standard(s)	**Manifest glaucoma:** reproducible, at least 2 consecutive, glaucomatous VF defects (defined as a PSD with P < 5% or GHT outside normal limits). **Visual field testing:** Humphrey Field Analyzer, 24-2 SITA-standard strategy (Carl Zeiss Meditec, Dublin, CA, USA). VF reliability criteria included fixation losses, false-positive and false-negative rates of < 33%. Optic disc appearance was not part of the reference standard.
Flow and timing	Reference standard and the index test were performed within 1 month. 3 participants (< 10%, 2 normal, 1 glaucoma) were excluded from the analysis because the GPS model could not had been calculated. Patients suspected of having glaucoma were enrolled but not included in the analysis, with no explanation reported.
Comparative	
Notes	None.

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Was a case-control design avoided?	No		

Harizman 2006 *(Continued)*

Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 2: Index Test (All tests)			
If a threshold was used, was it pre-specified?	Yes		
Were imaging test's quality assessed?	Yes		
Were any conflict of interest avoided	Yes		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Unclear		
Did all patients receive a reference standard	Yes		

Harizman 2006 *(Continued)*

Could the patient flow have introduced bias?	Unclear risk

Hoesl 2013

Study characteristics	
Patient Sampling	Healthy control participants and glaucoma patients were enrolled. Controls were recruited from the general population, as well as from the staff and employees of the University Erlangen-Nuremberg. Glaucoma participants were selected from those included in 'The Erlangen Glaucoma Registry', a clinical registry for cross-sectional and longitudinal observational study of patients with open-angle glaucoma or glaucoma suspect. One eye per person was randomly selected.
Patient characteristics and setting	**Sample size**: 134 eyes of 134 participants enrolled (102 glaucoma, 32 controls). Glaucoma patients were divided based on TSS value: 33 had TSS = 100, 31 had TSS ≥ 80 and ≤ 99, 38 had TSS < 80. **Age**:TSS = 100 glaucoma eyes: mean ± SD, 57.1 ± 10.3 years; 99 ≥, TSS ≥ 80 glaucoma: 60.0 ± 9.8 years; TSS < 80 glaucoma: 60.3 ± 11.1 years; controls 57.2 ± 6.1 years. **Sex**: 72 men (54 glaucoma, 18 controls) and 62 women (48 glaucoma, 16 controls). **Ethnicity**: not specified. **Country**: Germany. **Setting**: Department of Ophthalmology, University of Erlangen- Nuremberg, Erlangen. **Ocular comorbidities**: patients with diabetes, any eye diseases other than glaucoma, or myopic refractive error > 7 D or equivalent sphere > D diopter of astigmatism were excluded. **Spectrum of glaucoma severity**: the mean ± SD MD and PSD on the VF test were 7.3 ± 6.3 dB and 6.4 ± 2.5 dB, respectively for TSS = 100 glaucoma group; 7.4 ± 5.3 dB and 6.8 ± 2.9 dB respectively for 99 ≥ TSS and ≥ 80 glaucoma group; 7.4 ± 5.5 dB and 6.2 ± 2.8 dB respectively for TSS < 80 glaucoma group. Control participants: IOP ≤ 21 mmHg, normal optic disc and normal VF result.
Index tests	**Scanning laser polarimetry:** GDx VCC (software version 5.5.0; Carl Zeiss Meditec). Only images with a centred optic disc, well-illuminated and a scan score > 8 were accepted. The authors stated no conflict of interested.
Target condition and reference standard(s)	**Manifest Glaucoma:** IOP > 21 mmHg, open angle at gonioscopy, glaucomatous ONH appearance (defined as neuroretinal rim thinning, notching, visibility of localised RNFL defects, or an unusually small neuroretinal rim area in relation to the optic disc size and cup-to-disc ratios that were larger vertically than horizontally) and glaucomatous VF defects (defined as the presence of 3 adjacent test points with P < 0.05 or 2 adjacent test points with P < 0.01 in the pattern deviation map). **Visual field testing:** Octopus 500 (Haag-Streit; Peridata software, version 2.2.3). Reliability criteria were false-positive and false-negative rates < 12%. **Optic disc evaluation**: 15° colour photographs (Zeiss telecentric fundus camera, Germany).
Flow and timing	The time interval between index tests and reference standard was not reported.

Hoesl 2013 *(Continued)*

	No patients were reported as excluded from the analysis by the authors.
Comparative	
Notes	None.

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Was a case-control design avoided?	No		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (All tests)			
If a threshold was used, was it prespecified?	Yes		
Were imaging test's quality assessed?	Yes		
Were any conflict of interest avoided	Yes		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	

Hoesl 2013 *(Continued)*

Are there concerns that the target condition as defined by the reference standard does not match the question?		Low concern
DOMAIN 4: Flow and Timing		
Was there an appropriate interval between index test and reference standard?	Unclear	
Did all patients receive the same reference standard?	Yes	
Were all patients included in the analysis?	Yes	
Did all patients receive a reference standard	Yes	
Could the patient flow have introduced bias?		Unclear risk

Hong 2007

Study characteristics	
Patient Sampling	Primary open-angle glaucoma patients with early VF defects and healthy controls were included. One eye per person was randomly selected.
Patient characteristics and setting	**Sample size**: 120 eyes of 120 participants (72 glaucoma and 48 healthy controls). **Age:** glaucoma patients mean ± SD, 37.8 ± 15.6 years; controls 38.7 ± 13.6 years. **Sex:** 54 men (34 glaucoma and 20 controls); 66 women (38 glaucoma and 28 controls). **Country:** not specified. **Ocular comorbidities**: no significant cataract, BCVA < 20/40, SE > ±5 D, ocular diseases other than glaucoma, previous intraocular surgery, or narrow angle. **Spectrum of glaucoma severity:** mean ± SD MD and PSD on the VF test were -2.9 ± 1.12 dB and 3.26 ± 0.76 dB, respectively. Control participants**:** no VF loss by SAP, IOP < 21 mmHg, no ONH/RNFL changes suggestive of glaucoma.
Index tests	**Scanning laser polarimetry:** GDx VCC (Laser Diagnostic Technologies, Inc. San Diego, CA, USA). No author had conflict of interest.
Target condition and reference standard(s)	**Manifest glaucoma**: optic disc damage (defined as excavation, notching, focal or diffuse atrophy of neuroretinal rim area, vertical cup-to-disc ratio

Hong 2007 *(Continued)*

	more than 0.6, cup-to-disc asymmetry between fellow eyes more than 0.2, disc haemorrhage, baring of circumlinear blood vessels, or localised defect of the RNFL) and VF loss (defined as GHT outside normal limits or PSD with P < 5% or 3+ adjacent points below the 5% level on the pattern deviation plot, with at least 1 point below the 1% level). **Visual field testing:** Humphrey Field Analyzer, model II, 30-2 SITA-standard strategy (Carl Zeiss Meditec Inc., Dublin, CA, USA). VF reliability criteria were not reported.
Flow and timing	Reference standard and visual field were performed within 1 week. No patient was reported as excluded from the analysis.
Comparative	
Notes	None.

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		
Was a case-control design avoided?	No		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 2: Index Test (All tests)			
If a threshold was used, was it pre-specified?	Yes		
Were imaging test's quality assessed?	Unclear		
Were any conflict of interest avoided	Yes		
Could the conduct or interpretation of the index test have introduced bias?		Unclear risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		

Hong 2007 *(Continued)*

Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
Did all patients receive a reference standard	Yes		
Could the patient flow have introduced bias?		Low risk	

Huang 2010

Study characteristics	
Patient Sampling	Glaucoma patients and healthy controls who had sought treatment at the department of ophthalmology, were enrolled. One eye per person was selected.
Patient characteristics and setting	**Sample size**: 165 eyes of 165 participants (79 glaucoma, 86 healthy controls). **Age:** glaucoma patients mean ± SD, 44.3 ± 14.72 years; controls 40.2 ± 15.54. **Sex:** 82 men (42 glaucoma, 40 controls) and 83 women (37 glaucoma, 46 controls). **Ethnicity:** Taiwan Chinese. **Country:** China. **Ocular comorbidities**: no co-existing retinal disease, BCVA < 20/40, SE > ±5 D, uveitis, or non-glaucomatous optic neuropathy. **Setting:** Department of Ophthalmology, China Medical University Hospital, Taiwan. **Spectrum of glaucoma severity:** mean ± SD MD and PSD on the VF test were -5.6 ± 4.23 dB and 2.38 ± 3.15 dB respectively. **Control participants:** IOP < 21 mmHg, normal optic nerve appearance, and a normal VF result (MD and PSD within 95% confidence limits, and GHT within normal limits).
Index tests	**Scanning laser polarimetry:** GDx-VCC, software version 5.5.0 (Laser Diagnostic Technologies, Inc. San Diego, CA, USA). All measurements were ob-

Huang 2010 *(Continued)*

	tained by the same trained technician. The images had to be of high quality (a well-focused, even, centred optic disc without any motion artefact) and with a scan quality score > 7. No author had conflict of interest.
Target condition and reference standard(s)	**Manifest glaucoma:** repeatable (2 consecutive) glaucomatous VF defects (defined as a PSD outside the 95% normal confidence limits, or a GHT result outside 99% normal confidence limits). **Visual field testing:** Humphrey Field Analyzer, model 750, 30-2 programme (Carl Zeiss Meditec, Dublin, CA, USA). Optic disc appearance was not part of the reference standard.
Flow and timing	Reference standard and index test were performed within 3 months. No patients were reported as excluded from the analysis.
Comparative	
Notes	None.

Methodological quality

Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Was a case-control design avoided?	No		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (All tests)			
If a threshold was used, was it pre-specified?	Yes		
Were imaging test's quality assessed?	Yes		
Were any conflict of interest avoided	Yes		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		

Huang 2010 *(Continued)*

Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	No		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
Did all patients receive a reference standard	Yes		
Could the patient flow have introduced bias?		High risk	

Huang 2011

Study characteristics	
Patient Sampling	Glaucoma patients were retrospectively collected from the clinical database of the Glaucoma Service, where patients received OCT imaging as part of routine management. The control group was enrolled prospectively, between June 2008 and September 2009. One eye per person was randomly selected.
Patient characteristics and setting	**Sample size**: 220 eyes of 220 participants (146 glaucoma and 74 healthy controls). **Age:** glaucoma patients mean ± SD, 64.34 ± 8.28 years; controls 61.49 ± 9.91 years. **Sex:** 59 men (25 controls, 34 glaucoma), 82 women (49 controls, 33 glaucoma). **Ethnicity:** 75 white (48 glaucoma, 27 controls), 22 African-American (17 glaucoma, 5 controls), 118 Asian (73 glaucoma, 35 controls) and 15 Hispanic (8 glaucoma, 7 controls). **Country:** USA. **Ocular comorbidities**: no retinal disorders, BCVA < 20/40, SE < -6 D or > +3 D, optic nerve disorders other than glaucoma, previous intraocular surgery, diabetes or central nervous system disorders. **Setting:** Glaucoma Service at Beckman Vision Center, University of California, San Francisco. **Spectrum of glaucoma severity:** mean ± SD MD and PSD on the VF test were -3.3 ± 2.64 dB and 4.65 ± 3.01 dB respectively. Control participants**:** vertical cup-to-disc ratio ≤ 0.5, IOP ≤ 21 mmHg, and a normal VF (MD > 0 dB).

Huang 2011 *(Continued)*

Index tests	**Optical coherence tomography:** RTVue 100 OCT, software version 3.5 (Optovue, Fremont, CA, USA). The ganglion cells complex scan and nerve head map 4 mm scans were acquired. A single grader was assigned to redraw the disc margin and determine the anchoring points of the retinal pigmented epithelium layer. OCT image had signal strength > 45 for GCC scan and 30 for Nerve Head Map 4 mm scan. No author had conflict of interest.
Target condition and reference standard(s)	**Manifest glaucoma:** glaucomatous VF defects (defined as the presence of > 3 contiguous points lower than P < 0.05 and > 1 of these points below P < 0.01) and vertical cup-to-disc ratio large > 0.5. **Visual field testing:** Humphrey Field Analyzer, Model II, 30-2 SITA-standard strategy (Zeiss Meditec, Dublin, CA, USA). Severity of VF defects was graded by a masked grader. VF reliability criteria included fixation losses, false-positive and false-negative rates of < 20%. **Optic disc evaluation:** vertical cup-to-disc ratio was estimated by an experienced glaucoma specialist.
Flow and timing	Reference standard and index test were performed within 3 months. 1459 eyes from 810 participants received the reference and the index tests during enrolment period; 220 eyes of 220 participants were finally enrolled on the basis of inclusion criteria.
Comparative	
Notes	None.

Methodological quality

Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	No		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (All tests)			
If a threshold was used, was it pre-specified?	Yes		
Were imaging test's quality assessed?	Yes		
Were any conflict of interest avoided	Yes		

Huang 2011 *(Continued)*

Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	No		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
Did all patients receive a reference standard	Yes		
Could the patient flow have introduced bias?		High risk	

Hwang 2012

Study characteristics	
Patient Sampling	Healthy control participants and glaucoma patients (matched based on age, spherical equivalent and optic disc size) were recruited consecutively between May 2009 and September 2011. One eye per person was randomly selected.
Patient characteristics and setting	**Sample size**: 160 eyes of 160 participants enrolled (80 glaucoma, 80 controls). **Age**: glaucoma eyes mean ± SD, 53.94 ± 11.17 years; controls 55.39 ± 11.15 years. **Sex**: all men. **Ethnicity**: Korean. **Country**: Korea.

Hwang 2012 *(Continued)*

	Setting: Department of Ophthalmology, Armed Forces Capital Hospital, Seongnam. **Ocular comorbidities**: patients with spherical equivalent > ±2 D, BCVA < 20/30, history of ocular inflammation, trauma, previous ocular surgery or laser, presence of concurrent retinal disease or optic nerve disease other than glaucoma, or brain disorder that could influence VF results, were excluded. **Spectrum of glaucoma severity**: the mean ± SD MD and PSD on the VF test were -6.90 ± 4.79 dB and 7.44 ± 3.73 dB, respectively for glaucoma. Control participants: IOP ≤ 21 mmHg, open angle at gonioscopy, normal ONH, no RNFL defect on red-free fundus photography and normal VF result.
Index tests	**Optical coherence tomography:** Cirrus HD-OCT (software version 5.1.0.96; Carl Zeiss Meditec, Dublin, CA, USA). The Optic Disc Scan cube 200 x 200 was used. Images with poor quality (signal strength ≤ 6, incorrect identification of the vitreoretinal surface detection algorithm, misidentification of Bruch's membrane and prominent saccade during the scan) were excluded. The authors stated no conflict of interested.
Target condition and reference standard(s)	**Manifest Glaucoma:** open angle on gonioscopy, glaucomatous ONH changes (as increased cup–disc ratio and narrowing of the neuroretinal rim), RNFL defect (defined as a dark wedge-shaped area with its apex touching the optic disc border in the brightly striated pattern of the surrounding RNFL or generalised loss of RNFL visibility in the upper or lower retina), glaucomatous VF defects (defined as a cluster of 3 points with P < 5% on the pattern deviation map in at least 1 hemifield, including at least 1 point with P < 1%, or a cluster of 2 points with a P < 1% and GHT results outside normal limits, or a PSD outside 95% of normal limits). **Visual field testing:** Humphrey Field Analyzer (30-2 SITA standard programme, Carl Zeiss Meditec, Inc, Dublin, CA, USA). Reliability criteria were fixation losses, false-positive and false-negative rates < 15%. **Optic disc evaluation**: fundus examination with a +90 D and red-free fundus photograph using a Zeiss FF450 fundus camera (Carl Zeiss Meditec).
Flow and timing	The time interval between index tests and reference standard was not reported. No patients were reported as excluded from the analysis by the authors.
Comparative	
Notes	None.

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	No		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		High risk	

Hwang 2012 *(Continued)*

Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (All tests)			
If a threshold was used, was it pre-specified?	Yes		
Were imaging test's quality assessed?	Yes		
Were any conflict of interest avoided	Yes		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Unclear		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
Did all patients receive a reference standard	Yes		
Could the patient flow have introduced bias?		Unclear risk	

Iester 2008

Study characteristics	
Patient Sampling	Prospective, cross-sectional study. Patients were consecutively recruited. One eye per person was selected.
Patient characteristics and setting	**Sample size**: 214 eyes of 214 participants (95 glaucoma, 119 healthy controls). **Age**: glaucoma patients mean ± SD, 68.1 ± 11.9 years; controls 63.7 ± 12.3 years. **Country**: Italy. **Ocular comorbidities**: no ocular disease other than glaucoma, spherical refraction > ±8 D or secondary cause for glaucoma. **Setting**: Clinica Oculistica, Department of Neurological Sciences, Ophthalmology, Genetic, University of Genoa, Italy; Division of Ophthalmology, Ospedale S. Andrea, University La Sapienza II, Roma, Italy. **Spectrum of glaucoma severity**: mean ± SD MD and PSD on the VF test were -3.33 ± 4.92 dB and 3.82 ± 2.85 dB, respectively. Control participants: IOP < 21 mmHg, normal VF, normal ONH and RNFL on clinical examination.
Index tests	**Confocal scanning laser tomography**: HRT 3, software version 3.0 (Heidelberg Engineering, Heidelberg, Germany). Only high-quality images with acquisition sensitivity > 80% were included in the study. ONHs were analysed using 2 different methods: either the observer drew the contour line around the ONH or the system analysed the shape of the ONH without any user input. No author had conflict of interest.
Target condition and reference standard(s)	**Manifest glaucoma:** glaucomatous VF defects (defined as 3 adjacent points being depressed by 5 dB, with 1 of the points being depressed by at least 10 dB or 2 adjacent points being depressed by 10 dB or a 10 dB difference across the nasal horizontal meridian in 2 adjacent points) and/or a typical abnormal ONH (defined as notching, diffuse/generalised loss of optic rim tissue, vertical cup/disk diameter ratio asymmetry and disc haemorrhage), open angle at gonioscopy, IOP > 21 mmHg with no treatment. **Visual field testing:** Humphrey Field Analyzer, model 750, 24-2 SITA-standard strategy (HFA, Humphrey Inc, San Leandro, CA, USA). VF reliability criteria included fixation losses of < 20% and false-negative rates of < 30%.
Flow and timing	The time interval between reference standard and index test was specified. All patients enrolled were included in the analysis.
Comparative	
Notes	None.

Methodological quality			
Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		

Iester 2008 *(Continued)*

Was a case-control design avoided?	No		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (All tests)			
If a threshold was used, was it pre-specified?	Yes		
Were imaging test's quality assessed?	Yes		
Were any conflict of interest avoided	Yes		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Unclear		
Did all patients receive the same reference standard?	No		
Were all patients included in the analysis?	Yes		
Did all patients receive a reference standard	Yes		
Could the patient flow have introduced bias?		High risk	

Jeoung 2010

Study characteristics	
Patient Sampling	Eyes with preperimetric localised RNFL defects and normal control eyes meeting the eligibility criteria were consecutively enrolled from May 2008 to October 2008. One eye per person was randomly selected.
Patient characteristics and setting	**Sample size**: 110 eyes of 110 participants (55 preperimetric glaucoma and 55 healthy controls). **Age:** preperimetric glaucoma patients mean ± SD, 54.1 ± 10.4 years; controls 53.4 ± 10.6 years. **Sex:** 60 men (30 glaucoma and 30 controls) and 50 women (25 glaucoma and 25 controls). **Country:** Korea. **Ethnicity**: not specified. **Ocular comorbidities**: no uveitis, BCVA < 20/40, SE > ±5 D, ocular surgery other than cataract extraction, or diseases that may affect the peripapillary area. **Setting:** Glaucoma Clinic of Seoul National University Hospital, Korea. **Spectrum of glaucoma severity:** mean ± SD MD and PSD on the VF test -0.74 ± 0.96 dB and 1.85 ± 0.39 dB, respectively. **Control participants:** IOP ≤ 21 mmHg (with no history of increased IOP), absence of glaucomatous disc appearance (defined as intact neuroretinal rim without peripapillary haemorrhages, notches, or localised pallor), no visible RNFL defect according to red-free RNFL photography, and a normal SAP result.
Index tests	**Optical coherence tomography:** Cirrus HD-OCT, Optic Disc cube 200 x 200 programme, software version 3.0 (Carl Zeiss Meditec, Inc.). Patients were imaged after pupil dilation. The image quality scans were assessed by 2 experienced examiners masked to the clinical information. The minimum acceptable signal strength score was 6 and the examiners assessed subjectively the quality of the image evaluating the en-face image for eye movements. No author had conflict of interest.
Target condition and reference standard(s)	**Preperimetric glaucoma:** localised wedge-shaped RNFL defect clearly visible by red-free fundus photography with normal SAP results (defined as MD and PSD within 95% confidence limits and a GHT within normal limits) and open angle by gonioscopy. **Red-free fundus photography:** Digital fundus camera. 60°, wide-angle views of the optic disc, carefully focused on the retina using the built-in split-line focusing device were obtained and reviewed on an LCD monitor by 2 experienced observers. Localised RNFL defects were determined when their width at a 1-disc diameter distance from the edge of the disc was larger than that of a major retinal vessel, diverging in an arcuate or wedge shape and reaching the edge of the disc. **Visual field testing:** Humphrey Field Analyzer, model II 750, 30-2 SITA-standard strategy (Carl Zeiss Meditec, Dublin, CA, USA).
Flow and timing	Time interval between reference standard and index test was not reported. 171 eyes were initially enrolled. 19 eyes were excluded due to poor quality images. Of the 96 control eyes, only 55 eyes age- and sex-matched with glaucoma eyes, were selected for the analysis.

Jeoung 2010 *(Continued)*

Comparative	
Notes	None.

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	No		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (All tests)			
If a threshold was used, was it prespecified?	Yes		
Were imaging test's quality assessed?	Yes		
Were any conflict of interest avoided	Yes		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
Are there concerns that the target condition as defined by the refer-			Low concern

Jeoung 2010 *(Continued)*

ence standard does not match the question?		
DOMAIN 4: Flow and Timing		
Was there an appropriate interval between index test and reference standard?	Unclear	
Did all patients receive the same reference standard?	Yes	
Were all patients included in the analysis?	No	
Did all patients receive a reference standard	Yes	
Could the patient flow have introduced bias?		High risk

Jeoung 2013

Study characteristics	
Patient Sampling	Healthy controls and glaucoma patients were among participants in the Macular Ganglion Cell Imaging Study, an ongoing prospective study of glaucoma patients and healthy individuals at the Glaucoma Clinic of Seoul National University Hospital. One eye per person was randomly selected.
Patient characteristics and setting	**Sample size**: 545 eyes of 545 participants initially considered, 425 eyes eventually included in the analysis (306 glaucoma, 119 controls). 164 eyes with early glaucoma, 142 with moderate-to-advanced glaucoma. **Age**: early glaucoma eyes mean ± SD, 58.7 ± 10.2 years; moderate-to-advanced glaucoma eyes mean ± SD, 59.2 ± 13.1 years; controls 57.1 ± 12.3 years. **Sex**: 213 men (160 glaucoma, 53 controls) and 212 women (146 glaucoma, 66 controls). **Ethnicity**: not specified. **Country**: Korea. **Setting**: Glaucoma Clinic of Seoul National University Hospital. **Ocular comorbidities**: patients with BCVA < 20/40 in the study eye, refractive > ±6 D equivalent sphere and ±3 D astigmatism, retinal disease (diabetic retinopathy, macular degeneration, retinal detachment, epiretinal membrane) or non-glaucomatous optic nerve diseases, treatment that might be toxic to the retina or optic nerve, laser therapy, or ocular surgery except non-complicated cataract surgery were excluded. **Spectrum of glaucoma severity**: the mean ± SD MD and PSD on the VF test were -2.68 ± 1.76 dB and 5.47 ± 2.8 db, respectively for early glaucoma, -12.41 ± 5.92 dB and 12.20 ± 3.16 dB for moderate-to-severe glaucoma. **Control participants**: IOP ≤ 21 mmHg with no history of increased IOP, normal ONH appearance, no RNFL defect on red-free fundus photography and normal VF result.

Jeoung 2013 *(Continued)*

Index tests	**Optical coherence tomography:** Cirrus HD-OCT (software version 6.0, Carl Zeiss Meditec, Dublin, CA, USA). The macular cube 200 x 200 and optic disc cube 200 x 200 scanning protocols were used. The authors stated no conflict of interest.
Target condition and reference standard(s)	**Manifest Glaucoma:** glaucomatous optic disc cupping (defined as neuroretinal rim thinning, notching, excavation, or RNFL defect) and corresponding VF defect (defined as the presence of a cluster of 3+ non-edge points on the pattern deviation plot with a P < 5%, with 1 of these points having a P < 1%, a PSD with P < 5% or a GHT outside normal limits). **Visual field testing:** Humphrey Field Analyzer (model II 750, 30-2 SITA standard programme, Carl Zeiss Meditec, Inc, Dublin, CA, USA). Reliability criteria were fixation losses < 20, false-positive and false-negative rates < 33%. **Optic disc and RNFL evaluation**: colour disc photography, red-free RNFL photography (TRC-50IX; Topcon Corporation, Tokyo, Japan), evaluated independently by 2 observers in a random order and masked fashion, without knowledge of the clinical information.
Flow and timing	Index test and reference standard were performed within 1 month. 92 eyes were excluded because of diabetic retinopathy (n = 36), macular degeneration (n = 28), epiretinal membrane (n = 20), and ocular surgery history (n = 8). 28 eyes were excluded from the analysis due to poor-quality images.
Comparative	
Notes	Supported by Grant No. A121615 from the Korea Health technology R&D Project, Ministry of Health & Welfare, Republic of Korea, and by Grant No. 2009-0091931 from the National Research Foundation of Korea (NRF) funded by the Korea government (MEST).

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Was a case-control design avoided?	No		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (All tests)			
If a threshold was used, was it prespecified?	Yes		
Were imaging test's quality assessed?	Yes		

Jeoung 2013 *(Continued)*

Were any conflict of interest avoided	Yes		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	No		
Did all patients receive a reference standard	Yes		
Could the patient flow have introduced bias?		High risk	

Jindal 2010

Study characteristics

Jindal 2010 *(Continued)*

Patient Sampling	Healthy participants and patients with early-to-moderate primary open-angle glaucoma were enrolled prospectively. One eye per person was randomly selected.
Patient characteristics and setting	**Sample size**: 100 eyes of 100 participants (50 glaucoma, 50 healthy controls). **Age:** glaucoma patients mean ± SD, 58.78 ± 11.08 years, controls 44.74 ± 8.88 years. **Country:** not specified. **Ocular comorbidities**: no significant media opacity (corneal, lenticular), BCVA < 20/40, SE > ±5 D or other intraocular/neurological diseases affecting the RNFL, optic disc, or VF. **Spectrum of glaucoma severity:** mean ± SD MD and PSD on the VF test were -6.45 ± 2.47 dB and 5.71 ± 3.23 dB, respectively. Patients included were early or moderate glaucoma, according to Hodapp et al. grading scale. **Control participants:** IOP ≤ 21 mmHg, open angles by gonioscopy, normal clinical evaluation, and a normal VF test.
Index tests	**Confocal scanning laser ophthalmoscopy:** HRT 3, version 3.0. All images obtained were of good quality, defined as having a topographic SD of < 30 µm and had no floaters or opaque areas. The contour line was drawn by a single operator. No author had conflict of interest.
Target condition and reference standard(s)	**Manifest glaucoma:** IOP > 21 mmHg at diagnosis, open angle by gonioscopy, glaucomatous ONH changes and VF glaucomatous defects (defined as 3 contiguous non-edge points depressed with P < 5%, 1 of which had P < 1%, all being not contiguous with the blind spot and GHT outside normal limits and PSD < 5%). **Visual field testing:** Humphrey Field Analyzer, model II, 30-2 SITA-standard strategy. No details about VF reliability criteria were reported. **Optic disc evaluation**: stereoscopic dilated fundus examination.
Flow and timing	The time interval between reference standard and index test was not reported. No patients were reported as excluded from the analysis .
Comparative	
Notes	None.

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Was a case-control design avoided?	No		

Jindal 2010 *(Continued)*

Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (All tests)			
If a threshold was used, was it pre-specified?	Yes		
Were imaging test's quality assessed?	Yes		
Were any conflict of interest avoided	Yes		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Unclear		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Unclear risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Unclear		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
Did all patients receive a reference standard	Yes		
Could the patient flow have introduced bias?		Unclear risk	

Kanamori 2006

Study characteristics

Kanamori 2006 *(Continued)*

Patient Sampling	Retrospective study, performed between April 2003 and November 2003. Normal, ocular hypertensive, suspected/preperimetric glaucoma and manifest perimetric glaucoma eyes were enrolled. One eye per person was randomly selected.
Patient characteristics and setting	**Sample size**: 241 eyes of 201 participants (67 perimetric glaucoma, 55 preperimetric glaucoma, 26 OHT and 93 healthy controls). **Age:** perimetric glaucoma patients mean ± SD, 48.9 ± 12.6 years; preperimetric glaucoma patients mean ± SD, 48.5 ± 12.3 years; hypertensive mean ± SD 46.4 ± 11.4 years; controls 45 ± 15.5 years. **Sex:** 119 men (30 perimetric glaucoma, 22 preperimetric glaucoma,14 OHT, 53 controls) and 122 women (37 perimetric glaucoma, 33 preperimetric glaucoma, 12 OHT, 40 controls). **Country:** Japan. **Ocular comorbidities**: no previous ocular surgeries, BCVA < 20/40, cylinder refraction > ±4 D, retinal disease, significant vitreous opacity or diabetes. **Setting:** Department of Ophthalmology of the Kobe University Hospital. **Spectrum of glaucoma severity:** mean ± SD MD and PSD on the VF test were -3.55 ± 1.76 dB and 6.26 ± 10.82 dB for the perimetric glaucomatous eyes; -1.14 ± 1.41 dB and 1.46 ± 0.98 dB for the preperimetric glaucomatous eyes; -0.63 ± 1.11 dB and 1.24 ± 0.88 dB respectively for OHT eyes. **Control participants:** no family history of glaucoma, normal optic disc appearance, and normal IOP.
Index tests	**Scanning laser polarimetry:** GDx VCC, software version 5.3.2 (Laser Diagnostic Technologies, Inc., San Diego, CA, USA). Images were taken from each eye without pupillary dilation. Images were accepted only if the quality score was > 7. No details about authors' conflict interest were reported.
Target condition and reference standard(s)	**Manifest perimetric glaucoma:** glaucomatous optic neuropathy (vertical cup-disc asymmetry between fellow eyes of 0.2 or more and neuroretinal rim damages such as excavation, rim thinning, and notches) and associated VF loss (2+ contiguous points with a pattern deviation sensitivity loss of P < 0.01, or 3+ contiguous points with sensitivity loss of P < 0.05, in the superior or inferior arcuate areas, or a 10 dB difference across the nasal horizontal midline at 2+ adjacent locations and a GHT outside normal limits). **Manifest preperimetric glaucoma:** glaucomatous optic neuropathy (vertical cup-disc asymmetry between fellow eyes of 0.2 or more and neuroretinal rim damages such as excavation, rim thinning, and notches) with normal VF result. **Ocular hypertensive:** IOP > 21 mmHg (on 2 separate occasions), normal optic disc appearance and normal VF result. **Visual field testing:** Humphrey Field Analyzer, 30-2 SITA-standard strategy (Humphrey-Zeiss Instruments, Dublin, CA, USA). VF reliability criteria included fixation losses of < 20% and false-negative rates of < 25%. **Optic disc evaluation**: stereoscopic examination with slit-lamp biomicroscopy by glaucoma expert masked to the index test result.
Flow and timing	Reference standard and index tests were performed within 6 months. 32 eyes (> 10%) were excluded due to poor-quality image.
Comparative	

Kanamori 2006 *(Continued)*

Notes	None.

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Was a case-control design avoided?	No		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (All tests)			
If a threshold was used, was it prespecified?	Yes		
Were imaging test's quality assessed?	Yes		
Were any conflict of interest avoided	Unclear		
Could the conduct or interpretation of the index test have introduced bias?		Unclear risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern

Kanamori 2006 *(Continued)*

DOMAIN 4: Flow and Timing	
Was there an appropriate interval between index test and reference standard?	No
Did all patients receive the same reference standard?	Yes
Were all patients included in the analysis?	No
Did all patients receive a reference standard	Yes
Could the patient flow have introduced bias?	High risk

Kang 2012

Study characteristics	
Patient Sampling	Healthy control participants and glaucoma patients were recruited prospectively, in a consecutive manner, between March 2009 and February 2010. One eye per person was randomly selected. Only people with VF loss confined to 1 side of the horizontal median were enrolled.
Patient characteristics and setting	**Sample size**: 112 eyes of 112 participants initially enrolled. 108 eyes finally included in the analysis (54 glaucoma, 54 controls). **Age**: glaucoma eyes mean ± SD, 56.4 ± 11.8 years; controls 55.1 ± 6.90 years. **Sex**: 56 men (28 glaucoma, 28 controls) and 49 women (23 glaucoma, 26 controls). **Ethnicity**: not specified. **Country**: Korea. **Setting**: Glaucoma Clinic of Asan Medical Center, Seoul. **Ocular comorbidities**: eyes had to have BCVA ≥ 20/30, a spherical equivalent within ±5 D and a cylinder correction within +3 D. Patients with any ophthalmic disease other than glaucoma that could result in an HFA defect, or with histories of intraocular surgery or diabetes mellitus were excluded. **Spectrum of glaucoma severity**: the mean ± SD MD and PSD on the VF test were -5.12 ± 3.44 dB and 6.55 ± 3.73 dB, respectively for glaucoma. **Control participants:** IOP < 22 mmHg with no history of increased IOP, normal ONH appearance and normal VF result.
Index tests	**Optical coherence tomography:** Cirrus OCT (software version 3.0.0.50). Optic disc cube scan 200 x 200 mode. Images with poor quality (signal strength < 7, overt misalignment of the surface detection algorithm, overt displacement of the measurement circle) or horizontal eye motion observed within the measurement circle. The authors stated no conflict of interest.

Kang 2012 *(Continued)*

Target condition and reference standard(s)	**Manifest Glaucoma:** glaucomatous VF defect (defined as a GHT result outside 97% of normal limits, a PSD outside 95% of normal limits, and a cluster of 3+ points in the pattern deviation plot in a single hemifield (superior or inferior) with P < 0.05, 1 of which had a P < 0.01) regardless of the ONH or RNFL appearance). Glaucomatous VF loss was confined to 1 side of the horizontal meridian, as defined by 3+ adjacent points with P < 0.05 in a PD probability map, or 2+ adjacent points with P < 0.02 in a superior or inferior hemifield; and the hemifield of the other side had no clusters of 3 points with P < 0.05 and no clusters of 2 points with P < 0.02 on either total deviation or pattern deviation probability maps. **Visual field testing:** Humphrey Field Analyzer (24-2 SITA standard programme, Carl Zeiss Meditec, Inc, Dublin, CA,USA). Reliability criteria were fixation losses < 20, false-positive and false-negative rates < 15%.
Flow and timing	The time interval between index test and reference standard was not reported. 4 glaucoma eyes (< 10%) were excluded from the analysis: 1 eye for poor VF reliability test, 3 eyes due to poor-quality index test result.
Comparative	
Notes	None.

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	No		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (All tests)			
If a threshold was used, was it prespecified?	Yes		
Were imaging test's quality assessed?	Yes		
Were any conflict of interest avoided	Yes		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	

Kang 2012 *(Continued)*

Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Unclear		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
Did all patients receive a reference standard	Yes		
Could the patient flow have introduced bias?		Unclear risk	

Kim 2011

Study characteristics	
Patient Sampling	Glaucoma patients with or without high myopia were consecutively enrolled from January 2009 to June 2009. Normal controls were sequentially matched. One eye per person was randomly selected.
Patient characteristics and setting	**Sample size**: 196 participants examined, 150 included in the analysis. The participants were divided into 2 groups: a highly-myopic group (spherical equivalent < -6.00 D) and a non-highly myopic group (spherical equivalent > -6.00 D): 45 highly-myopic participants (21 glaucoma, 24 controls) and 105 non-highly myopic (56 glaucoma, 49 controls). **Age**: glaucoma highly-myopic eyes mean ± SD, 42.67 ± 16.32 years; highly-myopic controls 41.83 ± 12.44 years; glaucoma non-highly myopic eyes mean ± SD, 56.02 ± 14.90 years; non-highly myopic controls 52.39 ± 15.55 years;

Kim 2011 *(Continued)*

	Sex: 76 men (45 glaucoma, 31 controls) and 74 women (32 glaucoma, 42 controls). **Ethnicity**: Asian. **Country**: Korea. **Setting**: Glaucoma-Cataract Clinic of Severance Hospital, Seoul. **Ocular comorbidities**: highly-myopic eyes with any atypical non-glaucomatous field defect and eyes with a narrow angle, media opacity, prior history of ocular surgery, diabetes mellitus, or other diseases affecting the VF were excluded. **Spectrum of glaucoma severity**: the mean ± SD MD and PSD on the VF test were -8.56 ± 5.82 dB and 7.85 ± 4.76 dB respectively, for glaucoma highly-myopic eyes; were -9.49 ± 7.41 dB and 7.75 ± 4.16 dB, respectively for glaucoma non-highly myopic eyes. **Control participants:** IOP < 21 mmHg, normal appearance of ONH and normal VF test result.
Index tests	**Optical coherence tomography**: RTVue-100 (software version: 4.0.5.39, Optovue, Fremont, CA, USA). The nerve head map 4 mm diameter (NHM4) and the MM7 scanning protocols were used. Images with a poor quality (SSI < 35, overt misalignment of the surface detection algorithm or overt decentration of the measurement circle location) were excluded. No authors had conflict of interest.
Target condition and reference standard(s)	**Manifest glaucoma:** glaucomatous VF defects (defined as having 3+ significant (P < 0.05) non-edge contiguous points with at least 1 at the P < 0.01 level on the same side of the horizontal meridian in the pattern deviation plot, classified as outside normal limits in the GHT) and glaucomatous appearance of the ONH not otherwise described. **Visual field testing:** Humphrey Field Analyzer, 20-2 SITA standard programme (Carl Zeiss Meditec). VF reliability criteria were not reported. **Optic disc and RNFL evaluation**: stereoscopic optic disc photography or red-free RNFL photography.
Flow and timing	Index test and reference standard were performed on the same day. 46 eyes (> 10%) were excluded from the final analysis: 36 because of poor OCT image (low signal strength (11), improper scan decentration (14), presence of epiretinal membrane (2), erroneous RNFL or GCC profile (5) algorithm of the GCC failure(4)); 3 because of unacceptable stereoscopic fundus photography, and 7 due to unreliable VF.
Comparative	
Notes	This research was supported by the Basic Science Research Programme through the National Research Foundation of Korea (NRF) funded by the Ministry of Education, Science and Technology (No 2009-0076736).

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	No		

Kim 2011 *(Continued)*

Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (All tests)			
If a threshold was used, was it prespecified?	Yes		
Were imaging test's quality assessed?	Yes		
Were any conflict of interest avoided	Yes		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	No		

Kim 2011 *(Continued)*

Did all patients receive a reference standard	Yes
Could the patient flow have introduced bias?	High risk

Kim 2013a

Study characteristics	
Patient Sampling	Participants were enrolled consecutively from January 2009 to June 2009. NTG were sequentially enrolled as they presented. Primary open-angle glaucoma patients were randomly matched by age, sex, and visual field sensitivities to those of NTG group. Healthy controls were recruited from the hospital staff, nurses, the spouses or friends of patients, and patients referred for routine visual acuity examination, matched by age and sex with glaucoma patients. One eye per person was randomly selected.
Patient characteristics and setting	**Sample size**: 161 eyes of 161 participants included(52 with POAG, 51 with NTG, 58 controls). **Age**: POAG eyes mean ± SD, 57.02 ± 15.74 years; NTG 55.55 ± 14.50 years; controls 55.78 ± 10.98 years. **Sex**: 78 men (30 POAG, 22 NTG, 26 controls) and 83 women (22 POAG, 29 NTG, 32 controls). **Ethnicity**: Asian. **Country**: Korea. **Setting**: Glaucoma-Cataract Clinic of Severance Hospital in the Yonsei University College of Medicine, Seoul. **Ocular comorbidities**: patients with media opacity, history of ocular surgery (other than uncomplicated glaucoma and cataract surgery), or other diseases affecting the VF were excluded. **Spectrum of glaucoma severity**: the mean ± SD MD and PSD on the VF test were -7.09 ± 5.36 db and 6.41 ± 4.31 dB respectively, for NTG, -7.70 ± 4.40 and 7.67 ± 4.43 respectively, for POAG. **Control participants**: IOP < 21 mmHg, normal ONH appearance and normal VF results. BCVA ≥ 20/40 and refractive error between +3 and -8 D.
Index tests	**Optical coherence tomography**: RTVue-100 Fourier-Domain OCT (software version: 4.0.5.39; Optovue Inc, Fremont, CA, USA). NHM4 and MM7 scanning protocols were used. Images with signal strength index < 35, overt misalignment of the surface detection algorithm or overt decentration of the measurement location, were excluded. No authors had conflict of interest.
Target condition and reference standard(s)	**Manifest glaucoma:** glaucomatous VF defects (defined as having 3+ significant (P < 0.05) non-edge contiguous points with at least 1 at the P < 0.01 level on the same side of the horizontal meridian in the pattern deviation plot, and GHT outside normal limits) and glaucomatous ONH appearance (defined as cup-to-disc ratio > 0.7, inter-eye cup asymmetry > 0.2 or neuroretinal rim notching, focal thinning, disc haemorrhage, or vertical elongation of the optic cup). Glaucoma patients were classify in 2 subgroups: **OAG**: IOP before treatment > 21 mmHg based on 3 measurements on different days.

Kim 2013a *(Continued)*

	NTG: untreated peak IOP < 21 mmHg on repeated 3 measurements taken at different times. **Visual field testing:** Humphrey Field Analyzer, 20-2 SITA standard programme (Carl Zeiss Meditec). VF reliability criteria were fixation losses < 20% and false-positive and false-negative rates < 15%. **Optic disc evaluation**: slit-lamp biomicroscopy.
Flow and timing	Index test and reference standard were performed on the same day. Authors stated that "Data were discarded if the scan quality did not satisfy the criteria described above", but no patients were reported as excluded from the analysis.
Comparative	
Notes	None.

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	No		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (All tests)			
If a threshold was used, was it prespecified?	Yes		
Were imaging test's quality assessed?	Yes		
Were any conflict of interest avoided	Yes		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			

Kim 2013a *(Continued)*

Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Unclear		
Did all patients receive a reference standard	Yes		
Could the patient flow have introduced bias?		Unclear risk	

Kim 2013b

Study characteristics	
Patient Sampling	Healthy participants and patients with a RFNL defects were recruited in an observational case-control design study. No other details were reported. One eye per person was randomly selected.
Patient characteristics and setting	**Sample size**: 94 participants enrolled, 90 eyes of 90 participants included in the analysis (48 with RNFL defects, 42 controls). **Age**: eyes with RNFL defects mean ± SD, 55.4 ± 11.6 years; controls 51.0 ± 12.7 years. **Sex**: 35 men (18 with RNFL defects, 17 controls) and 55 women (30 with RNFL defects, 25 controls). **Ethnicity**: not reported. **Country**: Korea. **Setting**: Department of Ophthalmology, Seoul National University Hospital. **Ocular comorbidities**: patients with retinal abnormality, previous retinal laser or intraocular surgery other than a cataract extraction or neurologic diseases were excluded. All pa-

Kim 2013b *(Continued)*

	tient had to have BCVA ≥ 20/40, a spherical equivalent within ±5.00 D, and an open anterior chamber angle. **Spectrum of glaucoma severity**: the mean ± SD MD and PSD on the VF test were -3.1 ± 3.3 dB and 5.3 ± 3.4 dB respectively,for eyes with a localised RNFL defect. **Control participants:** IOP < 21 mmHg, normal ONH appearance, normal VF results and no RNFL defect visible on red-free RNFL photograph.
Index tests	**Optical coherence tomography**: Cirrus HD-OCT model 4000 (software version 5.1.1.6; Carl Zeiss Meditec). The optic disc cube scan was used. To be included all images had to have a signal strength > 6, good centring of the optic disc, and the absence of motion artefacts. **Optical coherence tomography**: 3D OCT-2000 (software version 7.20; Topcon Medical Systems). The circumpapillary and macular cube scans were used. All images had to have a Q factor score > 60. Authors' conflicts of interest were not reported.
Target condition and reference standard(s)	**Manifest glaucoma:** patients with a localised RNFL defect defined as a well-outlined, dark wedge-shaped area in the brightly-striated pattern of the surrounding healthy RNFL with its tip touching the optic disc border. Patients with a localised RNFL defect included those with perimetric glaucoma with corresponding VF defects and those with preperimetric glaucoma with a normal VF. **Visual field testing:** Humphrey Field Analyzer II (30-2 SITA standard programme (Carl Zeiss Meditec). Visual field reliability criteria were fixation losses <20% and false positive and false negative <15%. **RNFL evaluation**: red-free fundus photography (VX- 10; Kowa Optimed, Tokyo, Japan). Two trained specialists evaluated the photograph independently and in a masked fashion.
Flow and timing	Index tests were performed on the same day but the time interval between index tests sand reference standard was not reported. 4 participants (<10%) were excluded due to unacceptable OCT quality scans.
Comparative	
Notes	The work was supported by Grant number 3020110090 from the Seoul National University Hospital Research Fund donated by Mr. Bong Joo Kim.

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Was a case-control design avoided?	No		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		High risk	

Kim 2013b *(Continued)*

Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (All tests)			
If a threshold was used, was it prespecified?	Yes		
Were imaging test's quality assessed?	Unclear		
Were any conflict of interest avoided	Yes		
Could the conduct or interpretation of the index test have introduced bias?		Unclear risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Unclear		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
Did all patients receive a reference standard	Yes		
Could the patient flow have introduced bias?		Unclear risk	

Kim 2014a

Study characteristics	
Patient Sampling	Case-control study including participants in an ongoing study of glaucoma and healthy individuals. If both eyes eligible only one eye per person was randomly chosen.
Patient characteristics and setting	**Sample size**: 184 eyes of 205 participants (92 preperimetric glaucoma, 92 normal controls). **Age**: preperimetric glaucoma mean ± SD, 57.8 ± 11.4 years; controls, 57.6 ± 11.3 years. **Sex**: 95 men (45 preperimetric glaucoma, 50 controls) and 89 women (47 preperimetric glaucoma, 42 controls). **Ethnicity**: Korean **Setting:** Glaucoma Clinic of Seoul National University Hospital, Seoul. **Country**: South Korea. **Ocular comorbidities**: eyes with history of amblyopia, uveitis, intraocular surgery (excepting uncomplicated cataract surgery), diabetes, ocular diseases possibly affecting the peripapillary area (e.g., large peripapillary atrophy), or macular area (e.g., epiretinal membrane), and any other ocular or systemic diseases affecting the VF (e.g., retinal vein occlusion, ischaemic optic neuropathy), were excluded. **Spectrum of glaucoma severity**: the mean ± SD mean deviation and PSD on the VF test were -0.16 ± 1.61 and 1.99 ± 0.86 respectively, for preperimetric glaucoma. **Control participants:** IOP ≤ 21 mmHg with no history of increased IOP, an absence of glaucomatous disc appearance, no visible RNFL defect on red-free fundus photography, and a normal VF result.
Index tests	**Optical coherence tomography**: Cirrus OCT (Carl Zeiss Meditec, Dublin, CA, USA); software version 6.0. Only images that were well centred on the optic disc or fovea with signal strength of ≥ 6 were included in the analyses. GCA and optic disc cube 200 x 200 scanning protocols were used. The authors declare no conflict of interest.
Target condition and reference standard(s)	**Manifest glaucoma:** 1+ localised RNFL defects associated with a glaucomatous disc appearance (e.g. notching or thinning of neuroretinal rim), which have documented evidence of progression (e.g. focal or diffuse narrowing of neuroretinal rim, increased excavation, increased width or depth of RNFL defects) through stereoscopic disc photography (SDP) or red-free fundus photography performed at least 6 months before enrolment, and normal VF result (PSD > 5% and GHT within normal limits). **Visual field test:** Humphrey Field Analyzer II (Carl Zeiss Meditec, Inc.) with 30-2 SITA-algorithm. VF exams were considered reliable when fixation loss < 20%, false-positive and false negative rates < 33%. **RNFL evaluation:** red-free fundus photography (VX-10; Kowa Optimed, Tokyo, Japan). 2 glaucoma specialists independently evaluated the red-free fundus photographs without knowledge of the participant's clinical information.
Flow and timing	209 eyes were initially involved (117 eyes with glaucoma and 92 normal control eyes). After excluding 4 eyes for ambiguous RNFL defects and age-matching the two groups, 184 eyes of 184 subjects (92 preperimetric glaucoma and 92 age-matched healthy control participants) were included in the analysis.

Kim 2014a *(Continued)*

	More than 10% of the enrolled eyes were excluded from the analysis. No details reported about time interval between index and reference test.
Comparative	
Notes	None.

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Was a case-control design avoided?	No		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (All tests)			
If a threshold was used, was it prespecified?	Yes		
Were imaging test's quality assessed?	Yes		
Were any conflict of interest avoided	Yes		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes		

Kim 2014a *(Continued)*

Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Unclear		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	No		
Did all patients receive a reference standard	Yes		
Could the patient flow have introduced bias?		High risk	

Kim 2014b

Study characteristics	
Patient Sampling	Retrospective case-control study including early glaucoma, preperimetric glaucoma and healthy controls. If both eligible, one eye per person was randomly selected.
Patient characteristics and setting	**Sample size**: 204 eyes of 204 participants (72 early glaucoma, 68 preperimetric glaucoma, 64 normal controls) **Age**: early glaucoma mean ± SD, 56.83 ± 12.73 years; preperimetric glaucoma, 53.12 ± 10.69 years; controls, 51.77 ± 14.44 years; **Sex**: not reported. **Ethnicity**: not reported. **Setting:** general healthcare clinic or glaucoma clinic of the Guri Hanyang University Medical Center from September 2011 through May 2013. **Country**: South Korea. **Ocular comorbidities**: patients with co-existing retinal disease, uveitis, or non-glaucomatous optic disc neuropathy were excluded. **Spectrum of glaucoma severity**: the mean ± SD MD and PSD on the VF test were -3.08 ± 1.61 and 4.29 ± 2.64 respectively, for early glaucoma; -1.02 ± 1.29 and 1.87 ± 0.5 respectively for preperimetric glaucoma.

Kim 2014b *(Continued)*

	Control participants: first-degree relatives with glaucoma, no history or evidence of intraocular surgery, IOP < 22 mmHg, a normal optic disc appearance and ophthalmic findings.
Index tests	**Optical coherence tomography**: Cirrus OCT (Carl Zeiss Meditec, Dublin, CA, USA); software version 6.0. Poor-quality OCT images such as those with low signal strength (< 70), motion artefact, or decentration were excluded. 7 x 7 mm scanning disc protocol was used to analyse RNFL and GCC parameters. The authors declare no conflict of interest
Target condition and reference standard(s)	**Manifest perimetric glaucoma:** glaucomatous VF results (defined as a cluster of 3 points with P < 5% on the pattern deviation map in at least 1 hemifield, including at least 1 point with P < 1%; or a cluster of 2 points with P < 1%, and GHT or PSD outside normal limits) and glaucomatous ONH/RNFL appearance (neuroretinal rim loss or notching, focal thinning of the NFL, disc haemorrhages, or vertical elongation of the optic cup). **Manifest preperimetric glaucoma**: glaucomatous ONH/RNFL appearance (neuroretinal rim loss or notching, focal thinning of the NFL, disc haemorrhages, or vertical elongation of the optic cup) with normal VF results. **Visual field test:** Humphrey Field Analyzer (Carl Zeiss Meditec, Dublin, CA, USA) 30-2 SITA standard programme. The fixation losses < 20 %, and false-positive and false-negative errors < 15 %, were considered as reliable. **Optic disc/RNFL evaluation:** dilated funduscopy using a 78-D lens and stereoscopic optic disc photography.
Flow and timing	No details reported about exclusion and time interval between index and reference test.
Comparative	
Notes	None.

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Was a case-control design avoided?	No		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (All tests)			

Kim 2014b *(Continued)*

If a threshold was used, was it pre-specified?	Yes		
Were imaging test's quality assessed?	Yes		
Were any conflict of interest avoided	Yes		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Unclear		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Unclear		
Did all patients receive a reference standard	Yes		
Could the patient flow have introduced bias?		Unclear risk	

Kita 2013

Study characteristics	
Patient Sampling	Case-control study including glaucoma, and healthy controls. preperimetric glaucoma and healthy controls. One eye per person was randomly selected.

Kita 2013 *(Continued)*

Patient characteristics and setting	**Sample size**: 134 eyes of 134 participants (33 advanced glaucoma, 66 early glaucoma, 35 normal controls). **Age**: mean ± SD: advanced glaucoma 56.6 ± 10.5 years; early glaucoma 54.3 ± 10.9 years; controls, 50.7 ± 12.2 years. **Sex**: 52 men (12 advanced glaucoma, 23 early glaucoma, 17 controls) and 82 women (21 advanced glaucoma, 43 early glaucoma, 18 controls). **Ethnicity**: Japanese. **Setting:** Department of Ophthalmology, Toho University Ohashi Medical Center, Tokyo, between October 2009 and March 2011. **Country**: Japan **Ocular comorbidities**: patients with diseases that affected the visual field (e.g. pituitary lesions, demyelinating diseases, or diabetic retinopathy), retinal pathology, previous retinal laser procedures, or if they had any previous ocular surgeries, neurological disease, or a history of diabetes, were excluded. **Spectrum of glaucoma severity**: the mean ± SD MD on the VF test were -10.69 ± 3.7, for advanced glaucoma; -2.89 ± 1.74 for early glaucoma. **Control participants:** IOP < 21 mmHg, a normal ONH appearance, normal open anterior chamber angles, normal VF results for the GHT.
Index tests	**Optical coherence tomography**: RTVue-100 (software version 4.0.5.39; Optovue Inc., Fremont, CA, USA). Images with a signal strength < 45 due to media opacity, patient positioning, or excessive eye movement were excluded. GCC and ONH scanning protocol were used for the analysis. One authors received research support from manufacturer.
Target condition and reference standard(s)	**Manifest perimetric glaucoma:** glaucomatous optic neuropathy (defined as a neuroretinal rim narrowing of the optic disc margin with notching, excavation, or a visible RNFL defect) and VF glaucomatous defects (defined as a cluster of 3+ contiguous points in the pattern deviation plot with P < 5%, with at least 1 P < 1%, and GHT outside normal limits). **Visual field test:** Humphrey Field Analyzer (Carl Zeiss Meditec, Dublin, CA, USA) 30-2 and 24-2 SITA standard programme. The fixation losses < 20 %, and false-positive and false-negative errors were < 25 %, were considered as reliable. **Optic disc/RNFL evaluation: s**tereoscopic fundus examination.
Flow and timing	No details about exclusion reported. Index test and reference standard were performed within 3 months.
Comparative	
Notes	None.

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			

Kita 2013 *(Continued)*

Was a consecutive or random sample of patients enrolled?	Unclear		
Was a case-control design avoided?	No		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (All tests)			
If a threshold was used, was it pre-specified?	Yes		
Were imaging test's quality assessed?	Yes		
Were any conflict of interest avoided	No		
Could the conduct or interpretation of the index test have introduced bias?		High risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	No		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Unclear		

Kita 2013 *(Continued)*

Did all patients receive a reference standard	Yes
Could the patient flow have introduced bias?	High risk

Koh 2014

Study characteristics	
Patient Sampling	Case-control study in which glaucoma patients seen by a glaucoma specialist were consecutively enrolled during the period from May 2012 to October 2012 at the glaucoma clinic at Kim's Eye Hospital. Healthy control were recruited from among those who visited the clinic during the enrolment period for an annual health examination. One eye per person was included.
Patient characteristics and setting	**Sample size**: 110 eyes of 110 participants (60 glaucoma and 50 healthy controls). **Age**: glaucoma mean ± SD, 60.7 ± 13.9 years; controls, 58.5 ± 14.9 years. **Sex**: 50 men (27 glaucoma, 23 controls) and 60 women (33 glaucoma, 27 controls). **Ethnicity**: not reported. **Clinical Setting:** Glaucoma clinic at Kim's Eye Hospital, Seul. **Country**: Korea. **Ocular comorbidities**: patients with concurrent retinal disease (i.e. secondary to a vascular disorder, macular degeneration), optic nerve disease other than glaucoma, or a brain disorder that could influence VF results, or media opacity, were excluded. **Spectrum of glaucoma severity**: the median (1st and 3rd quartiles) MD and PSD on the VF test were -7.64 (-10.69 to -3.84) and 6.92 (4.75 to 8.81) respectively, for glaucomatous eyes. **Control participants:** IOP < 21 mmHg, normal anterior chamber and open angle, a normal ONH without glaucomatous changes; no RNFL defect on red-free fundus photography; and normal reliable VF test results.
Index tests	**Optical coherence tomography**: Cirrus HD-OCT (Carl Zeiss Meditec, Dublin, CA, USA). Optic disc cube 200 x 200 scan protocol was used for the analysis. **Optical coherence tomography**: Spectral OCT/scanning laser ophthalmoscopy (OPKO/OTI, Miami, FL, USA). Scan circle centred on the optic disc. All images had to have signal strength ≥ 6 and no motion artefacts. The authors report no conflicts of interest.
Target condition and reference standard(s)	**Manifest glaucoma:** normal anterior segment on slit-lamp examination, glaucomatous ONH appearance (increased cup-disc ratio and narrowing of the neuroretinal rim), RNFL defects on red-free fundus photography(dark wedge-shaped area with its apex touching the optic disc border in the brightly-striated pattern of the surrounding RNFL or a generalised loss of RNFL visibility in the upper or lower retina) and glaucomatous VF defects (a cluster of 3 points with P < 5% on the PD map in at least 1 hemifield, including at least 1 point with P < 1% or a cluster of 2 points with P < 1%, or GHT outside normal limits, or a PSD with P < 5%).

Koh 2014 *(Continued)*

	Visual field test: Humphrey Field Analyzer (Carl Zeiss Meditec); 24-2 SITA–standard strategy. All exams had fixation losses and false-positive and false-negative rates of < 15%.
Flow and timing	No details about exclusion were reported. The index and reference test were performed on the same day
Comparative	
Notes	None.

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	No		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (All tests)			
If a threshold was used, was it pre-specified?	Yes		
Were imaging test's quality assessed?	Yes		
Were any conflict of interest avoided	Yes		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		

Koh 2014 *(Continued)*

Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk
Are there concerns that the target condition as defined by the reference standard does not match the question?		Low concern
DOMAIN 4: Flow and Timing		
Was there an appropriate interval between index test and reference standard?	Yes	
Did all patients receive the same reference standard?	Yes	
Were all patients included in the analysis?	Unclear	
Did all patients receive a reference standard	Yes	
Could the patient flow have introduced bias?		Unclear risk

Kook 2005

Study characteristics	
Patient Sampling	Cases were recruited prospectively in a consecutive manner and examined between April 2003 and September 2004. The control group consisted of clinic staff, friends or spouses of patients, or volunteers from other specialty clinics. One eye per person was selected.
Patient characteristics and setting	**Sample size**: 136 eyes of 136 participants (70 glaucoma, 66 healthy controls). **Age:** glaucoma patients mean ± SD, 55.11 ± 10.49 years; controls 52.15 ± 11.81. **Sex:** 60 men (39 glaucoma, 21 controls) and 76 women (31 glaucoma, 45 controls). **Country:** Korea. **Ocular comorbidities**: no retinal pathology, BCVA < 20/30, spherical refraction > ±5 D, cylinder refraction > ±3 D, history of laser or intraocular surgery, intracranial abnormalities, or a lesion revealed by neurological examination. **Setting:** Asian Medical Center, University of Ulsan, Seoul. **Spectrum of glaucoma severity:** mean ± SD MD and PSD on the VF test were −4.59 ± 3.25 dB and 6.72 ± 3.08 dB, respectively. **Control participants:** normal VF, absence of glaucomatous ONH appearance, multiple IOPs < 21 mmHg.
Index tests	**Scanning laser polarimetry:** GDx VCC, software version 5.3.1 (Laser Diagnostic Technologies, Dublin, CA, USA). Only scans of high quality were used in the study (centred optic disc, well-focused even illumination throughout the fundus image, and no motion artefacts). Only eyes with a scan quality score of 8+ were analysed. Index tests were reviewed independently by 2 glaucoma specialists in a blinded fashion.

Kook 2005 *(Continued)*

	No author had conflict of interest.
Target condition and reference standard(s)	**Manifest glaucoma:** glaucomatous optic nerve appearance (excavation, neuroretinal rim thinning or notching, or asymmetry of the vertical cup-to-disc ratio of > 0.2), reproducible VF defects (defined as a GHT test result outside normal limits or as a CPSD outside 95% of normal limits) with localised VF loss confined to 1 side of the horizontal meridian on the HFA (more than 3 adjacent points with P < 0.05 in a pattern deviation probability map or > 2 adjacent points with P < 0.02, only in 1 side of the horizontal meridian) and normal anterior chambers on gonioscopy. **Visual field testing:** Humphrey Field Analyzer, 24-2 full threshold test strategy (Zeiss-Humphrey, Dublin, CA, USA). VF reliability criteria included fixation losses rates of < 20% and false-negative and false-positive rates of < 15%. **Optic disc evaluation:** simultaneous stereophotographs were assessed by 2 independent graders. Reference standard tests were review in a blind fashion.
Flow and timing	The time interval between reference standard and index test was not reported. 16 participants had poor-quality index or reference test results and were excluded from the analysis.
Comparative	
Notes	None.

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	No		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (All tests)			
If a threshold was used, was it pre-specified?	Yes		
Were imaging test's quality assessed?	Yes		
Were any conflict of interest avoided	Yes		

Kook 2005 *(Continued)*

Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Unclear		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	No		
Did all patients receive a reference standard	Yes		
Could the patient flow have introduced bias?		High risk	

Kotowski 2012

Study characteristics	
Patient Sampling	Healthy, glaucoma suspect and glaucoma patients were selected among those recruited in the 'Pittsburgh Imaging Technology Trial study' (a prospective longitudinal study designed to assess ocular structure over time). No details about methods of patient selection. Right eye was selected for each patient fitting the inclusion criteria.
Patient characteristics and setting	**Sample size**: 166 participants evaluated, 163 eyes of 163 participants included in the analysis (63 glaucoma, 49 glaucoma suspects, 51 controls). **Age**: glaucoma eyes mean 64.3 years; glaucoma suspects mean 61.6 years; controls 54.8 years.

Kotowski 2012 *(Continued)*

	Sex: 61 men (24 glaucoma, 18 glaucoma suspects, 19 controls) and 102 women (39 glaucoma, 31 glaucoma suspects, 32 controls). **Ethnicity**: not reported. **Country**: USA. **Setting**: University of Pittsburgh Medical Center Eye Center, Pittsburgh, PA. **Ocular comorbidities**: patients with history of diabetes, any macular pathology, conditions affecting VF other than glaucoma, previous ocular trauma or surgery other than glaucoma interventions or uncomplicated cataract extraction were excluded. Participants had to have visual acuity ≥ 20/40, refractive error between −6 and +3 D, and no visually significant media opacities. **Spectrum of glaucoma severity**: the median (IQR) MD and PSD on the VF test were -2.21 (-6.92 to -0,35) dB and 2.99 (1.65 to 8.84) dB respectively,for glaucoma. **Control participants:** normal findings on ocular exam, no history of elevated IOP and normal VF result (defined as MD and PSD within 95% limits of the normal population, and GHT within normal limit).
Index tests	**Optical coherence tomography**: Cirrus HD-OCT (software version 5.0; Carl Zeiss Meditec). The macular cube 200 x 200 and optic disc cube 200 x 200 were used. Image with signal strength < 7, motion artefacts or with segmentation errors were excluded. One author had potential conflict of interest.
Target condition and reference standard(s)	**Manifest glaucoma:** glaucomatous VF result (defined as a PSD outside of the 95% limits of the normal population or GHT outside normal limits) associated with abnormal optic disc appearance (rim notching, cup asymmetry, vertical cup to disc ratio > 0.7), RNFL defect or IOP > 21 mmHg. **Visual field testing:** Humphrey Field Analyzer (24-2 SITA standard programme (Carl Zeiss Meditec, Dublin, CA, USA). VF reliability criteria were fixation losses, false-positive and false-negative rates < 30%. **Optic disc evaluation:** not reported.
Flow and timing	Index test and reference standard were performed at the same visit. 3 eyes (< 10%) were excluded due to failure of the segmentation algorithm.
Comparative	
Notes	Supported in part by the National Institute of Health grants R01-EY13178 and P30-EY08098 (Bethesda, MD), The Eye and Ear Foundation (Pittsburgh, PA) and an unrestricted grant from Research to Prevent Blindness (New York, NY).

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Was a case-control design avoided?	No		
Did the study avoid inappropriate exclusions?	Yes		

Kotowski 2012 *(Continued)*

Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (All tests)			
If a threshold was used, was it prespecified?	Yes		
Were imaging test's quality assessed?	Yes		
Were any conflict of interest avoided	No		
Could the conduct or interpretation of the index test have introduced bias?		High risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Unclear risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
Did all patients receive a reference standard	Yes		

Kotowski 2012 *(Continued)*

Could the patient flow have introduced bias?	Low risk

Kratz 2014

Study characteristics	
Patient Sampling	Case-control study enrolling glaucoma and healthy participants recruited from January 2010 to December 2010 at the Sydney Eye Hospital, Sydney, Australia. One eye from each person was selected randomly if both eyes were eligible.
Patient characteristics and setting	**Sample size**: 173 eyes of 173 participants (85 glaucoma and 88 healthy controls). **Age**: glaucoma mean ± SD, 69.96 ± 1.13 years; controls, 67.38 ± 11.97 years. **Sex**: 90 men (50 glaucoma, 40 controls) and 83 women (35 glaucoma, 48 controls) **Ethnicity**: not reported. **Clinical Setting:** Sydney Eye Hospital, Sydney. **Country**: Australia. **Ocular comorbidities**: patient with clinical evidence of macular disease, past refractive or retinal surgery, neurologic pathology or diabetes were excluded. **Spectrum of glaucoma severity**: the mean ± SD MD and PSD on the VF test were -7.89 ± 7.03 and 6.45 ± 3.64 respectively, for glaucomatous eyes. **Control participants:** normal VF, and no history of IOP > 21 mmHg.
Index tests	**Optical coherence tomography**: Cirrus HD-OCT software (Version 5.1.0.96, Carl Zeiss Meditec, Inc., Dublin, CA, USA). Optic disc cube 200 x 200 scan protocol was used for the analysis. Scans with movement artefact or signal strength < 7 were excluded. **Confocal scanning laser ophthalmoscopy**: HRT3 (HRT; Heidelberg Engineering, GmbH, Dossenheim, Germany) Experienced examiners outlined the optic disc margin on the mean topographic image. All participants had image quality SD < 30 µm. The authors stated no conflicts of interest.
Target condition and reference standard(s)	**Manifest perimetric glaucoma:** glaucomatous VF defect, defined as GHT outside normal limits, or PSD with P < 5% or a cluster of 3+ points in the PD plot in a single hemifield (superior or inferior) with P < 5%, 1 needed a P < 1%. **Visual field test:** Humphrey Field Analyzer II (Carl Zeiss Meditec, Dublin, CA, USA); 24-2 SITA–standard strategy. All exams had fixation losses and false-positive and false-negative rates of < 20%. Imaging and VF tests were performed by trained technicians masked to other clinical information at the same visit.
Flow and timing	The index and reference test were performed on the same day.
Comparative	
Notes	None.
Methodological quality	

Kratz 2014 *(Continued)*

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	No		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (All tests)			
If a threshold was used, was it pre-specified?	Yes		
Were imaging test's quality assessed?	Yes		
Were any conflict of interest avoided	Yes		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		

Kratz 2014 *(Continued)*

Were all patients included in the analysis?	Yes
Did all patients receive a reference standard	Yes
Could the patient flow have introduced bias?	Low risk

Lee 2010

Study characteristics	
Patient Sampling	Healthy and glaucomatous participants who met the eligibility criteria were recruited prospectively between March 2008 and March 2009. One eye per person was randomly selected.
Patient characteristics and setting	**Sample size**: 165 eyes of 165 participants (88 glaucoma, 77 controls). **Age:** glaucoma patients mean ± SD, 53.7 ± 10.8 years; controls 51.7 ± 11.4. **Sex:** 87 men (39 controls, 48 glaucoma), and 78 women (38 controls, 40 glaucoma). **Ethnicity:** Korean. **Country:** Korea. **Ocular comorbidities**: no ocular pathologies other than glaucoma, BCVA < 20/30, spherical refraction > ±5 D, cylinder refraction > ±3 D, diabetes or closed angle at gonioscopy. **Setting:** Asan Medical Center, Seoul. **Spectrum of glaucoma severity:** mean ± SD MD and PSD on the VF test were -6.33 ± 4.79 dB and 6.7 ± 4.12 dB, respectively. **Control participants:** normal optic disc appearance, normal VF result, and IOP < 22 mmHg.
Index tests	**Scanning laser polarimetry:** GDx VCC (Carl Zeiss Meditec, Inc, Dublin, CA, USA). All images were acquired by a single well-trained operator. The pupils were dilated if their diameter was < 3 mm. All poor-quality scans, defined as those with a quality score grade < 8 and an atypical retardation pattern with a typical scan score of < 80 were excluded. **Optical coherence tomography**: Cirrus HD-OCT, "optic disc cube" scan (Carl Zeiss Meditec, Inc, Dublin, CA, USA). All images were acquired by a single well-trained operator. The pupils were dilated if their diameter was < 3 mm. Images with signal strength < 6, overt misalignment of the surface detection algorithm on at least 15% of consecutive A-scans or 20% of cumulative A-scans or overt decentration of the measurement circle location, were excluded. No details about authors' conflict of interest were reported.
Target condition and reference standard(s)	**Manifest glaucoma**: glaucomatous VF defect (defined as a cluster of 3 points with a P < 5% on a pattern deviation map in at least 1 hemifield, including at least 1 point with a P < 1% or a cluster of 2 points with a probability of < 1% and a GHT or PSD outside 99% normal limits) and a glaucomatous ONH appearance (vertical cup disc ratio > 0.7, or a vertical cup–disc ratio asymmetry > 0.2 between eyes, or diffuse/focal neural rim thinning or haemorrhage).

Lee 2010 *(Continued)*

	Visual field testing: Humphrey Field Analyzer, 24-2 SITA-standard strategy (Carl Zeiss Meditec, Dublin, CA, USA). VF reliability criteria included fixation losses rates of < 20% and false-negative and false-positive rates of < 15%. **Optic disc evaluation:** stereoscopic optic nerve photography.
Flow and timing	The time interval between reference standard and index tests was < 2 weeks. 19 (> 10%) eyes were excluded due to poor SD-OCT or GDx VCC quality images.
Comparative	
Notes	None.

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Was a case-control design avoided?	No		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (All tests)			
If a threshold was used, was it pre-specified?	Yes		
Were imaging test's quality assessed?	Yes		
Were any conflict of interest avoided	Unclear		
Could the conduct or interpretation of the index test have introduced bias?		Unclear risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		

Lee 2010 *(Continued)*

Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	No		
Did all patients receive a reference standard	Yes		
Could the patient flow have introduced bias?		High risk	

Leite 2011

Study characteristics	
Patient Sampling	Participants were recruited from the longitudinal Diagnostic Innovations in Glaucoma Study and the African Descent and Evaluation Study. Healthy participants were recruited from the general population. No other details on methods of patient selection were reported. Both eyes of some participants were included in the study.
Patient characteristics and setting	**Sample size**: 233 eyes (126 glaucoma, 107 controls) of 149 participants (91 glaucoma, 58 controls). **Age**: glaucoma eyes mean ± SD 70 ± 10 years; controls 50 ± 19 years. **Sex**: 97 men (58 glaucoma, 39 controls) and 136 women (68 glaucoma, 68 controls). **Ethnicity**: 76 African-American (49 glaucoma, 27 controls) **Country**: USA. **Setting**: Hamilton Glaucoma Center, University of California, San Diego. **Ocular comorbidities**: patients with co-existing retinal disease, uveitis, or non-glaucomatous optic neuropathy were excluded. All eyes had to have BCVA ≥ 20/40, spherical refraction within ±5.0 D, cylinder correction within ±3.0 D, and open angles on gonioscopy. **Spectrum of glaucoma severity**: the mean (first, third quartile) MD and PSD on the VF test were -5.85 (-7.59, -2.16) dB and 5.36 (2.15, 7.95) dB respectively, for glaucoma. **Control participants:** IOP < 22 mmHg with no history of elevated IOP and at least 2 reliable normal VFs (defined as PSD within 95% confidence limits and a GHT result within normal limits).

Leite 2011 *(Continued)*

Index tests	**Optical coherence tomography**: Spectralis OCT (Spectralis HRA-OCT; software version 5.2.0.3) The RNFL 3.45 mm scan was used. Only images with well-centred scan and a signal strength > 15 dB were included. **Optical coherence tomography**: Cirrus (software version 4.5, Carl Zeiss Meditec Inc.). The optic disc cube scan was used. Only images with a well-centred scan, a signal strength > 6 dB and the absence of movement artefacts were included. **Optical coherence tomography**: RTVue (software version 4.0.5.39). The ONH map scan was used. Only images with a signal strength ≥ 30 were included. Some authors had conflict of interest.
Target condition and reference standard(s)	**Manifest glaucoma:** glaucomatous VF result defined as a PSD outside the 95% normal limits or a GHT result outside the 99% normal limits. ONH appearance was not part of the reference standard. **Visual field testing:** Humphrey Field Analyzer, 24-2 SITA standard programme (Carl Zeiss Meditec, Dublin, CA, USA). All VFs were reviewed by the "visual field reading center", in order to check for artefacts or inappropriate fixation.
Flow and timing	Index tests were performed on the same day but the time interval between index tests and reference standard was not reported. No patients were reported by the authors excluded from the analysis .
Comparative	
Notes	Supported in part by National Eye Institute R01-EY08208 (FAM) and R01-11008 (LMZ), and CAPES grant BEX1327/09-7 (MTL). Participant retention incentive grants in the form of glaucoma medication at no cost (Alcon Laboratories Inc., Allergan, Pfizer Inc., and SANTEN Inc.).

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Was a case-control design avoided?	No		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (All tests)			
If a threshold was used, was it prespecified?	Yes		

Leite 2011 *(Continued)*

Were imaging test's quality assessed?	Yes		
Were any conflict of interest avoided	No		
Could the conduct or interpretation of the index test have introduced bias?		High risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Unclear		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
Did all patients receive a reference standard	Yes		
Could the patient flow have introduced bias?		Unclear risk	

Leung 2010

Study characteristics	
Patient Sampling	Normal participants and glaucoma patients were enrolled consecutively from August 2008 to February 2009. One eye per person was randomly selected.
Patient characteristics and setting	**Sample size**: 223 eyes of 223 participants (121 glaucoma, 102 healthy controls).

Leung 2010 *(Continued)*

	Age: perimetric glaucoma patients mean ± SD 54 ± 14.6 years; controls 50.3 ± 10.3 years. **Ethnicity:** Chinese. **Country:** China. **Ocular comorbidities**: no macular diseases, BCVA < 20/40, spherical refraction < -8 D or > +4 D, refractive or retinal surgery, neurologic diseases, or diabetes. **Setting:** University Eye Center at the Chinese University of Hong Kong. **Spectrum of glaucoma severity:** mean ± SD MD and PSD on the VF test were -8.99 ± 8.16 dB and 6.86 ± 4.12 dB, respectively. According to the Hodapp et al. grading scale, 63 eyes had early glaucoma, 58 moderate to advanced. **Control participants:** normal VF and no history of IOP > 21 mmHg.
Index tests	**Optical coherence tomography**: Cirrus HD-OCT, "optic disc cube" scan protocol software version 3.0 (Carl Zeiss Meditec Inc.). All the OCT scans had a signal strength of > 7. Saccadic eye movement was detected in the line-scanning ophthalmoscope overlaid with OCT en face during OCT imaging. Images with motion artefact were re-scanned at the same visit. Some authors had conflict of interest.
Target condition and reference standard(s)	**Manifest glaucoma:** glaucomatous VF defects (defined as ≥ 3 significant (P < 0.05) non-edge contiguous points with ≥ 1 at the P < 0.01 level on the same side of horizontal meridian in the pattern deviation plot and confirmed with ≥ 2 consecutive examinations). **Visual field testing:** Humphrey Field Analyzer, model II, 24-2 SITA-standard strategy (Carl Zeiss Meditec, Inc., Dublin, CA). VF reliability criteria included fixation losses rates, false-negative and false-positive rates of < 20%. Reference standard was performed by investigators masked to other clinical information. Optic disc appearance was not part of the reference standard.
Flow and timing	Reference standard and index tests were performed at the same visit. A total of 223 participants (102 normal subjects and 121 glaucoma patients) were enrolled consecutively. Authors stated that 5 subjects were excluded in the study (3 had low strength in Cirrus HD-OCT imaging and 2 had an epiretinal membrane at the macula evident in the OCT scan) but still 223 participants were included in the analysis.
Comparative	
Notes	None.

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	No		
Did the study avoid inappropriate exclusions?	Yes		

Leung 2010 *(Continued)*

Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (All tests)			
If a threshold was used, was it pre-specified?	Yes		
Were imaging test's quality assessed?	Yes		
Were any conflict of interest avoided	No		
Could the conduct or interpretation of the index test have introduced bias?		High risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Unclear		
Did all patients receive a reference standard	Yes		
Could the patient flow have introduced bias?		Unclear risk	

Lisboa 2013

Study characteristics	
Patient Sampling	A cohort of participants suspected of having glaucoma was selected from the Diagnostic Innovations in Glaucoma Study database, and followed for at least 5 years. A documented evidence of progressive glaucomatous change in the appearance of the optic disc was used as reference standard. Participants with progressive optic disc damage and no visual field loss were included in the preperimetric glaucoma group. Patients followed untreated for about 14 years without any evidence of progressive change in the appearance of the optic disc or visual field loss were used as the control group. Both eyes were selected for some patients.
Patient characteristics and setting	**Sample size**: 142 eyes (48 glaucoma, 94 controls) of 91 participants. **Age**: glaucoma eyes mean ± SD 65.9 ± 9.1 years; controls 64.2 ± 11.2 years. **Sex**: glaucoma: male 53%; controls: male 31% **Ethnicity**: 12 African-American (8 glaucoma, 4 controls). **Country**: USA. **Setting**: Hamilton Glaucoma Center, University of California, San Diego. **Ocular comorbidities**: patients with co-existing retinal disease, uveitis, or non-glaucomatous optic neuropathy were excluded. All eyes had to have BCVA ≥ 20/40, spherical refraction within ±5.0 D, cylinder correction within ±3.0 D, and open angles on gonioscopy. **Spectrum of glaucoma severity**: the mean (first, third quartile) MD and PSD on the VF test were -0.81 (-1.82, 0.12) dB and 1.75 (1.46, 1.84) dB respectively, for glaucoma. **Control participants:** participants followed untreated for a long period (13.6 ± 3.6 years) without any evidence of progressive change in the appearance of the optic disc or VF loss in both eyes.
Index tests	**Optical coherence tomography**: RTVue (software version 6.1.0.4; Optovue, Inc., Fremont, CA, USA). The ONH protocol and ganglion cell complex scanning protocols were used. Only good-quality images, as defined by a signal strength index ≥ 28 for RNFL and ONH measurements, and ≥ 32 for macular measurements were included in the analysis. Some authors had potential conflict of interest
Target condition and reference standard(s)	**Manifest preperimetric glaucoma:** documented evidence of progressive glaucomatous change in the appearance of the optic disc (based on focal or diffuse thinning of the neuroretinal rim, increased excavation, or enlargement of the RNFL defects) and normal VF result (defined as a MD and PSD within 95% confidence limits and a GHT result within normal limits). **Visual field testing:** 24-2 SITA standard programme (Carl Zeiss Meditec, Dublin, CA, USA). **Optic disc and RNFL evaluation:** stereoscopic optic disc photographs (TRC-SS, Topcon Instrument Corp. of America, Paramus, NJ). Stereoscopic sets of slides were examined using a stereoscopic viewer (Asahi, Pentax, Tokyo, Japan). 2 experienced graders, masked to the participant's identity, to other test results, and to the chronological sequence of the photographs, evaluated the stereophotographs.
Flow and timing	Reference standard was performed before index test but time interval between index test and reference standard was not reported. Index test different scanning protocols were performed within 6 months. No patients were reported by the authors as excluded from the analysis.
Comparative	
Notes	Supported in part by National Institutes of Health/National Eye Institute Grants EY021818 (FAM), EY11008 (LMZ), and EY14267 (LMZ); Coordena ̧ca ̃o de Aperfei ̧coamento de Pessoal de N´ıvel Superior (CAPES) grant Bolsas no Exterior (BEX) 1066/11-0; an unrestricted grant from

Lisboa 2013 *(Continued)*

Research to Prevent Blindness (New York, New York); and grants for participants' glaucoma medications from Alcon, Allergan, Pfizer, Merck, and Santen.

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		Low risk	
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 2: Index Test (All tests)			
If a threshold was used, was it prespecified?	Yes		
Were imaging test's quality assessed?	Yes		
Were any conflict of interest avoided	No		
Could the conduct or interpretation of the index test have introduced bias?		High risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes		

Lisboa 2013 *(Continued)*

Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	No		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
Did all patients receive a reference standard	Yes		
Could the patient flow have introduced bias?		High risk	

Mai 2007

Study characteristics	
Patient Sampling	Healthy controls and glaucoma patients were recruited. Controls were recruited consecutively either from an ongoing longitudinal follow-up study or from staff members, their friends and spouses, partners of the patients, or volunteers. No details on glaucoma patient selection method. One eye per person was randomly selected.
Patient characteristics and setting	**Sample size**: 133 eyes of 133 participants (92 glaucoma, 41 controls). **Age:** glaucoma patients mean ± SD, 65.4 ± 10.9 years; controls 61.2 ± 12.0. **Sex**: 73 men, 60 women. **Ethnicity:** white. **Country:** Netherland. **Ocular comorbidities**: no ocular disease other than glaucoma, BCVA < 20/40, spherical refraction < -7 D or > +3 D, intraocular surgery (except uncomplicated cataract surgery), diabetes mellitus or arterial hypertension. **Setting:** Rotterdam Eye Hospital, Rotterdam. **Spectrum of glaucoma severity:** mean ± SD of MD and PSD on the VF test for glaucoma were –9.4 ± 7.4 dB and 8.1 ± 3.9 dB, respectively. According to Hodapp et al. grading score 59 eyes had mild and moderate glaucoma, 33 severe.

Mai 2007 *(Continued)*

	Control participants: IOP < 21 mmHg, normal VF results (MD and PSD within 95% confidence limits and GHT within normal limits) and healthy-appearing ONH (no diffuse/local rim thinning, cupping, or optic disc haemorrhages).
Index tests	**Scanning laser polarimetry:** GDx VCC, software version 5.4.0, GDx-ECC, software version 5.5.0.11 (Carl Zeiss Meditec, Inc., Dublin, CA, USA). Images were acquired through undilated pupils, by 2 trained and experienced technicians following a standard protocol. Only images of high quality (with quality scan score ≥ 7) were selected. Some authors had conflict of interest.
Target condition and reference standard(s)	**Manifest glaucoma:** glaucomatous optic disc appearance (diffuse or local rim thinning or cupping), abnormal VF result (confirmed on 2 consecutive occasions and defined as 2 or more adjacent points at a P ≤ 0.01 level, or 3+ adjacent points at a P ≤ 0.05 level in the total deviation plot, or GHT outside normal limits) and open angle by gonioscopy. **Visual field testing:** Humphrey Field Analyzer, full threshold strategy (126 eyes), 24-2 SITA standard strategy (5 eyes), or 24-2 SITA-fast (2 eyes) (Carl Zeiss Meditec, Inc.). Reliability criteria included fixation losses rates of < 25% and false-positive rates of < 20%. Acceptable false-positive rate was < 20% and < 33% for controls and glaucoma respectively.
Flow and timing	The time interval between reference standard and index test was not reported. No patients were reported by the authors as excluded from the analysis.
Comparative	
Notes	None.

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Was a case-control design avoided?	No		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (All tests)			
If a threshold was used, was it pre-specified?	Yes		
Were imaging test's quality assessed?	Yes		

Mai 2007 *(Continued)*

Were any conflict of interest avoided	No		
Could the conduct or interpretation of the index test have introduced bias?		High risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Unclear		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
Did all patients receive a reference standard	Yes		
Could the patient flow have introduced bias?		Unclear risk	

Mansoori 2011

Study characteristics	
Patient Sampling	Cross-sectional study involving healthy and glaucoma participants. Glaucoma patients were recruited from patients attending glaucoma outpatient department, healthy controls were recruited from the staff of the same institute. One eye per person was randomly selected.
Patient characteristics and setting	**Sample size**: 178 eyes of 178 participants (83 glaucoma, 95 controls). **Age**: glaucoma eyes mean ± SD 57.1 ± 6.1 years; controls 56.9 ± 11 years. **Sex**: 79 men (40 glaucoma, 39 controls) and 99 women (43 glaucoma, 56 controls).

Mansoori 2011 *(Continued)*

	Ethnicity: Indian. **Country**: India. **Setting**: Department of Glaucoma, Pushpagiri Eye Institute, Andhra Pradesh. **Ocular comorbidities**: patients with family history of glaucoma, uveitis, corneal, retinal or macular pathology, neurological disease or abnormal disc appearance such as tilted disc or discs with peripapillary atrophy were excluded. All eyes had to have BCVA ≥ 20/30, spherical refraction within ±4.0 D, cylinder correction within ±2.0 D, clear ocular media and open angles on gonioscopy. **Spectrum of glaucoma severity**: the mean ± SD MD and PSD on the VF test were -4.6 ± 0.3 and 5.2 ± 0.7 respectively,for glaucoma. All glaucoma had MD > -6 dB. **Control participants:** IOP ≤ 21 mmHg, no past history of Increased IOP, normal optic disc and RNFL appearance and normal VF result (MD and PSD within 95% confidence limits and GHT within normal limits).
Index tests	**Optical coherence tomography**: OCT/SLO (OPKO/ OTI, Miami FL, USA). The RNFL scanning protocol after pupil dilation was used. A good-quality image required a signal strength > 7, a clear SLO image allowing optic disc and scan circle visibility, a dense colour saturation throughout all retinal layers and no algorithm failure. The authors stated no conflict of interest.
Target condition and reference standard(s)	**Manifest early glaucoma:** glaucomatous optic nerve damage and consistent VF loss (defined as the presence of a cluster of 3+ adjacent points on pattern deviation plot with a P < 5% with 1+ points with P < 1% and GHT outside normal limits), and IOP > 21 mmHg in > 2 occasions. **Visual field testing:** 24-2 SITA standard programme (Carl Zeiss Meditec, Dublin, CA, USA). **Optic disc and RNFL evaluation:** dilated fundus and optic disc examination with a +78 D lens.
Flow and timing	The time interval between index test and reference standard was not reported. No patients were reported by the authors as excluded from the analysis.
Comparative	
Notes	None.

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Was a case-control design avoided?	No		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		High risk	

Mansoori 2011 *(Continued)*

Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (All tests)			
If a threshold was used, was it pre-specified?	Yes		
Were imaging test's quality assessed?	Yes		
Were any conflict of interest avoided	Yes		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Unclear		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
Did all patients receive a reference standard	Yes		
Could the patient flow have introduced bias?		Unclear risk	

Medeiros 2004a

Study characteristics	
Patient Sampling	Patients' data were selected retrospectively from a research database, containing patients included in a prospective, longitudinal study designed to evaluate optic nerve structure and visual function in glaucoma. One eye per person was randomly selected.
Patient characteristics and setting	**Sample size**: 100 eligible patients, 114 included in the analysis (42 glaucoma patients, 32 glaucoma suspects and 40 healthy controls) **Age:** glaucoma patients mean ± SD 67 ± 11 years, glaucoma suspects 61 ± 12 years, controls 65 ± 11 years. **Ethnicity:** not specified **Country:** USA. **Ocular comorbidities**: no co-existing retinal disease, uveitis, or non-glaucomatous optic neuropathy. BCVA ≥ 20/40, spherical refraction within ±5.0 D, cylinder correction within ±3.0 D, and open angles on gonioscopy. **Setting:** Hamilton Glaucoma Center, University of California. **Spectrum of glaucoma severity:** mean MD on the VF test were -4.92 dB for glaucoma patient; According to the Hodapp- Parrish-Anderson grading scale, 27 patients were classified as having early defects, 9 had moderate defects and 6 had severe VF defects. **Control participants:** IOP ≤ 22 mmHg,with no history of increased IOP, a normal VF result and a healthy appearance of the optic disc and RNFL.
Index tests	**Scanning laser polarimetry:** GDx VCC, software version 5.0.1 (Laser Diagnostic Technologies Inc, San Diego, CA, USA). Good-quality image required a focused and evenly-illuminated reflectance image with a centred optic disc. Quality assessment was evaluated by an experienced examiner masked to the participant's identity and results of the other tests. One author had conflict of interest.
Target condition and reference standard(s)	**Manifest glaucoma:** repeatable (2 consecutive) abnormal VF test results, defined as a PSD outside the 95% normal GHT results outside 99% normal confidence limits, regardless of the appearance of the optic disc. **Glaucoma suspect**: ocular hypertension (IOP > 22 mmHg on more than 2 separate visits) or glaucomatous appearance of the optic disc (defined as neuroretinal rim thinning, excavation, notching, or characteristic RNFL defects). **Visual field testing:** Humphrey Field Analyzer, 24-2 full-threshold standard automated perimetry or SITA-standard programme (Carl Zeiss Meditec, Inc., Dublin, CA, USA). VF reliability criteria were not reported. **Optic disc evaluation:** stereoscopic optic disc photography.
Flow and timing	17 patients (> 10%) were not included in the final analysis due to poor-quality RNFL photograph or SLP image. All index tests were performed within 3 months, but no details about the time interval between index and reference test.
Comparative	
Notes	None.
Methodological quality	

Medeiros 2004a *(Continued)*

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Was a case-control design avoided?	No		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (All tests)			
If a threshold was used, was it pre-specified?	Yes		
Were imaging test's quality assessed?	Yes		
Were any conflict of interest avoided	No		
Could the conduct or interpretation of the index test have introduced bias?		High risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Unclear		

Medeiros 2004a *(Continued)*

Did all patients receive the same reference standard?	Yes	
Were all patients included in the analysis?	No	
Did all patients receive a reference standard	Yes	
Could the patient flow have introduced bias?		High risk

Medeiros 2004b

Study characteristics	
Patient Sampling	Patients were included in a prospective longitudinal study designed to evaluate optic nerve structure and visual function in glaucoma (Diagnostic Innovations in Glaucoma Study) from April 2002 to November 2003. All patients who met the inclusion criteria were enrolled in this study. One eye per person was randomly selected.
Patient characteristics and setting	**Sample size**: 183 eyes of 183 participants were enrolled, 141 eyes included in the analysis (75 glaucoma, 66 healthy controls). **Age:** glaucoma patients mean ± SD, 68 ± 10 years; controls 65 ± 8 years. **Country**: USA. **Setting:** Hamilton Glaucoma Center, University of California, San Diego. **Ocular comorbidities**: no co-existing retinal disease, close angle by gonioscopy, BCVA < 20/40, spherical refraction > ±5 D, cylinder refraction > ±3 D, uveitis, or non-glaucomatous optic neuropathy. **Spectrum of glaucoma severity:** mean ± SD MD on the VF test was −4.89 ± 3.9 dB. According to the Hodapp et al. grading scale, 53 eyes had early glaucoma,11 moderate and 11 severe. **Control participants:** IOP ≤ 22 mmHg, normal VF result (MD and PSD within 95% confidence limits and GHT within normal limits) and healthy ONH/RNFL appearance (no diffuse/focal rim thinning, cupping, optic disc haemorrhage, or RNFL defects).
Index tests	**Scanning laser polarimetry:** GDx VCC, software version 5.0.1 (Laser Diagnostic Technologies Inc, San Diego, CA, USA). Assessment of image quality was performed by an experienced examiner masked to the participant's identity and results from the other tests. Good-quality images required a focused and evenly-illuminated reflectance image with a centred optic disc, a residual anterior segment retardation of 15 nm or less and an atypical scan score < 25. One author had conflict of interest.
Target condition and reference standard(s)	**Manifest glaucoma:** repeated (2 consecutive) glaucomatous VF loss defined as a PSD with P < 5% or a GHT outside normal limits. **Visual field testing:** Humphrey Field Analyzer, 24-2 SITA-standard strategy (Carl Zeiss Meditec, Inc). VF reliability criteria were not reported. Optic disc appearance was not part of the reference standard.
Flow and timing	Reference standard and index tests were performed within 6 months.

Medeiros 2004b *(Continued)*

	42 of 183 participants(> 10%) had unacceptable-quality imaging scans and were not included in the analysis.
Comparative	
Notes	None.

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Was a case-control design avoided?	No		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (All tests)			
If a threshold was used, was it pre-specified?	Yes		
Were imaging test's quality assessed?	Yes		
Were any conflict of interest avoided	No		
Could the conduct or interpretation of the index test have introduced bias?		High risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern

Medeiros 2004b *(Continued)*

DOMAIN 4: Flow and Timing		
Was there an appropriate interval between index test and reference standard?	No	
Did all patients receive the same reference standard?	Yes	
Were all patients included in the analysis?	No	
Did all patients receive a reference standard	Yes	
Could the patient flow have introduced bias?		High risk

Medeiros 2005

Study characteristics	
Patient Sampling	Patients' data were selected from a research database, containing patients included in a prospective, longitudinal study designed to evaluate optic nerve structure and visual function in glaucoma. Normal participants were recruited from the staff and employees of the University of California, as well as from the general population. One eye per person was randomly selected.
Patient characteristics and setting	**Sample size**: 136 patients (41 perimetric glaucoma, 30 preperimetric glaucoma, 65 healthy controls). **Age:** perimetric glaucoma patients mean ± SD, 65 ± 9 years, preperimetric glaucoma 70 ± 11 years, controls 66 ± 11 years. **Ethnicity:** not specified. **Country:** USA. **Ocular comorbidities**: no co-existing retinal disease, uveitis, or non-glaucomatous optic neuropathy. BCVA ≥ 20/40, spherical refraction within ±5.0 D, cylinder correction within ±3.0 D. **Setting:** Hamilton Glaucoma Center, University of California. **Spectrum of glaucoma severity:** mean ± SD MD and PSD on the VF test were -7.53 ± 6.58 dB and 7.13 ± 3.60 dB for perimetric glaucoma, -2.07 ± 1.65 dB and 1.65 ± 0.3 dB for preperimetric glaucoma, -0.59 ± 1.13 dB and 1.59 ± 0.38 dB for control group, respectively. **Control participants:** IOP ≤ 22 mmHg,with no history of increased IOP, a normal VF result and a normal clinical examination.
Index tests	**Scanning laser polarimetry:** GDx VCC, software version 5.0.1 (Laser Diagnostic Technologies Inc, San Diego, CA, USA). To be acceptable each image required a focused and evenly-illuminated reflectance image with a centred optic disc, residual anterior segment retardation ≤ 15 nm and an atypical scan score > 25. Quality assessment was performed by an experienced examiner masked to the participant's identity and results of the other tests. No details about conflict of interest were reported.
Target condition and reference standard(s)	**Manifest perimetric glaucoma:** evidence of progressive glaucomatous change in the appearance of the optic disc (as assessed by simultaneous stereoscopic optic disc photographs and defined by focal or diffuse thinning of the neuroretinal rim, increased excavation, or enlarge-

Medeiros 2005 *(Continued)*

	ment of RNFL defects) and abnormal VF result (GHT outside normal limits or a PSD with P < 5%). **Manifest preperimetric glaucoma:** evidence of progressive glaucomatous change in the appearance of the optic disc (as assessed by simultaneous stereoscopic optic disc photographs and defined by focal or diffuse thinning of the neuroretinal rim, increased excavation, or enlargement of RNFL defects) and normal VF result. **Optic disc evaluation**: stereoscopic optic disc photographs were acquired with TRC-SS (Topcon, Paramus, New Jersey, USA) and included only if had a good quality. For each participant, the most recent stereophotograph was compared with the oldest available (at least 1 year time interval) by 2 experienced graders masked to the participant's identity and to the temporal sequence of the photographs. **Visual field testing:** Humphrey Field Analyzer,24-2 SITA standard (Zeiss-Humphrey, Dublin, CA, USA).
Flow and timing	No patients were reported by the authors as excluded from the analysis. The GDx VCC imaging date was always after the date of the optic disk stereophotograph that showed progression.
Comparative	
Notes	None.

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Was a case-control design avoided?	No		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (All tests)			
If a threshold was used, was it prespecified?	Yes		
Were imaging test's quality assessed?	Yes		
Were any conflict of interest avoided	Unclear		

Medeiros 2005 *(Continued)*

Could the conduct or interpretation of the index test have introduced bias?		Unclear risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Unclear		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
Did all patients receive a reference standard	Yes		
Could the patient flow have introduced bias?		Unclear risk	

Moreno 2011

Study characteristics	
Patient Sampling	Healthy controls and early glaucoma patients were prospectively and consecutively enrolled. One eye per person was randomly selected.

Moreno 2011 *(Continued)*

Patient characteristics and setting	**Sample size**: 123 eyes of 123 participants (67 glaucoma, 56 controls). **Age**: glaucoma eyes mean ± SD 64.3 ± 11.8 years; controls 56.5 ± 12.9 years. **Sex**: 49 men (27 glaucoma, 22 controls) and 74 women (40 glaucoma, 34 controls). **Ethnicity**: 65 white (36 glaucoma, 29 controls), 35 African descent (19 glaucoma, 16 controls), 23 mixed (12 glaucoma, 11 controls). **Country**: Brazil. **Setting**: not specified. **Ocular comorbidities**: patients with previous ocular surgery or trauma, spherical equivalent > ±4.0 D, history of using oral or topical steroids, and any ocular disease other than glaucoma including moderate or advanced cataract, were excluded. **Spectrum of glaucoma severity**: the mean ± SD MD on the VF test were -2.5 ± 1.6 dB,for glaucoma. All glaucoma patients had MD > -6 dB. **Control participants:** IOP < 21 mmHg, normal VF results and no glaucomatous optic neuropathy.
Index tests	**Optical coherence tomography**: RTVue-100 OCT (software version A4, Optovue, Fremont, CA, USA). The GCC and RNFL 3.45 mm scanning protocols were used. Images with signal strength indices < 40 or not well centred were excluded. All images were acquired by a single experienced operator who was masked to patients' clinical data. The authors stated no conflict of interest.
Target condition and reference standard(s)	**Manifest early glaucoma:** glaucomatous optic neuropathy (defined as a vertical cup-to-disc ratio of ≥ 0.6, asymmetry of cup-to-disc ratio ≥ 0.2 between eyes, and presence of localised RNFL defects or neuroretinal rim defects or both) and glaucomatous VF defects (defined as 3+ points in clusters, with a P < 5% on the pattern deviation plot (excluding those on the edge of the field or directly above or below the blind spot), a PSD with a P < 5%, or a GHT results outside the normal limits). **Visual field testing:** Humphrey Field Analzyer(24-2 SITA standard programme (Carl Zeiss Meditec, Dublin, CA, USA). Reliability criteria were not reported. **Optic disc evaluation:** funduscopy and stereophotograph assessment.
Flow and timing	The time interval between index test and reference standard was not reported. No patient were reported by the authors as excluded from the analysis.
Comparative	
Notes	None.

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	No		

Moreno 2011 *(Continued)*

Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (All tests)			
If a threshold was used, was it pre-specified?	Yes		
Were imaging test's quality assessed?	Yes		
Were any conflict of interest avoided	Yes		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Unclear		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
Did all patients receive a reference standard	Yes		
Could the patient flow have introduced bias?		Unclear risk	

Moreno-Montañés 2008

Study characteristics	
Patient Sampling	Healthy, ocular hypertensive and glaucoma participants were consecutively enrolled. One eye per person was randomly selected.
Patient characteristics and setting	**Sample size**: 182 eyes of 182 participants (83 glaucoma, 40 OHT, 59 healthy controls). **Age:** glaucoma patients mean (range), 68 (60 to 73) years; hypertensive 63.5 (57 to 70.5); controls 56 (47 to 67). **Sex:** 87 men (45 glaucoma, 16 OHT, 26 controls) and 95 women (38 glaucoma, 24 hypertensive, 33 controls). **Ethnicity:** white. **Country:** Spain. **Ocular comorbidities**: no corneal/retinal disease, BCVA < 20/40, spherical equivalent > ±5 D, no substantial media opacity. **Setting:** Department of Ophthalmology, Clínica Universitaria de Navarra, Pamplona; Institut Catalá de la Retina, Barcelona. **Spectrum of glaucoma severity:** mean (range) MD/PSD on the VF test were -4.94 (-12.58 to -2.67)/4.29 (2.15 to 8.34) dB, for glaucoma eyes; -0.99 (-2.52 to -0.29)/1.5(1.40 to 1.87) dB for OHT eyes. **Control participants:** IOP ≤ 21 mmHg, normal VF, and no familiar glaucoma.
Index tests	**Confocal scanning laser ophthalmoscopy**: HRT 3, software version 3.0 (Heidelberg Engineering, Dossenheim, Germany). All images were acquired after pupil dilation and were of good quality, defined as having a topographic SD of ≤ 30 µm. Contour lines were placed in the margin of the optic disk by experienced users and were reviewed by 2 authors. No author had conflict of interest.
Target condition and reference standard(s)	**Manifest glaucoma:** glaucomatous defects reproducible in at least 3 reliable and consecutive VFs (defined as at least 3 contiguous locations were outside the 95% normal limits of the pattern deviation plot and 1 was outside the 99% normal limits), with open angle at gonioscopy. **OHT:** IOP > 21 mmHg on 3 different days, with 3 consecutive normal VFs. **Visual field testing:** Humphrey Field Analyzer, 24-2 SITA-standard strategy (Carl Zeiss Meditec, Dublin, CA,USA). Reliability criteria included fixation losses rates, false-positive and false-negative rates of < 30%. The optic disc appearance was not part of the reference standard.
Flow and timing	Reference standard and index test were performed on the same day. A total of 182 eyes were enrolled. Authors stated that in 7 eyes (3 normal, 1 ocular hypertensive, 3 glaucomatous) the GPS failed to provide a sectorial classification and were excluded from the enrolled group but still 182 participants were reported and included in the analysis.
Comparative	
Notes	None.
Methodological quality	

Moreno-Montañés 2008 *(Continued)*

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	No		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (All tests)			
If a threshold was used, was it pre-specified?	Yes		
Were imaging test's quality assessed?	Yes		
Were any conflict of interest avoided	Yes		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		

Moreno-Montañés 2008 *(Continued)*

Did all patients receive the same reference standard?	Yes	
Were all patients included in the analysis?	Unclear	
Did all patients receive a reference standard	Yes	
Could the patient flow have introduced bias?		Unclear risk

Moreno-Montañés 2010

Study characteristics	
Patient Sampling	Normal eyes and eyes with glaucoma were recruited prospectively. Normal group included patients consecutively recruited from hospital staff, nurses, relatives of patients, and patients referred for a routine visual acuity examination without ocular diseases. One eye per person was randomly selected.
Patient characteristics and setting	**Sample size**: 216 eyes of 216 participants (86 glaucoma, 130 healthy controls). **Age:** glaucoma patients mean ± SD, 60.12 ± 12.45 years; controls 58.22 ± 10.85 years. **Sex:** 109 men and 107 women. **Ethnicity:** white. **Country:** Spain. **Ocular comorbidities**: no corneal/retinal disease, BCVA < 20/40, spherical equivalent > ±5 D or substantial media opacity. **Setting:** Department of Ophthalmology, Clínica Universidad de Navarra,Pamplona. **Spectrum of glaucoma severity:** according to the 'glaucoma staging system', 35 eyes had early glaucoma (stage 1; mean ± SD MD of -3.0 ± 1.21 dB), 21 eyes had moderate (stage 2, mean ± SD MD of -7.81 ± 2.01 dB), 14 eyes had advanced (stage 3, mean ± SD MD of -14.7 ± 1.32 dB), 16 eyes had severe (stage 4, mean ± SD MD of -26.14 ± 2.88 dB). **Control participants:** IOP ≤ 21 mmHg, normal VFs, and no familiar glaucoma history.
Index tests	**Optical coherence tomography**: Cirrus HD-OCT, OCT volume scan, software version 3.0 (Carl Zeiss Meditec, Dublin, CA, USA). The OCT examinations were performed after pupil dilation by an experienced operator who was different from the examiner who performed the VF testing and was masked to the other findings. Only cases with signal strength of > 6 were included in the analysis. No author had conflict of interest.
Target condition and reference standard(s)	**Manifest glaucoma:** IOP > 21 mmHg on at least 3 different days, open-angle at gonioscopy and defects reproducible in at least 3 reliable and consecutive VFs performed on different days (according to the 'glaucoma staging system'). **Visual field testing:** Humphrey Field Analyzer, 24-2 SITA standard strategy (Carl Zeiss Meditec, Inc., Dublin, CA, USA). No details were reported about VF reliability criteria.

Moreno-Montañés 2010 *(Continued)*

	Optic disc appearance was not part of the reference standard
Flow and timing	The index tests were performed on the same day but the time interval between reference standard and index test was not specified. 216 participants were enrolled. 50 participants (> 10%) were excluded due to OCT scan's signal strength < 6 and 166 were actually included in the analysis.
Comparative	
Notes	None.

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Was a case-control design avoided?	No		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (All tests)			
If a threshold was used, was it pre-specified?	Yes		
Were imaging test's quality assessed?	Yes		
Were any conflict of interest avoided	Yes		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	

Moreno-Montañés 2010 *(Continued)*

Are there concerns that the target condition as defined by the reference standard does not match the question?		Low concern
DOMAIN 4: Flow and Timing		
Was there an appropriate interval between index test and reference standard?	Unclear	
Did all patients receive the same reference standard?	Yes	
Were all patients included in the analysis?	No	
Did all patients receive a reference standard	Yes	
Could the patient flow have introduced bias?		High risk

Mwanza 2012

Study characteristics	
Patient Sampling	Helthy controls and early glaucoma patients were recruited in this cross-sectional multicentre study from January to March 2011. One eye per person was randomly selected.
Patient characteristics and setting	**Sample size**: 157 participants enrolled, 154 eyes of 154 participants included in the analysis (55 glaucoma, 99 controls). **Age**: glaucoma eyes mean ± SD 64.4 ± 9.6 years; controls 62.3 ± 9.6 years. **Sex**: not reported. **Ethnicity**: not reported. **Setting**: 4 glaucoma practices were involved in this multicentre study. Bascom Palmer Eye Institute, Miami Miller School of Medicine, University of Miami, Miami, Florida; Department of Ophthalmology, Stanford University, Palo Alto, California; Eye Institute of Utah, Salt Lake City, Utah; Department of Ophthalmology, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina. **Ocular comorbidities**: patients in the glaucoma group with a BCVA < 20/40, spherical refraction error outside the interval < -12 D or >+8 D, cylinder correction > 3 D, previous or current vitreoretinal diseases or surgery, active infection of the anterior or posterior segment of either eye, diabetic retinopathy or macular oedema, history of dementia, multiple sclerosis, or a life-threatening or debilitating disease were excluded. No detail about control group comorbidities. **Spectrum of glaucoma severity**: the mean ± SD MD on the VF test were -3.2 ± 1.8 dB,for glaucoma. All glaucoma patients had MD ≥ -6 dB. **Control participants:** No details were reported.
Index tests	**Optical coherence tomography**: Cirrus HD-OCT (Carl Zeiss Meditec). The macular cube 200 x 200 and the Optic disc cube 200 x 200 scanning protocols were used to acquire the images. Only good-quality scans (signal strength ≥ 6, no RNFL discontinuity or misalignment, involuntary saccade or blinking artefacts, and absence of algorithm segmentation failure) were used for analysis.

Mwanza 2012 *(Continued)*

	Some authors had conflict of interest.
Target condition and reference standard(s)	**Manifest early glaucoma:** glaucomatous optic disc changes and glaucomatous VF defects, defined as GHT outside normal limits or PSD with a P < 5%, or a cluster of > 3 points in the pattern deviation plot in a single hemifield (superior or inferior) with a P < 5%, 1 with a P < 1%. **Visual field testing:** Humphrey Field Analzyer(SITA standard programme (Carl Zeiss Meditec, Dublin, CA, USA). Reliability criteria were not reported. **Optic disc evaluation:** dilated fundus examination. No details about how the reference standard was conducted and interpreted in the control group.
Flow and timing	The reference standard was conducted within 6 months of enrolment. 3 glaucoma patients were excluded due to repeated segmentation failure on the index test examination.
Comparative	
Notes	None.

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Was a case-control design avoided?	No		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (All tests)			
If a threshold was used, was it pre-specified?	Yes		
Were imaging test's quality assessed?	Yes		
Were any conflict of interest avoided	No		
Could the conduct or interpretation of the index test have introduced bias?		High risk	

Mwanza 2012 *(Continued)*

Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Unclear risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Unclear
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	No		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
Did all patients receive a reference standard	Yes		
Could the patient flow have introduced bias?		High risk	

Mwanza 2013

Study characteristics	
Patient Sampling	Case-control study including data of participants previously enrolled in 2 earlier glaucoma SD-OCT imaging studies and 1 ongoing study. Only one randomly selected eye per person was used.
Patient characteristics and setting	**Sample size**: 253 subjects (104 early glaucoma, 149 controls). Modelling set (69 early glaucoma, 100 controls), plus a validation set (34 early glaucoma, 49 controls) **Age**: modelling set: glaucoma mean ± SD, 66.0 ± 11.85, controls 62.8 ± 9.47 years. Validation set: glaucoma mean ± SD, 67.9 ± 12.56, controls 61.7 ± 9.56 years. **Sex**: not reported **Ethnicity**: not specified.

Mwanza 2013 *(Continued)*

	Clinical setting: glaucoma clinic of the Anne Bates Leach Eye Hospital, Department of Ophthalmology, University of Miami Miller School of Medicine. **Country**: USA. **Ocular comorbidities**: patients with spherical diopters or < 3 cylindrical diopters, active infection of the anterior or posterior segment of either eye, previous or current vitreoretinal diseases or surgery in the study eye, or evidence of diabetic retinopathy or macular oedema were excluded. **Spectrum of glaucoma severity**: the mean ± SD MD on the VF test were -3.19 ± 1.69 for glaucoma patients. **Control participants:** No details reported.
Index tests	**Optical coherence tomography**: Cirrus HD-OCT (Carl Zeiss Meditec, Inc.). Macular cube 200 x 200 and optic disc cube 200 x 200 protocol were used for the analysis. Images with signal strength < 6, RNFL misalignment or discontinuity, blinking or involuntary saccade artefacts, and algorithm segmentation failure were excluded. The authors declare no conflict of interest.
Target condition and reference standard(s)	**Manifest glaucoma:** glaucomatous optic disc changes (defined as cup-to-disc ratio > 0.5 in either eye, or cup to disc asymmetry ≥ 0.2, or focal thinning of the rim in either eye) with corresponding VF defects (GHT outside normal limits, PSD with P < 5% or a cluster 3+ points in the pattern deviation plot in a single hemifield with P < 5%, one having P < 1%. **Visual field test:** Humphrey Field Analyzer (Carl Zeiss Meditec, Dublin, CA, USA). No details about criteria for including healthy controls **Optic disc/RNFL evauation:** dilated ophthalmoscopic examination and retinal photograph evaluation.
Flow and timing	No details about exclusion and time interval between index and reference test. Controls did not undergo one of the reference tests used (VF test).
Comparative	
Notes	None.

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Was a case-control design avoided?	No		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		High risk	

Mwanza 2013 *(Continued)*

Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (All tests)			
If a threshold was used, was it pre-specified?	Yes		
Were imaging test's quality assessed?	Yes		
Were any conflict of interest avoided	Yes		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Unclear		
Did all patients receive the same reference standard?	No		
Were all patients included in the analysis?	Unclear		
Did all patients receive a reference standard	No		
Could the patient flow have introduced bias?		High risk	

Mwanza 2014

Study characteristics

Mwanza 2014 *(Continued)*

Patient Sampling	Prospective, case-control study including early glaucoma and healthy controls data of participants previously enrolled. Only one randomly-selected eye per person was used.
Patient characteristics and setting	**Sample size**: 99 participants (50 early glaucoma, 49 controls). The diagnosis of early glaucoma was based on a visual field MD ≥ -6 dB. **Age**: glaucoma mean ± SD, 63.1 ± 0.1 (range, 45.6 to 83.09, controls 66.4 ± 10.8 years (range 45.8 to 89.3). **Sex**: 40 men (22 glaucoma, 18 controls) and 59 women (28 glaucoma, 31 controls). **Ethnicity**: not specified. **Clinical setting**: Bascom Palmer Eye Institute in Miami, Florida; the Glaucoma Associates of Texas in Dallas, Texas; Stanford University in Palo Alto, California. **Country**: USA. **Ocular comorbidities**: patients with media opacities, non-glaucomatous optic neuropathy (i.e. multiple sclerosis, trauma), past or current retinal disease (i.e. retinal detachment, diabetic or infectious retinopathy, age-related macular degeneration), history of retinal surgery, laser or radiation therapy, or systemic medication that may induce optic neuropathy, were excluded. **Spectrum of glaucoma severity**: the mean ± SD MD on the VF test were -2.96 ± 1.93 for glaucoma. **Control participants:** IOP ≤ 21 mmHg, normal-looking ONH without cupping, asymmetry in cup-to-disc ratio of < 0.2, notching, or disc haemorrhage. VF not performed.
Index tests	**Optical coherence tomography**: Cirrus HD-OCT (Carl Zeiss Meditec, Inc). Optic disc cube 200 x 200 protocol and macular cube 516 x 258 protocols, were used. Only scans with a signal strength ≥ 6 and without motion (blinking or saccades) artefacts, segmentation failure caused by algorithm dysfunction, vitreous floaters, were used for analysis. The authors declare no conflict of interest.
Target condition and reference standard(s)	**Manifest glaucoma:** typical ONH cupping associated with glaucomatous VF deficits. No further details reported. **Visual field test:** Humphrey Field Analyzer (Carl Zeiss Meditec, Dublin, CA, USA), 24-2 SITA standard programme. **Optic disc/RNFL evauation:** ophthalmoscopy.
Flow and timing	No details about exclusion and time interval between index and reference test. Controls did not undergo one of the reference tests used (VF test).
Comparative	
Notes	None.

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			

Mwanza 2014 *(Continued)*

Was a consecutive or random sample of patients enrolled?	Unclear		
Was a case-control design avoided?	No		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (All tests)			
If a threshold was used, was it pre-specified?	Yes		
Were imaging test's quality assessed?	Yes		
Were any conflict of interest avoided	Yes		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Unclear		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Unclear risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Unclear
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Unclear		
Did all patients receive the same reference standard?	No		
Were all patients included in the analysis?	Unclear		
Did all patients receive a reference standard	No		

Mwanza 2014 *(Continued)*

Could the patient flow have introduced bias?	High risk

Na 2013a

Study characteristics	
Patient Sampling	Prospective, case-control study including consecutive preperimetric glaucoma and healthy controls. One eye was randomly selected if both eyes were eligible.
Patient characteristics and setting	**Sample size**: 173 participants (105 preperimetric glaucoma, 68 controls). **Age**: preperimetric glaucoma mean ± SD, 51.2 ± 10.7, controls 52.3 ± 12.6 years. **Sex**: 86 men (59 glaucoma, 27 controls) and 87 women (46 glaucoma, 41 controls). **Ethnicity**: Asian. **Clinical setting**: Asian Medical Center, Seoul, between July 2010 and February 2011. **Country**: South Korea. **Ocular comorbidities**: patients with evidence of any intracranial or otolaryngeal lesion, a history of massive haemorrhage or haemodynamic crisis, any other ophthalmic disease that could affect ONH or RNFL evaluation, any condition that might bias SD-OCT measurements (peripapillary atrophy, chorioretinal coloboma or posterior staphyloma or both), or a history of diabetes mellitus or eye surgery/laser treatment, were excluded. **Spectrum of glaucoma severity**: the mean ± SD MD and pattern SD on the VF test were -0.34 ± 1.31 and 1.63 ± 0.3 respectively for preperimetric glaucoma. **Control participants:** IOP < 22 mmHg, no history of IOP elevation, normal VF results, intact neuroretinal, no disc haemorrhage, notches or any localised RNFL defect.
Index tests	**Optical coherence tomography**: RTVue SD-OCT (Optovue, Inc.). Software version A4.0.5.100. ONH and GCC scanning protocols were used for analysis. Images with signal strength index values of the ONH or GCC maps < 45 were excluded. The authors declare no conflict of interest.
Target condition and reference standard(s)	**Manifest glaucoma:** localised RNFL defects (present if their width at a 1-disc diameter distance from the edge of the disc was larger than a major retinal vessel and if they diverged in an arcuate or wedge shape reaching the edge of the disc) and normal VF test result (defined as the absence of a cluster of 3 points with P < 5% 5% or a cluster of 2 points with P < 1%on the pattern deviation plot, and a GHT within normal limits). **Visual field test:** Humphrey Field Analyzer (Carl Zeiss Meditec, Dublin, CA, USA), 24-2 SITa standard programme. Reliable examinations had false-positive error < 15%, a false-negative error < 15% and a fixation loss < 20%. **RNFL evaluation:** digital fundus camera (TRC-50IX; Topcon, Tokyo, Japan, and MegaPlus 1.4i, Kodak, Rochester, New York, USA).
Flow and timing	6 participants (< 10%) were excluded because of unacceptable image quality. Index and reference test were performed on the same day.
Comparative	

Na 2013a *(Continued)*

Notes	None.

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	No		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (All tests)			
If a threshold was used, was it pre-specified?	Yes		
Were imaging test's quality assessed?	Yes		
Were any conflict of interest avoided	Yes		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			

Na 2013a *(Continued)*

Was there an appropriate interval between index test and reference standard?	Yes	
Did all patients receive the same reference standard?	Yes	
Were all patients included in the analysis?	Yes	
Did all patients receive a reference standard	Yes	
Could the patient flow have introduced bias?		Low risk

Na 2013b

Study characteristics	
Patient Sampling	Case-control study enrolling consecutive glaucoma patients between September 2010 and February 2012, at the Asian Medical Center, Seoul, Korea. Control group consisted of clinic staff, friends or spouses of patients, and volunteers from other specialty clinics. One eye per person was included in the analysis.
Patient characteristics and setting	**Sample size**: 84 eyes of 84 participants (42 glaucoma and 42 healthy controls). **Age**: glaucoma mean ± SD, 50.69 ± 10.34 years; controls, 50.76 ± 9.77 years. **Sex**: 40 men (21 glaucoma, 19 controls) and 44 women (21 glaucoma, 23 controls). **Ethnicity**: not reported. **Clinical Setting:** Asian Medical Center, Seoul. **Country**: Korea. **Ocular comorbidities**: patients with intracranial or otolaryngeal lesion, with a history of massive haemorrhage or haemodynamic crisis, who presented with any other ophthalmic disease that could result in VF defects, or with diabetes mellitus or eye surgery/laser treatment, were excluded. **Spectrum of glaucoma severity**: the mean ± SD MD and PSD on the VF test were -4.19 ± 2.06 and 6.04 ± 3.45 respectively, for glaucomatous eyes. **Control participants:** IOP < 22 mmHg, no history of IOP elevation above 21 mmHg, absence of ONH abnormality, and a normal VF result.
Index tests	**Optical coherence tomography**: Cirrus OCT (Carl Zeiss Meditec, Inc., Dublin, CA, USA). Optic disc cube 200 x 200 scan protocol was used for the analysis. Scans had signal strengths > 6, and no motion artefact. **Scanning laser polarimetry**: GDx VCC (Carl Zeiss Meditec, Inc., Dublin, CA, USA); software version 5.6.0.8. Accepted images had a centred optic disc, were well focused and adequately illuminated over the entire image, and did not show motion artefacts. Images with TSS < 80 were excluded. The authors stated no conflicts of interest.
Target condition and reference standard(s)	**Manifest perimetric glaucoma:** localised VF loss (defined as 3+ adjacent points with P < 0.05 in a PD probability map, or 2+ adjacent points with P < 0.02 in a superior or in-

Na 2013b *(Continued)*

	ferior hemifield, and the hemifield of the other side had no clusters of 3 points with P < 0.05 and no clusters of 2 points with P < 0.02 on either total deviation or PD probability maps) confined to one side of the horizontal meridian, GHT outside normal limits, a PSD with P < 5%, and a cluster of 3+ points in the PD plot in a single hemifield (superior or inferior) with P < 0.05, one with P < 0.01, and open angle by gonioscopy. **Visual field test:** Humphrey Field Analyzer (Carl Zeiss Meditec); 24-2 SITA–standard strategy. All exams had fixation losses < 20% and false-positive and false-negative rates of < 15%.
Flow and timing	6 glaucoma (< 10%) were excluded due to low-quality images. The time interval between index and reference test was not reported.
Comparative	
Notes	None.

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	No		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (All tests)			
If a threshold was used, was it pre-specified?	Yes		
Were imaging test's quality assessed?	Yes		
Were any conflict of interest avoided	Yes		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			

Na 2013b *(Continued)*

Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Unclear		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
Did all patients receive a reference standard	Yes		
Could the patient flow have introduced bias?		Unclear risk	

Nakatani 2011

Study characteristics	
Patient Sampling	Normal participants, preperimetric and perimetric primary open-angle glaucoma were enrolled. One eye per person was selected.
Patient characteristics and setting	**Sample size**: 64 eyes of 64 participants (32 early glaucoma (13 preperimetric and 19 perimetric glaucoma) and 32 healthy controls). **Age:** glaucoma patients mean ± SD, 61.5 ± 7.7 years; controls 57.3 ± 10.9 years. **Sex:** 33 men (14 glaucoma, 19 controls) and 31 women (18 glaucoma, 13 controls). **Country:** Japan. **Ocular comorbidities**: no cataract, BCVA < 20/40, spherical refraction > ±6 D, cylinder refraction > ±2 D, close angle by gonioscopy and ocular pathology other than glaucoma. **Setting:** Himi Municipal Hospital. **Spectrum of glaucoma severity:** mean ± SD MD and PSD on the VF test were -2.14 ± 1.77 dB and 3.86 ± 2.66 dB. All glaucoma patients had MD > -6 dB.

Nakatani 2011 *(Continued)*

	Control participants: normal ONH appearance, IOP < 21 mmHg, and normal SAP results.
Index tests	**Optic Coherence Tomography** : 3D-OCT- 1000 Mark II, 3D scan and RNFL 3.4 mm protocol (Topcon, Tokyo, Japan). 3 consecutive scans with no obvious misalignment between the centre of the scans and the optic disc or the fovea were acquired after pupil dilatation and by the same operator. A mean of 3 scans was used for the analysis. The authors stated no source of support.
Target condition and reference standard(s)	**Manifest glaucoma:** comprised perimetric glaucoma eyes defined as glaucomatous optic disc abnormalities with a localised RNFL at areas of rim thinning and glaucomatous VF defects (defined as a cluster of 3+ non-edge points with P < 5% and at least 1 point with P < 1% in the pattern deviation probability plot or PSD with P < 5% or GHT outside normal limits) and preperimetric glaucoma eyes defined as glaucomatous optic disc abnormalities with localised RNFL defect at areas of rim thinning, without glaucomatous VF defects. **Visual field testing:** Humphrey Field Analyzer, 30-2 SITA strategy (Carl Zeiss Meditec Inc., Dublin, CA, USA). Reliability criteria included fixation losses rates < 20%, and false-positive and false-negative rates of < 33%. **Optic disc evaluation**: dilated fundus biomicroscopy using 78-diopter lens, stereoscopic optic disc photography.
Flow and timing	The reference standard and index test were performed on the same day. All participants enrolled were included in the analysis.
Comparative	
Notes	None.

Methodological quality

Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Was a case-control design avoided?	No		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (All tests)			
If a threshold was used, was it pre-specified?	Yes		
Were imaging test's quality assessed?	Yes		

Nakatani 2011 *(Continued)*

Were any conflict of interest avoided	Yes		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
Did all patients receive a reference standard	Yes		
Could the patient flow have introduced bias?		Low risk	

Nouri-Mahdavi 2013

Study characteristics	
Patient Sampling	Case-control study in which glaucoma and normal participants were prospectively recruited between December 2010 and October 2012. Both eyes of some participants were included in the analysis.
Patient characteristics and setting	**Sample size**: 150 eyes of 99 participants (59 eyes of 47 subjects with early glaucoma, 91 eyes of 52 normal healthy controls). **Age**: glaucoma mean ± SD, 66.1 ± 6.0, controls 58.6 ± 9.2 years. **Sex**: 56 men (23 glaucoma, 33 controls) and 94 women (36 glaucoma, 58 controls).

Nouri-Mahdavi 2013 *(Continued)*

	Ethnicity: 75 white (36 glaucoma, 39 controls); 10 African-American (6 glaucoma, 4 controls); 5 Hispanic (2 glaucoma, 3 controls); 9 Asian (3 glaucoma, 6 controls). **Clinical setting**: University of California, Los Angeles (UCLA) and Glaucoma Clinic, Jules Stein Eye Institute, between December 2010 and October 2012. **Country**: USA. **Ocular comorbidities**: patients with evidence of retinal or neurologic diseases or prior glaucoma surgery were excluded. **Spectrum of glaucoma severity**: the mean ± SD MD and PSD on the VF test were -2.5 ± 1.9 and 4.5 ± 2.2 respectively for early glaucoma. All glaucoma has MD≥ -6 dB. **Control participants:** normal eye examinations, including normal VFs, and not having definitive evidence of glaucomatous damage at the level of the ONH.
Index tests	**Optical coherence tomography**: Cirrus HD-OCT, (Carl Zeiss Meditec, Dublin, CA, USA). Software version 6.0. Optic disc cube 200 x 200 and macular cube 200 x 200 scanning protocols were used for analysis. Images with signal strength < 7, lost data on the peripapillary ring, obvious motion artefact, or incorrect segmentation, were excluded. The authors declare no conflict of interest.
Target condition and reference standard(s)	**Manifest glaucoma:** glaucomatous VF test results, defined as GHT outside normal limits and the presence of ≥ 4 abnormal test locations on a pattern deviation plot, with P < 5% both confirmed at least one. **Visual field test:** standard automated perimetry or short-wavelength automated perimetry. Only eyes with reliable visual fields (false-positive rate of 15% or less) were included.
Flow and timing	Only eyes with reliable visual fields were included but no further details on number of exclusions were reported. Index and reference tests were performed on the same day.
Comparative	
Notes	None.

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	No		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		High risk	

Nouri-Mahdavi 2013 *(Continued)*

Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (All tests)			
If a threshold was used, was it pre-specified?	Yes		
Were imaging test's quality assessed?	Yes		
Were any conflict of interest avoided	Yes		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Unclear		
Did all patients receive a reference standard	Yes		
Could the patient flow have introduced bias?		Unclear risk	

Oddone 2008

Study characteristics	
Patient Sampling	A series of consecutive normal and POAG participants from the population attending the glaucoma clinics were enrolled. Normal controls were people attending the outpa-

Oddone 2008 *(Continued)*

	tient clinics, spouses and friends of the recruited patients, or volunteers from the hospital staff. One eye per person was randomly selected.
Patient characteristics and setting	**Sample size**: 242 eyes of 242 participants enrolled; 236 included in the analysis (99 glaucoma,137 healthy controls). **Age:** glaucoma patients mean ± SD, 62.7 ± 11 years; controls 60.9 ± 13 years. **Sex:** 105 men (45 glaucoma, 60 controls) and 131 women (54 glaucoma, 77 controls). **Country:** Italy. **Ocular comorbidities**: no neuro-ophthalmologic/retinal diseases, BCVA < 20/40, spherical refraction > ±5 D, cylinder refraction > ±3 D, uveitis, close angle by gonioscopy, ocular surgery or laser treatments, ocular trauma, rheumatologic systemic diseases and diabetes. **Setting:** University of Rome Tor Vergata, Rome; University of Milan San Paolo, Milan; University of Genoa, Genoa. **Spectrum of glaucoma severity:** according to the VF defect severity: 42 eyes were at stage 1 (MD > -6 dB), 29 eyes at stage 2 (MD < -6 dB and > -12 dB), 28 at stage 3 (MD < -12 dB). Mean ± SD MD/CPSD on the VF test were respectively -3.74 ± 1.29 dB/4.67 ± 1.72 dB (stage 1), -8.35 ± 1.83 dB/7.5 ± 2.41 dB (stage 2), -18.07 ± 4.93/10.4 ± 2.88 dB (stage 3). **Control participants:** IOP < 22 mmHg and a normal VF test result.
Index tests	**Confocal scanning laser tomography**: HRT 3, software version 3.0 (Heidelberg Engineering GmbH, Dossenheim, Germany). After scanning, a contour line was manually placed around the ONH edge by 3 experienced investigators masked to the participant's diagnosis. Only high-quality images with acquisition sensitivity > 90% and a SD < 40 were considered acceptable. No author had conflict of interest.
Target condition and reference standard(s)	**Manifest glaucoma:** history of IOP > 24 mmHg in the hospital notes and glaucomatous VF defects (defined as GHT outside normal limits, MD and PSD outside 95% confidence limits and a cluster of at least 3 points with P < 0.05 in the pattern deviation plot, one of each with P < 0.01 affecting the same hemifield). **Visual field testing:** Humphrey Field Analyzer, 24-2 SITA-standard strategy (Carl Zeiss Meditec, Inc., Dublin, CA, USA). VF reliability criteria were not specified. Optic disc appearance was not part of the reference standard.
Flow and timing	The time interval between reference standard and index test was not reported. 268 participants were initially screened, 242 were enrolled. In 6 eyes (2.5%, 4 glaucoma and 2 controls) the GPS was unable to provide a classification, and were excluded from the analysis.
Comparative	
Notes	None.

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			

Oddone 2008 *(Continued)*

Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	No		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (All tests)			
If a threshold was used, was it pre-specified?	Yes		
Were imaging test's quality assessed?	Yes		
Were any conflict of interest avoided	Yes		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Unclear		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		

Oddone 2008 *(Continued)*

Did all patients receive a reference standard	
Could the patient flow have introduced bias?	Unclear risk

Oddone 2011

Study characteristics	
Patient Sampling	Healthy participants and glaucoma patients were consecutively enrolled from the population attending the glaucoma clinics. Normal controls were either people attending the outpatient clinics, spouses and friends of the recruited patients, or volunteers from the hospital staff. One eye per person was enrolled.
Patient characteristics and setting	**Sample size**: 136 participants screened, 130 enrolled, 120 eye of 120 participants finally included in the analysis (70 glaucoma, 50 controls). **Age**: glaucoma eyes mean ± SD 66.2 ± 8.6 years; controls 64.3 ± 6.0 years. **Sex**: 71 men (42 glaucoma, 29 controls) and 49 women (28 glaucoma, 21 controls). **Ethnicity**: not reported. **Country:** Italy. **Setting**: G.B. Bietti Eye Foundation, Rome, and University of Rome Tor Vergata, Rome. **Ocular comorbidities**: patients with history of neuro-ophthalmologic or retinal diseases, uveitis, previous ocular surgery or laser treatments, history of ocular trauma, rheumatologic systemic diseases, and diabetes were excluded. All eyes had to have BCVA ≥ 20/40, a spherical refraction within ±5 D, astigmatism within ±3 D, and an open angle by gonioscopy. **Spectrum of glaucoma severity**: the mean ± SD MD and PSD on the VF test were -8.4 ± 6.8 dB and 7.2 ± 4.5 dB, respectively for glaucoma. **Control participants:** IOP < 22 mmHg in both eyes with no history of IOP > 21 mmHg, a GHT within normal limits and a MD and a PSD within 95% confidence limits confirmed in 2 reliable, consecutive VF tests.
Index tests	**Optical coherence tomography**: Cirrus HD-OCT (software version 3.0). The optic disc cube 200 x 200 was used to acquire the images. All images were acquired by a single, well-trained investigator during the same visit. Only scans with a signal strength of ≥ 6, without RNFL discontinuity or misalignments, eye movements, or blinking artefacts were included in the analysis. **Scanning laser polarimetry:** GDx VCC (Carl Zeiss Meditec, Dublin, CA, USA). Only high-quality images (well-focused and uniformly illuminated reflectance image, with a centred optic disc and a quality score > 8) and without an atypical retardation pattern were included. **Confocal scanning laser ophthalmoscopy:** HRT3 (software version 3.0, Heidelberg Engineering GmbH, Dossenheim, Germany). The contour line was manually placed around the ONH edge by one experienced investigator masked to the subset diagnosis. Only high-quality images (acquisition sensitivity > 90% and a SD > 30) were included in the analysis. None of the authors had conflict of interest.
Target condition and reference standard(s)	**Manifest glaucoma:** documented history of IOP > 24 mmHg and glaucomatous VF damage defined as a GHT outside normal limits, MD and PSD outside 95% confidence limits, and a cluster of > 3 points with P < 5% in the pattern deviation plot, one with P < 1% affecting the same

Oddone 2011 *(Continued)*

	hemifield (the cluster had not to be contiguous with the blind spot and had not to cross the horizontal midline). **Visual field testing:** Humphrey Field Analzyer (24-2 SITA standard programme (Carl Zeiss Meditec). VF reliability criteria were not reported. Optic nerve appearance was not part of the reference standard.
Flow and timing	The time interval between index tests and reference standard was not reported. 10 participants (< 10%) were excluded from the analysis: 6 eyes due to atypical pattern on GDx VCC and in 4 eyes the HRT3 GPS analysis was unable to provide a classification.
Comparative	
Notes	None.

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	No		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (All tests)			
If a threshold was used, was it prespecified?	Yes		
Were imaging test's quality assessed?	Yes		
Were any conflict of interest avoided	Yes		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern

Oddone 2011 *(Continued)*

DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Unclear		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
Did all patients receive a reference standard	Yes		
Could the patient flow have introduced bias?		Unclear risk	

Pablo 2010

Study characteristics	
Patient Sampling	Healthy and glaucoma eyes were consecutive enrolled. Normal eyes were recruited from patients referred for refraction that underwent routine examination without abnormal ocular findings, hospital staff, and relatives of hospital patients. One eye per person was randomly selected.
Patient characteristics and setting	**Sample size**: 118 participants enrolled, 105 eyes of 105 participants included in the analysis (43 POAG, 10 pseudo-exfoliative glaucoma, 1 pigmentary glaucoma and 51 healthy controls). **Age:** glaucoma patients mean ± SD, 61.9 ± 6.8 years; controls 59.1 ± 9.6 years. **Ethnicity:** white. **Country:** Spain

Pablo 2010 *(Continued)*

	Ocular comorbidities: no previous intraocular surgery, BCVA < 20/30, spherical refraction > ±5 D, cylinder refraction > ±2 D history of ocular or neurologic disease, diabetes or other systemic diseases. **Setting:** Miguel Servet University Hospital, Zaragoza. **Spectrum of glaucoma severity:** mean ± SD MD and PSD on the VF test were -2.91 ± 2.28 dB and 3.16 ± 2.07 dB, respectively. **Control participants:** IOP < 21 mmHg, no history of increased IOP, and a normal SAP.
Index tests	**Confocal scanning laser ophthalmoscopy:** HRT3 (Heidelberg Engineering, Dossenheim, Germany). Images were obtained through dilated pupils. All scans had to have an interscan SD < 30 mm. The margin of the optic discs was manually traced by the same glaucoma specialist, who was masked to the patients' identity and clinical history. No author had conflict of interest.
Target condition and reference standard(s)	**Manifest glaucoma:** IOP > 21 mmHg (on at least 3 readings on different days) and glaucomatous VF defects (defined as PSD with P < 5% and/or a GHT outside normal limits). **Visual field testing:** Humphrey Field Analyzer, model 750, 24-2 SITA-standard strategy (Carl Zeiss Meditec, Dublin, CA, USA). Reliability criteria included fixation losses, false-positive and false-negative rates of < 20%. Optic disc appearance was not part of the reference standard.
Flow and timing	The time interval between reference standard and index tests was < 6 weeks. 13 pre-selected participants were not included in the analysis (2 did not provide informed consent, 6 did not complete all of the required tests and in 5 GPS analyses produced only a global result or no results).
Comparative	
Notes	None.

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	No		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (All tests)			

Pablo 2010 *(Continued)*

If a threshold was used, was it pre-specified?	Yes		
Were imaging test's quality assessed?	Yes		
Were any conflict of interest avoided	Yes		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	No		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	No		
Did all patients receive a reference standard	Yes		
Could the patient flow have introduced bias?		High risk	

Pueyo 2006

Study characteristics	
Patient Sampling	Healthy, ocular hypertensive and glaucoma eyes were enrolled. All high-IOP patients and those affected by glaucoma were consecutively selected amongst the patients seen in consulting rooms and who fulfilled the inclusion criteria set for this study. One eye per person was selected.
Patient characteristics and setting	**Sample size**: 427 eyes of 427 participants (74 glaucoma, 287 ocular hypertensive and 66 healthy controls).

Pueyo 2006 *(Continued)*

	Age: glaucoma patients mean ± SD, 64.79 ± 9.31 years; OHT patients 55.10 ± 11.63 years; controls 58.95 ± 11.74years. **Country:** Spain. **Ocular comorbidities**: no retinal disease, BCVA < 32/40, spherical refraction > ±5 D, cylinder refraction > ±3 D, angular abnormalities, history of traumatism or ocular surgery, or neuro-ophthalmologic disease. **Setting:** Ophthalmology Service, University Hospital Miguel Servet, Zaragoza, Ophthalmology Service, San Carlos Hospital, Madrid. **Spectrum of glaucoma severity:** mean ± SD MD and PSD on the VF test were -6.56 ± 6.07 dB and 5.11 ± 3.66 dB for glaucoma eyes, -0.30 ± 1.12 dB and 0.97 ± 0.75 dB for OHT eyes. **Control participants:** IOP ≤ 21 mmHg, automated perimetry and optic nerve appearance compatible with normality.
Index tests	**Scanning laser polarimetry:** GDx VCC (Laser Diagnostic Technologies, San Diego). No details about images quality assessment or conflict of interest were reported.
Target condition and reference standard(s)	**Manifest glaucoma:** IOP > 21 mmHg, glaucomatous VF defects (defined as CPSD with P < 2%, and/or group of 3+ adjoining points with a probability level < 1% and/or altered GHT) and glaucomatous optic nerve signs (defined as focal/diffuse thinning of the neuroretinal ring, papillar haemorrhages, asymmetry in the proportion excavation/vertical disc above 0.2 between both eyes). **Ocular Hypertension:** IOP > 21 mmHg with automated perimetry compatible with normality, without considering papillar morphology. **Visual field testing:** Humphrey Field Analyzer, model 745, 24-2 full threshold strategy. VF reliability criteria included fixation losses rates < 20, false-positive and false-negative rates of < 33%. **Optic disc evaluation:** papilla assessment was done by a glaucoma specialist.
Flow and timing	The time interval between reference standard and index tests was not reported. The authors stated that all the patients for whom it was not possible to obtain good-quality images of all the structural analysis techniques were excluded from the study, but no other details were specified.
Comparative	
Notes	None.

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	No		
Did the study avoid inappropriate exclusions?	Yes		

Pueyo 2006 *(Continued)*

Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (All tests)			
If a threshold was used, was it pre-specified?	Yes		
Were imaging test's quality assessed?	Unclear		
Were any conflict of interest avoided	Unclear		
Could the conduct or interpretation of the index test have introduced bias?		High risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Unclear		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Unclear		
Did all patients receive a reference standard	Yes		
Could the patient flow have introduced bias?		Unclear risk	

Rao 2010a

Study characteristics	
Patient Sampling	Normal and glaucoma participants seen in a tertiary eye care centre between July 2004 and February 2006 were enrolled. Consecutively-seen patients with glaucoma formed the study group, whereas the normal participants were from among those referred for refraction without any abnormal ocular findings, patients' relatives, or hospital staff. One eye per person was randomly selected.
Patient characteristics and setting	**Sample size**: 177 eyes of 177 participants (98 glaucoma, 79 healthy controls). **Age:** glaucoma patients mean ± SD, 55.2 ± 9.1 years; controls 51.9 ± 10.6 years. **Sex:** 111 men (62 glaucoma, 49 controls) and 66 women (36 glaucoma, 30 controls). **Ethnicity:** Indian. **Country:** India. **Ocular comorbidities**: no intraocular surgery within the previous 6 months, BCVA < 20/40, spherical refraction > ±5 D, cylinder refraction > ±3 D, any retinal or neurologic diseases that could confound the results of VF examination. **Setting:** Eye care centre in Hyderabad, Central India, between July. **Spectrum of glaucoma severity:** mean ± SD MD and PSD on the VF test were -7.3 ± 6.7 dB and 5.33 ± 3.86 dB, respectively. **Control participants:** IOP < 22 mmHg in both eyes, no history of increased IOP or family history of glaucoma, normal VF result and optic disc appearance.
Index tests	**Confocal scanning laser ophthalmoscopy:** HRT2 (Heidelberg Engineering, Dossenheim, Germany). After scan, data were exported to HRT3 to be processed without altering the location of the contour line. A single experienced operator had acquired 3 scans and drawn the disc margin in each scan. Only images with inter-scan SD of ≤ 50 µm were included. No author had conflict of interest.
Target condition and reference standard(s)	**Manifest glaucoma:** glaucomatous ONH appearance (defined as focal/diffuse neuroretinal rim thinning, localised notching, or nerve fibre layer defects) with correlating VF defects (presence of a cluster of 3 points on pattern deviation probability plot with P < 5%, one of which had P < 1%, or a PSD with P < 5%, or a GHT outside normal limits). **Visual field testing:** Humphrey Field Analyzer, 24-2 SITA-standard strategy (Zeiss-Humphrey Systems, Dublin, CA, USA). Reliability criteria included fixation losses rates, false-positive and false-negative rates of < 20%. **Optic disc evaluation:** dilated fundus examination by 2 glaucoma specialists.
Flow and timing	The time interval between reference standard and index test was not reported. 10 participants (6 glaucoma and 4 normal) were excluded due to poor-quality images (< 10%).
Comparative	
Notes	None.
Methodological quality	

Rao 2010a *(Continued)*

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	No		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (All tests)			
If a threshold was used, was it pre-specified?	Yes		
Were imaging test's quality assessed?	Yes		
Were any conflict of interest avoided	Yes		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Unclear		

Rao 2010a *(Continued)*

Did all patients receive the same reference standard?	Yes	
Were all patients included in the analysis?	Yes	
Did all patients receive a reference standard	Yes	
Could the patient flow have introduced bias?		Unclear risk

Rao 2010b

Study characteristics	
Patient Sampling	Glaucoma patients were enrolled in a prospective, longitudinal study designed to evaluate optic nerve structure and visual function in glaucoma (Diagnostic Innovations in Glaucoma Study). Healthy participants were recruited from the general population through advertisement, as well as from the staff and employees of the University of California. When both eyes of participants satisfied the inclusion criteria, both were included.
Patient characteristics and setting	**Sample size**: 74 eyes of 44 normal participants and 140 eyes of 106 glaucoma patients. **Age:** glaucoma patients mean ± SD, 68.34 ± 10.54 years; controls 62.34 ± 12.04 years. **Ethnicity:** 119 white (40 controls, 79 glaucoma); 31 African-American (4 controls, 27 glaucoma). **Country:** USA. **Ocular comorbidities**: no co-existing retinal disease, close angle by gonioscopy, BCVA < 20/40, spherical refraction > ±5 D, cylinder refraction > ±3 D, uveitis, or non-glaucomatous optic neuropathy. **Setting:** Hamilton Glaucoma Center, University of California, San Diego. **Spectrum of glaucoma severity:** mean (with 1st and 3rd quartile values) MD and PSD on the VF test were -3.67 (-2.05, -7.07) dB and 4.03 (2.58, 9.10) dB. **Control participants:** IOP < 21 mmHg, with no history of increased IOP and a normal VF result (MD and PSD within the 95% confidence limits, and a GHT within normal limits).
Index tests	**Optic Coherence Tomography**: RTVue-100, software version 4.0.5.39 (Optovue Inc, Fremont, CA, USA). The ONH and GCC scan protocols were acquired. Only high-quality images, as defined by a signal strength index > 30 were used for analysis. Some authors had conflict of interest.
Target condition and reference standard(s)	**Manifest glaucoma:** repeatable (> 2 consecutive), glaucomatous VF result (defined as a PSD outside the 95% confidence limits or a GHT outside normal limits, or both). **Visual field testing:** Humphrey Field Analyzer, 24-2 SITA-standard strategy (Carl Zeiss Meditec Inc.). VF reliability criteria were not specified.

Rao 2010b *(Continued)*

	Optic disc appearance was not part of the reference standard.
Flow and timing	The time interval between reference standard and index test was < 1 year. 3 participants (2 normal and 1 glaucoma patient, < 10%) were excluded from the analysis due to incorrect baseline disc drawing.
Comparative	
Notes	None.

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		
Was a case-control design avoided?	No		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (All tests)			
If a threshold was used, was it pre-specified?	Yes		
Were imaging test's quality assessed?	Yes		
Were any conflict of interest avoided	No		
Could the conduct or interpretation of the index test have introduced bias?		High risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	

Rao 2010b *(Continued)*

Are there concerns that the target condition as defined by the reference standard does not match the question?		Low concern
DOMAIN 4: Flow and Timing		
Was there an appropriate interval between index test and reference standard?	No	
Did all patients receive the same reference standard?	Yes	
Were all patients included in the analysis?	Yes	
Did all patients receive a reference standard	Yes	
Could the patient flow have introduced bias?		High risk

Rao 2012a

Study characteristics	
Patient Sampling	Cross-sectional study comprised consecutive early glaucoma patients and 2 cohort of healthy controls. One cohort (1) recruited from people who attend a tertiary eye-care clinic for a routine eye examination, patients' relatives, and hospital staff. Another cohort (2) including consecutive patients who were referred to tertiary clinic centre by general ophthalmologists as glaucoma suspects based on optic disc morphology but confirmed by glaucoma experts to be non-glaucomatous.
Patient characteristics and setting	**Sample size**: 260 eyes of 147 participants (65 eyes of 46 glaucoma patients, 119 eyes of 60 controls for cohort 1 and 76 eyes of 41 controls for cohort 2). **Age**: glaucoma eyes mean ± SD 51.9 ± 13.2 years; controls cohort 1, 47.1 ± 12.8 years; controls cohort 2, 50.2 ± 14.7 years. **Sex**: 94 men (33 glaucoma, 32 controls cohort 1, 29 controls cohort 2) and 53 women (13 glaucoma, 28 controls cohort 1, 12 controls cohort 2). **Country**: India. **Ethnicity**: not reported. **Setting**: Glaucoma Center L. V. Prasad Eye Institute, Banjara Hills, Hyderabad. **Ocular comorbidities**: patients with any media opacities, intraocular surgery within the previous 6 months, and any retinal or neurologic diseases other than glaucoma that could confound the results of VF examination and structural measurements with SD-OCT were excluded. All eyes had to have BCVA ≥ 20/40, refractive error within ±5.0 D sphere and ±3 D cylinder. **Spectrum of glaucoma severity**: the mean ± SD MD and PSD on the VF test were -3.2 ± 1.5 dB and 2.8 ± 1.8 dB, respectively for glaucoma. All glaucoma eyes had early stage of disease according to Hodapp et al. classification. **Control participants:** 2 cohorts of participants were used as control group: **- Cohort 1**: IOP < 22 mmHg with no history of increased IOP, no family history of glaucoma, no optic disc morphology suspicious for glaucoma and normal visual field result.

Rao 2012a *(Continued)*

	- Cohort 2: IOP < 22 mmHg with no history of increased IOP, no family history of glaucoma, and normal visual fields. They were referred to clinical centre by general ophthalmologists as glaucoma suspects based on optic disc morphology but their optic discs were confirmed on clinical examination by the glaucoma experts to be non-glaucomatous, but physiological variations of normal.
Index tests	**Optical coherence tomography**: RTVue (software version 4.0.5.39; Optovue Inc., Fremont, CA, USA). The ONH and GCC scanning protocols were used to acquire the images. Only well-centred images with a signal strength index of ≥ 30 were included in the analysis. One author had conflict of interest.
Target condition and reference standard(s)	**Manifest glaucoma:** glaucomatous optic disc changes (defined as focal or diffuse neuroretinal rim thinning, localised notching or nerve fibre layer defects) and corresponding VF defects (defined by 2 of the following criteria: the presence of a cluster of 3 points on a pattern deviation probability plot with P < 5%, one of which had P < 1%; a PSD with P < %5; or a GHT result outside normal limits. **Visual field testing:** Humphrey Field Analzyer, model 750 (24-2 SITA standard programme (Carl Zeiss Meditec). VF reliability criteria were fixation losses, false-positive and false-negative rates < 20%. **Optic disc evaluation:** digital optic disc photographs (Visupac 4.2.2; Carl Zeiss Meditec Systems, GmbH, Pirmasens, Germany). The photographs were evaluated by 2 experts who were masked to the clinical examination and VF and OCT results of the participants.
Flow and timing	All participants had both protocols as well as the VF testing performed on the same day. The authors stated that "Eyes in which the segmentation algorithm failed were excluded" but no participants were reported as excluded from the analysis.
Comparative	
Notes	None.

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	No		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (All tests)			

Rao 2012a *(Continued)*

If a threshold was used, was it pre-specified?	Yes		
Were imaging test's quality assessed?	Yes		
Were any conflict of interest avoided	No		
Could the conduct or interpretation of the index test have introduced bias?		High risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Unclear		
Did all patients receive a reference standard	Yes		
Could the patient flow have introduced bias?		Unclear risk	

Rao 2012b

Study characteristics	
Patient Sampling	Consecutive early glaucoma patients and healthy controls were enrolled at a tertiary eye-care facility between August 2008 and June 2009. The normal participants were recruited from among those who came for a routine eye examination, patients' relatives and hospital staff. Both eyes were included for some participants.
Patient characteristics and setting	**Sample size**: 216 (91 early glaucoma, 125 control) eyes of 123 participants (59 early glaucoma, 64 control) were enrolled and included in the analysis. **Age**: glaucoma eyes mean ± SD 51.8 ± 13.4 years; controls 47.7 ± 13.4 years. **Sex**: not reported. **Ethnicity**: Indian. **Country**: India. **Setting**: glaucoma Center L. V. Prasad Eye Institute, Banjara Hills, Hyderabad. **Ocular comorbidities**: patients with any media opacities, intraocular surgery within the previous 6 months, and any retinal or neurologic diseases other than glaucoma that could confound the results of VF examination and structural measurements with SD-OCT were excluded. All eyes had to have BCVA ≥ 20/40, refractive error within ±5.0 D sphere and ±3 D cylinder. **Spectrum of glaucoma severity**: the mean ± SD MD and PSD on the VF test were -2.6 ± 1.8 dB and 2.4 ± 1.5 dB, respectively for glaucoma. All glaucoma eyes had early stage of disease according to Hodapp et al. classification. **Control participants:** IOP < 22 mmHg with no history of increased IOP, no family history of glaucoma, no optic disc morphology suspicious for glaucoma (focal or diffuse neuroretinal rim thinning, localised notching or nerve fibre layer defects) and normal VF result.
Index tests	**Optical coherence tomography**: RTVue (software version 4.0.5.39; Optovue Inc., Fremont, CA, USA). The ONH and GCC scanning protocols were used to acquire the images. Only well-centred images with a signal strength index of ≥ 30 were included in the analysis. One author had conflict of interest.
Target condition and reference standard(s)	**Manifest glaucoma:** glaucomatous optic disc changes (defined as focal or diffuse neuroretinal rim thinning, localised notching or nerve fibre layer defects) and corresponding VF defects (defined by 2 of the following criteria: the presence of a cluster of 3 points on a pattern deviation probability plot with P < 5%, one of which had P < 1% or a PSD with P < %5 or a GHT result outside normal limits). **Visual field testing:** Humphrey Field Analzyer, model 750 (24-2 SITA standard programme (Carl Zeiss Meditec). Reliability criteria were fixation losses, false-positive and false-negative rates < 20%. **Optic disc evaluation:** dilated fundus examination by 2 glaucoma specialists.
Flow and timing	The index test and reference standard were performed on the same day. No participants were reported by the authors as excluded from the analysis.
Comparative	

Rao 2012b *(Continued)*

Notes	None.

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	No		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (All tests)			
If a threshold was used, was it pre-specified?	Yes		
Were imaging test's quality assessed?	Yes		
Were any conflict of interest avoided	No		
Could the conduct or interpretation of the index test have introduced bias?		High risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			

Rao 2012b *(Continued)*

Was there an appropriate interval between index test and reference standard?	Yes	
Did all patients receive the same reference standard?	Yes	
Were all patients included in the analysis?	Yes	
Did all patients receive a reference standard	Yes	
Could the patient flow have introduced bias?		Low risk

Rao 2013

Study characteristics	
Patient Sampling	Case-control study including preperimetric glaucoma and 2 different control group (patient referred by general ophthalmologist as glaucoma suspects for optic disc appearance and healthy controls not suspected of having glaucoma), evaluated at a tertiary eye-care facility between January 2010 and December 2012. One eye was randomly selected if both eyes were eligible.
Patient characteristics and setting	**Sample size**: 166 eyes of 166 participants (34 eyes of 34 preperimetric glaucoma, 72 eyes of 72 controls with optic disc appearance suspected of having glaucoma and 60 eyes of 60 healthy controls with no optic disc appearance suspected for glaucoma). **Age**: glaucoma mean (range), 54 (41 to 61), controls group 1 52 (41 to 62) years, controls group 2 50 (38 to 57). **Sex**: not reported. **Ethnicity**: not reported. **Clinical setting**:L. V. Prasad Eye Institute, Banjara Hills, Hyderabad, between January 2010 and December 2012. **Country**: India. **Ocular comorbidities**: patients with any media opacities that prevented good-quality optic disc photographs and SDOCT imaging and any retinal (including macular) disease other than glaucoma that could confound the evaluations, were excluded. **Spectrum of glaucoma severity**: the mean (range) MD and PSD on the VF test were -2.14 (-4.25 to -0.98) and 1.82 (1.44 to 2.18) respectively for preperimetric glaucoma. **Control participants:** **- control group 1**: patient referred by general ophthalmologists, as glaucoma suspects based on the optic disc morphology. Their optic discs were confirmed on masked evaluation of disc photographs by the glaucoma experts to be non-glaucomatous with large physiologic cupping. All patients had IOP < 22 mmHg in both eyes, no past history of increased IOP, no family history of glaucoma, and normal VF. - c**ontrol group 2:** no suspicious findings for glaucoma, a normal ocular examination, IOP < 22 mmHg in both eyes, no past history of increased IOP, no family history of glaucoma, and normal VF.

Rao 2013 *(Continued)*

Index tests	**Optical coherence tomography**: RTVue SD-OCT (Optovue, Inc., Fremont, CA, USA), software version 5.1.0.90. ONH and GCC scanning protocols were used for analysis. Optic disc cube 200 x 200 and macular cube 200 x 200 scanning protocols were used for analysis. Images not well centred with a signal strength index < 30, or segmentation algorithm failed, were excluded. One of the author had conflict of interest with the manufacturer.
Target condition and reference standard(s)	**Manifest glaucoma:** glaucomatous optic neuropathy (neuroretinal rim thinning, notching, and/or RNFL defects), and normal VF results (PSD with P < 5% or the GHT within normal limits). **Visual field test:** Humphrey Field Analyzer, model 750 (Zeiss Humphrey Systems, Dublin, CA, USA), 24-2 SITA standard programme. Reliable exams had fixation losses, false-positive and false-negative response rates < 20%. **Optic disc evaluation**: digital optic disc photographs (450plus with VISUPAC 4.2.2; Carl Zeiss Meditec Systems GmbH, Pirmasens, Germany). Optic disc photographs were evaluated independently by 2 glaucoma experts who were masked to other clinical examination results.
Flow and timing	Quote: "Eyes in which the segmentation algorithm failed were excluded", but no further details about exclusion reported. Index and reference test were performed on the same day.
Comparative	
Notes	None.

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	No		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (All tests)			
If a threshold was used, was it prespecified?	Yes		
Were imaging test's quality assessed?	Yes		

Rao 2013 *(Continued)*

Were any conflict of interest avoided	No		
Could the conduct or interpretation of the index test have introduced bias?		High risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Unclear		
Did all patients receive a reference standard	Yes		
Could the patient flow have introduced bias?		Unclear risk	

Rao 2014

Study characteristics	
Patient Sampling	Consecutive participants referred by general ophthalmologists to a tertiary eye-care facility between September 2010 and November 2012 for a glaucoma evaluation. The control

Rao 2014 *(Continued)*

	group against which the glaucoma cohort was evaluated consisted of people referred to the Institute from general ophthalmologists as glaucoma suspects based on the optic disc appearance.
Patient characteristics and setting	**Sample size**: 215 eyes of 165 participants (106 eyes of 79 glaucoma patients and 109 eyes of 86 controls). **Age**: glaucoma median (interquartile range), 53 (48, 59) years; controls, 54 (45, 62) years. **Sex**: not reported. **Ethnicity**: not reported. **Clinical Setting:** Glaucoma Center, L. V. Prasad Eye Institute, Hyderabad. **Country**: India. **Ocular comorbidities**: patients with any media opacities that prevented good imaging and any retinal (including macular) or neurological diseases other than glaucoma which could confound the VF or imaging examination were excluded. **Spectrum of glaucoma severity**: the median (interquartile range) MD and PSD on the VF test were -6.86 (-12.11, -4.65) and 7.58 (3.99, 10.49) respectively, for glaucomatous eyes. **Control participants:** normal optic disc appearance and normal VF result.
Index tests	**Optical coherence tomography** RTVue (software version 5.1.0.90; Optovue Inc, Fremont, CA, USA). ONH scanning protocol used for the analysis. Only well-centred images with a signal strength index of ≥ 30 and no segmentation algorithm failure were used for the analysis. **Scanning laser polarimetry**: GDxPRO (version 1.1.1; Carl Zeiss Meditec, Inc.). Only well-focused, centred and illuminated images with a quality score of ≥ 7, a typical scan score > 80, and a residual anterior segment retardation of < 4 were included for analysis. The authors stated no conflicts of interest.
Target condition and reference standard(s)	**Manifest perimetric glaucoma:** glaucomatous optic disc appearance (based on the presence of focal or diffuse neuroretinal rim thinning, localised notching or nerve fibre layer defects) and VF defect (defined as PSD with P < 5% and GHT outside normal limits). **Visual field test:** Humphrey Field Analyzer model 750 (Zeiss Humphrey Systems, Dublin, CA, USA); 24-2 SITA–standard strategy. All exams had fixation losses, false-positive and false-negative rates of < 20%. **Optic disc evaluation**: Digital photographs (Visupac 4.2.2; Carl Zeiss Meditec Systems GmbH, Pirmasens, Germany). 2 experts, masked to the clinical examination results of the participants, evaluated independently the photographs.
Flow and timing	69 eyes were (> 10%) were excluded due to poor-quality imaging tests or unreliable reference standard. 61 eyes were excluded due to the optic disc classification not correlating with visual fields. The index and reference test were performed on the same day.
Comparative	
Notes	None.

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns

Rao 2014 *(Continued)*

DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Unclear		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		Unclear risk	
Are there concerns that the included patients and setting do not match the review question?			Unclear
DOMAIN 2: Index Test (All tests)			
If a threshold was used, was it prespecified?	Yes		
Were imaging test's quality assessed?	Yes		
Were any conflict of interest avoided	Yes		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		

Rao 2014 *(Continued)*

Did all patients receive the same reference standard?	Yes	
Were all patients included in the analysis?	No	
Did all patients receive a reference standard	Yes	
Could the patient flow have introduced bias?		High risk

Reus 2004

Study characteristics	
Patient Sampling	Glaucoma patients were recruited consecutively from an ongoing longitudinal follow-up study. Healthy participants were recruited either consecutively from an ongoing longitudinal follow-up study or from employees of The Rotterdam Eye Hospital and their spouses and friends. One eye per healthy participant was selected randomly. One eye per glaucoma patient was selected, choosing the eye with the more positive MD at VF, if both were eligible.
Patient characteristics and setting	**Sample size**: 239 eyes of 239 participants were enrolled, 219 eyes were actually included in the analysis (146 glaucoma, 73 healthy controls). **Age:** glaucoma patients mean ± SD, 61 ± 10 years; controls 59 ± 11. **Sex:** 115 men (81 glaucoma, 34 controls) and 104 women (65 glaucoma, 39 controls). **Ethnicity:** white. **Country:** Netherland. **Ocular comorbidities**: no history of ocular disease (as posterior segment eye disease and corneal disease), BCVA < 20/40, previous intraocular surgery (except for uncomplicated cataract surgery), systemic hypertension or diabetes. **Setting:** Rotterdam Eye Hospital. **Spectrum of glaucoma severity:** mean ± SD MD and PSD on the VF test were -8.45 ± 6.81 dB and 8.13 ± 3.88 dB. According to the Hodapp et al. grading scale: 37 eyes had mild glaucoma, 28 moderate, 81 severe. **Control participants:** IOP < 21 mmHg in both eyes, normal visual fields (GHT within normal limits and no nerve fibre bundle VF defects in the total or pattern deviation probability plots or both) and healthy-looking ONH.
Index tests	**Scanning laser polarimetry**: GDx VCC (Laser Diagnostic Technologies, Inc., San Diego, CA, USA). All scans were acquired through undilated pupils, and were of high quality (i.e. with a centred optic disc, well focused, even and just illuminated throughout the field). No author had conflict of interest.
Target condition and reference standard(s)	**Manifest glaucoma:** glaucomatous ONH appearance with a corresponding glaucomatous nerve fibre bundle abnormality on the total and/or pattern deviation probability plots with SAP and open angle by gonioscopy.

Reus 2004 *(Continued)*

	Visual field testing: Humphrey Field Analyzer, 24-2 SITA-Standard strategy (Carl Zeiss Meditec, Dublin, CA, USA). VF reliability criteria were not reported.
Flow and timing	The time interval between reference standard and index test was not reported. 20 (< 10%) participants were excluded from the analysis: in 4 healthy and 12 glaucoma participants the GDx VCC software flagged measurements as "results may not be compatible with normative database", in 4 glaucoma patients high-quality images could not be obtained.
Comparative	
Notes	None.

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	No		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (All tests)			
If a threshold was used, was it pre-specified?	Yes		
Were imaging test's quality assessed?	Yes		
Were any conflict of interest avoided	Unclear		
Could the conduct or interpretation of the index test have introduced bias?		Unclear risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		

Reus 2004 *(Continued)*

Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Unclear		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
Did all patients receive a reference standard	Yes		
Could the patient flow have introduced bias?		Unclear risk	

Reus 2007

Study characteristics	
Patient Sampling	Healthy participants and glaucoma patients were selected from a cohort of patients and controls who had been originally recruited for an ongoing longitudinal glaucoma study. Healthy participants had been recruited from spouses and friends of patients and from employees of the Rotterdam Eye Hospital and their spouses and friends. One eye per person was randomly selected.
Patient characteristics and setting	**Sample size**: 94 eyes of 94 participants (48 glaucoma, 6 ocular hypertensive, 40 healthy controls). **Age:** glaucoma patients mean, 61 years; controls 59. **Sex:** 45 men (26 glaucoma, 19 controls) and 43 women (22 glaucoma, 21 controls). **Ethnicity:** white. **Country:** Netherland. **Ocular comorbidities**: no co-existing ocular diseases, BCVA < 20/40, previous intraocular surgery (except for any uncomplicated cataract surgery or, if applicable, glaucoma surgery), or diabetes mellitus. **Setting:** Rotterdam Eye Hospital. **Spectrum of glaucoma severity:** mean ± SD MD and PSD on the VF test were -6.56 ± 6.32 dB and 7.71 ± 4.03 dB, respectively. **Control participants:** IOP < 21 mmHg in both eyes, normal VFs (GHT within normal limits and no nerve fibre bundle VF defects in the total or pattern deviation probability plots or both) and healthy-looking ONH.

Reus 2007 *(Continued)*

Index tests	**Scanning laser polarimetry**: GDx VCC, software version 5.4.0 (Carl Zeiss Meditec AG, Jena, Germany). Only high-quality scans, i.e. with a centred ONH, well focused, evenly and justly illuminated throughout the image, and without any motion artefacts, were accepted. Some authors had conflict of interest.
Target condition and reference standard(s)	**Manifest glaucoma:** glaucomatous optic disc appearance (with notching or thinning of the neuroretinal rim) and a reproducible corresponding nerve fibre bundle VF defect. **Visual field testing:** Humphrey Field Analyzer, II (Carl Zeiss Meditec AG, Jena, Germany). The details about reliability criteria were not specified.
Flow and timing	The time interval between reference standard and index tests was not reported. All participant selected were included in the analysis.
Comparative	
Notes	None.

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Was a case-control design avoided?	No		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (All tests)			
If a threshold was used, was it pre-specified?	Yes		
Were imaging test's quality assessed?	Yes		
Were any conflict of interest avoided	No		
Could the conduct or interpretation of the index test have introduced bias?		High risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			

Reus 2007 *(Continued)*

Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Unclear		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
Did all patients receive a reference standard	Yes		
Could the patient flow have introduced bias?		Unclear risk	

Rho 2014

Study characteristics	
Patient Sampling	Case-control study enrolling consecutive early glaucoma eyes and age-matched healthy control eyes in 2013 at CHA Bundang Medical Center, Seongnam, Republic of Korea. One eye per person was randomly selected.
Patient characteristics and setting	**Sample size**: 120 eyes of 120 participants (58 early glaucoma and 62 healthy controls). **Age**: glaucoma mean (range), 53.31 (19 to 76) years; controls, 52.05 (20 to 70) years. **Sex**: 52 men (24 glaucoma, 28 controls) and 68 women (34 glaucoma, 34 controls) **Ethnicity**: Korean. **Clinical Setting:** CHA Bundang Medical Center, CHA University, Seongnam. **Country**: Korea. **Ocular comorbidities**: patients with retinal disease, neuro-ophthalmologic disease, history of refractive or retinal surgery within 3 months, or closed iridocorneal angle and refractive error more than 68.0 diopters and 63.0 diopters of cylinder were excluded.

Rho 2014 *(Continued)*

	Spectrum of glaucoma severity: the mean (range) MD on the VF test were -1.60 (-5.65 to 0.50), for glaucomatous eyes of the validation group. All glaucoma had MD > -6. **Control participants:** IOP < 21 mmHg, normal optic disc, or normal VF test results.
Index tests	**Optical coherence tomography**: Spectralis OCT (Heidelberg Engineering GmbH, Heidelberg, Germany), software version 5.4.7.0. Peripapillary RNFL thickness-es scanning protocol was used for the analysis. Only images with image quality scores > 22 were accepted. The authors had no disclosure to be declared.
Target condition and reference standard(s)	**Manifest glaucoma:** glaucomatous optic disc contours and a corresponding ab-normal VF result (defined as 1+ of the following criteria: GHT outside normal limits; PSD with P < 0.05; 3+ non-edged points in a cluster decreased with P < 0.05, with 1 of these with P < 0.01) or RNFL defects plus a corresponding abnormal VF result. **Visual field test:** Humphrey Field Analyzer II (Carl Zeiss Meditec, Inc., Dublin, CA, USA); 24-2 SITA–standard strategy. Reliable VF were defined by fixation loss < 30%, and false-positive and false-negative rates of < 20%. **RNFL evaluation**: fundus photography and red-free photography with a fundus camera (VX-10i; Kowa, Nagoya, Japan).
Flow and timing	The time interval between index and reference test was not specified.
Comparative	
Notes	None.

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of pa-tients enrolled?	Yes		
Was a case-control design avoided?	No		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have intro-duced bias?		High risk	
Are there concerns that the included pa-tients and setting do not match the review question?			High
DOMAIN 2: Index Test (All tests)			
If a threshold was used, was it pre-specified?	Yes		
Were imaging test's quality assessed?	Yes		
Were any conflict of interest avoided	Yes		

Rho 2014 *(Continued)*

Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
Could the reference standard, its conduct, or its interpretation have introduced bias?		High risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Unclear		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
Did all patients receive a reference standard	Yes		
Could the patient flow have introduced bias?		Unclear risk	

Roberti 2014

Study characteristics	
Patient Sampling	Prospective, case-control study, conducted at the IRCCS-Fondazione G. B. Bietti, Rome. One eye per person was included.
Patient characteristics and setting	**Sample size**: 104 eyes of 104 participants (46 glaucoma and 58 controls). **Age**: glaucoma mean ± SD, 61 ± 12.9 years; controls, 58.5 ± 11.3 years. **Sex**: not reported. **Ethnicity**: not reported. **Clinical setting**: IRCCS-Fondazione G. B. Bietti, Rome. **Country**: Italy.

Roberti 2014 *(Continued)*

	Ocular comorbidities: patients with any active or past retinal pathologies (including diabetic retinopathy or age-related macular degeneration), opacities of optic media, history of ocular surgery (except for uncomplicated cataract or glaucoma surgery), were excluded. **Spectrum of glaucoma severity**: the mean ± SD MD and PSD on the VF test were -7 ± 5.9 and 6.9 ± 4.6 respectively, for glaucomatous eyes. **Control participants:** IOP < 22mmHg, normal-appearing optic disc, and normal VF test result.
Index tests	**Confocal scanning laser ophthalmoscopy:** Heidelberg Retina Tomograph (HRT3; Heidelberg Engineering GmbH, Heidelberg, Germany). Software version 3.0. Only images with acquisition sensitivity of > 90% and a SD < 40, were used for the study. **Optical coherence tomography**: RTVue-100, software version 5.1.0.90 (Carl Zeiss Meditec Systems GmbH, Pirmasens, Germany). Only images with signal strength index >50 were accepted. The authors reported no conflict of interest.
Target condition and reference standard(s)	**Manifest glaucoma:** glaucomatous VF defect, defined as the consistent presence of a cluster of 3+ non-edge points on the pattern deviation plot with a probability of occurring in < 5% of the normal population with one of these points having the probability of occurring in < 1% of the normal population, a PSD with P < 5%, or a GHT result outside normal limits. **Visual field test:** Humphrey Field Analyzer, SITA-standard 24-2 (Carl Zeiss Meditec, Dublin, CA). Reliable VF were defined by fixation loss and false-negative rate < 25%, and false-positive < 15%.
Flow and timing	The time interval between index and reference test was not specified.
Comparative	
Notes	None.

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	No		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (All tests)			

Roberti 2014 *(Continued)*

If a threshold was used, was it pre-specified?	Yes		
Were imaging test's quality assessed?	Yes		
Were any conflict of interest avoided	Yes		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Unclear		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
Did all patients receive a reference standard	Yes		
Could the patient flow have introduced bias?		Unclear risk	

Rolle 2011

Study characteristics	
Patient Sampling	Glaucoma preperimetric patients consecutively enrolled and sex- and age-matched normal control participants from normal healthy population were recruited between October 2009 and September 2010. One eye per person was randomly selected.
Patient characteristics and setting	**Sample size**: 178 eyes of 178 participants (126 preperimetric glaucoma, 52 healthy controls)

Rolle 2011 *(Continued)*

	Age: preperimetric glaucoma eyes mean ± SD 58.1 ± 6.91 years; controls 57.8 ± 6.71 years. **Sex**: 76 men, 102 women. **Ethnicity**: white. **Setting**: Eye Clinic, Section of Ophthalmology, Department of Clinical Physiopathology, University of Turino. **Country:** Italy. **Ocular comorbidities**: patients with previous intraocular surgery, diabetic retinopathy or other diseases that could cause VF loss or optic disc abnormalities were excluded. All eyes had to have BCVA ≥ 20/40, spherical equivalent refractive error ≤ +3 D or ≥ -6 D. **Spectrum of glaucoma severity**: the mean ± SD MD on the VF test were 1.41 ± 0.7 dB, for glaucoma. **Control participants:** IOP < 21 mmHg, normal VF test (MD and PSD within 95% limits of the normal reference and a GHT within 97% limits), normal ONH/RNFL appearance (intact neuroretinal rim without peripapillary haemorrhages, notches, localised pallor, or RNFL defects) and open angle by gonioscopy.
Index tests	**Optical coherence tomography**: FD-OCT RTVue-100 (software version A4, 5, 0, 59, Optovue Inc, Fremont, CA, USA). The ONH and GCC scanning protocols were used to acquire the images. Images with motion artefacts, segmentation errors and signal strength index < 45 were excluded. Authors stated no conflict of interest.
Target condition and reference standard(s)	**Manifest glaucoma:** glaucomatous optic disc changes (defined as optic rim notch or diffuse/generalised loss of optic rim tissue; vertical cup/disc diameter ratio asymmetry, unexplained by side differences in optic disc size), disc haemorrhages in conjunction with the finding of IOP > 21 mmHg and normal VF result. **Visual field testing:** Humphrey Field Analzyer 24-2 SITA standard programme (Carl Zeiss Meditec). Reliability criteria were fixation losses, false-positive and false-negative rates ≤ 25%. **Optic disc evaluation:** slit-lamp biomicroscopy.
Flow and timing	The time interval between index test and reference standard was not reported. No patients were reported by the authors as excluded from the analysis.
Comparative	
Notes	None.

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	No		
Did the study avoid inappropriate exclusions?	Yes		

Rolle 2011 *(Continued)*

Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (All tests)			
If a threshold was used, was it pre-specified?	Yes		
Were imaging test's quality assessed?	Yes		
Were any conflict of interest avoided	Yes		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Unclear		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
Did all patients receive a reference standard	Yes		
Could the patient flow have introduced bias?		Unclear risk	

Schrems 2010

Study characteristics	
Patient Sampling	Healthy participants, ocular hypertensive and glaucoma (preperimetric and perimetric) patients were recruited from the Erlangen Glaucoma Registry (a clinical registry for cross-sectional and longitudinal observation study of patients with open-angle glaucoma or glaucoma suspect).
Patient characteristics and setting	**Sample size**: 386 participants (95 perimetric glaucoma, 89 preperimetric glaucoma, 145 ocular hypertensive, 57 controls). **Age:** preperimetric glaucoma mean ± SD, 55.7 ± 11.3 years; perimetric glaucoma mean ± SD 56.4 ± 11.2 years; ocular hypertensive mean ± SD 53.9 ± 12years; controls 49.9 ± 13 years. **Sex:** 179 men (30 preperimetric glaucoma,56 perimetric glaucoma, 76 OHT, 17 controls) and 207 women (59 preperimetric glaucoma, 39 perimetric glaucoma, 69 OHT, 40 controls) **Country:** Germany. **Ocular comorbidities**: no ocular diseases other than glaucoma,BCVA < 16/40, spherical refraction > ±8 D, diabetes. **Setting:** Department of Ophthalmology, University of Erlangen-Nuremberg, Schwabachanlage, Erlangen. **Spectrum of glaucoma severity:** mean ± SD MD/PSD on the VF test were -0.37 ± 1.3/2.48 ± 1.72 dB, for preperimetric glaucoma; -6.26 ± 5.26/32.6 ± 28.8 dB for perimetric glaucoma; 0.44 ± 1.4/2.19 ± 1.88 dB for OHT. **Control participants:** normal VFs and normal clinical examination.
Index tests	**Scanning laser polarimetry:** GDx VCC (Carl Zeiss Meditec Inc, Dublin, CA, USA). A score ≥ 7 was the minimum standard for good-quality scans in this study. No author had conflict of interest.
Target condition and reference standard(s)	**Manifest perimetric glaucoma:** IOP > 21 mmHg, abnormal appearance of the optic disc (unusually small neuroretinal rim area in relation to the optic disc size and cup/disc ratios being higher vertically compared with horizontally or notching, or localised/diffuse RNFL loss) and glaucomatous VF defects (defined by a reproducible reduction in sensitivity of at least 10 dB in a cluster of ≥ 2 contiguous locations and/or a deterioration of at least 5 dB in a cluster of ≥ 3 contiguous locations with at least one of those with ≥ 10 dB), with open angle by gonioscopy. **Manifest preperimetric glaucoma:** IOP > 21 mmHg, glaucomatous optic disc appearance without any corresponding VF loss. **OHT**: IOP > 21 mmHg, with normal optic disc appearance and VF test result. **Visual field test:** No details were reported about how VF testing was conducted. **Optic disc evaluation:** 15° colour photographs (Zeiss telecentric fundus camera, Germany). The analyses were independently performed by 2 glaucoma specialists.
Flow and timing	The time interval between reference standard and index tests was not reported. All participants recruited were included in the analysis.
Comparative	

Schrems 2010 *(Continued)*

Notes	None.

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Was a case-control design avoided?	No		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (All tests)			
If a threshold was used, was it pre-specified?	Yes		
Were imaging test's quality assessed?	Yes		
Were any conflict of interest avoided	Yes		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Unclear		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Unclear risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			

Schrems 2010 *(Continued)*

Was there an appropriate interval between index test and reference standard?	Unclear
Did all patients receive the same reference standard?	Yes
Were all patients included in the analysis?	Yes
Did all patients receive a reference standard	Yes
Could the patient flow have introduced bias?	Unclear risk

Sehi 2007

Study characteristics	
Patient Sampling	Healthy volunteers (such as office employees and friends or family members of patients with glaucoma) and patients with glaucoma who met the eligibility criteria were prospectively enrolled. One eye per person was randomly selected.
Patient characteristics and setting	**Sample size**: 158 eyes of 158 participants (63 glaucoma, 95 healthy controls) stratified into 2 groups based on the TSS obtained with SLP-VCC. **Age:** glaucoma patients mean ± SD, 63.3 ± 9.0 years; controls 54.6 ± 10.5 years. **Sex:** 53 men (25 glaucoma, 28 normal,) and 105 women (38 glaucoma, 67 control). **Ethnicity:** 137 white non-Hispanic, 11 black, 6 Asian, 2 Pacific Islander and 2 Hispanic. **Country:** USA. **Ocular comorbidities**: no previous intraocular surgery (except for uncomplicated cataract extraction), BCVA < 20/40, SE > ±5 D, ocular disease other than glaucoma or cataract, peripapillary atrophy, or retinal disease. **Setting:** Institutes involved in the AIG study: Oregon health and science university; University of Southern California; Bascom Palmer Eye Institute, University of Miami; Eye Center, University of Pittsburgh Clinical Center. **Spectrum of glaucoma severity:** mean ± SD MD and PSD on the VF test were -4.2 ± 4.3 dB and 5.4 ± 4.3 dB, respectively. **Control participants:** IOP ≤ 21 mmHg, normal optic disc appearance and normal VF results (GHT within normal limits, and MD and PSD of P > 5%).
Index tests	**Scanning laser polarimetry**: GDx-ECC and GDx VCC, software version 5.5.0 (Carl Zeiss Meditec Dublin, CA, USA). 3 consecutive scans were obtained through undilated pupils, with VCC and ECC on the same day by the same examiner. The average of 3 measurements was used for the analysis. Images that were obtained during eye movement were excluded, as well as unfocused, poorly-centred images or images with a quality scan score < 8. One author had conflict of interest.

Sehi 2007 *(Continued)*

Target condition and reference standard(s)	**Manifest glaucoma:** glaucomatous optic nerve damage (defined as either cup-to-disc asymmetry between fellow eyes of > 0.2, rim thinning, notching, excavation, or RNFL defect) and corresponding abnormal SAP result (GHT and PSD outside 95% normal limits). **Visual field testing:** Humphrey Field Analyzer, 24-2 SITA-standard strategy (Carl-Zeiss Meditec, Inc., Dublin, CA, USA). SAP reliability criteria included fixation losses rates, false-positive and false-negative rates of < 33%. **Optic disc evaluation:** dilated stereoscopic examination.
Flow and timing	Index tests were performed on the same day but no details about the time interval between reference standard and index test was reported. No patients were reported as excluded from the analysis.
Comparative	
Notes	None.

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Was a case-control design avoided?	No		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (All tests)			
If a threshold was used, was it pre-specified?	Yes		
Were imaging test's quality assessed?	Yes		
Were any conflict of interest avoided	No		
Could the conduct or interpretation of the index test have introduced bias?		High risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			

Sehi 2007 *(Continued)*

Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Unclear		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
Did all patients receive a reference standard	Yes		
Could the patient flow have introduced bias?		Unclear risk	

Seong 2010

Study characteristics	
Patient Sampling	Glaucoma patients were recruited prospectively, in a consecutive manner. The controls consisted of hospital staff, staff family members, spouses of patients, or volunteers. One eye per person was randomly selected.
Patient characteristics and setting	**Sample size**: 167 eyes of 167 participants (102 normal tension glaucoma, 65 healthy controls). **Age:** glaucoma patients mean ± SD, 54.9 ± 11.4 years; controls 52.7 ± 12.1 years. **Sex:** 82 men (49 glaucoma, 33 controls) and 85 women (53 glaucoma, 32 controls). **Country:** South Korea. **Ocular comorbidities**: no ocular diseases other than glaucoma, BCVA < 20/30, spherical refraction > ±5 D, cylinder refraction > ±3 D, close angle by gonioscopy, neurological diseases, or diabetes. **Setting:** Glaucoma clinic of the Asan Medical Center, Seoul. **Spectrum of glaucoma severity:** according to the Hodapp et al. grading scale, 56 eyes had early glaucoma, 46 eyes moderate-to-advanced. Mean ± SD MD/PSD on the VF test were -2.62 ± 1.72/3.43 ± 2.03 dB, for early glaucoma; -12.1 ± 4.4/10.1 ± 3.55 dB for moderate to advanced glaucoma.

Seong 2010 *(Continued)*

	Control participants: IOP < 22 mmHg, no history of IOP elevation and no perimetric defects.
Index tests	**Optical coherence tomography**: RTVue-100 OCT, software version 4.0.0.143 (Optovue, Inc.). The GCC, NHM4 and RNFL 3.45 scan protocols were acquired after pupil dilation by a single well-trained operator who was masked to the diagnosis. Images with signal strength index < 40 or with overt misalignment of the surface detection algorithm were excluded from the analysis. No author had conflict of interest.
Target condition and reference standard(s)	**Normal Tension Glaucoma:** a maximum IOP < 22 mmHg before any antiglaucoma therapy, open angle by gonioscopy, glaucomatous VF defects (defined as a cluster of 3 points with P < 5% on the pattern deviation map in at least 1 hemifield, including at least 1 point with P < 1%; or a cluster of 2 points with P < 1% and a GHT result outside 99% of normal limits; or a PSD outside 95% of normal limits), and glaucomatous optic disc appearance (increased cupping or a difference in vertical cup-disc ratio of > 0.2 between eyes, or diffuse/focal neural rim thinning, disc haemorrhage, or RNFL defects). **Visual field testing:** Humphrey Field Analyzer, 24-2 SITA-standard strategy (Carl Zeiss Meditec, Dublin, CA, USA). Reliability criteria included fixation losses rates < 20%, false-positive and false-negative rates < 15%.
Flow and timing	All index test images were acquired during the same patient visit but no details about the time interval between reference standard and index test was reported. 12 eyes (< 10%) were excluded from the analysis due to poor image quality.
Comparative	
Notes	None.

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	No		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (All tests)			
If a threshold was used, was it pre-specified?	Yes		
Were imaging test's quality assessed?	Yes		

Seong 2010 *(Continued)*

Were any conflict of interest avoided	Yes		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Unclear		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
Did all patients receive a reference standard	Yes		
Could the patient flow have introduced bias?		Unclear risk	

Shah 2006

Study characteristics	
Patient Sampling	All participants were selected among people enrolled prospectively in the longitudinal Diagnostic Innovations in Glaucoma Study. One eye per person was randomly chosen.
Patient characteristics and setting	**Sample size**: 123 eyes of 123 participants were enrolled. 101 eyes (43 glaucoma, 58 controls) in the first analysis (functional definition of glaucoma). 114 eyes (65 glaucoma, 49 controls) in the second analysis (structural definition of glaucoma). **Age:** glaucoma patients mean ± SD, 68.3 ± 3.5 years, controls 58.6 ± 2 years, for the first analysis; glaucoma patients mean ± SD, 65.5 ± 3 years, controls 60.1 ± 3.5 years, for the second analysis.

Shah 2006 *(Continued)*

	Sex: 44 men (22 glaucoma, 22 controls) and 57 women (21 glaucoma, 36 controls) in the first analysis; 35 men (27 glaucoma, 18 controls) and 69 women (38 glaucoma, 31 controls) in the second analysis. **Ethnicity:** 88 white (47 glaucoma, 41 controls) in the first analysis; 99 white (57 glaucoma, 42 controls) in the second analysis. **Country:** USA. **Ocular comorbidities**: no previous intraocular surgery (except for uncomplicated cataract or glaucoma surgery), BCVA < 20/40, spherical refraction > ±5 D, cylinder refraction > ±3 D, ocular diseases other than glaucoma, close angle by gonioscopy, non-glaucomatous secondary causes of elevated IOP. **Setting:** Hamilton Glaucoma Center, University of California, San Diego. **Spectrum of glaucoma severity:** no details reported **Control participants:** no glaucomatous VF damage and no history of IOP > 22 mmHg.
Index tests	**Scanning laser polarimetry:** GDx VCC,software version 5.5.0.14 (Carl ZeissMeditec, Inc., Dublin, CA, USA). Only images of good quality as assessed by an expert examiner(focused and evenly-illuminated reflectance image with a centred optic disc, a residual anterior segment retardation of < 15 nm, and a typical scan score of > 25) were included. Some authors had conflict of interest.
Target condition and reference standard(s)	**Manifest glaucoma:** 2 parallel analyses were conducted on 2 sample of patients partly overlapping (some patients were included in both analyses): the first one using a functional definition of glaucoma (repeatable glaucomatous field loss by SAP, defined as PSD outside the 95% normal confidence limits or GHT outside normal limits) and the second one using a structural definition of glaucomatous optic neuropathy based on assessment of optic disc stereophotographs (defined as focal rim notching, rim thinning, or RNFL abnormality). **Visual field testing:** Humprheys Field Analyzer, 24-2 SITA-standard strategy (Carl Zeiss Meditec). Reliability criteria included fixation losses rates, false-positive and false-negative rates < 25%. **Optic disc evaluation:** each stereoscopic optic disc photographs was evaluated by 2 expert graders in a masked fashion. Adjudication by a third expert grader was completed in cases of disagreement. We extracted data only for analysis using a functional definition of the reference standard.
Flow and timing	The time interval between reference standard and index tests was < 6 months. Of 123 eyes, 101 were included in the SAP analysis group, and 114 were included in the stereophotograph analysis group.
Comparative	
Notes	None.

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		

Shah 2006 *(Continued)*

Was a case-control design avoided?	No		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (All tests)			
If a threshold was used, was it prespecified?	Yes		
Were imaging test's quality assessed?	Yes		
Were any conflict of interest avoided	No		
Could the conduct or interpretation of the index test have introduced bias?		High risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	No		
Did all patients receive the same reference standard?	Yes		

Shah 2006 *(Continued)*

Were all patients included in the analysis?	No	
Did all patients receive a reference standard	Yes	
Could the patient flow have introduced bias?		High risk

Shin 2013

Study characteristics	
Patient Sampling	Glaucoma patients with localised RNFL defects and normal controls who visited the glaucoma centre from September 2010 to August 2011 were enrolled. One eye per person was randomly selected.
Patient characteristics and setting	**Sample size**: 136 eyes of 136 participants enrolled (64 glaucoma, 72 healthy controls). **Age**: not reported. **Sex**: not reported. **Ethnicity**: not reported. **Country**: Korea. **Setting**: Glaucoma Center at Hanyang University Medical Center, Seoul. **Ocular comorbidities**: patients with any ophthalmic or neurological disease known to affect RNFL thickness or BCVA < 20/40, spherical equivalent refractive errors < -8.0 D or > +4.0 D, were excluded. **Spectrum of glaucoma severity**: The mean ± SD MD on the VF test were -6.26 ± 4.16 dB for glaucoma. **Control participants:** no history of IOP > 21 mmHg, a normal ONH and RNFL appearance on cSLO RNFL photographs and normal VF test result.
Index tests	**Optical coherence tomography**: 3D OCT-2000 (software version 7.11; Topcon, Tokyo, Japan). The 3D disc scanning protocol was used to acquire the images. All images had to have quality score > 50. The authors stated no conflict of interest.
Target condition and reference standard(s)	**Manifest glaucoma:** the presence of localised RNFL defects on cSLO RNFL photographs associated with glaucomatous optic nerve appearance (defined as increased cupping, neuroretinal rim notching, optic disc haemorrhage, or cup-to-disc ratio > 0.2 between the eyes) and corresponding VF defects. **Visual field testing:** Humphrey Field Analzyer (Carl Zeiss Meditec). **RNFL evaluation:** Wide-angle (60°) red-free RNFL photographs were obtained with a confocal scanning laser ophthalmoscope (cSLO, F-10; Nidek, Gamagori, Japan) using the blue reflectance imaging technique. All topographic measurements of RNFL defects were performed by 2 masked examiners.
Flow and timing	The index test and reference standard were performed on the same day.

Shin 2013 *(Continued)*

	No patients were reported by the authors as excluded from the analysis.
Comparative	
Notes	None.

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Was a case-control design avoided?	No		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (All tests)			
If a threshold was used, was it pre-specified?	Yes		
Were imaging test's quality assessed?	Yes		
Were any conflict of interest avoided	Yes		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			

Shin 2013 *(Continued)*

Was there an appropriate interval between index test and reference standard?	Yes	
Did all patients receive the same reference standard?	Yes	
Were all patients included in the analysis?	Yes	
Did all patients receive a reference standard	Yes	
Could the patient flow have introduced bias?		Low risk

Sullivan-Mee 2013

Study characteristics	
Patient Sampling	Glaucoma patients and normal controls were selected among those involved in a prospective, longitudinal, observational glaucoma research study. Patients fitting the inclusion criteria were selected from the study database. "Both eyes per person were considered and the lowest of the paired eye RNFL measurements was used for determining the ability of the measured parameters to identify early glaucoma in a patient."
Patient characteristics and setting	**Sample size**: 128 fitting inclusion criteria, 100 participants finally included in the analysis (50 glaucoma, 50 healthy controls). **Age**: glaucoma eyes mean ± SD 68.9 ± 9.1 years; controls 66.2 ± 9.4 years. **Sex**: 95 men (47 glaucoma, 48 controls) and 5 women (3 glaucoma, 2 controls). **Ethnicity**: 44 white non-Hispanic (22 glaucoma, 22 controls), 48 Hispanic (22 glaucoma, 26 controls), 5 black (1 glaucoma, 4 controls) and 3 American Indian (1 glaucoma, 2 controls). **Country**: USA. **Setting**: New Mexico Veterans Administration Health Care System, Albuquerque, New Mexico. **Ocular comorbidities**: patients with corneal or scleral pathologic conditions, prior refractive, corneal, or incisional glaucoma surgery, secondary glaucoma diagnoses, VF loss resulting from non-glaucomatous pathologic features (including retinal, optic nerve, or visual pathway disorders), refractive error > ±5 D, and astigmatism > ±3 D, were excluded. **Spectrum of glaucoma severity**: the mean ± SD MD and PSD on the VF test for glaucoma were -0.92 ± 1.74 dB and 2.78 ± 1.30 dB respectively for the right eye, -1.29 ± 1.54 dB and 2.74 ± 1.29 dB respectively for the left eye. All eyes had early glaucoma, according to the Hodapp et al. classification. **Control participants:** IOP < 22 mmHg, normal optic nerve appearance and normal VF.
Index tests	**Optical coherence tomography**: Spectralis SD-OCT (Heidelberg Engineering, Carlsbad, CA, USA). The RNFL 3.45 mm and the posterior pole asymmetry analysis scanning protocols were used to acquire the images. Images with poor quality (poor centration, segmentation errors, scan quality <15, more than 4 of 61 raster scans had significant segmentation errors, image signal prevented accurate boundary detection for Bruch's membrane or internal limiting membrane in all or part of 4 scans or more, significant retinal or vitreoretinal pathologic features were evident) were excluded. Authors stated no conflict of interest.

Sullivan-Mee 2013 *(Continued)*

Target condition and reference standard(s)	**Manifest glaucoma:** glaucomatous optic neuropathy (defined as thinning, excavation, rim erosion, or notch of the neuroretinal rim) and glaucomatous VF defect (defined as GHT results outside normal limits, the presence of at least 3 contiguous test points on the pattern deviation plot with P < 1% and at least 1 at P < 5%, not including points on the edge of the field, or both. **Visual field testing:** Humphrey Field Analzyer (24-2 SITA standard programme (Carl Zeiss Meditec, Inc, Dublin, CA, USA). Reliability VF criteria were fixation losses < 33%, false positive and false negative < 15%. **Optic nerve evaluation:** dilated fundus examination.
Flow and timing	The time interval between index test and reference standard was not reported. 28 participants (> 10%) were excluded from the analysis due to poor OCT scan quality or confounding retinal abnormalities (epiretinal membrane, vitreomacular traction syndrome, or large drusen).
Comparative	
Notes	The work was supported by the Veterans Administration Office of Research and Development. This support included a new investigator grant from the regional Veterans Integrated Service Network (VISN 18).

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Was a case-control design avoided?	No		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (All tests)			
If a threshold was used, was it pre-specified?	Yes		
Were imaging test's quality assessed?	Yes		
Were any conflict of interest avoided	Yes		
Could the conduct or interpretation of the index test have introduced bias?		High risk	

Sullivan-Mee 2013 *(Continued)*

Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Unclear		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	No		
Did all patients receive a reference standard	Yes		
Could the patient flow have introduced bias?		High risk	

Sung 2013

Study characteristics	
Patient Sampling	Case-control study including early and preperimetric glaucoma and healthy normal controls. One eye was randomly selected if both eyes were eligible.
Patient characteristics and setting	**Sample size**: 204 patients enrolled, 179 eyes of 179 participants included in the analysis (70 early glaucoma, 37 preperimetric glaucoma and 72 normal controls).

Sung 2013 *(Continued)*

	Age: early glaucoma mean ± SD, 53.97 ± 12.36 years; preperimetric 54.22 ± 12.70 years, controls group 50.68 ± 13.73 years. **Sex**: 99 men (41 early glaucoma, 17 preperimetric glaucoma, 41 controls) and 80 women (29 early glaucoma, 20 preperimetric glaucoma, 31 controls) **Ethnicity**: Korean. **Clinical setting**: Department of Ophthalmology, Chonnam National University Medical School and Hospital, between February 1, 2012 and July 30, 2012. **Country**: Korea. **Ocular comorbidities**: patients with history of other eye diseases like neurological disease leading to VF abnormality, or diabetic retinopathy or macular oedema or histories of intraocular surgery other than uncomplicated cataract surgery, were excluded. **Spectrum of glaucoma severity**: the median (first and third quartile values) MD and PSD on the VF test were -3.54 (-5.23, -2.08) and 2.83 (2.19, 4.81) respectively for early glaucoma; -0.72 (-1.51, -0.13) and 1.8 (1.53, 2.04) respectively for preperimetric glaucoma. **Control participants:** no family history of glaucoma, no previous intraocular surgery, IOP ≤ 21 mmHg, non-glaucomatous optic disc appearance and normal VF.
Index tests	**Optical coherence tomography**: Cirrus HD-OCT (Carl Zeiss Meditec Inc., Dublin, CA, USA). Macular cube 200 x 200 protocol and optic disc cube 200 x 200 protocol scans were used for analysis. Images with image quality factor < 6 and with eye movements or blinking artefacts or showing an apparent segmentation error were excluded. The authors declared no conflict of interest.
Target condition and reference standard(s)	**Manifest early glaucoma:** glaucomatous optic nerve damage (defined as the vertical cup-to-disc ratio ≥ 0.7 or > 0.2 asymmetry between the vertical cup-to-disc ratio of both eyes or focal neural rim notching or generalised loss of the neural rim) and VF loss with MD ≥ -6 (defined as having ≥ 3 non-edge, contiguous points with P < 0.05 and ≥ 1 points with P < 0.01 on the same side of horizontal meridian in the pattern SD plot and confirmed in at least 2 consecutive examinations. **Manifest preperimetric glaucoma:** normal VF with progressive glaucomatous optic nerve damage. **Visual field test:** Humphrey Field Analyzer (Carl Zeiss Meditec Inc.) 30-2 SITA standard. Reliable exams had fixation losses ≥ 20%, false positive and false negative response ≥ 33%. **Optic disc evaluation:** disc photography and red-free RNFL photography.
Flow and timing	10 eyes were excluded for low image quality, 7 eyes for intraretinal segmentation error, and 8 eyes for unreliable VF results (> 10%). The time interval between index and reference test was not reported.
Comparative	
Notes	None.

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			

Sung 2013 *(Continued)*

Was a consecutive or random sample of patients enrolled?	Unclear		
Was a case-control design avoided?	No		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (All tests)			
If a threshold was used, was it prespecified?	Yes		
Were imaging test's quality assessed?	Yes		
Were any conflict of interest avoided	Yes		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Unclear		

Sung 2013 *(Continued)*

Did all patients receive the same reference standard?	Yes	
Were all patients included in the analysis?	No	
Did all patients receive a reference standard	Yes	
Could the patient flow have introduced bias?		High risk

Takahashi 2008

Study characteristics	
Patient Sampling	Healthy participants and glaucoma patients were enrolled at the outpatient clinic. One eye per person was randomly selected.
Patient characteristics and setting	**Sample size**: 170 eyes of 170 participants (47 glaucoma, 38 glaucoma with diabetes, 40 with diabetes (without glaucoma) and 45 healthy controls). **Age:** glaucoma patients mean ± SD, 69.2 ± 8.3years; glaucoma patients with diabetes mean ± SD 71.3 ± 7.5years; diabetes patients mean ± SD, 66.2 ± 7.8 years ;controls 68.9 ± 5.9 years. **Country:** Japan. **Ocular comorbidities**: no neuro-ophthalmologic disease, BCVA < 32/40, spherical refraction > ±5 D, cylinder refraction > ±3 D uveitis, macular/retinal disease, or previous refractive or intraocular surgery. **Setting:** Senshokai Eye Institute in Kyoto. **Spectrum of glaucoma severity:** mean ± SD MD on the VF test was 6.56 ± 1.6 dB for glaucoma eyes (without diabetes) and 7.58 ± 2.1 dB for glaucoma with diabetes. **Control participants:** IOP < 22 mmHg, no history of diabetes or elevated IOP, a healthy optic disc, and no repeatable abnormal VF results.
Index tests	**Scanning laser polarimetry:** GDx VCC, software version 5.5.1 (Carl Zeiss Meditec, Dublin, CA, USA). Only high-quality images (defined as a well-focused and uniformly illuminated reflectance image with a centred optic disc that had minimal residual anterior segment retardation without an atypical retardation pattern) were included. No author had conflict of interest.
Target condition and reference standard(s)	**Manifest glaucoma:** glaucomatous ONH changes (defined as undermining of the cup, notching, focal/diffuse thinning of the rim area, nasal shifting of the retinal vessels or asymmetric enlargement of the cup (cup-to-disc asymmetries > 0.2)) and glaucomatous VF defect (defined as 3 consecutive point depressions exceeding 5 dB more than the age-matched controls and at least one of 3 consecutive points with a depression > 10 dB or 2 consecutive points depressed > 10 dB and 2 adjacent points across the nasal horizontal meridian with a difference of > 5 dB). **Visual field testing:** Octopus visual field analyzer, Octopus 301, version 2.04, full-threshold (G1) programme (Interzeag, Schlieren, Switzerland). No details about VF reliability criteria were reported.

Takahashi 2008 *(Continued)*

	Optic disc evaluation: 45° high-quality fundus colour photography (CF-PU2; Canon Inc., Tokyo, Japan). Two experienced graders measured each fundus colour photograph independently and were masked to the test results of the other.
Flow and timing	The time interval between reference standard and index tests was <3 months. Poor images from 36 participants were considered unacceptable and were excluded from this study.
Comparative	
Notes	None.

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Was a case-control design avoided?	No		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (All tests)			
If a threshold was used, was it pre-specified?	Yes		
Were imaging test's quality assessed?	Yes		
Were any conflict of interest avoided	Yes		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes		

Takahashi 2008 *(Continued)*

Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	No		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	No		
Did all patients receive a reference standard	Yes		
Could the patient flow have introduced bias?		High risk	

Weinreb 2003

Study characteristics	
Patient Sampling	Healthy and glaucoma consecutive patients who met the diagnostic inclusion criteria were enrolled. One eye per person was randomly selected.
Patient characteristics and setting	**Sample size**: 94 eyes of 94 participants (54 glaucoma, 40 healthy controls). **Age:** glaucoma patients mean ± SD, 68.7 ± 9.2 years; controls 64.0 ± 10.4. **Sex:** 41 men, 53 women. **Ethnicity:** 79 white, 5 Hispanic, 3 African-American, 2 Asian-American, 2 Indo-European, and 3 unknown. **Country:** USA. **Ocular comorbidities**: no co-existing retinal disease, BCVA < 20/40, uveitis, or non-glaucomatous optic neuropathy. **Clinical setting:** Hamilton Glaucoma Center, University of California, San Diego. **Spectrum of glaucoma severity:** mean ± SD MD on the VF test for glaucoma was -6.49 ± 4.94 dB. **Control participants:** no history of increased IOP, healthy appearance of the ONH/RNFL (no diffuse/focal rim thinning, cupping, or RNFL defects), and normal SAP results (MD and CPSD within 95% confidence limits, GHT within normal limits).

Weinreb 2003 *(Continued)*

Index tests	**Scanning laser polarimetry:** GDx Nerve Fibre Analyzer, version 2.0.01 modified with a VCC (Laser Diagnostic Technologies, San Diego, CA, USA). No details about quality images assessment were reported. One author had conflict of interest.
Target condition and reference standard(s)	**Manifest glaucoma:** repeatable (2 consecutive) glaucomatous VF test results (defined as a CPSD outside the 95% normal limits or a GHT outside the normal limits). **Visual field testing:** Humphrey Field Analyzer, 24-2 SITA-standard or full-threshold strategy (Zeiss-Humphrey Systems, Dublin, CA, USA). No details about VF reliability criteria were reported. **Optic disc evaluation:** dilated stereoscopic fundus examination.
Flow and timing	The time interval between reference standard and index test was not specified. No patients were reported as excluded from the analysis.
Comparative	
Notes	None.

Methodological quality

Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	No		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (All tests)			
If a threshold was used, was it pre-specified?	Yes		
Were imaging test's quality assessed?	Unclear		
Were any conflict of interest avoided	No		
Could the conduct or interpretation of the index test have introduced bias?		High risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern

Weinreb 2003 *(Continued)*

DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Unclear		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
Did all patients receive a reference standard	Yes		
Could the patient flow have introduced bias?		Unclear risk	

Wu 2012

Study characteristics	
Patient Sampling	Glaucoma patients and normal controls were prospectively recruited between January 2009 and July 2009. No more details about methods of selection were reported. One eye per person was randomly selected.
Patient characteristics and setting	**Sample size**: 146 eyes of 146 participants (61 glaucoma, 85 healthy controls). **Age**: glaucoma eyes mean ± SD 69.2 ± 13.0 years; controls 63.5 ± 14.0years. **Sex**: 65 men (25 glaucoma, 40 controls) and 81 women (36 glaucoma, 45 controls). **Ethnicity**: 104 white (41 glaucoma, 63 controls). **Country**: USA. **Setting**: Glaucoma Service, Massachusetts Eye and Ear Infirmary, Department of Ophthalmology, Harvard Medical School, Boston, Massachusetts. **Ocular comorbidities**: patients with congenital anomalies of the anterior chamber, corneal scarring or opacities, diabetic proliferative or severe nonproliferative retinopathy, visual field loss due to a non-glaucoma condition, were excluded. All eyes had to have BCVA ≥ 20/40 and spherical equivalent within ±5 D. **Spectrum of glaucoma severity**: the mean ± SD MD and PSD on the VF test were -9.61 ± 8.76 dB and 6.14 ± 3.43 dB respectively, for glaucoma.

Wu 2012 *(Continued)*

	Control participants: no ocular disease, except for mild cataracts, and normal VF test results, as defined by PSD > 5% and GHT results within normal limits.
Index tests	**Optical coherence tomography**: Spectralis OCT (software version, 4.0, Heidelberg Engineering, Inc, Heidelberg, Germany). The circular RNFl 3.45 mm was used to acquire the images. All the images without good quality (signal strength < 15, a clear fundus image with good optic disc and scan circle visibility, RNFL visible and without interruptions, and a continuous scan pattern without missing or blank areas) were excluded from the analysis. One author had conflict of interest.
Target condition and reference standard(s)	**Manifest glaucoma:** glaucomatous optic nerve changes and corresponding glaucomatous VF defect, defined as 3+ contiguous test locations in the PSD plot with P < 5%, with at least 1 with P < 1% on the same side of the horizontal meridian. **Visual field testing:** Humphrey Field Analzyer 750 (24-2 SITA standard programme ((Carl Zeiss Meditec, Inc, Dublin, CA, USA). Reliability criteria were fixation losses < 33%, false positive and false negative < 20%. **Optic nerve evaluation:** dilated ophthalmoscopy.
Flow and timing	Index test and reference standard were performed on the same day. No patients were reported by the authors as excluded from the analysis.
Comparative	
Notes	Supported in part by grant R01 EY14975-01 from the National Institutes of Health, Bethesda, Maryland.

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Was a case-control design avoided?	No		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (All tests)			
If a threshold was used, was it pre-specified?	Yes		
Were imaging test's quality assessed?	Yes		
Were any conflict of interest avoided	No		

Wu 2012 *(Continued)*

Could the conduct or interpretation of the index test have introduced bias?		High risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
Did all patients receive a reference standard	Yes		
Could the patient flow have introduced bias?		Low risk	

Yamada 2014

Study characteristics	
Patient Sampling	Retrospective case-control study comprised patients (preperimetric glaucoma, perimetric glaucoma and controls) who were screened for glaucoma at the Kyoto University Hospital from March 7, 2011, through November 19, 2012. One eye per person was randomly selected
Patient characteristics and setting	**Sample size:** 122 eyes of 122 participants (31 advanced glaucoma, 31 early glaucoma, 30 preperimetric glaucoma and 30 healthy controls). **Age**: advanced glaucoma ± SD, 63.0 ± 14.4 years; early glaucoma, 61.8 ± 11.5 years; preperimetric glaucoma, 56.9 ± 14.7 years; controls, 56.9 ± 17.3 years. **Sex**: 69 men (32 perimetric glaucoma, 17 preperimetric glaucoma, 20 controls) and 53 women (30 perimetric glaucoma, 13 preperimetric glaucoma, 10 controls).

Yamada 2014 *(Continued)*

	Ethnicity: not specified. **Clinical Setting:** Kyoto University Hospital, Kyoto. **Country**: Japan. **Ocular comorbidities**: patients with opaque media, diabetic retinopathy or another ophthalmic disease that could cause VF defects or fundus abnormalities, or a history of eye trauma or intraocular surgery, as well as patients with a history of systemic or neurologic disease that could affect the visual field, were excluded. **Spectrum of glaucoma severity**: the mean ± SD MD on the VF test were -0.7 ± 1.0 for preperimetric glaucoma, -2.1 ± 1.5 for early perimetric glaucoma, -15.7 ± 7.8, fro advanced perimetric glaucoma. Early glaucoma had MD > -6, advanced MD < -6. **Control participants:** IOP of ≤ 21 mmHg, a normal-appearing optic disc, and normal VF test results.
Index tests	**Optical coherence tomography**: Spectralis HRA+OCT system (Heidelberg Engineering, Heidelberg, Germany). Macular scanning protocol was used for the analysis. The authors reported no conflict of interest.
Target condition and reference standard(s)	**Manifest perimetric glaucoma:** glaucomatous optic disc appearance (defined as the presence of localised or diffuse neuroretinal rim thinning) and/or RNFL defects (classified as glaucomatous when its width at a 1-disc-diameter distance from the edge of the disc was larger than that of a major retinal vessel, it diverged from the edge of the optic disc in an arcuate or wedge shape) and typical reproducible VF defects (defined as the presence of GHT outside normal limits and a PSD with P < 5%; or a cluster of 3+ adjacent non-edge points in typical glaucomatous locations that did not cross the horizontal meridian, all of which were depressed on the PD plot with P < 5%, and 1 of which was depressed with P < 1%, on at least 2 consecutive examinations). **Manifest preperimetric glaucoma:** glaucomatous optic disc appearance and normal VF results. **Visual field test:** Humphrey Field Analyzer (Carl Zeiss Meditec); 24-2 SITA–standard strategy. Reliable VF were defined by fixation loss, and false-positive and false-negative rates of < 15%. **Optic disc evaluation:** Stereo disc photograph **(**3-Dx simultaneous stereo disc camera; Nidek, Gamagori, Japan). **RNFL evaluation**: Red-free fundus photograph (Heidelberg Retina Angiography 2 [HRA2]; Heidelberg Engineering, Heidelberg, Germany).
Flow and timing	38 eyes (> 10%) were excluded on the basis of ocular or systemic disease history or because OCT images were of poor quality. The time interval between index and reference test was not specified.
Comparative	
Notes	None.

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			

Yamada 2014 *(Continued)*

Was a consecutive or random sample of patients enrolled?	No		
Was a case-control design avoided?	No		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (All tests)			
If a threshold was used, was it prespecified?	Yes		
Were imaging test's quality assessed?	Yes		
Were any conflict of interest avoided	Yes		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			

Yamada 2014 *(Continued)*

Was there an appropriate interval between index test and reference standard?	Unclear
Did all patients receive the same reference standard?	Yes
Were all patients included in the analysis?	No
Did all patients receive a reference standard	Yes
Could the patient flow have introduced bias?	High risk

Yang 2014

Study characteristics	
Patient Sampling	Case-control study included participants in the Diagnostic Innovations in Glaucoma Study (a prospective longitudinal study designed to evaluate optic nerve structure and visual function in glaucoma). Healthy participants were recruited from the general population through advertisements or from the staff and employees at the University of California, San Diego. Both eyes were used for some participants.
Patient characteristics and setting	**Sample size:** 210 eyes of 148 participants (144 eyes from 106 glaucoma, and 66 eyes from 42 healthy controls). **Age**: glaucoma ± SD, 71.4 ± 10.2 years; controls, 60.1 ± 12.8 years. **Sex**: 71 men (56 glaucoma, 15 controls) and 77 women (50 glaucoma, 27 controls). **Ethnicity**: not specified. **Clinical Setting:** University of California, San Diego, CA. **Country**: USA. **Ocular comorbidities**: patients with ocular or systemic disease that could affect the optic nerve or visual field were excluded. **Spectrum of glaucoma severity**: the mean ± SD MD on the VF test was -5.9 ± 6.4 for glaucoma. **Control participants:** IOP < 22 mmHg with no history of increased IOP and normal VF result in both eyes.
Index tests	**Optical coherence tomography**: Swept-source Deep Range Imaging-OCT (DRI-OCT-1, Topcon). 2 Deep Range Imaging-OCT scan modes, a wide-angle scan and a 3-dimensional horizontal disc circle grid scan, were acquired. The quality of each scan and the accuracy of the segmentation algorithm were reviewed independently by masked reviewers. **Optical coherence tomography:** Spectralis SD-OCT (software v 5.3.0.7, Heidelberg Engineering, Heidelberg, Germany) RNFL circle scan was used for the analysis. Images with the signal strength < 15 dB, with artefacts, inverted or clipped and those that had co-existent retinal pathologic abnormalities, were excluded.

Yang 2014 *(Continued)*

	The authors declared conflict of interest with manufacturer.
Target condition and reference standard(s)	**Manifest perimetric glaucoma:** glaucomatous VF result (defined as a PSD with P < 5% or a GHT outside normal limit, or both) or documented evidence of progressive optic disc changes on masked grading of stereophotographs, with or without abnormal SAP results. **Visual field test:** Humphrey Field Analyzer II (Carl Zeiss Meditec, Dublin, CA, USA); 24-2 SITA–standard strategy. Reliable VF were defined by fixation losses or false-negative errors < 33%, and false-positive errors < 15%. All VFs were evaluated by the Visual Field Assessment Center. **Optic disc evaluation:** Stereo optic disc photograph (Kowa Nonmyd WX3D, v. VK27E, Kowa, Tokyo Japan). Progression was assessed by experienced graders who were masked to the participants' identities and to other test results at the Optic Disc Reading Center.
Flow and timing	44 eyes (> 10%) were excluded due to image-quality scores < 50 or clipped/poorly-focused images or images with segmentation failure and motion artefacts. The time interval between index and reference test was not specified.
Comparative	
Notes	None.

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		
Was a case-control design avoided?	No		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (All tests)			
If a threshold was used, was it prespecified?	Yes		
Were imaging test's quality assessed?	Yes		
Were any conflict of interest avoided	No		
Could the conduct or interpretation of the index test have introduced bias?		High risk	

Yang 2014 *(Continued)*

Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Unclear		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	No		
Did all patients receive a reference standard	Yes		
Could the patient flow have introduced bias?		High risk	

Yoshida 2014

Study characteristics	
Patient Sampling	Case-control study comprised patients with open-angle glaucoma who were enrolled between January 2009 and March 2010 and healthy controls. If both eyes fulfilled the inclusion criteria, the eye with a better data quality factor in the SD-OCT examination was included in the study.
Patient characteristics and setting	**Sample size:** 210 eyes of 210 participants (126 glaucoma, and 84 healthy controls). **Age**: glaucoma ± SD, 60.1 ± 13.1 years; controls, 52.6 ± 15.6 years. **Sex**: 100 men (53 glaucoma, 47 controls) and 110 women (73 glaucoma, 37 controls) **Ethnicity**: not specified.

Yoshida 2014 *(Continued)*

	Clinical Setting: University of Tokyo Hospital or the Tajimi Iwase eye clinic **Country**: Japan. **Ocular comorbidities**: patients with ocular diseases that could affect the results of SD-OCT examinations, such as diabetic retinopathy or age-related macular degeneration, were excluded. **Spectrum of glaucoma severity**: the mean ± SD MD on the VF test was -5.6 ± 5.2 for glaucoma. **Control participants:** no abnormal findings on biomicroscopy, gonioscopy and funduscopy, and normal VF test results according to Anderson-Patella's criteria.
Index tests	**Optical coherence tomography**: 3D OCT-1000 (Topcon Corp., Tokyo, Japan) for controls, 3D OCT-1000 (68 eyes) or 3D OCT-2000 (Topcon Corp., Tokyo, Japan) (58 eyes) for glaucoma. Raster scan protocol was used for analysis. Images influenced by involuntary blinking or saccade, and those with quality factor < 60% were excluded. The OCT models used in controls and glaucoma were different. The authors declared no conflict of interest.
Target condition and reference standard(s)	**Manifest perimetric glaucoma:** glaucomatous ONH appearance (as a rim notch with a rim width ≤ 0.1, a vertical cup-to-disc ratio of > 0.7 and/or a RNFL defect (with its edge at the ONH margin greater than a major retinal vessel) diverging in an arcuate or wedge shape) and glaucomatous VF defects (fulfilling at least one of Anderson- Patella's criteria: a cluster of ≥ 3 points non-edge in the pattern deviation plot in a single hemifield (superior/inferior) with P < 5%, one of which must have been P < 1%, a GHT outside of normal limits, or an abnormal PSD with P < 5%). **Visual field test:** Humphrey Field Analyzer (Carl Zeiss Meditec); 24-2 SITA–standard strategy for controls; 24-2 or 30-2, for glaucoma. Reliable VF were defined by fixation losses < 25%, and false-negative errors and false-positive errors < 15%. **Optic disc evaluation:** optic disc stereophotograph.
Flow and timing	The time interval between index and reference test was 3 months.
Comparative	
Notes	None.

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		
Was a case-control design avoided?	No		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		High risk	

Yoshida 2014 *(Continued)*

Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (All tests)			
If a threshold was used, was it pre-specified?	Yes		
Were imaging test's quality assessed?	Yes		
Were any conflict of interest avoided	Yes		
Could the conduct or interpretation of the index test have introduced bias?		High risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	No		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
Did all patients receive a reference standard	Yes		
Could the patient flow have introduced bias?		High risk	

Zelefsky 2006

Study characteristics	
Patient Sampling	Normals participants, glaucoma suspects, and glaucoma patients were enrolled. One eye per person was randomly selected.
Patient characteristics and setting	**Sample size**: 220 eyes of 220 participants (84 glaucoma, 136 healthy controls). **Age:** mean age was 51 ± 13 years for blacks (53M/71F) and 50 ± 16 years for whites (35M/61F). **Sex:** 88 men (53 blacks, 35 whites) and 132 women (71 blacks, 61 whites) **Ethnicity:** 96 whites (32 glaucoma, 64 controls) and 124 blacks (52 glaucoma, 72 controls). **Country**: not specified. **Setting**: not specified. **Ocular comorbidities**: no narrow angles, BCVA < 20/40, refractive spherical refraction < ±5 D/cylinder refraction > ±3 D, retinal disease, significant ocular surface disease, non-glaucomatous optic neuropathy, or history of intraocular surgery other than uncomplicated cataract surgery. **Spectrum of glaucoma severity:** mean ± SD MD on the VF test was -7.3 ± 6.7 dB (-8.45 ± 7.21 dB for blacks, -5.45 ± 5.18 dB for white). **Control participants:** normal visual fields (PSD > 5% and GHT within 97% normal limits) and a normal clinical examination.
Index tests	**Confocal scanning laser ophthalmoscopy:** HRT 2 (Heidelberg Engineering GmbH, Dossenheim, Germany). Data result were exported to the HRT3 software after the acquisition. Good image quality was defined as follows: acquisition sensitivity < 90%, topography SD < 40 mm, > ¾ of the disc within the target circle, minimal movement during the acquisition movie, no floaters over the disc. A trained technician, relying on stereophotographs of the respective optic disc, outlined the optic disc margin on the mean topographic image. No author had conflict of interest.
Target condition and reference standard(s)	**Manifest glaucoma:** glaucomatous VF defect on 2 consecutive fields (defined as PSD < 5% or GHT outside normal limits, or both). **Visual field testing:** Humphrey Field Analyzer, model II, 24-2 SITA-standard strategy (Carl Zeiss Meditec, Dublin, CA, USA). Reliability criteria included fixation losses rates, false-positive and false-negative rates < 33%. Optic disc appearance was not part of the reference standard.
Flow and timing	Reference standard and index tests were performed within 1 month. No patients were reported by the authors as excluded from the analysis.
Comparative	
Notes	None.

Methodological quality			
Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**

Zelefsky 2006 *(Continued)*

DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Was a case-control design avoided?	No		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (All tests)			
If a threshold was used, was it pre-specified?	Yes		
Were imaging test's quality assessed?	Yes		
Were any conflict of interest avoided	Yes		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		

Zelefsky 2006 *(Continued)*

Did all patients receive a reference standard	Yes
Could the patient flow have introduced bias?	Low risk

Zeppieri 2010

Study characteristics	
Patient Sampling	Healthy, ocular hypertensive and glaucoma patients were recruited consecutively. Normal participants were recruited from staff members and volunteers. One eye per person was randomly selected.
Patient characteristics and setting	**Sample size**: 319 eyes of 319 participants (75 perimetric glaucoma, 67 preperimetric glaucoma, 87 ocular hypertensive and 90 healthy controls). **Age:** glaucoma perimetric patients mean ± SD, 65.9 ± 11 years, glaucoma preperimetric patients mean ± SD, 63.9 ± 9.3 years, OHT patients mean ± SD, 63.6 ± 10.3 years, controls 53.4 ± 13.2. **Country:** Italy, USA, Argentina. **Ocular comorbidities**: no secondary causes of glaucoma, media opacity, SE > ±5 D, angle alterations, large peripapillary atrophy, diabetes, neurological disorders or previous intraocular surgery (excluding cataract surgery performed at least 6 months prior). **Setting:** S. Maria della Misericordia Hospital, Udine, Italy; Discoveries in Sight, Devers Eye Institute, Portland, Oregon; Centro Oftalmologico Sampaolesi y Fundacion Argentina Oftalmologica, Buenos Aires, Argentina. **Spectrum of glaucoma severity:** mean ± SD MD/PSD on the VF test were -2.1 ± 1.5/2.7 ± 0.9 dB, for perimetric glaucoma; -0.9 ± 1.3/1.7 ± 0.5 dB, for preperimetric glaucoma, -0.3 ± 1.4/1.5 ± 0.4 dB for OHT. **Control participants:** normal IOP, optic nerve/RNFL appearance and SAP results.
Index tests	**Scanning Laser Polarimetry:** GDx VCC, software version 5.1.0 (Carl Zeiss Meditec Inc.). All images had quality scores > 8, residual anterior segment retardation < 15 nm and typical scan score > 75. No authors had conflict of interest.
Target condition and reference standard(s)	**Manifest perimetric glaucoma**: IOP > 21 mmHg before medication and reproducible glaucomatous VF defects (defined by the Anderson and Patella criteria). **Manifest preperimetric glaucoma:** IOP > 21 mmHg before medication, glaucomatous optic disc/RNFL appearance (excavation or notching involving > 2 clock hours or focal/diffuse atrophy of neural rim area involving > 2 clock hours or disc haemorrhage or focal/generalised RNFL atrophy) and co-existing normal VF test result. **Ocular hypertensive:** IOP > 21 mmHg without medication, normal optic disc/RNFL appearance, and normal VF test result. **Visual field testing**: Humphrey Field Analyzer, model II 750, 30-2 SITA-standard strategy (Carl Zeiss Meditec, Dublin, CA, USA). Reliability criteria included fixation losses rates < 20%, false-positive < 15% and false-negative rates < 33%. **Optic disc evaluation:** slit-lamp indirect ophthalmoscopy and a 78-D lens. The eyes were classified on the basis of masked consensus by 2 expert graders. Adjudication by a third expert grader was completed in cases of disagreement.

Zeppieri 2010 *(Continued)*

Flow and timing	The time interval between reference standard and index tests was < 4 months. 9 participants were not included in the analysis due to poor-quality images or unreliable SAP test.
Comparative	
Notes	None.

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	No		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (All tests)			
If a threshold was used, was it prespecified?	Yes		
Were imaging test's quality assessed?	Yes		
Were any conflict of interest avoided	Yes		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes		

Zeppieri 2010 *(Continued)*

Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	No		
Did all patients receive the same reference standard?	No		
Were all patients included in the analysis?	Yes		
Did all patients receive a reference standard	Yes		
Could the patient flow have introduced bias?		High risk	

Zhang 2014

Study characteristics	
Patient Sampling	Case-control study of participants from the Diagnostic Innovations in Glaucoma Study at the University of California (San Diego) including manifest glaucoma, glaucoma suspects and healthy controls. For some participants, both eyes were enrolled.
Patient characteristics and setting	**Sample size**: 390 eyes of 224 participants (159 eyes of 93 glaucoma, 154 eyes of 89 glaucoma suspects, 77 of 42 normal controls). **Age**: glaucoma mean ± SD, 70.87 ± 12.19 years; glaucoma suspects 66.03 ± 12.48 years, controls group 50.68 ± 13.73 years. **Sex**: 107 men (45 glaucoma, 45 glaucoma suspects, 17 controls) and 117 women (48 glaucoma, 44 glaucoma suspects, 25 controls) **Ethnicity**: 145 European descent, 63 African-American, 16 other. **Clinical setting**: University of California, San Diego (UCSD). **Country**: USA. **Ocular comorbidities**: patients with any other ocular or systemic disease that could affect the optic nerve or the visual field, were excluded. **Spectrum of glaucoma severity**: the mean ± SD mean deviation and on the visual field test were -5.06 ± 5.43 for glaucoma; 0.47 ± 1.73 for glaucoma suspects. **Control participants:** IOP of ≤ 21 mmHg, with no history of increased IOP.

Zhang 2014 *(Continued)*

Index tests	**Optical coherence tomography**: Cirrus HD-OCT (Carl Zeiss Meditec Inc., Dublin, CA, USA), software version 6.5. Macular cube 200 x 200 protocol and optic disc cube 200 x 200 protocol) scans were used for analysis. Images with image quality factor < 7, movement artefacts, segmentation errors or not centred on the optic disc of fovea were excluded. Some of the authors had conflict of interest.
Target condition and reference standard(s)	**Manifest glaucoma:** glaucomatous VF results (defined as PSD with P < 5% or a GHT outside the normal limits) or progressive glaucomatous optic disc changes on stereo photo. **Glaucoma suspects:** optic disc appearance of glaucoma and normal VF results. **Visual field test:** Humphrey Field Analyzer (Carl Zeiss Meditec Inc.) 30-2 SITA standard programme. Reliable exams had fixation losses ≥ 33%, false-positive and false-negative response ≥ 15%. **Optic disc evaluation:** optic disc stereophotography.
Flow and timing	No details about exclusion. The time interval between Index and reference test was 6 months.
Comparative	
Notes	None.

Methodological quality

Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Was a case-control design avoided?	No		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (All tests)			
If a threshold was used, was it pre-specified?	Yes		
Were imaging test's quality assessed?	Yes		
Were any conflict of interest avoided	No		
Could the conduct or interpretation of the index test have introduced bias?		High risk	

Zhang 2014 *(Continued)*

Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	No		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Unclear		
Did all patients receive a reference standard	Yes		
Could the patient flow have introduced bias?		High risk	

Zheng 2008

Study characteristics	
Patient Sampling	Glaucoma patients and healthy controls were enrolled from June 2005 to June 2006. Both eyes per each participant were included in the study.
Patient characteristics and setting	**Sample size**: 300 eyes of 190 participant (220 glaucoma, 80 healthy controls). **Age: perimetric** glaucoma patients mean ± SD, 57.4 ± 9.33 years, controls 53.35 ± 11.38 years. **Ethnicity:** not specified. **Country:** China. **Ocular comorbidities**: No history of ocular disease, no history of diabetes. **Setting:** Beijing Hospital. **Spectrum of glaucoma severity:** mean ± SD MD on the VF test were 1.76 ± 1.71 dB for early glaucoma group, 12.38 ± 6.05 dB for advanced glaucoma group.

Zheng 2008 *(Continued)*

	Control participants: IOP < 21 mmHg, normal VF result, no abnormalities for routine eye checks the optimal vision correction ≥ 1.0 and diopter range ±6.00.
Index tests	**Scanning laser polarimetry:** GDx VCC, (Laser Diagnostic Technologies Inc, San Diego, CA, USA).
Target condition and reference standard(s)	**Primary open angle glaucoma:** IOP > 21 mmHg, characteristic VF defects for glaucoma (such as paracentral scotoma, nasal step, arcuate scotoma), specific changes under ophthalmo-fundoscope such as damages to the optic disc, and focal and diffusive RNFL defects, and wide angle of anterior chamber. **Primary angle closure glaucoma:** eye-anatomical changes for angle-closure glaucoma, history of acute increase of IOP or repetitive mild-to-moderate increase of IOP with or without symptoms, narrow angle of anterior chamber, closed angle when IOP increases, and changes of discus opticus and visual field. **Visual field testing:** Octopus 101 (Interzeag Inc., Switzerland) A type III light cursor was used, the persistence time was 100 ms, and the background lightness was 4 apostilbs. The programmes G2 or tG2 were used to measure 59 – 72 testing sites within the centre of 30°.
Flow and timing	The time interval between index and reference test was not reported. No patient were reported by the authors as excluded from the analysis.
Comparative	
Notes	None.

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Was a case-control design avoided?	No		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (All tests)			
If a threshold was used, was it pre-specified?	Yes		
Were imaging test's quality assessed?	Unclear		
Were any conflict of interest avoided	Unclear		
Could the conduct or interpretation of the index test have introduced bias?		Unclear risk	

Zheng 2008 *(Continued)*

Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Unclear		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
Did all patients receive a reference standard	Yes		
Could the patient flow have introduced bias?		Unclear risk	

BCVA: best corrected visual acuity
CPSD: corrected pattern standard deviation
IOP: intraocular pressure
GHT: glaucoma hemifield test
MD: mean deviation
NS: not specified
NTG: normal tension glaucoma
ONH: optic nerve head
POAG: primary open angle glaucoma
PACG: primary angle closure glaucoma
PSD: pattern standard deviation
RNFL: retinal nerve fibre layer
SAP: standard automated perimetry
VF: visual field

DATA

Presented below are all the data for all of the tests entered into the review.

Table Tests. Data tables by test

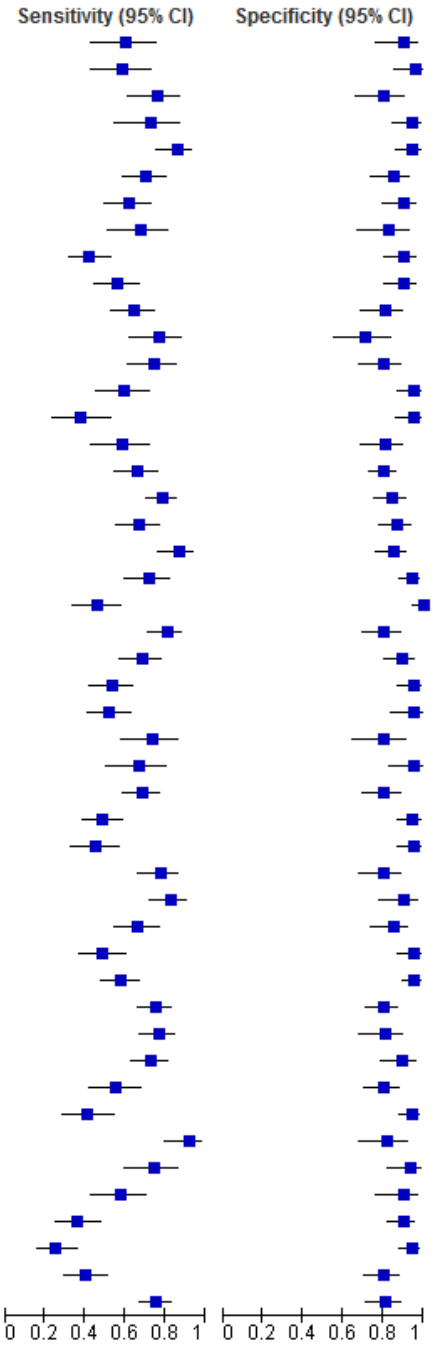
Test	No. of studies	No. of participants
1 GDx: Inferior average	30	6788
2 GDx: NFI	35	7193
3 GDx: Superior average	30	6788
4 GDx: TSNIT average	30	6535
5 OCT: RNFL average	57	13153
6 OCT: RNFL inferior quadrant	45	10599
7 OCT: RNFL nasal quadrant	30	6836
8 OCT: RNFL superior quadrant	43	10372
9 OCT: RNFL temporal quadrant	30	6836
10 HRT: Bathija function	1	214
11 HRT: Cup area	7	1882
12 HRT: C/D area ratio	9	2905
13 HRT: vertical C/D ratio	8	2622
14 HRT: Cup shape measure	6	1778
15 HRT: Cup volume	9	2905
16 HRT: FSM discriminant function o Mikelberg function	6	1650
17 HRT: MRA	8	1395
18 HRT: Rim area	9	2904
19 HRT: RB discriminant function	6	1642
20 HRT: Rim Volume	7	1882
21 OCT: GCC RTVue average thickness	19	5314
22 OCT: GCC RTVue superior thickness	16	4772
23 OCT: GCC RTVue inferior thickness	16	4772
24 OCT: GCC RTVue FLV	13	3899
25 OCT: GCC RTVue GLV	12	3695
26 OCT: GCC 3DTopcon average thickness	4	656
27 OCT: GCC 3DTopcon superior thickness	3	494

Test	No. of studies	No. of participants
28 OCT: GCC 3DTopcon inferior thickness	3	494
29 OCT: GCIPL Cirrus average thickness	11	2433
30 OCT: GCIPL Cirrus minimum thickness	9	1739
31 OCT: GCIPL Cirrus superior thickness	8	1571
32 OCT: GCIPL Cirrus inferior thickness	8	1571
33 OCT: ONH Disc area	7	1913
34 OCT: ONH Cup area	9	2562
35 OCT: ONH Rim area	17	4648
36 OCT: ONH Rim volume	6	1743
37 OCT: ONH Nerve head volume	4	1451
38 OCT: ONH Cup volume	9	3013
39 OCT: ONH C/D area ratio	17	4648
40 OCT: ONH horizontal C/D ratio	6	1971
41 OCT: ONH vertical C/D ratio	15	4085
42 OCT: GCIPL Cirrus Inferonasal quadrant	8	1571
43 OCT: GCIPL Cirrus Inferotemporal quadrant	8	1571
44 OCT: GCIPL Cirrus Superonasal quadrant	8	1571
45 OCT: GCIPL Cirrus Superotemporal quadrant	8	1571
46 OCT: GCC Spectralis average thickness	0	0
47 Direct comparison: GDx NFI	8	1090
48 Direct comparison: OCT RNFL average	8	1090

Test 1. GDx: Inferior average

GDx: Inferior average

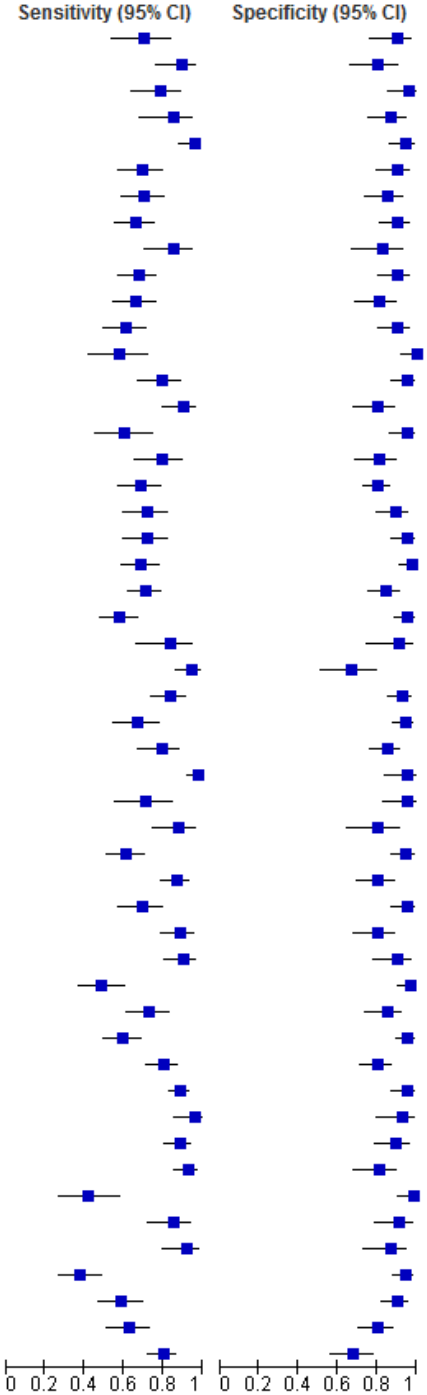
Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Aptel 2010	24	4	16	36	0.60 [0.43, 0.75]	0.90 [0.76, 0.97]
Badala 2007	27	2	19	44	0.59 [0.43, 0.73]	0.96 [0.85, 0.99]
Badala 2007	35	9	11	37	0.76 [0.61, 0.87]	0.80 [0.66, 0.91]
Benitez-del-Castillo 2011	24	3	9	52	0.73 [0.54, 0.87]	0.95 [0.85, 0.99]
Bertuzzi 2014	60	4	10	67	0.86 [0.75, 0.93]	0.94 [0.86, 0.98]
Borque 2008	50	9	21	51	0.70 [0.58, 0.81]	0.85 [0.73, 0.93]
Borque 2008	44	6	27	54	0.62 [0.50, 0.73]	0.90 [0.79, 0.96]
Brusini 2005	27	7	13	33	0.68 [0.51, 0.81]	0.82 [0.67, 0.93]
Brusini 2006a	40	6	55	56	0.42 [0.32, 0.53]	0.90 [0.80, 0.96]
Chen 2007	46	6	36	56	0.56 [0.45, 0.67]	0.90 [0.80, 0.96]
Chen 2007	53	12	29	50	0.65 [0.53, 0.75]	0.81 [0.69, 0.90]
Chen 2008	36	13	11	32	0.77 [0.62, 0.88]	0.71 [0.56, 0.84]
Da Pozzo 2005	44	13	15	52	0.75 [0.62, 0.85]	0.80 [0.68, 0.89]
Da Pozzo 2005	35	3	24	62	0.59 [0.46, 0.72]	0.95 [0.87, 0.99]
Da Pozzo 2006	18	3	30	59	0.38 [0.24, 0.53]	0.95 [0.87, 0.99]
Da Pozzo 2006	28	12	20	50	0.58 [0.43, 0.72]	0.81 [0.69, 0.90]
De Leon-Ortega 2006	52	30	27	119	0.66 [0.54, 0.76]	0.80 [0.73, 0.86]
Garudadri 2012	98	15	27	80	0.78 [0.70, 0.85]	0.84 [0.75, 0.91]
Huang 2010	53	11	26	75	0.67 [0.56, 0.77]	0.87 [0.78, 0.93]
Kanamori 2006	58	14	9	79	0.87 [0.76, 0.94]	0.85 [0.76, 0.92]
Kanamori 2006	48	5	19	88	0.72 [0.59, 0.82]	0.95 [0.88, 0.98]
Kook 2005	32	0	38	66	0.46 [0.34, 0.58]	1.00 [0.95, 1.00]
Lee 2010	71	15	17	62	0.81 [0.71, 0.88]	0.81 [0.70, 0.89]
Lee 2010	60	8	28	69	0.68 [0.57, 0.78]	0.90 [0.81, 0.95]
Lee 2010	47	4	41	73	0.53 [0.42, 0.64]	0.95 [0.87, 0.99]
Mai 2007	48	2	44	39	0.52 [0.42, 0.63]	0.95 [0.83, 0.99]
Medeiros 2004a	31	8	11	32	0.74 [0.58, 0.86]	0.80 [0.64, 0.91]
Medeiros 2004a	28	2	14	38	0.67 [0.50, 0.80]	0.95 [0.83, 0.99]
Medeiros 2004b	73	15	34	61	0.68 [0.59, 0.77]	0.80 [0.70, 0.89]
Medeiros 2004b	52	4	55	72	0.49 [0.39, 0.58]	0.95 [0.87, 0.99]
Medeiros 2005	32	3	39	62	0.45 [0.33, 0.57]	0.95 [0.87, 0.99]
Medeiros 2005	55	13	16	52	0.77 [0.66, 0.87]	0.80 [0.68, 0.89]
Oddone 2011	58	5	12	45	0.83 [0.72, 0.91]	0.90 [0.78, 0.97]
Pueyo 2006	49	10	25	56	0.66 [0.54, 0.77]	0.85 [0.74, 0.92]
Pueyo 2006	36	3	38	63	0.49 [0.37, 0.61]	0.95 [0.87, 0.99]
Rao 2014	61	5	45	104	0.58 [0.48, 0.67]	0.95 [0.90, 0.98]
Rao 2014	80	22	26	87	0.75 [0.66, 0.83]	0.80 [0.71, 0.87]
Schrems 2010	73	11	22	46	0.77 [0.67, 0.85]	0.81 [0.68, 0.90]
Schrems 2010	69	6	26	51	0.73 [0.63, 0.81]	0.89 [0.78, 0.96]
Sehi 2007	35	19	28	76	0.56 [0.42, 0.68]	0.80 [0.71, 0.88]
Sehi 2007	26	5	37	90	0.41 [0.29, 0.54]	0.95 [0.88, 0.98]
Takahashi 2008	43	8	4	37	0.91 [0.80, 0.98]	0.82 [0.68, 0.92]
Takahashi 2008	35	3	12	42	0.74 [0.60, 0.86]	0.93 [0.82, 0.99]
Weinreb 2003	31	4	23	36	0.57 [0.43, 0.71]	0.90 [0.76, 0.97]
Zeppieri 2010	29	9	51	81	0.36 [0.26, 0.48]	0.90 [0.82, 0.95]
Zeppieri 2010	20	5	60	86	0.25 [0.16, 0.36]	0.95 [0.88, 0.98]
Zeppieri 2010	32	18	48	72	0.40 [0.29, 0.52]	0.80 [0.70, 0.88]
Zheng 2008	98	15	32	65	0.75 [0.67, 0.83]	0.81 [0.71, 0.89]



Test 2. GDx: NFI

GDx: NFI

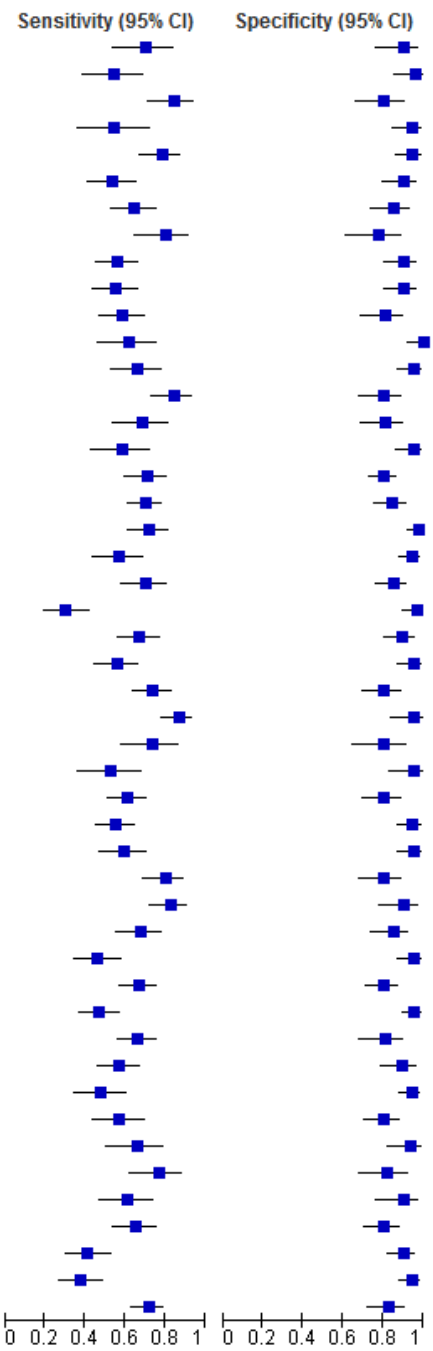
Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Aptel 2010	28	4	12	36	0.70 [0.53, 0.83]	0.90 [0.76, 0.97]
Badala 2007	41	9	5	37	0.89 [0.76, 0.96]	0.80 [0.66, 0.91]
Badala 2007	36	2	10	44	0.78 [0.64, 0.89]	0.96 [0.85, 0.99]
Benitez-del-Castillo 2011	28	7	5	48	0.85 [0.68, 0.95]	0.87 [0.76, 0.95]
Bertuzzi 2014	67	4	3	67	0.96 [0.88, 0.99]	0.94 [0.86, 0.98]
Borque 2008	49	6	22	54	0.69 [0.57, 0.79]	0.90 [0.79, 0.96]
Borque 2008	50	9	21	51	0.70 [0.58, 0.81]	0.85 [0.73, 0.93]
Bowd 2005	61	7	31	65	0.66 [0.56, 0.76]	0.90 [0.81, 0.96]
Brusini 2005	34	7	6	33	0.85 [0.70, 0.94]	0.82 [0.67, 0.93]
Brusini 2006a	64	6	31	56	0.67 [0.57, 0.77]	0.90 [0.80, 0.96]
Chen 2007	54	12	28	50	0.66 [0.55, 0.76]	0.81 [0.69, 0.90]
Chen 2007	50	6	32	56	0.61 [0.50, 0.72]	0.90 [0.80, 0.96]
Chen 2008	27	0	20	45	0.57 [0.42, 0.72]	1.00 [0.92, 1.00]
Da Pozzo 2005	47	3	12	62	0.80 [0.67, 0.89]	0.95 [0.87, 0.99]
Da Pozzo 2005	53	13	6	52	0.90 [0.79, 0.96]	0.80 [0.68, 0.89]
Da Pozzo 2006	29	3	19	59	0.60 [0.45, 0.74]	0.95 [0.87, 0.99]
Da Pozzo 2006	38	12	10	50	0.79 [0.65, 0.90]	0.81 [0.69, 0.90]
De Leon-Ortega 2006	54	30	25	119	0.68 [0.57, 0.78]	0.80 [0.73, 0.86]
Essock 2005	48	7	19	60	0.72 [0.59, 0.82]	0.90 [0.80, 0.96]
Essock 2005	48	3	19	64	0.72 [0.59, 0.82]	0.96 [0.87, 0.99]
Garas 2012	68	2	31	76	0.69 [0.59, 0.78]	0.97 [0.91, 1.00]
Garudadri 2012	89	15	36	80	0.71 [0.62, 0.79]	0.84 [0.75, 0.91]
Gonzales de la Rosa 2013	60	5	44	97	0.58 [0.48, 0.67]	0.95 [0.89, 0.98]
Hoesl 2013	26	3	5	29	0.84 [0.66, 0.95]	0.91 [0.75, 0.98]
Hong 2007	68	16	4	32	0.94 [0.86, 0.98]	0.67 [0.52, 0.80]
Huang 2010	66	6	13	80	0.84 [0.74, 0.91]	0.93 [0.85, 0.97]
Kanamori 2006	45	5	22	88	0.67 [0.55, 0.78]	0.95 [0.88, 0.98]
Kanamori 2006	53	14	14	79	0.79 [0.67, 0.88]	0.85 [0.76, 0.92]
Mai 2007	90	2	2	39	0.98 [0.92, 1.00]	0.95 [0.83, 0.99]
Medeiros 2004a	30	2	12	38	0.71 [0.55, 0.84]	0.95 [0.83, 0.99]
Medeiros 2004a	37	8	5	32	0.88 [0.74, 0.96]	0.80 [0.64, 0.91]
Medeiros 2004b	65	4	42	72	0.61 [0.51, 0.70]	0.95 [0.87, 0.99]
Medeiros 2004b	93	15	14	61	0.87 [0.79, 0.93]	0.80 [0.70, 0.89]
Medeiros 2005	49	3	22	62	0.69 [0.57, 0.79]	0.95 [0.87, 0.99]
Medeiros 2005	63	13	8	52	0.89 [0.79, 0.95]	0.80 [0.68, 0.89]
Oddone 2011	63	5	7	45	0.90 [0.80, 0.96]	0.90 [0.78, 0.97]
Pueyo 2006	36	3	38	83	0.49 [0.37, 0.61]	0.97 [0.90, 0.99]
Pueyo 2006	54	10	20	56	0.73 [0.61, 0.83]	0.85 [0.74, 0.92]
Rao 2014	63	5	43	104	0.59 [0.49, 0.69]	0.95 [0.90, 0.98]
Rao 2014	85	22	21	87	0.80 [0.71, 0.87]	0.80 [0.71, 0.87]
Reus 2004	144	4	18	73	0.89 [0.83, 0.93]	0.95 [0.87, 0.99]
Reus 2007	46	3	2	37	0.96 [0.86, 0.99]	0.93 [0.80, 0.98]
Schrems 2010	84	6	11	51	0.88 [0.80, 0.94]	0.89 [0.78, 0.96]
Schrems 2010	88	11	7	46	0.93 [0.85, 0.97]	0.81 [0.68, 0.90]
Shah 2006	18	1	25	57	0.42 [0.27, 0.58]	0.98 [0.91, 1.00]
Takahashi 2008	40	4	7	41	0.85 [0.72, 0.94]	0.91 [0.79, 0.98]
Takahashi 2008	43	6	4	39	0.91 [0.80, 0.98]	0.87 [0.73, 0.95]
Zeppieri 2010	30	5	50	86	0.38 [0.27, 0.49]	0.95 [0.88, 0.98]
Zeppieri 2010	47	9	33	81	0.59 [0.47, 0.70]	0.90 [0.82, 0.95]
Zeppieri 2010	50	18	30	72	0.63 [0.51, 0.73]	0.80 [0.70, 0.88]
Zheng 2008	104	26	26	54	0.80 [0.72, 0.86]	0.68 [0.56, 0.78]



Test 3. GDx: Superior average

GDx: Superior average

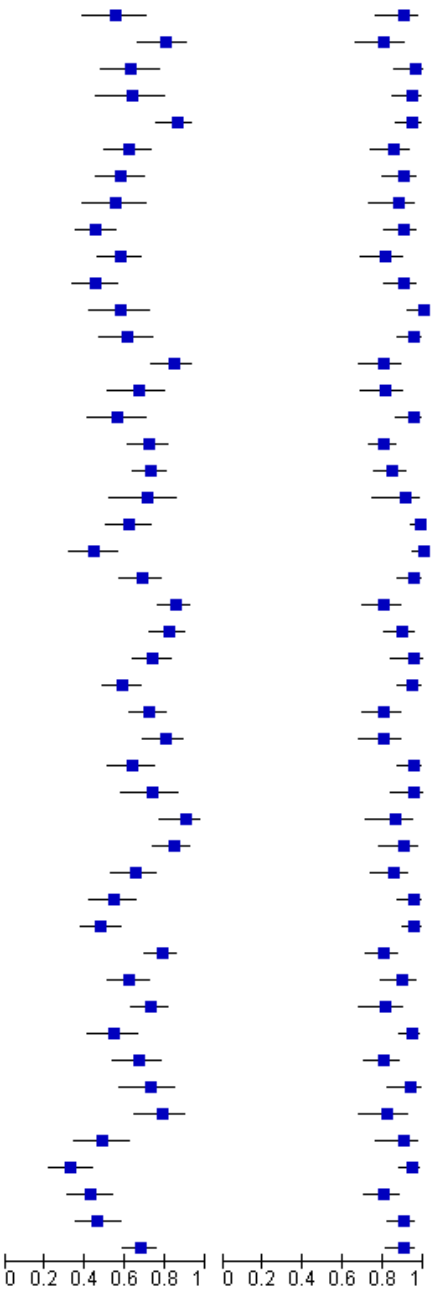
Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Aptel 2010	28	4	12	36	0.70 [0.53, 0.83]	0.90 [0.76, 0.97]
Badala 2007	25	2	21	44	0.54 [0.39, 0.69]	0.96 [0.85, 0.99]
Badala 2007	39	9	7	37	0.85 [0.71, 0.94]	0.80 [0.66, 0.91]
Benitez-del-Castillo 2011	18	3	15	52	0.55 [0.36, 0.72]	0.95 [0.85, 0.99]
Bertuzzi 2014	55	4	15	67	0.79 [0.67, 0.87]	0.94 [0.86, 0.98]
Borque 2008	38	6	33	54	0.54 [0.41, 0.65]	0.90 [0.79, 0.96]
Borque 2008	46	9	25	51	0.65 [0.53, 0.76]	0.85 [0.73, 0.93]
Brusini 2005	32	9	8	31	0.80 [0.64, 0.91]	0.78 [0.62, 0.89]
Brusini 2006a	53	6	42	56	0.56 [0.45, 0.66]	0.90 [0.80, 0.96]
Chen 2007	45	6	37	56	0.55 [0.43, 0.66]	0.90 [0.80, 0.96]
Chen 2007	48	12	34	50	0.59 [0.47, 0.69]	0.81 [0.69, 0.90]
Chen 2008	29	0	18	45	0.62 [0.46, 0.75]	1.00 [0.92, 1.00]
Da Pozzo 2005	39	3	20	62	0.66 [0.53, 0.78]	0.95 [0.87, 0.99]
Da Pozzo 2005	50	13	9	52	0.85 [0.73, 0.93]	0.80 [0.68, 0.89]
Da Pozzo 2006	33	12	15	50	0.69 [0.54, 0.81]	0.81 [0.69, 0.90]
Da Pozzo 2006	28	3	20	59	0.58 [0.43, 0.72]	0.95 [0.87, 0.99]
De Leon-Ortega 2006	56	30	23	119	0.71 [0.60, 0.81]	0.80 [0.73, 0.86]
Garudadri 2012	88	15	37	80	0.70 [0.62, 0.78]	0.84 [0.75, 0.91]
Huang 2010	57	2	22	84	0.72 [0.61, 0.82]	0.98 [0.92, 1.00]
Kanamori 2006	38	5	29	88	0.57 [0.44, 0.69]	0.95 [0.88, 0.98]
Kanamori 2006	47	14	20	79	0.70 [0.58, 0.81]	0.85 [0.76, 0.92]
Kook 2005	21	2	49	64	0.30 [0.20, 0.42]	0.97 [0.89, 1.00]
Lee 2010	59	8	29	69	0.67 [0.56, 0.77]	0.90 [0.81, 0.95]
Lee 2010	49	4	39	73	0.56 [0.45, 0.66]	0.95 [0.87, 0.99]
Lee 2010	65	15	23	62	0.74 [0.63, 0.83]	0.81 [0.70, 0.89]
Mai 2007	80	2	12	39	0.87 [0.78, 0.93]	0.95 [0.83, 0.99]
Medeiros 2004a	31	8	11	32	0.74 [0.58, 0.86]	0.80 [0.64, 0.91]
Medeiros 2004a	22	2	20	38	0.52 [0.36, 0.68]	0.95 [0.83, 0.99]
Medeiros 2004b	65	15	42	61	0.61 [0.51, 0.70]	0.80 [0.70, 0.89]
Medeiros 2004b	59	4	48	72	0.55 [0.45, 0.65]	0.95 [0.87, 0.99]
Medeiros 2005	42	3	29	62	0.59 [0.47, 0.71]	0.95 [0.87, 0.99]
Medeiros 2005	57	13	14	52	0.80 [0.69, 0.89]	0.80 [0.68, 0.89]
Oddone 2011	58	5	12	45	0.83 [0.72, 0.91]	0.90 [0.78, 0.97]
Pueyo 2006	50	10	24	56	0.68 [0.56, 0.78]	0.85 [0.74, 0.92]
Pueyo 2006	34	3	40	63	0.46 [0.34, 0.58]	0.95 [0.87, 0.99]
Rao 2014	71	22	35	87	0.67 [0.57, 0.76]	0.80 [0.71, 0.87]
Rao 2014	50	5	56	104	0.47 [0.37, 0.57]	0.95 [0.90, 0.98]
Schrems 2010	63	11	32	46	0.66 [0.56, 0.76]	0.81 [0.68, 0.90]
Schrems 2010	54	6	41	51	0.57 [0.46, 0.67]	0.89 [0.78, 0.96]
Sehi 2007	30	5	33	90	0.48 [0.35, 0.61]	0.95 [0.88, 0.98]
Sehi 2007	36	19	27	76	0.57 [0.44, 0.70]	0.80 [0.71, 0.88]
Takahashi 2008	31	3	16	42	0.66 [0.51, 0.79]	0.93 [0.82, 0.99]
Takahashi 2008	36	8	11	37	0.77 [0.62, 0.88]	0.82 [0.68, 0.92]
Weinreb 2003	33	4	21	36	0.61 [0.47, 0.74]	0.90 [0.76, 0.97]
Zeppieri 2010	52	18	28	72	0.65 [0.54, 0.75]	0.80 [0.70, 0.88]
Zeppieri 2010	33	9	47	81	0.41 [0.30, 0.53]	0.90 [0.82, 0.95]
Zeppieri 2010	30	5	50	86	0.38 [0.27, 0.49]	0.95 [0.88, 0.98]
Zheng 2008	93	14	37	66	0.72 [0.63, 0.79]	0.82 [0.72, 0.90]



Test 4. GDx: TSNIT average

GDx: TSNIT average

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Aptel 2010	22	4	18	36	0.55 [0.38, 0.71]	0.90 [0.76, 0.97]
Badala 2007	37	9	9	37	0.80 [0.66, 0.91]	0.80 [0.66, 0.91]
Badala 2007	29	2	17	44	0.63 [0.48, 0.77]	0.96 [0.85, 0.99]
Benitez-del-Castillo 2011	21	3	12	52	0.64 [0.45, 0.80]	0.95 [0.85, 0.99]
Bertuzzi 2014	60	4	10	67	0.86 [0.75, 0.93]	0.94 [0.86, 0.98]
Borque 2008	44	9	27	51	0.62 [0.50, 0.73]	0.85 [0.73, 0.93]
Borque 2008	41	6	30	54	0.58 [0.45, 0.69]	0.90 [0.79, 0.96]
Brusini 2005	22	5	18	35	0.55 [0.38, 0.71]	0.88 [0.73, 0.96]
Brusini 2006a	43	6	52	56	0.45 [0.35, 0.56]	0.90 [0.80, 0.96]
Chen 2007	47	12	35	50	0.57 [0.46, 0.68]	0.81 [0.69, 0.90]
Chen 2007	37	6	45	56	0.45 [0.34, 0.57]	0.90 [0.80, 0.96]
Chen 2008	27	0	20	45	0.57 [0.42, 0.72]	1.00 [0.92, 1.00]
Da Pozzo 2005	36	3	23	62	0.61 [0.47, 0.73]	0.95 [0.87, 0.99]
Da Pozzo 2005	50	13	9	52	0.85 [0.73, 0.93]	0.80 [0.68, 0.89]
Da Pozzo 2006	32	12	16	50	0.67 [0.52, 0.80]	0.81 [0.69, 0.90]
Da Pozzo 2006	27	3	21	59	0.56 [0.41, 0.71]	0.95 [0.87, 0.99]
De Leon-Ortega 2006	57	30	22	119	0.72 [0.61, 0.82]	0.80 [0.73, 0.86]
Garudadri 2012	91	15	34	80	0.73 [0.64, 0.80]	0.84 [0.75, 0.91]
Hoesl 2013	22	3	9	29	0.71 [0.52, 0.86]	0.91 [0.75, 0.98]
Huang 2010	49	1	30	85	0.62 [0.50, 0.73]	0.99 [0.94, 1.00]
Kook 2005	31	0	39	66	0.44 [0.32, 0.57]	1.00 [0.95, 1.00]
Lee 2010	60	4	28	73	0.68 [0.57, 0.78]	0.95 [0.87, 0.99]
Lee 2010	75	15	13	62	0.85 [0.76, 0.92]	0.81 [0.70, 0.89]
Lee 2010	72	8	16	69	0.82 [0.72, 0.89]	0.90 [0.81, 0.95]
Mai 2007	68	2	24	39	0.74 [0.64, 0.83]	0.95 [0.83, 0.99]
Medeiros 2004b	63	4	44	72	0.59 [0.49, 0.68]	0.95 [0.87, 0.99]
Medeiros 2004b	77	15	30	61	0.72 [0.62, 0.80]	0.80 [0.70, 0.89]
Medeiros 2005	57	13	14	52	0.80 [0.69, 0.89]	0.80 [0.68, 0.89]
Medeiros 2005	45	3	26	62	0.63 [0.51, 0.75]	0.95 [0.87, 0.99]
Na 2013b	31	2	11	40	0.74 [0.58, 0.86]	0.95 [0.84, 0.99]
Na 2013b	38	6	4	36	0.90 [0.77, 0.97]	0.86 [0.71, 0.95]
Oddone 2011	59	5	11	45	0.84 [0.74, 0.92]	0.90 [0.78, 0.97]
Pueyo 2006	48	10	26	56	0.65 [0.53, 0.76]	0.85 [0.74, 0.92]
Pueyo 2006	40	3	34	63	0.54 [0.42, 0.66]	0.95 [0.87, 0.99]
Rao 2014	51	5	55	104	0.48 [0.38, 0.58]	0.95 [0.90, 0.98]
Rao 2014	83	22	23	87	0.78 [0.69, 0.86]	0.80 [0.71, 0.87]
Schrems 2010	59	6	36	51	0.62 [0.52, 0.72]	0.89 [0.78, 0.96]
Schrems 2010	69	11	26	46	0.73 [0.63, 0.81]	0.81 [0.68, 0.90]
Sehi 2007	34	5	29	90	0.54 [0.41, 0.67]	0.95 [0.88, 0.98]
Sehi 2007	42	19	21	76	0.67 [0.54, 0.78]	0.80 [0.71, 0.88]
Takahashi 2008	34	3	13	42	0.72 [0.57, 0.84]	0.93 [0.82, 0.99]
Takahashi 2008	37	8	10	37	0.79 [0.64, 0.89]	0.82 [0.68, 0.92]
Weinreb 2003	26	4	28	36	0.48 [0.34, 0.62]	0.90 [0.76, 0.97]
Zeppieri 2010	26	5	54	86	0.33 [0.22, 0.44]	0.95 [0.88, 0.98]
Zeppieri 2010	34	18	46	72	0.42 [0.32, 0.54]	0.80 [0.70, 0.88]
Zeppieri 2010	37	9	43	81	0.46 [0.35, 0.58]	0.90 [0.82, 0.95]
Zheng 2008	88	8	42	72	0.68 [0.59, 0.76]	0.90 [0.81, 0.96]



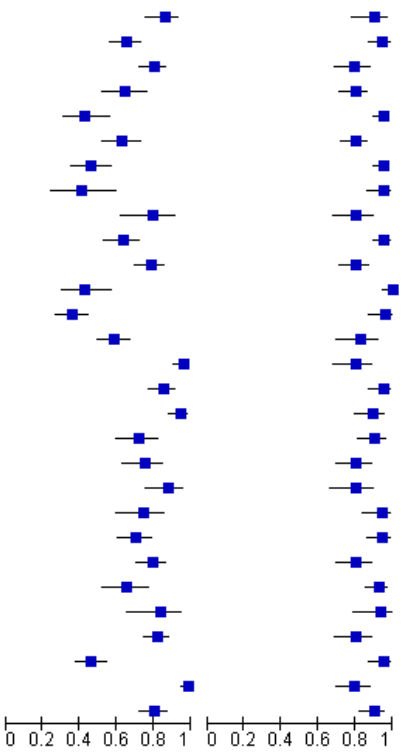
Test 5. OCT: RNFL average

OCT: RNFL average

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Akashi 2013	57	4	18	83	0.76 [0.65, 0.85]	0.95 [0.89, 0.99]
Aptel 2010	38	4	2	36	0.95 [0.83, 0.99]	0.90 [0.76, 0.97]
Arintawati 2013	29	3	52	65	0.36 [0.25, 0.47]	0.96 [0.88, 0.99]
Arintawati 2013	46	14	35	54	0.57 [0.45, 0.68]	0.79 [0.68, 0.88]
Barella 2013	16	5	41	41	0.28 [0.17, 0.42]	0.89 [0.76, 0.96]
Barella 2013	35	9	22	37	0.61 [0.48, 0.74]	0.80 [0.66, 0.91]
Begum 2014a	46	3	16	50	0.74 [0.62, 0.84]	0.94 [0.84, 0.99]
Begum 2014a	55	11	7	42	0.89 [0.78, 0.95]	0.79 [0.66, 0.89]
Benitez-del-Castillo 2011	24	1	9	54	0.73 [0.54, 0.87]	0.98 [0.90, 1.00]
Bertuzzi 2014	63	4	7	67	0.90 [0.80, 0.96]	0.94 [0.86, 0.98]
Chen 2013	20	0	15	52	0.57 [0.39, 0.74]	1.00 [0.93, 1.00]
Cho 2011	28	0	21	43	0.57 [0.42, 0.71]	1.00 [0.92, 1.00]
Choi 2013	37	1	17	55	0.69 [0.54, 0.80]	0.98 [0.90, 1.00]
Fang 2010	31	2	11	32	0.74 [0.58, 0.86]	0.94 [0.80, 0.99]
Fang 2010	33	5	9	29	0.79 [0.63, 0.90]	0.85 [0.69, 0.95]
Garas 2011	93	0	18	93	0.84 [0.76, 0.90]	1.00 [0.96, 1.00]
Garas 2012	65	1	34	77	0.66 [0.55, 0.75]	0.99 [0.93, 1.00]
Gonzales de la Rosa 2013	71	5	33	97	0.68 [0.58, 0.77]	0.95 [0.89, 0.98]
Hoesl 2013	30	3	1	29	0.97 [0.83, 1.00]	0.91 [0.75, 0.98]
Huang 2011	119	9	27	65	0.82 [0.74, 0.87]	0.88 [0.78, 0.94]
Hwang 2012	77	8	3	72	0.96 [0.89, 0.99]	0.90 [0.81, 0.96]
Jeoung 2010	24	11	31	44	0.44 [0.30, 0.58]	0.80 [0.67, 0.90]
Jeoung 2010	8	3	47	52	0.15 [0.06, 0.27]	0.95 [0.85, 0.99]
Jeoung 2013	82	4	82	115	0.50 [0.42, 0.58]	0.97 [0.92, 0.99]
Kang 2012	32	5	22	49	0.59 [0.45, 0.72]	0.91 [0.80, 0.97]
Kang 2012	40	11	14	43	0.74 [0.60, 0.85]	0.80 [0.66, 0.89]
Kang 2012	26	3	28	51	0.48 [0.34, 0.62]	0.94 [0.85, 0.99]
Kim 2011	40	5	16	44	0.71 [0.58, 0.83]	0.90 [0.78, 0.97]
Kim 2013a	41	6	11	52	0.79 [0.65, 0.89]	0.90 [0.79, 0.96]
Kim 2013b	30	2	18	40	0.63 [0.47, 0.76]	0.95 [0.84, 0.99]
Kim 2014a	43	15	49	77	0.47 [0.36, 0.57]	0.84 [0.75, 0.91]
Kim 2014b	42	7	26	57	0.62 [0.49, 0.73]	0.89 [0.79, 0.95]
Kita 2013	74	3	25	32	0.75 [0.65, 0.83]	0.91 [0.77, 0.98]
Koh 2014	46	3	14	48	0.77 [0.64, 0.87]	0.94 [0.84, 0.99]
Koh 2014	58	10	2	40	0.97 [0.88, 1.00]	0.80 [0.66, 0.90]
Kotowski 2012	52	10	11	41	0.83 [0.71, 0.91]	0.80 [0.67, 0.90]
Kotowski 2012	45	3	18	48	0.71 [0.59, 0.82]	0.94 [0.84, 0.99]
Lee 2010	75	8	13	69	0.85 [0.76, 0.92]	0.90 [0.81, 0.95]
Lee 2010	66	4	22	73	0.75 [0.65, 0.84]	0.95 [0.87, 0.99]
Lee 2010	77	15	11	62	0.88 [0.79, 0.94]	0.81 [0.70, 0.89]
Leite 2011	83	5	43	102	0.66 [0.57, 0.74]	0.95 [0.89, 0.98]
Leite 2011	101	21	25	86	0.80 [0.72, 0.87]	0.80 [0.72, 0.87]
Leung 2010	105	10	16	92	0.87 [0.79, 0.92]	0.90 [0.83, 0.95]
Lisboa 2013	38	19	10	75	0.79 [0.65, 0.90]	0.80 [0.70, 0.87]
Lisboa 2013	34	5	14	89	0.71 [0.56, 0.83]	0.95 [0.88, 0.98]
Mansoori 2011	71	10	12	86	0.86 [0.76, 0.92]	0.90 [0.82, 0.95]
Mansoori 2011	79	19	4	76	0.95 [0.88, 0.99]	0.80 [0.71, 0.88]
Moreno 2011	29	3	27	64	0.52 [0.38, 0.65]	0.96 [0.87, 0.99]
Moreno 2011	35	13	21	54	0.63 [0.49, 0.75]	0.81 [0.69, 0.89]
Moreno-Montañés 2010	32	5	29	100	0.52 [0.39, 0.65]	0.95 [0.89, 0.98]
Moreno-Montañés 2010	45	16	16	89	0.74 [0.61, 0.84]	0.85 [0.76, 0.91]
Mwanza 2012	47	7	11	92	0.81 [0.69, 0.90]	0.93 [0.86, 0.97]
Mwanza 2014	32	2	18	47	0.64 [0.49, 0.77]	0.96 [0.86, 1.00]
Na 2013a	69	3	36	65	0.66 [0.56, 0.75]	0.96 [0.88, 0.99]
Na 2013a	84	11	21	57	0.80 [0.71, 0.87]	0.84 [0.73, 0.92]
Na 2013a	78	5	27	63	0.74 [0.65, 0.82]	0.93 [0.84, 0.98]
Na 2013b	31	2	11	40	0.74 [0.58, 0.86]	0.95 [0.84, 0.99]
Na 2013b	38	6	4	36	0.90 [0.77, 0.97]	0.86 [0.71, 0.95]
Nakatani 2011	19	6	13	26	0.59 [0.41, 0.76]	0.81 [0.64, 0.93]
Nakatani 2011	16	3	16	29	0.50 [0.32, 0.68]	0.91 [0.75, 0.98]
Nouri-Mahdavi 2013	52	8	7	83	0.88 [0.77, 0.95]	0.91 [0.83, 0.96]
Oddone 2011	60	5	10	45	0.86 [0.75, 0.93]	0.90 [0.78, 0.97]
Rao 2010b	91	4	49	70	0.65 [0.56, 0.73]	0.95 [0.87, 0.99]

Test 5. (Continued)

Oddone 2011	60	5	10	45	0.86 [0.75, 0.93]	0.90 [0.78, 0.97]
Rao 2010b	91	4	49	70	0.65 [0.56, 0.73]	0.95 [0.87, 0.99]
Rao 2010b	112	15	28	59	0.80 [0.72, 0.86]	0.80 [0.69, 0.88]
Rao 2012a	42	24	23	95	0.65 [0.52, 0.76]	0.80 [0.71, 0.87]
Rao 2012a	28	6	37	113	0.43 [0.31, 0.56]	0.95 [0.89, 0.98]
Rao 2012b	57	25	34	100	0.63 [0.52, 0.73]	0.80 [0.72, 0.87]
Rao 2012b	42	6	49	119	0.46 [0.36, 0.57]	0.95 [0.90, 0.98]
Rao 2013	14	3	20	57	0.41 [0.25, 0.59]	0.95 [0.86, 0.99]
Rao 2013	27	12	7	48	0.79 [0.62, 0.91]	0.80 [0.68, 0.89]
Rao 2014	67	5	39	104	0.63 [0.53, 0.72]	0.95 [0.90, 0.98]
Rao 2014	83	22	23	87	0.78 [0.69, 0.86]	0.80 [0.71, 0.87]
Rho 2014	25	0	33	62	0.43 [0.30, 0.57]	1.00 [0.94, 1.00]
Rolle 2011	45	2	81	50	0.36 [0.27, 0.45]	0.96 [0.87, 1.00]
Rolle 2011	74	9	52	43	0.59 [0.50, 0.67]	0.83 [0.70, 0.92]
Seong 2010	98	13	4	52	0.96 [0.90, 0.99]	0.80 [0.68, 0.89]
Seong 2010	87	3	15	62	0.85 [0.77, 0.92]	0.95 [0.87, 0.99]
Seong 2010	96	7	6	59	0.94 [0.88, 0.98]	0.89 [0.79, 0.96]
Shin 2013	46	7	18	65	0.72 [0.59, 0.82]	0.90 [0.81, 0.96]
Shin 2013	48	14	16	58	0.75 [0.63, 0.85]	0.81 [0.70, 0.89]
Sullivan-Mee 2013	44	10	6	40	0.88 [0.76, 0.95]	0.80 [0.66, 0.90]
Sullivan-Mee 2013	37	3	13	48	0.74 [0.60, 0.85]	0.94 [0.84, 0.99]
Sung 2013	75	4	32	68	0.70 [0.60, 0.79]	0.94 [0.86, 0.98]
Sung 2013	85	14	22	58	0.79 [0.71, 0.87]	0.81 [0.70, 0.89]
Wu 2012	40	6	21	79	0.66 [0.52, 0.77]	0.93 [0.85, 0.97]
Yamada 2014	25	2	5	29	0.83 [0.65, 0.94]	0.94 [0.79, 0.99]
Yang 2014	118	13	26	53	0.82 [0.75, 0.88]	0.80 [0.69, 0.89]
Yang 2014	66	3	78	63	0.46 [0.38, 0.54]	0.95 [0.87, 0.99]
Yoshida 2014	124	17	2	67	0.98 [0.94, 1.00]	0.80 [0.70, 0.88]
Yoshida 2014	101	8	25	76	0.80 [0.72, 0.87]	0.90 [0.82, 0.96]



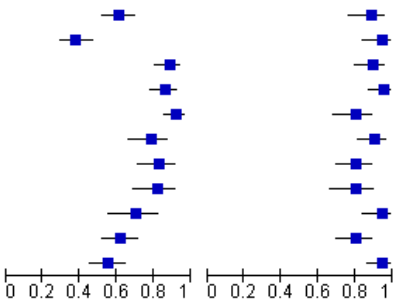
Test 6. OCT: RNFL inferior quadrant

OCT: RNFL inferior quadrant

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Aptel 2010	38	4	2	36	0.95 [0.83, 0.99]	0.90 [0.76, 0.97]
Arintawati 2013	26	3	55	65	0.32 [0.22, 0.43]	0.96 [0.88, 0.99]
Arintawati 2013	38	14	43	54	0.47 [0.36, 0.58]	0.79 [0.68, 0.88]
Barella 2013	36	9	21	37	0.63 [0.49, 0.76]	0.80 [0.66, 0.91]
Barella 2013	34	5	23	41	0.60 [0.46, 0.72]	0.89 [0.76, 0.96]
Begum 2014a	50	3	12	50	0.81 [0.69, 0.90]	0.94 [0.84, 0.99]
Begum 2014a	56	11	6	42	0.90 [0.80, 0.96]	0.79 [0.66, 0.89]
Benitez-del-Castillo 2011	27	6	6	49	0.82 [0.65, 0.93]	0.89 [0.78, 0.96]
Bertuzzi 2014	63	4	7	67	0.90 [0.80, 0.96]	0.94 [0.86, 0.98]
Cho 2011	31	0	18	43	0.63 [0.48, 0.77]	1.00 [0.92, 1.00]
Choi 2013	40	3	14	53	0.74 [0.60, 0.85]	0.95 [0.85, 0.99]
Fang 2010	27	2	15	32	0.64 [0.48, 0.78]	0.94 [0.80, 0.99]
Fang 2010	31	5	11	29	0.74 [0.58, 0.86]	0.85 [0.69, 0.95]
Garas 2011	93	2	18	91	0.84 [0.76, 0.90]	0.98 [0.92, 1.00]
Garas 2012	63	2	36	76	0.64 [0.53, 0.73]	0.97 [0.91, 1.00]
Hoesl 2013	30	3	1	29	0.97 [0.83, 1.00]	0.91 [0.75, 0.98]
Huang 2011	131	19	15	55	0.90 [0.84, 0.94]	0.74 [0.63, 0.84]
Hwang 2012	58	8	22	72	0.72 [0.61, 0.82]	0.90 [0.81, 0.96]
Jeoung 2010	20	11	35	44	0.36 [0.24, 0.50]	0.80 [0.67, 0.90]
Jeoung 2010	7	3	48	52	0.13 [0.05, 0.24]	0.95 [0.85, 0.99]
Jeoung 2013	101	6	63	113	0.62 [0.54, 0.69]	0.95 [0.89, 0.98]
Kim 2011	43	6	13	43	0.77 [0.64, 0.87]	0.88 [0.75, 0.95]
Kim 2013a	41	13	11	45	0.79 [0.65, 0.89]	0.78 [0.65, 0.87]
Kim 2013b	27	2	21	40	0.56 [0.41, 0.71]	0.95 [0.84, 0.99]
Kim 2014a	80	10	12	82	0.87 [0.78, 0.93]	0.89 [0.81, 0.95]
Kim 2014b	55	21	13	43	0.81 [0.70, 0.89]	0.67 [0.54, 0.78]
Koh 2014	53	3	7	48	0.88 [0.77, 0.95]	0.94 [0.84, 0.99]
Koh 2014	56	10	4	40	0.93 [0.84, 0.98]	0.80 [0.66, 0.90]
Lee 2010	83	15	5	62	0.94 [0.87, 0.98]	0.81 [0.70, 0.89]
Lee 2010	77	8	11	69	0.88 [0.79, 0.94]	0.90 [0.81, 0.95]
Lee 2010	73	4	15	73	0.83 [0.73, 0.90]	0.95 [0.87, 0.99]
Leite 2011	83	5	43	102	0.66 [0.57, 0.74]	0.95 [0.89, 0.98]
Leite 2011	100	21	26	86	0.79 [0.71, 0.86]	0.80 [0.72, 0.87]
Leung 2010	105	10	16	92	0.87 [0.79, 0.92]	0.90 [0.83, 0.95]
Lisboa 2013	24	5	24	89	0.50 [0.35, 0.65]	0.95 [0.88, 0.98]
Lisboa 2013	36	19	12	75	0.75 [0.60, 0.86]	0.80 [0.70, 0.87]
Mansoori 2011	59	19	24	76	0.71 [0.60, 0.81]	0.80 [0.71, 0.88]
Mansoori 2011	24	10	59	86	0.29 [0.19, 0.40]	0.90 [0.82, 0.95]
Moreno-Montañés 2010	37	5	24	100	0.61 [0.47, 0.73]	0.95 [0.89, 0.98]
Moreno-Montañés 2010	42	16	19	89	0.69 [0.56, 0.80]	0.85 [0.76, 0.91]
Mwanza 2012	54	1	4	98	0.93 [0.83, 0.98]	0.99 [0.95, 1.00]
Mwanza 2013	26	5	9	44	0.74 [0.57, 0.88]	0.90 [0.78, 0.97]
Mwanza 2013	23	2	12	47	0.66 [0.48, 0.81]	0.96 [0.86, 1.00]
Mwanza 2014	37	2	13	47	0.74 [0.60, 0.85]	0.96 [0.86, 1.00]
Na 2013a	65	5	40	63	0.62 [0.52, 0.71]	0.93 [0.84, 0.98]
Na 2013a	76	12	29	56	0.72 [0.63, 0.81]	0.82 [0.71, 0.91]
Na 2013a	58	3	47	65	0.55 [0.45, 0.65]	0.96 [0.88, 0.99]
Nakatani 2011	23	6	9	26	0.72 [0.53, 0.86]	0.81 [0.64, 0.93]
Nakatani 2011	17	3	15	29	0.53 [0.35, 0.71]	0.91 [0.75, 0.98]
Nouri-Mahdavi 2013	55	8	4	83	0.93 [0.84, 0.98]	0.91 [0.83, 0.96]
Oddone 2011	57	5	13	45	0.81 [0.70, 0.90]	0.90 [0.78, 0.97]
Rao 2010b	100	4	40	70	0.71 [0.63, 0.79]	0.95 [0.87, 0.99]
Rao 2010b	109	15	31	59	0.78 [0.70, 0.84]	0.80 [0.69, 0.88]
Rao 2012a	31	6	34	113	0.48 [0.35, 0.60]	0.95 [0.89, 0.98]
Rao 2012a	44	24	21	95	0.68 [0.55, 0.79]	0.80 [0.71, 0.87]
Rao 2012b	67	25	24	100	0.74 [0.63, 0.82]	0.80 [0.72, 0.87]
Rao 2012b	53	6	38	119	0.58 [0.47, 0.68]	0.95 [0.90, 0.98]
Rao 2013	28	12	6	48	0.82 [0.65, 0.93]	0.80 [0.68, 0.89]
Rao 2013	18	3	16	57	0.53 [0.35, 0.70]	0.95 [0.86, 0.99]
Rao 2014	57	5	49	104	0.54 [0.44, 0.64]	0.95 [0.90, 0.98]
Rao 2014	86	22	20	87	0.81 [0.72, 0.88]	0.80 [0.71, 0.87]
Rolle 2011	77	6	49	46	0.61 [0.52, 0.70]	0.88 [0.77, 0.96]
Rolle 2011	48	3	78	49	0.38 [0.30, 0.47]	0.94 [0.84, 0.99]

Test 6. (Continued)

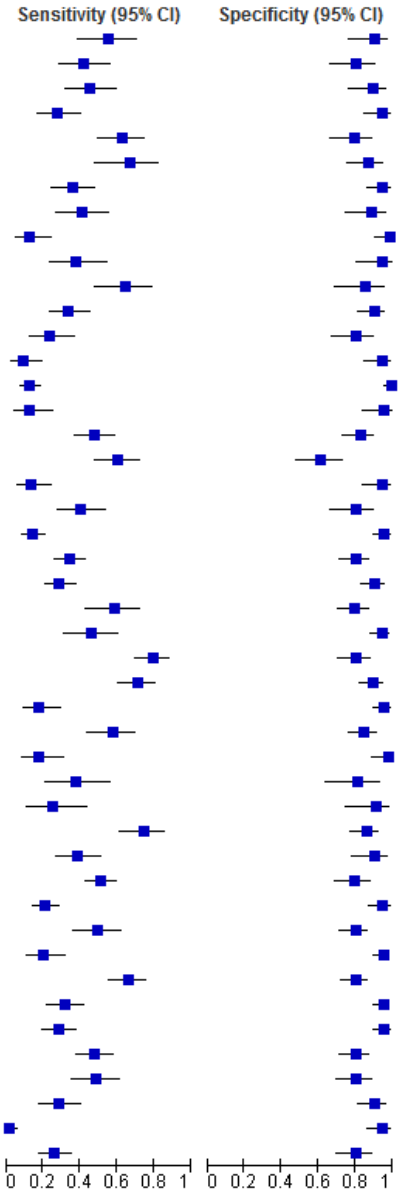
Rolle 2011	77	6	49	46	0.61 [0.52, 0.70]	0.88 [0.77, 0.96]
Rolle 2011	48	3	78	49	0.38 [0.30, 0.47]	0.94 [0.84, 0.99]
Seong 2010	90	7	12	59	0.88 [0.80, 0.94]	0.89 [0.79, 0.96]
Seong 2010	88	3	14	62	0.86 [0.78, 0.92]	0.95 [0.87, 0.99]
Seong 2010	94	13	8	52	0.92 [0.85, 0.97]	0.80 [0.68, 0.89]
Shin 2013	50	7	14	65	0.78 [0.66, 0.87]	0.90 [0.81, 0.96]
Shin 2013	53	14	11	58	0.83 [0.71, 0.91]	0.81 [0.70, 0.89]
Sullivan-Mee 2013	41	10	9	40	0.82 [0.69, 0.91]	0.80 [0.66, 0.90]
Sullivan-Mee 2013	35	3	15	48	0.70 [0.55, 0.82]	0.94 [0.84, 0.99]
Sung 2013	66	14	41	58	0.62 [0.52, 0.71]	0.81 [0.70, 0.89]
Sung 2013	59	4	48	68	0.55 [0.45, 0.65]	0.94 [0.86, 0.98]



Test 7. OCT: RNFL nasal quadrant

OCT: RNFL nasal quadrant

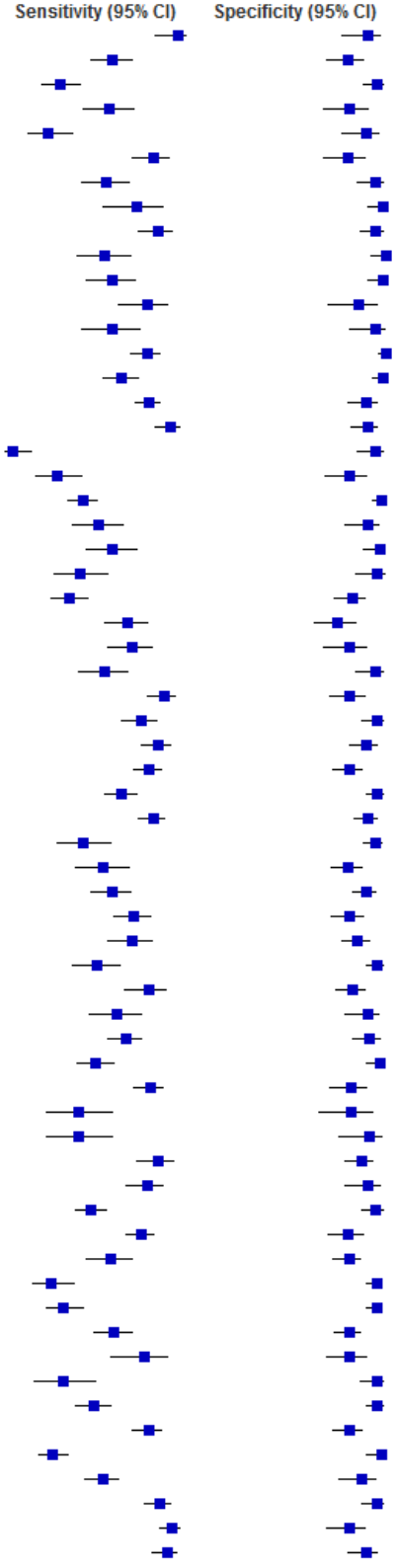
Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Aptel 2010	22	4	18	36	0.55 [0.38, 0.71]	0.90 [0.76, 0.97]
Barella 2013	24	9	33	37	0.42 [0.29, 0.56]	0.80 [0.66, 0.91]
Barella 2013	26	5	31	41	0.46 [0.32, 0.59]	0.89 [0.76, 0.96]
Begum 2014a	17	3	45	50	0.27 [0.17, 0.40]	0.94 [0.84, 0.99]
Begum 2014a	39	11	23	42	0.63 [0.50, 0.75]	0.79 [0.66, 0.89]
Benitez-del-Castillo 2011	22	7	11	48	0.67 [0.48, 0.82]	0.87 [0.76, 0.95]
Bertuzzi 2014	25	4	45	67	0.36 [0.25, 0.48]	0.94 [0.86, 0.98]
Cho 2011	20	5	29	38	0.41 [0.27, 0.56]	0.88 [0.75, 0.96]
Choi 2013	7	1	47	55	0.13 [0.05, 0.25]	0.98 [0.90, 1.00]
Fang 2010	16	2	26	32	0.38 [0.24, 0.54]	0.94 [0.80, 0.99]
Fang 2010	27	5	15	29	0.64 [0.48, 0.78]	0.85 [0.69, 0.95]
Hwang 2012	27	8	53	72	0.34 [0.24, 0.45]	0.90 [0.81, 0.96]
Jeoung 2010	13	11	42	44	0.24 [0.13, 0.37]	0.80 [0.67, 0.90]
Jeoung 2010	5	3	50	52	0.09 [0.03, 0.20]	0.95 [0.85, 0.99]
Jeoung 2013	21	1	143	118	0.13 [0.08, 0.19]	0.99 [0.95, 1.00]
Kim 2013b	6	2	42	40	0.13 [0.05, 0.25]	0.95 [0.84, 0.99]
Kim 2014a	44	16	48	76	0.48 [0.37, 0.58]	0.83 [0.73, 0.90]
Kim 2014b	41	25	27	39	0.60 [0.48, 0.72]	0.61 [0.48, 0.73]
Koh 2014	8	3	52	48	0.13 [0.06, 0.25]	0.94 [0.84, 0.99]
Koh 2014	24	10	36	40	0.40 [0.28, 0.53]	0.80 [0.66, 0.90]
Leite 2011	18	5	108	102	0.14 [0.09, 0.22]	0.95 [0.89, 0.98]
Leite 2011	43	21	83	86	0.34 [0.26, 0.43]	0.80 [0.72, 0.87]
Leung 2010	35	10	86	92	0.29 [0.21, 0.38]	0.90 [0.83, 0.95]
Lisboa 2013	28	19	20	75	0.58 [0.43, 0.72]	0.80 [0.70, 0.87]
Lisboa 2013	22	5	26	89	0.46 [0.31, 0.61]	0.95 [0.88, 0.98]
Mansoori 2011	66	19	17	76	0.80 [0.69, 0.88]	0.80 [0.71, 0.88]
Mansoori 2011	59	10	24	86	0.71 [0.60, 0.81]	0.90 [0.82, 0.95]
Moreno-Montañés 2010	11	5	50	100	0.18 [0.09, 0.30]	0.95 [0.89, 0.98]
Moreno-Montañés 2010	35	16	26	89	0.57 [0.44, 0.70]	0.85 [0.76, 0.91]
Mwanza 2014	9	1	41	48	0.18 [0.09, 0.31]	0.98 [0.89, 1.00]
Nakatani 2011	12	6	20	26	0.38 [0.21, 0.56]	0.81 [0.64, 0.93]
Nakatani 2011	8	3	24	29	0.25 [0.11, 0.43]	0.91 [0.75, 0.98]
Nouri-Mahdavi 2013	44	13	15	78	0.75 [0.62, 0.85]	0.86 [0.77, 0.92]
Oddone 2011	27	5	43	45	0.39 [0.27, 0.51]	0.90 [0.78, 0.97]
Rao 2010b	72	15	68	59	0.51 [0.43, 0.60]	0.80 [0.69, 0.88]
Rao 2010b	29	4	111	70	0.21 [0.14, 0.28]	0.95 [0.87, 0.99]
Rao 2012a	32	24	33	95	0.49 [0.37, 0.62]	0.80 [0.71, 0.87]
Rao 2012a	13	6	52	113	0.20 [0.11, 0.32]	0.95 [0.89, 0.98]
Rao 2012b	60	25	31	100	0.66 [0.55, 0.76]	0.80 [0.72, 0.87]
Rao 2012b	29	6	62	119	0.32 [0.22, 0.42]	0.95 [0.90, 0.98]
Rao 2014	30	5	76	104	0.28 [0.20, 0.38]	0.95 [0.90, 0.98]
Rao 2014	51	22	55	87	0.48 [0.38, 0.58]	0.80 [0.71, 0.87]
Shin 2013	31	14	33	58	0.48 [0.36, 0.61]	0.81 [0.70, 0.89]
Shin 2013	18	7	46	65	0.28 [0.18, 0.41]	0.90 [0.81, 0.96]
Sung 2013	2	4	105	68	0.02 [0.00, 0.07]	0.94 [0.86, 0.98]
Sung 2013	28	14	79	58	0.26 [0.18, 0.36]	0.81 [0.70, 0.89]



Test 8. OCT: RNFL superior quadrant

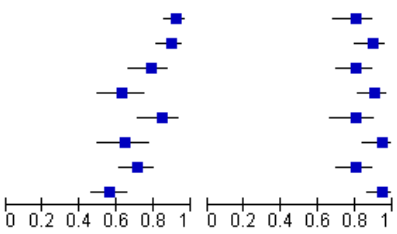
OCT: RNFL superior quadrant

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Aptel 2010	38	4	2	36	0.95 [0.83, 0.99]	0.90 [0.76, 0.97]
Arintawati 2013	48	14	33	54	0.59 [0.48, 0.70]	0.79 [0.68, 0.88]
Arintawati 2013	25	3	56	65	0.31 [0.21, 0.42]	0.96 [0.88, 0.99]
Barella 2013	33	9	24	37	0.58 [0.44, 0.71]	0.80 [0.66, 0.91]
Barella 2013	14	5	43	41	0.25 [0.14, 0.38]	0.89 [0.76, 0.96]
Begum 2014a	51	11	11	42	0.82 [0.70, 0.91]	0.79 [0.66, 0.89]
Begum 2014a	35	3	27	50	0.56 [0.43, 0.69]	0.94 [0.84, 0.99]
Benitez-del-Castillo 2011	24	1	9	54	0.73 [0.54, 0.87]	0.98 [0.90, 1.00]
Bertuzzi 2014	59	4	11	67	0.84 [0.74, 0.92]	0.94 [0.86, 0.98]
Cho 2011	27	0	22	43	0.55 [0.40, 0.69]	1.00 [0.92, 1.00]
Choi 2013	32	1	22	55	0.59 [0.45, 0.72]	0.98 [0.90, 1.00]
Fang 2010	33	5	9	29	0.79 [0.63, 0.90]	0.85 [0.69, 0.95]
Fang 2010	25	2	17	32	0.60 [0.43, 0.74]	0.94 [0.80, 0.99]
Garas 2011	87	0	24	93	0.78 [0.70, 0.86]	1.00 [0.96, 1.00]
Garas 2012	64	1	35	77	0.65 [0.54, 0.74]	0.99 [0.93, 1.00]
Huang 2011	116	8	30	66	0.79 [0.72, 0.86]	0.89 [0.80, 0.95]
Hwang 2012	73	8	7	72	0.91 [0.83, 0.96]	0.90 [0.81, 0.96]
Jeoung 2010	3	3	52	52	0.05 [0.01, 0.15]	0.95 [0.85, 0.99]
Jeoung 2010	16	11	39	44	0.29 [0.18, 0.43]	0.80 [0.67, 0.90]
Jeoung 2013	71	3	93	116	0.43 [0.36, 0.51]	0.97 [0.93, 0.99]
Kim 2011	29	5	27	44	0.52 [0.38, 0.65]	0.90 [0.78, 0.97]
Kim 2013a	31	2	21	56	0.60 [0.45, 0.73]	0.97 [0.88, 1.00]
Kim 2013b	20	2	28	40	0.42 [0.28, 0.57]	0.95 [0.84, 0.99]
Kim 2014a	33	17	59	75	0.36 [0.26, 0.47]	0.82 [0.72, 0.89]
Kim 2014b	46	17	22	47	0.68 [0.55, 0.78]	0.73 [0.61, 0.84]
Koh 2014	42	10	18	40	0.70 [0.57, 0.81]	0.80 [0.66, 0.90]
Koh 2014	33	3	27	48	0.55 [0.42, 0.68]	0.94 [0.84, 0.99]
Lee 2010	77	15	11	62	0.88 [0.79, 0.94]	0.81 [0.70, 0.89]
Lee 2010	66	4	22	73	0.75 [0.65, 0.84]	0.95 [0.87, 0.99]
Lee 2010	78	8	14	69	0.85 [0.76, 0.91]	0.90 [0.81, 0.95]
Leite 2011	100	21	26	86	0.79 [0.71, 0.86]	0.80 [0.72, 0.87]
Leite 2011	81	5	45	102	0.64 [0.55, 0.73]	0.95 [0.89, 0.98]
Leung 2010	99	10	22	92	0.82 [0.74, 0.88]	0.90 [0.83, 0.95]
Lisboa 2013	21	5	27	89	0.44 [0.29, 0.59]	0.95 [0.88, 0.98]
Lisboa 2013	26	19	22	75	0.54 [0.39, 0.69]	0.80 [0.70, 0.87]
Mansoori 2011	49	10	34	86	0.59 [0.48, 0.70]	0.90 [0.82, 0.95]
Mansoori 2011	59	19	24	76	0.71 [0.60, 0.81]	0.80 [0.71, 0.88]
Moreno-Montañés 2010	43	16	18	89	0.70 [0.57, 0.81]	0.85 [0.76, 0.91]
Moreno-Montañés 2010	31	5	30	100	0.51 [0.38, 0.64]	0.95 [0.89, 0.98]
Mwanza 2012	46	18	12	81	0.79 [0.67, 0.89]	0.82 [0.73, 0.89]
Mwanza 2014	31	5	19	44	0.62 [0.47, 0.75]	0.90 [0.78, 0.97]
Na 2013a	70	6	35	62	0.67 [0.57, 0.76]	0.91 [0.82, 0.97]
Na 2013a	53	2	52	66	0.50 [0.41, 0.60]	0.97 [0.90, 1.00]
Na 2013a	84	13	21	55	0.80 [0.71, 0.87]	0.81 [0.70, 0.89]
Nakatani 2011	13	6	19	26	0.41 [0.24, 0.59]	0.81 [0.64, 0.93]
Nakatani 2011	13	3	19	29	0.41 [0.24, 0.59]	0.91 [0.75, 0.98]
Nouri-Mahdavi 2013	50	12	9	79	0.85 [0.73, 0.93]	0.87 [0.78, 0.93]
Oddone 2011	55	5	15	45	0.79 [0.67, 0.87]	0.90 [0.78, 0.97]
Rao 2010b	67	4	73	70	0.48 [0.39, 0.56]	0.95 [0.87, 0.99]
Rao 2010b	105	15	35	59	0.75 [0.67, 0.82]	0.80 [0.69, 0.88]
Rao 2012a	38	24	27	95	0.58 [0.46, 0.71]	0.80 [0.71, 0.87]
Rao 2012a	17	6	48	113	0.26 [0.16, 0.39]	0.95 [0.89, 0.98]
Rao 2012b	30	6	61	119	0.33 [0.23, 0.44]	0.95 [0.90, 0.98]
Rao 2012b	55	25	36	100	0.60 [0.50, 0.71]	0.80 [0.72, 0.87]
Rao 2013	26	12	8	48	0.76 [0.59, 0.89]	0.80 [0.68, 0.89]
Rao 2013	11	3	23	57	0.32 [0.17, 0.51]	0.95 [0.86, 0.99]
Rao 2014	52	5	54	104	0.49 [0.39, 0.59]	0.95 [0.90, 0.98]
Rao 2014	84	22	22	87	0.79 [0.70, 0.87]	0.80 [0.71, 0.87]
Rolle 2011	34	1	92	51	0.27 [0.19, 0.36]	0.98 [0.90, 1.00]
Rolle 2011	68	7	58	45	0.54 [0.45, 0.63]	0.87 [0.74, 0.94]
Seong 2010	87	3	15	62	0.85 [0.77, 0.92]	0.95 [0.87, 0.99]
Seong 2010	94	13	8	52	0.92 [0.85, 0.97]	0.80 [0.68, 0.89]
Seong 2010	91	7	11	59	0.89 [0.82, 0.94]	0.89 [0.79, 0.96]



Test 8. (Continued)

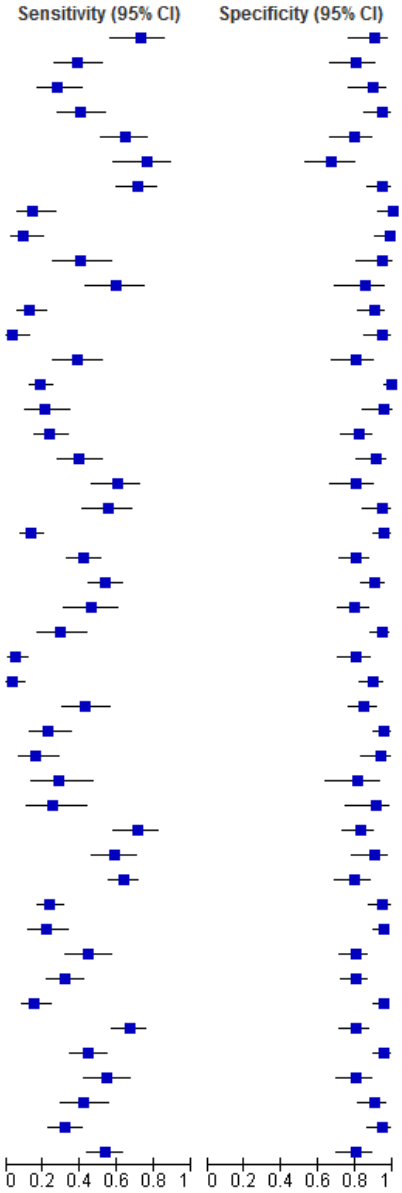
Seong 2010	94	13	8	52	0.92 [0.85, 0.97]	0.80 [0.68, 0.89]
Seong 2010	91	7	11	59	0.89 [0.82, 0.94]	0.89 [0.79, 0.96]
Shin 2013	50	14	14	58	0.78 [0.66, 0.87]	0.81 [0.70, 0.89]
Shin 2013	40	7	24	65	0.63 [0.50, 0.74]	0.90 [0.81, 0.96]
Sullivan-Mee 2013	42	10	8	40	0.84 [0.71, 0.93]	0.80 [0.66, 0.90]
Sullivan-Mee 2013	32	3	18	48	0.64 [0.49, 0.77]	0.94 [0.84, 0.99]
Sung 2013	76	14	31	58	0.71 [0.61, 0.79]	0.81 [0.70, 0.89]
Sung 2013	60	4	47	68	0.56 [0.46, 0.66]	0.94 [0.86, 0.98]



Test 9. OCT: RNFL temporal quadrant

OCT: RNFL temporal quadrant

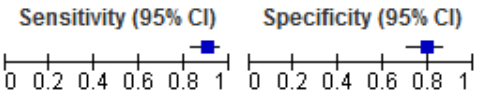
Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Aptel 2010	29	4	11	36	0.72 [0.56, 0.85]	0.90 [0.76, 0.97]
Barella 2013	22	9	35	37	0.39 [0.26, 0.52]	0.80 [0.66, 0.91]
Barella 2013	16	5	41	41	0.28 [0.17, 0.42]	0.89 [0.76, 0.96]
Begum 2014a	25	3	37	50	0.40 [0.28, 0.54]	0.94 [0.84, 0.99]
Begum 2014a	40	11	22	42	0.65 [0.51, 0.76]	0.79 [0.66, 0.89]
Benitez-del-Castillo 2011	25	18	8	37	0.76 [0.58, 0.89]	0.67 [0.53, 0.79]
Bertuzzi 2014	50	4	20	67	0.71 [0.59, 0.82]	0.94 [0.86, 0.98]
Cho 2011	7	0	42	43	0.14 [0.06, 0.27]	1.00 [0.92, 1.00]
Choi 2013	5	1	49	55	0.09 [0.03, 0.20]	0.98 [0.90, 1.00]
Fang 2010	17	2	25	32	0.40 [0.26, 0.57]	0.94 [0.80, 0.99]
Fang 2010	25	5	17	29	0.60 [0.43, 0.74]	0.85 [0.69, 0.95]
Hwang 2012	10	8	70	72	0.13 [0.06, 0.22]	0.90 [0.81, 0.96]
Jeoung 2010	2	3	53	52	0.04 [0.00, 0.13]	0.95 [0.85, 0.99]
Jeoung 2010	21	11	34	44	0.38 [0.25, 0.52]	0.80 [0.67, 0.90]
Jeoung 2013	30	1	134	118	0.18 [0.13, 0.25]	0.99 [0.95, 1.00]
Kim 2013b	10	2	38	40	0.21 [0.10, 0.35]	0.95 [0.84, 0.99]
Kim 2014a	22	17	70	75	0.24 [0.16, 0.34]	0.82 [0.72, 0.89]
Kim 2014b	27	6	41	58	0.40 [0.28, 0.52]	0.91 [0.81, 0.96]
Koh 2014	36	10	24	40	0.60 [0.47, 0.72]	0.80 [0.66, 0.90]
Koh 2014	33	3	27	48	0.55 [0.42, 0.68]	0.94 [0.84, 0.99]
Leite 2011	17	5	109	102	0.13 [0.08, 0.21]	0.95 [0.89, 0.98]
Leite 2011	53	21	73	86	0.42 [0.33, 0.51]	0.80 [0.72, 0.87]
Leung 2010	65	10	56	92	0.54 [0.44, 0.63]	0.90 [0.83, 0.95]
Lisboa 2013	22	19	26	75	0.46 [0.31, 0.61]	0.80 [0.70, 0.87]
Lisboa 2013	14	5	34	89	0.29 [0.17, 0.44]	0.95 [0.88, 0.98]
Mansoori 2011	4	19	79	76	0.05 [0.01, 0.12]	0.80 [0.71, 0.88]
Mansoori 2011	3	10	80	86	0.04 [0.01, 0.10]	0.90 [0.82, 0.95]
Moreno-Montañés 2010	26	16	35	89	0.43 [0.30, 0.56]	0.85 [0.76, 0.91]
Moreno-Montañés 2010	14	5	47	100	0.23 [0.13, 0.35]	0.95 [0.89, 0.98]
Mwanza 2014	8	3	42	46	0.16 [0.07, 0.29]	0.94 [0.83, 0.99]
Nakatani 2011	9	6	23	26	0.28 [0.14, 0.47]	0.81 [0.64, 0.93]
Nakatani 2011	8	3	24	29	0.25 [0.11, 0.43]	0.91 [0.75, 0.98]
Nouri-Mahdavi 2013	42	16	17	75	0.71 [0.58, 0.82]	0.82 [0.73, 0.90]
Oddone 2011	41	5	29	45	0.59 [0.46, 0.70]	0.90 [0.78, 0.97]
Rao 2010b	89	15	51	59	0.64 [0.55, 0.72]	0.80 [0.69, 0.88]
Rao 2010b	33	4	107	70	0.24 [0.17, 0.31]	0.95 [0.87, 0.99]
Rao 2012a	14	6	51	113	0.22 [0.12, 0.33]	0.95 [0.89, 0.98]
Rao 2012a	29	24	36	95	0.45 [0.32, 0.57]	0.80 [0.71, 0.87]
Rao 2012b	29	25	62	100	0.32 [0.22, 0.42]	0.80 [0.72, 0.87]
Rao 2012b	14	6	77	119	0.15 [0.09, 0.24]	0.95 [0.90, 0.98]
Rao 2014	71	22	35	87	0.67 [0.57, 0.76]	0.80 [0.71, 0.87]
Rao 2014	47	5	59	104	0.44 [0.35, 0.54]	0.95 [0.90, 0.98]
Shin 2013	35	14	29	58	0.55 [0.42, 0.67]	0.81 [0.70, 0.89]
Shin 2013	27	7	37	65	0.42 [0.30, 0.55]	0.90 [0.81, 0.96]
Sung 2013	34	4	73	68	0.32 [0.23, 0.41]	0.94 [0.86, 0.98]
Sung 2013	57	14	50	58	0.53 [0.43, 0.63]	0.81 [0.70, 0.89]



Test 10. HRT: Bathija function

HRT: Bathija function

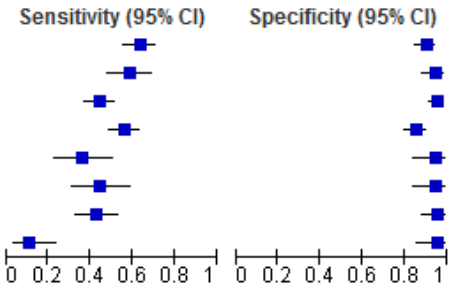
Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Iester 2008	86	25	9	94	0.91 [0.83, 0.96]	0.79 [0.71, 0.86]



Test 11. HRT: Cup area

HRT: Cup area

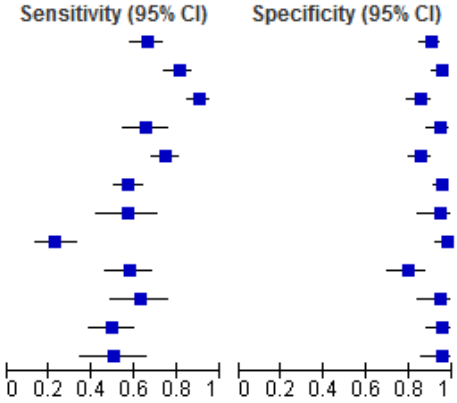
Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Bozkurt 2010	100	18	58	166	0.63 [0.55, 0.71]	0.90 [0.85, 0.94]
Ferreras 2008a	53	5	37	88	0.59 [0.48, 0.69]	0.95 [0.88, 0.98]
Ferreras 2008b	93	11	117	214	0.44 [0.37, 0.51]	0.95 [0.91, 0.98]
Ferreras 2008b	117	34	93	191	0.56 [0.49, 0.63]	0.85 [0.80, 0.89]
Jindal 2010	18	3	32	48	0.36 [0.23, 0.51]	0.94 [0.84, 0.99]
Pablo 2010	24	3	30	48	0.44 [0.31, 0.59]	0.94 [0.84, 0.99]
Rao 2010a	42	4	56	75	0.43 [0.33, 0.53]	0.95 [0.88, 0.99]
Roberti 2014	5	3	41	55	0.11 [0.04, 0.24]	0.95 [0.86, 0.99]



Test 12. HRT: C/D area ratio

HRT: C/D area ratio

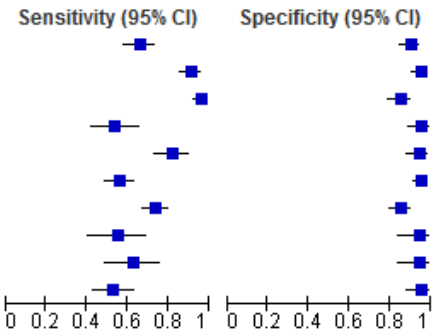
Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Bozkurt 2010	104	18	54	166	0.66 [0.58, 0.73]	0.90 [0.85, 0.94]
Calvo 2014	126	9	30	173	0.81 [0.74, 0.87]	0.95 [0.91, 0.98]
Calvo 2014	141	27	15	155	0.90 [0.85, 0.95]	0.85 [0.79, 0.90]
Ferreras 2008a	59	5	31	88	0.66 [0.55, 0.75]	0.95 [0.88, 0.98]
Ferreras 2008b	156	34	54	191	0.74 [0.68, 0.80]	0.85 [0.80, 0.89]
Ferreras 2008b	120	11	90	214	0.57 [0.50, 0.64]	0.95 [0.91, 0.98]
Jindal 2010	29	3	22	48	0.57 [0.42, 0.71]	0.94 [0.84, 0.99]
Kratz 2014	19	2	66	86	0.22 [0.14, 0.33]	0.98 [0.92, 1.00]
Kratz 2014	49	18	36	70	0.58 [0.46, 0.68]	0.80 [0.70, 0.87]
Pablo 2010	34	3	20	48	0.63 [0.49, 0.76]	0.94 [0.84, 0.99]
Rao 2010a	48	4	50	75	0.49 [0.39, 0.59]	0.95 [0.88, 0.99]
Roberti 2014	23	3	23	55	0.50 [0.35, 0.65]	0.95 [0.86, 0.99]



Test 13. HRT: vertical C/D ratio

HRT: vertical C/D ratio

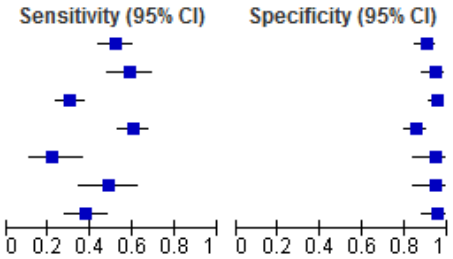
Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Bozkurt 2010	104	18	54	166	0.66 [0.58, 0.73]	0.90 [0.85, 0.94]
Calvo 2014	142	9	14	173	0.91 [0.85, 0.95]	0.95 [0.91, 0.98]
Calvo 2014	150	27	6	155	0.96 [0.92, 0.99]	0.85 [0.79, 0.90]
De Leon-Ortega 2007	42	4	36	85	0.54 [0.42, 0.65]	0.96 [0.89, 0.99]
Ferreras 2008a	74	5	16	88	0.82 [0.73, 0.89]	0.95 [0.88, 0.98]
Ferreras 2008b	117	11	93	214	0.56 [0.49, 0.63]	0.95 [0.91, 0.98]
Ferreras 2008b	154	34	56	191	0.73 [0.67, 0.79]	0.85 [0.80, 0.89]
Jindal 2010	28	3	23	48	0.55 [0.40, 0.69]	0.94 [0.84, 0.99]
Pablo 2010	34	3	20	48	0.63 [0.49, 0.76]	0.94 [0.84, 0.99]
Rao 2010a	52	4	46	75	0.53 [0.43, 0.63]	0.95 [0.88, 0.99]



Test 14. HRT: Cup shape measure

HRT: Cup shape measure

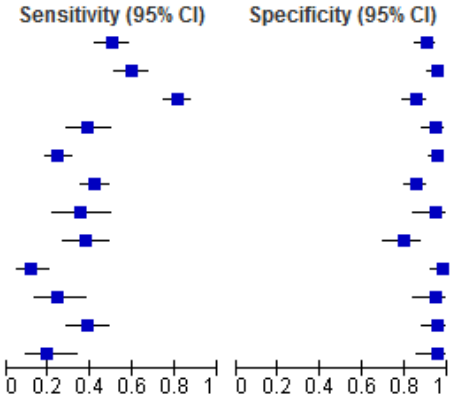
Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Bozkurt 2010	82	18	76	166	0.52 [0.44, 0.60]	0.90 [0.85, 0.94]
Ferreras 2008a	53	5	37	88	0.59 [0.48, 0.69]	0.95 [0.88, 0.98]
Ferreras 2008b	63	11	147	214	0.30 [0.24, 0.37]	0.95 [0.91, 0.98]
Ferreras 2008b	126	34	84	191	0.60 [0.53, 0.67]	0.85 [0.80, 0.89]
Jindal 2010	11	3	39	48	0.22 [0.12, 0.36]	0.94 [0.84, 0.99]
Pablo 2010	26	3	28	48	0.48 [0.34, 0.62]	0.94 [0.84, 0.99]
Rao 2010a	37	4	61	75	0.38 [0.28, 0.48]	0.95 [0.88, 0.99]



Test 15. HRT: Cup volume

HRT: Cup volume

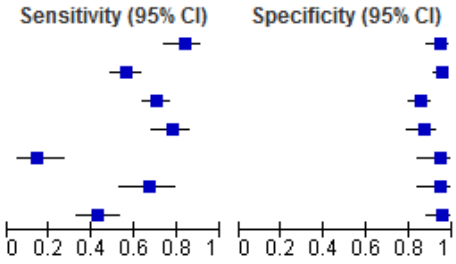
Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Bozkurt 2010	79	18	79	166	0.50 [0.42, 0.58]	0.90 [0.85, 0.94]
Calvo 2014	93	9	63	173	0.60 [0.51, 0.67]	0.95 [0.91, 0.98]
Calvo 2014	127	27	29	155	0.81 [0.74, 0.87]	0.85 [0.79, 0.90]
Ferreras 2008a	35	5	55	88	0.39 [0.29, 0.50]	0.95 [0.88, 0.98]
Ferreras 2008b	52	11	158	214	0.25 [0.19, 0.31]	0.95 [0.91, 0.98]
Ferreras 2008b	88	34	122	191	0.42 [0.35, 0.49]	0.85 [0.80, 0.89]
Jindal 2010	18	3	33	48	0.35 [0.22, 0.50]	0.94 [0.84, 0.99]
Kratz 2014	32	18	53	70	0.38 [0.27, 0.49]	0.80 [0.70, 0.87]
Kratz 2014	10	2	75	86	0.12 [0.06, 0.21]	0.98 [0.92, 1.00]
Pablo 2010	13	3	41	48	0.24 [0.13, 0.38]	0.94 [0.84, 0.99]
Rao 2010a	38	4	60	75	0.39 [0.29, 0.49]	0.95 [0.88, 0.99]
Roberti 2014	9	3	37	55	0.20 [0.09, 0.34]	0.95 [0.86, 0.99]



Test 16. HRT: FSM discriminant function o Mikelberg function

HRT: FSM discriminant function o Mikelberg function

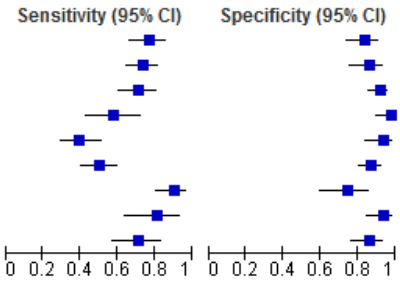
Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Ferreras 2008a	75	5	15	88	0.83 [0.74, 0.90]	0.95 [0.88, 0.98]
Ferreras 2008b	117	11	93	214	0.56 [0.49, 0.63]	0.95 [0.91, 0.98]
Ferreras 2008b	148	34	62	191	0.70 [0.64, 0.77]	0.85 [0.80, 0.89]
Iester 2008	74	16	21	103	0.78 [0.68, 0.86]	0.87 [0.79, 0.92]
Jindal 2010	7	3	43	48	0.14 [0.06, 0.27]	0.94 [0.84, 0.99]
Pablo 2010	36	3	18	48	0.67 [0.53, 0.79]	0.94 [0.84, 0.99]
Rao 2010a	42	4	56	75	0.43 [0.33, 0.53]	0.95 [0.88, 0.99]



Test 17. HRT: MRA

HRT: MRA

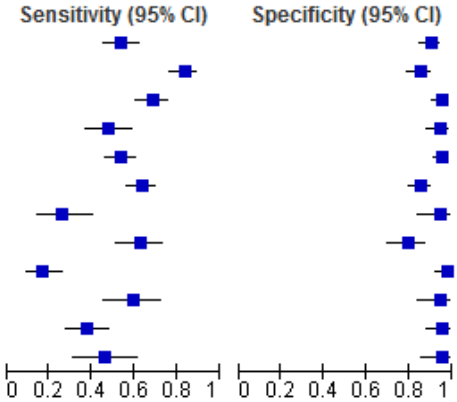
Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
De Leon-Ortega 2007	60	15	18	74	0.77 [0.66, 0.86]	0.83 [0.74, 0.90]
Ferreras 2007	85	10	30	61	0.74 [0.65, 0.82]	0.86 [0.76, 0.93]
Harizman 2006	59	11	24	123	0.71 [0.60, 0.81]	0.92 [0.86, 0.96]
Jindal 2010	29	1	21	49	0.58 [0.43, 0.72]	0.98 [0.89, 1.00]
Moreno-Montañés 2008	33	4	50	55	0.40 [0.29, 0.51]	0.93 [0.84, 0.98]
Oddone 2008	52	18	52	121	0.50 [0.40, 0.60]	0.87 [0.80, 0.92]
Oddone 2011	63	13	7	37	0.90 [0.80, 0.96]	0.74 [0.60, 0.85]
Zelefsky 2006	26	4	6	60	0.81 [0.64, 0.93]	0.94 [0.85, 0.98]
Zelefsky 2006	37	10	15	62	0.71 [0.57, 0.83]	0.86 [0.76, 0.93]



Test 18. HRT: Rim area

HRT: Rim area

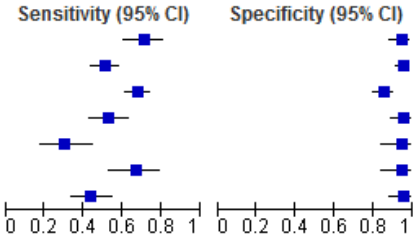
Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Bozkurt 2010	85	18	73	166	0.54 [0.46, 0.62]	0.90 [0.85, 0.94]
Calvo 2014	130	27	26	155	0.83 [0.77, 0.89]	0.85 [0.79, 0.90]
Calvo 2014	107	9	49	173	0.69 [0.61, 0.76]	0.95 [0.91, 0.98]
Ferreras 2008a	43	5	47	88	0.48 [0.37, 0.59]	0.95 [0.88, 0.98]
Ferreras 2008b	112	11	98	214	0.53 [0.46, 0.60]	0.95 [0.91, 0.98]
Ferreras 2008b	133	34	77	191	0.63 [0.56, 0.70]	0.85 [0.80, 0.89]
Jindal 2010	13	3	37	48	0.26 [0.15, 0.40]	0.94 [0.84, 0.99]
Kratz 2014	53	18	32	70	0.62 [0.51, 0.73]	0.80 [0.70, 0.87]
Kratz 2014	14	2	71	86	0.16 [0.09, 0.26]	0.98 [0.92, 1.00]
Pablo 2010	32	3	22	48	0.59 [0.45, 0.72]	0.94 [0.84, 0.99]
Rao 2010a	37	4	61	75	0.38 [0.28, 0.48]	0.95 [0.88, 0.99]
Roberti 2014	21	3	25	55	0.46 [0.31, 0.61]	0.95 [0.86, 0.99]



Test 19. HRT: RB discriminant function

HRT: RB discriminant function

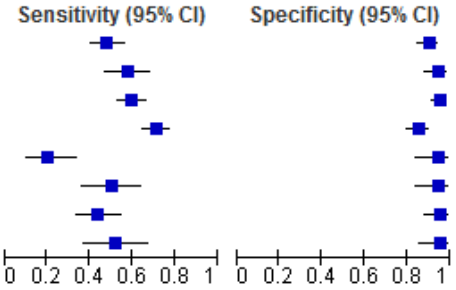
Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Ferreras 2008a	64	5	26	88	0.71 [0.61, 0.80]	0.95 [0.88, 0.98]
Ferreras 2008b	107	11	103	214	0.51 [0.44, 0.58]	0.95 [0.91, 0.98]
Ferreras 2008b	142	34	68	191	0.68 [0.61, 0.74]	0.85 [0.80, 0.89]
Gonzales de la Rosa 2013	55	5	49	97	0.53 [0.43, 0.63]	0.95 [0.89, 0.98]
Jindal 2010	15	3	35	48	0.30 [0.18, 0.45]	0.94 [0.84, 0.99]
Pablo 2010	36	3	18	48	0.67 [0.53, 0.79]	0.94 [0.84, 0.99]
Rao 2010a	43	4	55	75	0.44 [0.34, 0.54]	0.95 [0.88, 0.99]



Test 20. HRT: Rim Volume

HRT: Rim Volume

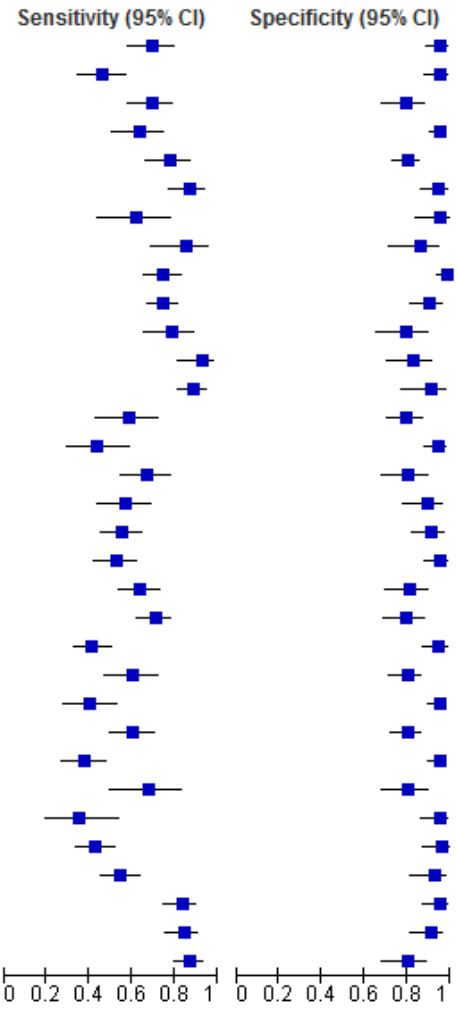
Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Bozkurt 2010	76	18	82	166	0.48 [0.40, 0.56]	0.90 [0.85, 0.94]
Ferreras 2008a	52	5	38	88	0.58 [0.47, 0.68]	0.95 [0.88, 0.98]
Ferreras 2008b	125	11	85	214	0.60 [0.53, 0.66]	0.95 [0.91, 0.98]
Ferreras 2008b	150	34	60	191	0.71 [0.65, 0.77]	0.85 [0.80, 0.89]
Jindal 2010	10	3	40	48	0.20 [0.10, 0.34]	0.94 [0.84, 0.99]
Pablo 2010	27	3	27	48	0.50 [0.36, 0.64]	0.94 [0.84, 0.99]
Rao 2010a	43	4	55	75	0.44 [0.34, 0.54]	0.95 [0.88, 0.99]
Roberti 2014	24	3	22	55	0.52 [0.37, 0.67]	0.95 [0.86, 0.99]



Test 21. OCT: GCC RTVue average thickness

OCT: GCC RTVue average thickness

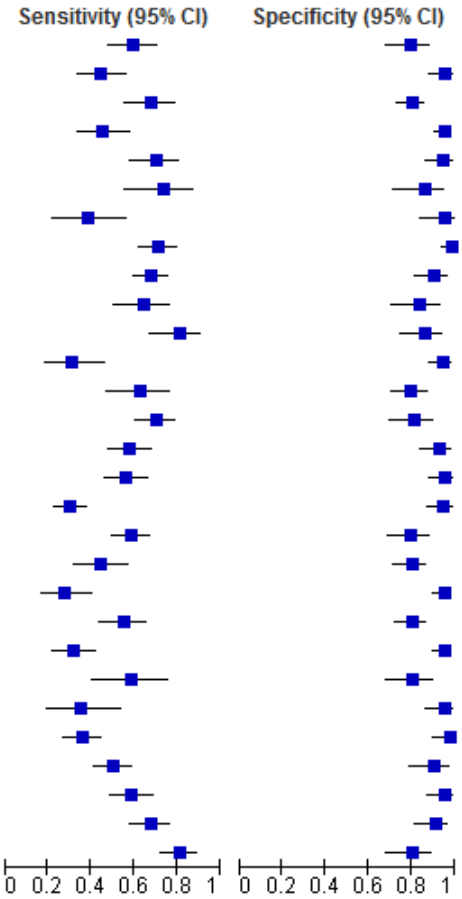
Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Akashi 2013	52	4	23	83	0.69 [0.58, 0.79]	0.95 [0.89, 0.99]
Arintawati 2013	37	3	44	65	0.46 [0.35, 0.57]	0.96 [0.88, 0.99]
Arintawati 2013	56	14	25	54	0.69 [0.58, 0.79]	0.79 [0.68, 0.88]
Begum 2014b	43	8	25	157	0.63 [0.51, 0.75]	0.95 [0.91, 0.98]
Begum 2014b	53	33	15	132	0.78 [0.66, 0.87]	0.80 [0.73, 0.86]
Bertuzzi 2014	61	4	9	67	0.87 [0.77, 0.94]	0.94 [0.86, 0.98]
Fang 2010	21	2	13	40	0.62 [0.44, 0.78]	0.95 [0.84, 0.99]
Fang 2010	29	6	5	36	0.85 [0.69, 0.95]	0.86 [0.71, 0.95]
Garas 2011	83	1	28	92	0.75 [0.66, 0.83]	0.99 [0.94, 1.00]
Huang 2011	109	7	37	67	0.75 [0.67, 0.81]	0.91 [0.81, 0.96]
Kim 2011	44	10	12	39	0.79 [0.66, 0.88]	0.80 [0.66, 0.90]
Kim 2013a	48	10	4	48	0.92 [0.81, 0.98]	0.83 [0.71, 0.91]
Kita 2013	88	3	11	32	0.89 [0.81, 0.94]	0.91 [0.77, 0.98]
Lisboa 2013	28	19	20	75	0.58 [0.43, 0.72]	0.80 [0.70, 0.87]
Lisboa 2013	21	5	27	89	0.44 [0.29, 0.59]	0.95 [0.88, 0.98]
Moreno 2011	45	11	22	45	0.67 [0.55, 0.78]	0.80 [0.68, 0.90]
Moreno 2011	38	6	29	50	0.57 [0.44, 0.69]	0.89 [0.78, 0.96]
Na 2013a	58	6	47	62	0.55 [0.45, 0.65]	0.91 [0.82, 0.97]
Na 2013a	55	3	50	65	0.52 [0.42, 0.62]	0.96 [0.88, 0.99]
Na 2013a	67	13	38	55	0.64 [0.54, 0.73]	0.81 [0.70, 0.89]
Rao 2010b	99	15	41	59	0.71 [0.62, 0.78]	0.80 [0.69, 0.88]
Rao 2010b	58	4	82	70	0.41 [0.33, 0.50]	0.95 [0.87, 0.99]
Rao 2012a	39	24	26	95	0.60 [0.47, 0.72]	0.80 [0.71, 0.87]
Rao 2012a	26	6	39	113	0.40 [0.28, 0.53]	0.95 [0.89, 0.98]
Rao 2012b	55	25	36	100	0.60 [0.50, 0.71]	0.80 [0.72, 0.87]
Rao 2012b	34	6	57	119	0.37 [0.27, 0.48]	0.95 [0.90, 0.98]
Rao 2013	23	12	11	48	0.68 [0.49, 0.83]	0.80 [0.68, 0.89]
Rao 2013	12	3	22	57	0.35 [0.20, 0.54]	0.95 [0.86, 0.99]
Rolle 2011	54	2	72	50	0.43 [0.34, 0.52]	0.96 [0.87, 1.00]
Rolle 2011	69	4	57	48	0.55 [0.46, 0.64]	0.92 [0.81, 0.98]
Seong 2010	85	3	17	62	0.83 [0.75, 0.90]	0.95 [0.87, 0.99]
Seong 2010	86	6	16	59	0.84 [0.76, 0.91]	0.91 [0.81, 0.97]
Seong 2010	89	13	13	52	0.87 [0.79, 0.93]	0.80 [0.68, 0.89]



Test 22. OCT: GCC RTVue superior thickness

OCT: GCC RTVue superior thickness

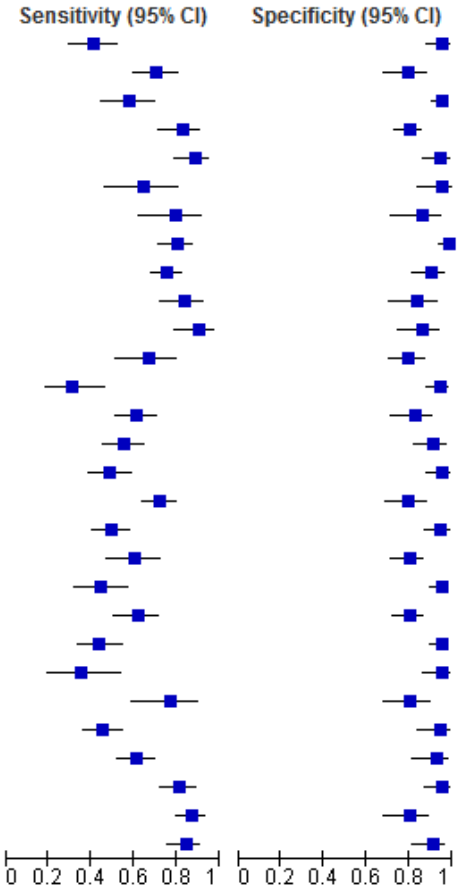
Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Arintawati 2013	48	14	33	54	0.59 [0.48, 0.70]	0.79 [0.68, 0.88]
Arintawati 2013	36	3	45	65	0.44 [0.33, 0.56]	0.96 [0.88, 0.99]
Begum 2014b	46	33	22	132	0.68 [0.55, 0.78]	0.80 [0.73, 0.86]
Begum 2014b	31	8	37	157	0.46 [0.33, 0.58]	0.95 [0.91, 0.98]
Bertuzzi 2014	49	4	21	67	0.70 [0.58, 0.80]	0.94 [0.86, 0.98]
Fang 2010	25	6	9	36	0.74 [0.56, 0.87]	0.86 [0.71, 0.95]
Fang 2010	13	2	21	40	0.38 [0.22, 0.56]	0.95 [0.84, 0.99]
Garas 2011	79	1	32	92	0.71 [0.62, 0.79]	0.99 [0.94, 1.00]
Huang 2011	99	7	47	67	0.68 [0.60, 0.75]	0.91 [0.81, 0.96]
Kim 2011	36	8	20	41	0.64 [0.50, 0.77]	0.84 [0.70, 0.93]
Kim 2013a	42	8	10	50	0.81 [0.67, 0.90]	0.86 [0.75, 0.94]
Lisboa 2013	15	5	33	89	0.31 [0.19, 0.46]	0.95 [0.88, 0.98]
Lisboa 2013	30	19	18	75	0.63 [0.47, 0.76]	0.80 [0.70, 0.87]
Na 2013a	74	13	31	55	0.70 [0.61, 0.79]	0.81 [0.70, 0.89]
Na 2013a	61	5	44	63	0.58 [0.48, 0.68]	0.93 [0.84, 0.98]
Na 2013a	59	3	46	65	0.56 [0.46, 0.66]	0.96 [0.88, 0.99]
Rao 2010b	42	4	98	70	0.30 [0.23, 0.38]	0.95 [0.87, 0.99]
Rao 2010b	82	15	58	59	0.59 [0.50, 0.67]	0.80 [0.69, 0.88]
Rao 2012a	29	24	36	95	0.45 [0.32, 0.57]	0.80 [0.71, 0.87]
Rao 2012a	18	6	47	113	0.28 [0.17, 0.40]	0.95 [0.89, 0.98]
Rao 2012b	50	25	41	100	0.55 [0.44, 0.65]	0.80 [0.72, 0.87]
Rao 2012b	29	6	62	119	0.32 [0.22, 0.42]	0.95 [0.90, 0.98]
Rao 2013	20	12	14	48	0.59 [0.41, 0.75]	0.80 [0.68, 0.89]
Rao 2013	12	3	22	57	0.35 [0.20, 0.54]	0.95 [0.86, 0.99]
Rolle 2011	45	1	81	51	0.36 [0.27, 0.45]	0.98 [0.90, 1.00]
Rolle 2011	63	5	63	47	0.50 [0.41, 0.59]	0.90 [0.79, 0.97]
Seong 2010	60	3	42	62	0.59 [0.49, 0.68]	0.95 [0.87, 0.99]
Seong 2010	69	6	33	59	0.68 [0.58, 0.77]	0.91 [0.81, 0.97]
Seong 2010	83	13	19	52	0.81 [0.72, 0.88]	0.80 [0.68, 0.89]



Test 23. OCT: GCC RTVue inferior thickness

OCT: GCC RTVue inferior thickness

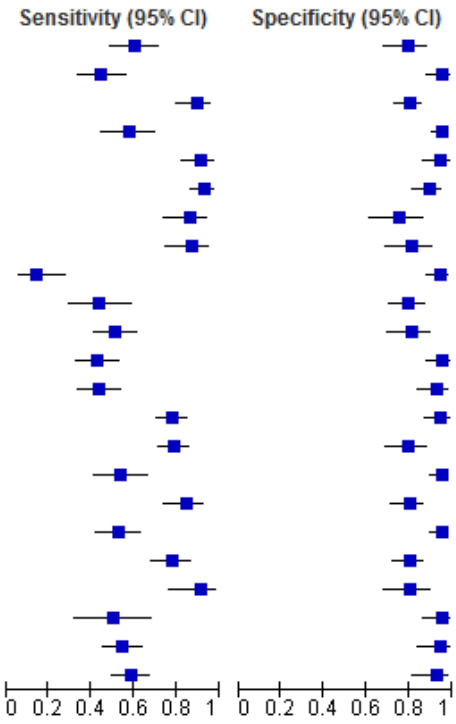
Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Arintawati 2013	33	3	48	65	0.41 [0.30, 0.52]	0.96 [0.88, 0.99]
Arintawati 2013	57	14	24	54	0.70 [0.59, 0.80]	0.79 [0.68, 0.88]
Begum 2014b	39	8	29	157	0.57 [0.45, 0.69]	0.95 [0.91, 0.98]
Begum 2014b	56	33	12	132	0.82 [0.71, 0.91]	0.80 [0.73, 0.86]
Bertuzzi 2014	62	4	8	67	0.89 [0.79, 0.95]	0.94 [0.86, 0.98]
Fang 2010	22	2	12	40	0.65 [0.46, 0.80]	0.95 [0.84, 0.99]
Fang 2010	27	6	7	36	0.79 [0.62, 0.91]	0.86 [0.71, 0.95]
Garas 2011	89	1	22	92	0.80 [0.72, 0.87]	0.99 [0.94, 1.00]
Huang 2011	110	7	36	67	0.75 [0.68, 0.82]	0.91 [0.81, 0.96]
Kim 2011	47	8	9	41	0.84 [0.72, 0.92]	0.84 [0.70, 0.93]
Kim 2013a	47	8	5	50	0.90 [0.79, 0.97]	0.86 [0.75, 0.94]
Lisboa 2013	32	19	16	75	0.67 [0.52, 0.80]	0.80 [0.70, 0.87]
Lisboa 2013	15	5	33	89	0.31 [0.19, 0.46]	0.95 [0.88, 0.98]
Na 2013a	64	12	41	56	0.61 [0.51, 0.70]	0.82 [0.71, 0.91]
Na 2013a	58	6	47	62	0.55 [0.45, 0.65]	0.91 [0.82, 0.97]
Na 2013a	51	3	54	65	0.49 [0.39, 0.59]	0.96 [0.88, 0.99]
Rao 2010b	101	15	39	59	0.72 [0.64, 0.79]	0.80 [0.69, 0.88]
Rao 2010b	69	4	71	70	0.49 [0.41, 0.58]	0.95 [0.87, 0.99]
Rao 2012a	39	24	26	95	0.60 [0.47, 0.72]	0.80 [0.71, 0.87]
Rao 2012a	29	6	36	113	0.45 [0.32, 0.57]	0.95 [0.89, 0.98]
Rao 2012b	56	25	35	100	0.62 [0.51, 0.72]	0.80 [0.72, 0.87]
Rao 2012b	40	6	51	119	0.44 [0.34, 0.55]	0.95 [0.90, 0.98]
Rao 2013	12	3	22	57	0.35 [0.20, 0.54]	0.95 [0.86, 0.99]
Rao 2013	26	12	8	48	0.76 [0.59, 0.89]	0.80 [0.68, 0.89]
Rolle 2011	57	3	69	49	0.45 [0.36, 0.54]	0.94 [0.84, 0.99]
Rolle 2011	77	4	49	48	0.61 [0.52, 0.70]	0.92 [0.81, 0.98]
Seong 2010	83	3	19	62	0.81 [0.72, 0.88]	0.95 [0.87, 0.99]
Seong 2010	89	13	13	52	0.87 [0.79, 0.93]	0.80 [0.68, 0.89]
Seong 2010	86	6	16	59	0.84 [0.76, 0.91]	0.91 [0.81, 0.97]



Test 24. OCT: GCC RTVue FLV

OCT: GCC RTVue FLV

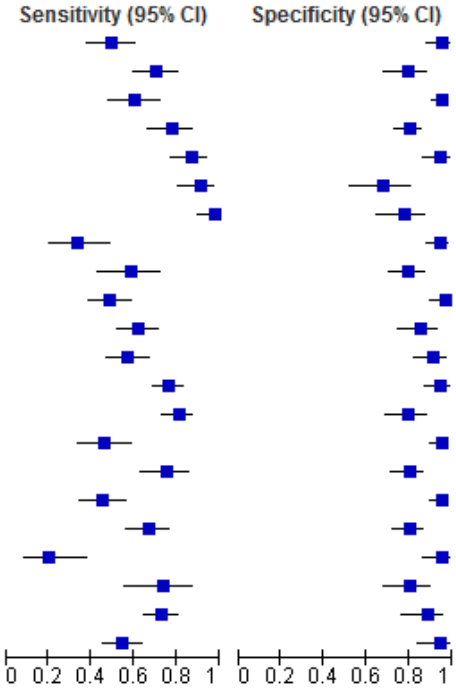
Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Arintawati 2013	49	14	32	54	0.60 [0.49, 0.71]	0.79 [0.68, 0.88]
Arintawati 2013	36	3	45	65	0.44 [0.33, 0.56]	0.96 [0.88, 0.99]
Begum 2014b	61	33	7	132	0.90 [0.80, 0.96]	0.80 [0.73, 0.86]
Begum 2014b	39	8	29	157	0.57 [0.45, 0.69]	0.95 [0.91, 0.98]
Bertuzzi 2014	64	4	6	67	0.91 [0.82, 0.97]	0.94 [0.86, 0.98]
Garas 2011	103	10	8	83	0.93 [0.86, 0.97]	0.89 [0.81, 0.95]
Kim 2011	48	12	8	37	0.86 [0.74, 0.94]	0.76 [0.61, 0.87]
Kim 2013a	45	11	7	47	0.87 [0.74, 0.94]	0.81 [0.69, 0.90]
Lisboa 2013	7	5	41	89	0.15 [0.06, 0.28]	0.95 [0.88, 0.98]
Lisboa 2013	21	19	27	75	0.44 [0.29, 0.59]	0.80 [0.70, 0.87]
Na 2013a	54	13	51	55	0.51 [0.41, 0.61]	0.81 [0.70, 0.89]
Na 2013a	45	3	60	65	0.43 [0.33, 0.53]	0.96 [0.88, 0.99]
Na 2013a	46	5	59	63	0.44 [0.34, 0.54]	0.93 [0.84, 0.98]
Rao 2010b	109	4	31	70	0.78 [0.70, 0.84]	0.95 [0.87, 0.99]
Rao 2010b	110	15	30	59	0.79 [0.71, 0.85]	0.80 [0.69, 0.88]
Rao 2012a	35	6	30	113	0.54 [0.41, 0.66]	0.95 [0.89, 0.98]
Rao 2012a	55	24	10	95	0.85 [0.74, 0.92]	0.80 [0.71, 0.87]
Rao 2012b	48	6	43	119	0.53 [0.42, 0.63]	0.95 [0.90, 0.98]
Rao 2012b	71	25	20	100	0.78 [0.68, 0.86]	0.80 [0.72, 0.87]
Rao 2013	31	12	3	48	0.91 [0.76, 0.98]	0.80 [0.68, 0.89]
Rao 2013	17	3	17	57	0.50 [0.32, 0.68]	0.95 [0.86, 0.99]
Rolle 2011	69	3	57	49	0.55 [0.46, 0.64]	0.94 [0.84, 0.99]
Rolle 2011	74	4	52	48	0.59 [0.50, 0.67]	0.92 [0.81, 0.98]



Test 25. OCT: GCC RTVue GLV

OCT: GCC RTVue GLV

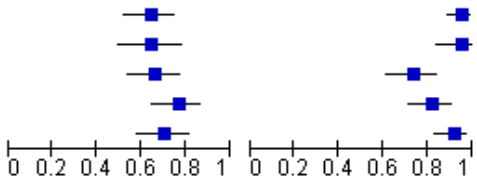
Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Arintawati 2013	40	3	41	65	0.49 [0.38, 0.61]	0.96 [0.88, 0.99]
Arintawati 2013	57	14	24	54	0.70 [0.59, 0.80]	0.79 [0.68, 0.88]
Begum 2014b	41	8	27	157	0.60 [0.48, 0.72]	0.95 [0.91, 0.98]
Begum 2014b	53	33	15	132	0.78 [0.66, 0.87]	0.80 [0.73, 0.86]
Bertuzzi 2014	61	4	9	67	0.87 [0.77, 0.94]	0.94 [0.86, 0.98]
Kim 2011	51	16	5	33	0.91 [0.80, 0.97]	0.67 [0.52, 0.80]
Kim 2013a	51	13	1	45	0.98 [0.90, 1.00]	0.78 [0.65, 0.87]
Lisboa 2013	16	5	32	89	0.33 [0.20, 0.48]	0.95 [0.88, 0.98]
Lisboa 2013	28	19	20	75	0.58 [0.43, 0.72]	0.80 [0.70, 0.87]
Na 2013a	51	2	54	66	0.49 [0.39, 0.59]	0.97 [0.90, 1.00]
Na 2013a	65	10	40	58	0.62 [0.52, 0.71]	0.85 [0.75, 0.93]
Na 2013a	60	6	45	62	0.57 [0.47, 0.67]	0.91 [0.82, 0.97]
Rao 2010b	107	4	33	70	0.76 [0.69, 0.83]	0.95 [0.87, 0.99]
Rao 2010b	113	15	27	59	0.81 [0.73, 0.87]	0.80 [0.69, 0.88]
Rao 2012a	30	6	35	113	0.46 [0.34, 0.59]	0.95 [0.89, 0.98]
Rao 2012a	49	24	16	95	0.75 [0.63, 0.85]	0.80 [0.71, 0.87]
Rao 2012b	41	6	50	119	0.45 [0.35, 0.56]	0.95 [0.90, 0.98]
Rao 2012b	61	25	30	100	0.67 [0.56, 0.77]	0.80 [0.72, 0.87]
Rao 2013	7	3	27	57	0.21 [0.09, 0.38]	0.95 [0.86, 0.99]
Rao 2013	25	12	9	48	0.74 [0.56, 0.87]	0.80 [0.68, 0.89]
Rolle 2011	92	6	34	46	0.73 [0.64, 0.81]	0.88 [0.77, 0.96]
Rolle 2011	69	3	57	49	0.55 [0.46, 0.64]	0.94 [0.84, 0.99]



Test 26. OCT: GCC 3DTopcon average thickness

OCT: GCC 3DTopcon average thickness

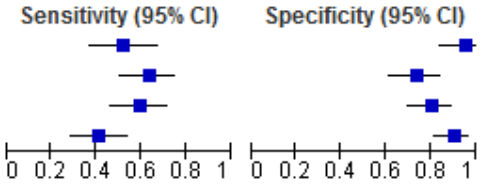
Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Akashi 2013	48	4	27	83	0.64 [0.52, 0.75]	0.95 [0.89, 0.99]
Kim 2013b	31	2	17	40	0.65 [0.49, 0.78]	0.95 [0.84, 0.99]
Kim 2014b	45	17	23	47	0.66 [0.54, 0.77]	0.73 [0.61, 0.84]
Shin 2013	49	13	15	59	0.77 [0.64, 0.86]	0.82 [0.71, 0.90]
Shin 2013	45	6	19	66	0.70 [0.58, 0.81]	0.92 [0.83, 0.97]



Test 27. OCT: GCC 3DTopcon superior thickness

OCT: GCC 3DTopcon superior thickness

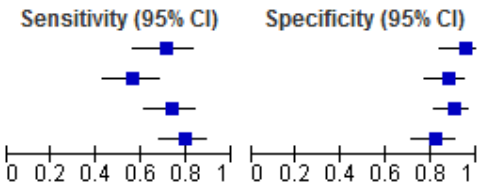
Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Kim 2013b	25	2	23	40	0.52 [0.37, 0.67]	0.95 [0.84, 0.99]
Kim 2014b	43	17	25	47	0.63 [0.51, 0.75]	0.73 [0.61, 0.84]
Shin 2013	38	14	26	58	0.59 [0.46, 0.71]	0.81 [0.70, 0.89]
Shin 2013	26	7	38	65	0.41 [0.29, 0.54]	0.90 [0.81, 0.96]



Test 28. OCT: GCC 3DTopcon inferior thickness

OCT: GCC 3DTopcon inferior thickness

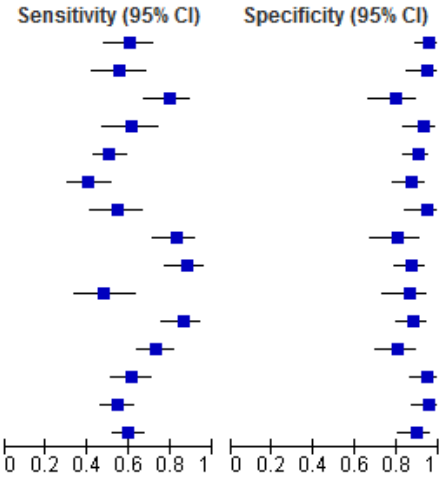
Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Kim 2013b	34	2	14	40	0.71 [0.56, 0.83]	0.95 [0.84, 0.99]
Kim 2014b	38	8	30	56	0.56 [0.43, 0.68]	0.88 [0.77, 0.94]
Shin 2013	47	7	17	65	0.73 [0.61, 0.84]	0.90 [0.81, 0.96]
Shin 2013	51	13	13	59	0.80 [0.68, 0.89]	0.82 [0.71, 0.90]



Test 29. OCT: GCIPL Cirrus average thickness

OCT: GCIPL Cirrus average thickness

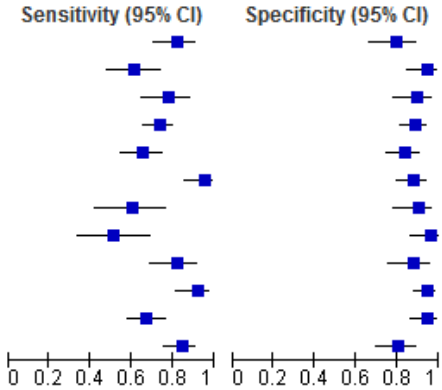
Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Akashi 2013	45	4	30	83	0.60 [0.48, 0.71]	0.95 [0.89, 0.99]
Begum 2014a	34	3	28	50	0.55 [0.42, 0.68]	0.94 [0.84, 0.99]
Begum 2014a	49	11	13	42	0.79 [0.67, 0.88]	0.79 [0.66, 0.89]
Choi 2013	33	4	21	52	0.61 [0.47, 0.74]	0.93 [0.83, 0.98]
Jeoung 2013	83	12	81	107	0.51 [0.43, 0.58]	0.90 [0.83, 0.95]
Kim 2014a	37	12	55	80	0.40 [0.30, 0.51]	0.87 [0.78, 0.93]
Kotowski 2012	34	3	29	48	0.54 [0.41, 0.67]	0.94 [0.84, 0.99]
Kotowski 2012	52	10	11	41	0.83 [0.71, 0.91]	0.80 [0.67, 0.90]
Mwanza 2012	51	13	7	86	0.88 [0.77, 0.95]	0.87 [0.79, 0.93]
Mwanza 2014	24	7	26	42	0.48 [0.34, 0.63]	0.86 [0.73, 0.94]
Nouri-Mahdavi 2013	51	11	8	80	0.86 [0.75, 0.94]	0.88 [0.79, 0.94]
Sung 2013	78	14	29	58	0.73 [0.63, 0.81]	0.81 [0.70, 0.89]
Sung 2013	65	4	42	68	0.61 [0.51, 0.70]	0.94 [0.86, 0.98]
Zhang 2014	86	4	73	73	0.54 [0.46, 0.62]	0.95 [0.87, 0.99]
Zhang 2014	95	8	64	69	0.60 [0.52, 0.67]	0.90 [0.81, 0.95]



Test 30. OCT: GCIPL Cirrus minimum thickness

OCT: GCIPL Cirrus minimum thickness

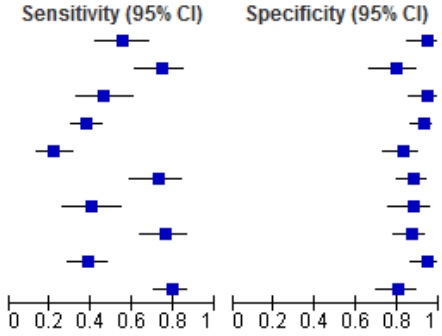
Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Begum 2014a	51	11	11	42	0.82 [0.70, 0.91]	0.79 [0.66, 0.89]
Begum 2014a	38	3	24	50	0.61 [0.48, 0.73]	0.94 [0.84, 0.99]
Choi 2013	42	6	12	50	0.78 [0.64, 0.88]	0.89 [0.78, 0.96]
Jeoung 2013	120	14	44	105	0.73 [0.66, 0.80]	0.88 [0.81, 0.93]
Kim 2014a	60	15	32	77	0.65 [0.55, 0.75]	0.84 [0.75, 0.91]
Mwanza 2012	55	12	3	87	0.95 [0.86, 0.99]	0.88 [0.80, 0.94]
Mwanza 2013	21	5	14	44	0.60 [0.42, 0.76]	0.90 [0.78, 0.97]
Mwanza 2013	18	2	17	47	0.51 [0.34, 0.69]	0.96 [0.86, 1.00]
Mwanza 2014	41	6	9	43	0.82 [0.69, 0.91]	0.88 [0.75, 0.95]
Nouri-Mahdavi 2013	54	5	5	86	0.92 [0.81, 0.97]	0.95 [0.88, 0.98]
Sung 2013	72	4	35	68	0.67 [0.58, 0.76]	0.94 [0.86, 0.98]
Sung 2013	90	14	17	58	0.84 [0.76, 0.90]	0.81 [0.70, 0.89]



Test 31. OCT: GCIPL Cirrus superior thickness

OCT: GCIPL Cirrus superior thickness

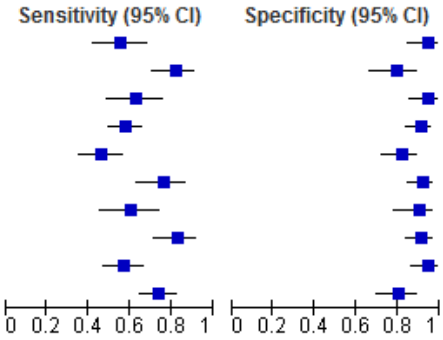
Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Begum 2014a	34	3	28	50	0.55 [0.42, 0.68]	0.94 [0.84, 0.99]
Begum 2014a	46	11	16	42	0.74 [0.62, 0.84]	0.79 [0.66, 0.89]
Choi 2013	25	3	29	53	0.46 [0.33, 0.60]	0.95 [0.85, 0.99]
Jeoung 2013	62	9	102	110	0.38 [0.30, 0.46]	0.92 [0.86, 0.96]
Kim 2014a	20	16	72	76	0.22 [0.14, 0.32]	0.83 [0.73, 0.90]
Mwanza 2012	42	12	16	87	0.72 [0.59, 0.83]	0.88 [0.80, 0.94]
Mwanza 2014	20	6	30	43	0.40 [0.26, 0.55]	0.88 [0.75, 0.95]
Nouri-Mahdavi 2013	45	12	14	79	0.76 [0.63, 0.86]	0.87 [0.78, 0.93]
Sung 2013	41	4	66	68	0.38 [0.29, 0.48]	0.94 [0.86, 0.98]
Sung 2013	85	14	22	58	0.79 [0.71, 0.87]	0.81 [0.70, 0.89]



Test 32. OCT: GCIPL Cirrus inferior thickness

OCT: GCIPL Cirrus inferior thickness

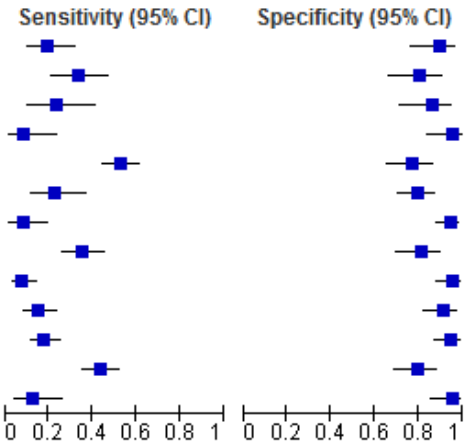
Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Begum 2014a	34	3	28	50	0.55 [0.42, 0.68]	0.94 [0.84, 0.99]
Begum 2014a	51	11	11	42	0.82 [0.70, 0.91]	0.79 [0.66, 0.89]
Choi 2013	34	3	20	53	0.63 [0.49, 0.76]	0.95 [0.85, 0.99]
Jeoung 2013	95	11	69	108	0.58 [0.50, 0.66]	0.91 [0.84, 0.95]
Kim 2014a	42	17	50	75	0.46 [0.35, 0.56]	0.82 [0.72, 0.89]
Mwanza 2012	44	8	14	91	0.76 [0.63, 0.86]	0.92 [0.85, 0.96]
Mwanza 2014	30	5	20	44	0.60 [0.45, 0.74]	0.90 [0.78, 0.97]
Nouri-Mahdavi 2013	49	8	10	83	0.83 [0.71, 0.92]	0.91 [0.83, 0.96]
Sung 2013	61	4	46	68	0.57 [0.47, 0.67]	0.94 [0.86, 0.98]
Sung 2013	79	14	28	58	0.74 [0.64, 0.82]	0.81 [0.70, 0.89]



Test 33. OCT: ONH Disc area

OCT: ONH Disc area

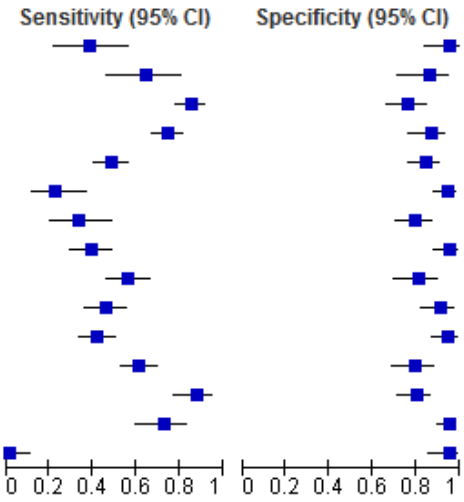
Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Barella 2013	11	5	46	41	0.19 [0.10, 0.32]	0.89 [0.76, 0.96]
Barella 2013	19	9	38	37	0.33 [0.21, 0.47]	0.80 [0.66, 0.91]
Fang 2010	8	6	26	36	0.24 [0.11, 0.41]	0.86 [0.71, 0.95]
Fang 2010	3	2	31	40	0.09 [0.02, 0.24]	0.95 [0.84, 0.99]
Huang 2011	77	17	69	57	0.53 [0.44, 0.61]	0.77 [0.66, 0.86]
Lisboa 2013	11	19	37	75	0.23 [0.12, 0.37]	0.80 [0.70, 0.87]
Lisboa 2013	4	5	44	89	0.08 [0.02, 0.20]	0.95 [0.88, 0.98]
Na 2013a	37	13	68	55	0.35 [0.26, 0.45]	0.81 [0.70, 0.89]
Na 2013a	8	3	97	65	0.08 [0.03, 0.14]	0.96 [0.88, 0.99]
Na 2013a	16	6	89	62	0.15 [0.09, 0.24]	0.91 [0.82, 0.97]
Rao 2010b	25	4	115	70	0.18 [0.12, 0.25]	0.95 [0.87, 0.99]
Rao 2010b	61	15	79	59	0.44 [0.35, 0.52]	0.80 [0.69, 0.88]
Roberti 2014	6	3	40	55	0.13 [0.05, 0.26]	0.95 [0.86, 0.99]



Test 34. OCT: ONH Cup area

OCT: ONH Cup area

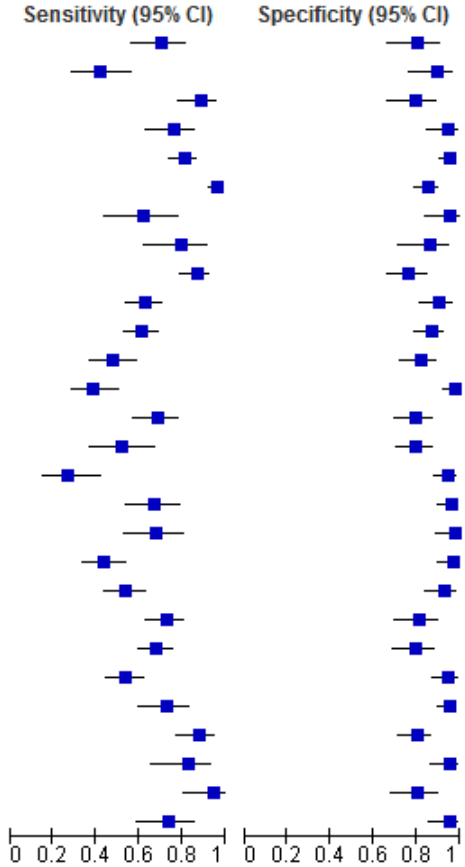
Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Fang 2010	13	2	21	40	0.38 [0.22, 0.56]	0.95 [0.84, 0.99]
Fang 2010	22	6	12	36	0.65 [0.46, 0.80]	0.86 [0.71, 0.95]
Garas 2011	95	22	16	71	0.86 [0.78, 0.92]	0.76 [0.66, 0.85]
Huang 2011	109	10	37	64	0.75 [0.67, 0.81]	0.86 [0.77, 0.93]
Jeoung 2013	79	19	85	100	0.48 [0.40, 0.56]	0.84 [0.76, 0.90]
Lisboa 2013	11	5	37	89	0.23 [0.12, 0.37]	0.95 [0.88, 0.98]
Lisboa 2013	16	19	32	75	0.33 [0.20, 0.48]	0.80 [0.70, 0.87]
Na 2013a	41	3	64	65	0.39 [0.30, 0.49]	0.96 [0.88, 0.99]
Na 2013a	59	13	46	55	0.56 [0.46, 0.66]	0.81 [0.70, 0.89]
Na 2013a	48	6	57	62	0.46 [0.36, 0.56]	0.91 [0.82, 0.97]
Rao 2010b	59	4	81	70	0.42 [0.34, 0.51]	0.95 [0.87, 0.99]
Rao 2010b	86	15	54	59	0.61 [0.53, 0.70]	0.80 [0.69, 0.88]
Rao 2012a	57	24	8	95	0.88 [0.77, 0.95]	0.80 [0.71, 0.87]
Rao 2012a	47	6	18	113	0.72 [0.60, 0.83]	0.95 [0.89, 0.98]
Roberti 2014	1	3	45	55	0.02 [0.00, 0.12]	0.95 [0.86, 0.99]



Test 35. OCT: ONH Rim area

OCT: ONH Rim area

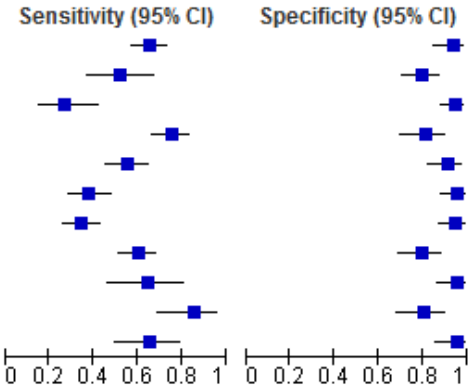
Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Barella 2013	40	9	17	37	0.70 [0.57, 0.82]	0.80 [0.66, 0.91]
Barella 2013	24	5	33	41	0.42 [0.29, 0.56]	0.89 [0.76, 0.96]
Begum 2014a	55	11	7	42	0.89 [0.78, 0.95]	0.79 [0.66, 0.89]
Begum 2014a	47	3	15	50	0.76 [0.63, 0.86]	0.94 [0.84, 0.99]
Calvo 2014	126	9	30	173	0.81 [0.74, 0.87]	0.95 [0.91, 0.98]
Calvo 2014	150	27	6	155	0.96 [0.92, 0.99]	0.85 [0.79, 0.90]
Fang 2010	21	2	13	40	0.62 [0.44, 0.78]	0.95 [0.84, 0.99]
Fang 2010	27	6	7	36	0.79 [0.62, 0.91]	0.86 [0.71, 0.95]
Garas 2011	96	22	15	71	0.86 [0.79, 0.92]	0.76 [0.66, 0.85]
Huang 2011	91	7	55	67	0.62 [0.54, 0.70]	0.91 [0.81, 0.96]
Jeoung 2013	100	16	64	103	0.61 [0.53, 0.68]	0.87 [0.79, 0.92]
Kim 2014a	44	17	48	75	0.48 [0.37, 0.58]	0.82 [0.72, 0.89]
Kratz 2014	33	2	52	86	0.39 [0.28, 0.50]	0.98 [0.92, 1.00]
Kratz 2014	58	18	27	70	0.68 [0.57, 0.78]	0.80 [0.70, 0.87]
Lisboa 2013	25	19	23	75	0.52 [0.37, 0.67]	0.80 [0.70, 0.87]
Lisboa 2013	13	5	35	89	0.27 [0.15, 0.42]	0.95 [0.88, 0.98]
Mwanza 2012	39	4	19	95	0.67 [0.54, 0.79]	0.96 [0.90, 0.99]
Mwanza 2014	34	1	16	48	0.68 [0.53, 0.80]	0.98 [0.89, 1.00]
Na 2013a	46	2	59	66	0.44 [0.34, 0.54]	0.97 [0.90, 1.00]
Na 2013a	56	5	49	63	0.53 [0.43, 0.63]	0.93 [0.84, 0.98]
Na 2013a	76	13	29	55	0.72 [0.63, 0.81]	0.81 [0.70, 0.89]
Rao 2010b	95	15	45	59	0.68 [0.59, 0.75]	0.80 [0.69, 0.88]
Rao 2010b	75	4	65	70	0.54 [0.45, 0.62]	0.95 [0.87, 0.99]
Rao 2012a	47	6	18	113	0.72 [0.60, 0.83]	0.95 [0.89, 0.98]
Rao 2012a	57	24	8	95	0.88 [0.77, 0.95]	0.80 [0.71, 0.87]
Rao 2013	28	3	6	57	0.82 [0.65, 0.93]	0.95 [0.86, 0.99]
Rao 2013	32	12	2	48	0.94 [0.80, 0.99]	0.80 [0.68, 0.89]
Roberti 2014	34	3	12	55	0.74 [0.59, 0.86]	0.95 [0.86, 0.99]



Test 36. OCT: ONH Rim volume

OCT: ONH Rim volume

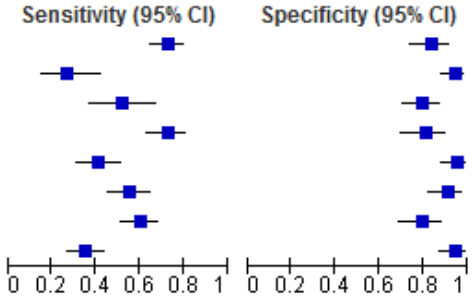
Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Huang 2011	95	5	51	69	0.65 [0.57, 0.73]	0.93 [0.85, 0.98]
Lisboa 2013	25	19	23	75	0.52 [0.37, 0.67]	0.80 [0.70, 0.87]
Lisboa 2013	13	5	35	89	0.27 [0.15, 0.42]	0.95 [0.88, 0.98]
Na 2013a	79	13	26	55	0.75 [0.66, 0.83]	0.81 [0.70, 0.89]
Na 2013a	58	6	47	62	0.55 [0.45, 0.65]	0.91 [0.82, 0.97]
Na 2013a	40	3	65	65	0.38 [0.29, 0.48]	0.96 [0.88, 0.99]
Rao 2010b	48	4	92	70	0.34 [0.26, 0.43]	0.95 [0.87, 0.99]
Rao 2010b	84	15	56	59	0.60 [0.51, 0.68]	0.80 [0.69, 0.88]
Rao 2013	22	3	12	57	0.65 [0.46, 0.80]	0.95 [0.86, 0.99]
Rao 2013	29	12	5	48	0.85 [0.69, 0.95]	0.80 [0.68, 0.89]
Roberti 2014	30	3	16	55	0.65 [0.50, 0.79]	0.95 [0.86, 0.99]



Test 37. OCT: ONH Nerve head volume

OCT: ONH Nerve head volume

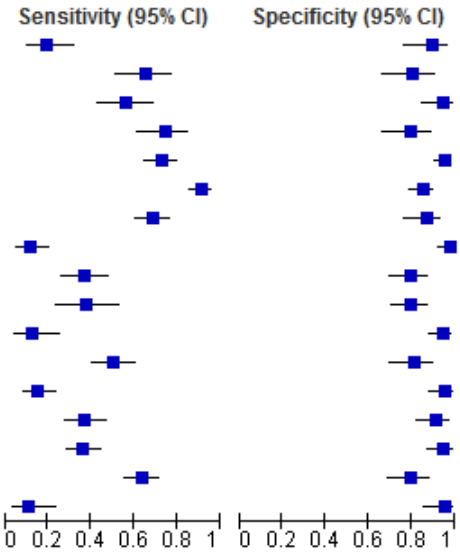
Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Huang 2011	106	12	40	62	0.73 [0.65, 0.80]	0.84 [0.73, 0.91]
Lisboa 2013	13	5	35	89	0.27 [0.15, 0.42]	0.95 [0.88, 0.98]
Lisboa 2013	25	19	23	75	0.52 [0.37, 0.67]	0.80 [0.70, 0.87]
Na 2013a	76	13	29	55	0.72 [0.63, 0.81]	0.81 [0.70, 0.89]
Na 2013a	43	3	62	65	0.41 [0.31, 0.51]	0.96 [0.88, 0.99]
Na 2013a	58	6	47	62	0.55 [0.45, 0.65]	0.91 [0.82, 0.97]
Rao 2010b	84	15	56	59	0.60 [0.51, 0.68]	0.80 [0.69, 0.88]
Rao 2010b	49	4	91	70	0.35 [0.27, 0.44]	0.95 [0.87, 0.99]



Test 38. OCT: ONH Cup volume

OCT: ONH Cup volume

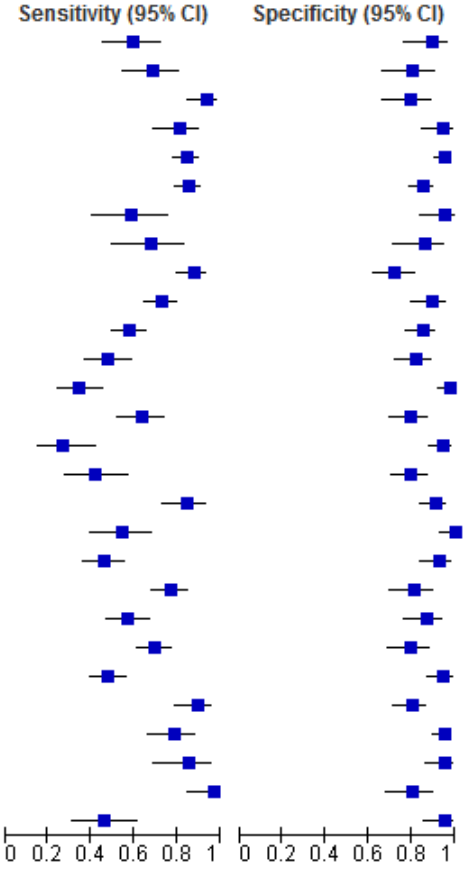
Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Barella 2013	11	5	46	41	0.19 [0.10, 0.32]	0.89 [0.76, 0.96]
Barella 2013	37	9	20	37	0.65 [0.51, 0.77]	0.80 [0.66, 0.91]
Begum 2014a	35	3	27	50	0.56 [0.43, 0.69]	0.94 [0.84, 0.99]
Begum 2014a	46	11	16	42	0.74 [0.62, 0.84]	0.79 [0.66, 0.89]
Calvo 2014	113	9	43	173	0.72 [0.65, 0.79]	0.95 [0.91, 0.98]
Calvo 2014	142	27	14	155	0.91 [0.85, 0.95]	0.85 [0.79, 0.90]
Huang 2011	100	10	46	64	0.68 [0.60, 0.76]	0.86 [0.77, 0.93]
Kratz 2014	10	2	75	86	0.12 [0.06, 0.21]	0.98 [0.92, 1.00]
Kratz 2014	31	18	54	70	0.36 [0.26, 0.48]	0.80 [0.70, 0.87]
Lisboa 2013	18	19	30	75	0.38 [0.24, 0.53]	0.80 [0.70, 0.87]
Lisboa 2013	6	5	42	89	0.13 [0.05, 0.25]	0.95 [0.88, 0.98]
Na 2013a	53	13	52	55	0.50 [0.41, 0.60]	0.81 [0.70, 0.89]
Na 2013a	16	3	89	65	0.15 [0.09, 0.24]	0.96 [0.88, 0.99]
Na 2013a	39	6	66	62	0.37 [0.28, 0.47]	0.91 [0.82, 0.97]
Rao 2010b	51	4	89	70	0.36 [0.28, 0.45]	0.95 [0.87, 0.99]
Rao 2010b	89	15	51	59	0.64 [0.55, 0.72]	0.80 [0.69, 0.88]
Roberti 2014	5	3	41	55	0.11 [0.04, 0.24]	0.95 [0.86, 0.99]



Test 39. OCT: ONH C/D area ratio

OCT: ONH C/D area ratio

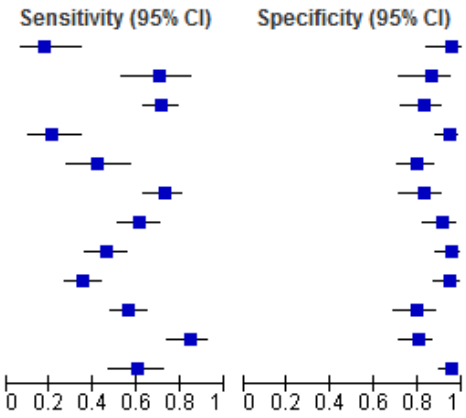
Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Barella 2013	34	5	23	41	0.60 [0.46, 0.72]	0.89 [0.76, 0.96]
Barella 2013	39	9	18	37	0.68 [0.55, 0.80]	0.80 [0.66, 0.91]
Begum 2014a	58	11	4	42	0.94 [0.84, 0.98]	0.79 [0.66, 0.89]
Begum 2014a	50	3	12	50	0.81 [0.69, 0.90]	0.94 [0.84, 0.99]
Calvo 2014	132	9	24	173	0.85 [0.78, 0.90]	0.95 [0.91, 0.98]
Calvo 2014	133	27	23	155	0.85 [0.79, 0.90]	0.85 [0.79, 0.90]
Fang 2010	20	2	14	40	0.59 [0.41, 0.75]	0.95 [0.84, 0.99]
Fang 2010	23	6	11	36	0.68 [0.49, 0.83]	0.86 [0.71, 0.95]
Garas 2011	97	26	14	67	0.87 [0.80, 0.93]	0.72 [0.62, 0.81]
Huang 2011	106	8	40	66	0.73 [0.65, 0.80]	0.89 [0.80, 0.95]
Jeoung 2013	95	18	69	101	0.58 [0.50, 0.66]	0.85 [0.77, 0.91]
Kim 2014a	44	17	48	75	0.48 [0.37, 0.58]	0.82 [0.72, 0.89]
Kratz 2014	29	2	56	86	0.34 [0.24, 0.45]	0.98 [0.92, 1.00]
Kratz 2014	54	18	31	70	0.64 [0.52, 0.74]	0.80 [0.70, 0.87]
Lisboa 2013	13	5	35	89	0.27 [0.15, 0.42]	0.95 [0.88, 0.98]
Lisboa 2013	20	19	28	75	0.42 [0.28, 0.57]	0.80 [0.70, 0.87]
Mwanza 2012	49	9	9	90	0.84 [0.73, 0.93]	0.91 [0.83, 0.96]
Mwanza 2014	27	0	23	49	0.54 [0.39, 0.68]	1.00 [0.93, 1.00]
Na 2013a	48	5	57	63	0.46 [0.36, 0.56]	0.93 [0.84, 0.98]
Na 2013a	81	13	24	55	0.77 [0.68, 0.85]	0.81 [0.70, 0.89]
Na 2013a	60	9	45	59	0.57 [0.47, 0.67]	0.87 [0.76, 0.94]
Rao 2010b	97	15	43	59	0.69 [0.61, 0.77]	0.80 [0.69, 0.88]
Rao 2010b	67	4	73	70	0.48 [0.39, 0.56]	0.95 [0.87, 0.99]
Rao 2012a	58	24	7	95	0.89 [0.79, 0.96]	0.80 [0.71, 0.87]
Rao 2012a	51	6	14	113	0.78 [0.67, 0.88]	0.95 [0.89, 0.98]
Rao 2013	29	3	5	57	0.85 [0.69, 0.95]	0.95 [0.86, 0.99]
Rao 2013	33	12	1	48	0.97 [0.85, 1.00]	0.80 [0.68, 0.89]
Roberti 2014	21	3	25	55	0.46 [0.31, 0.61]	0.95 [0.86, 0.99]



Test 40. OCT: ONH horizontal C/D ratio

OCT: ONH horizontal C/D ratio

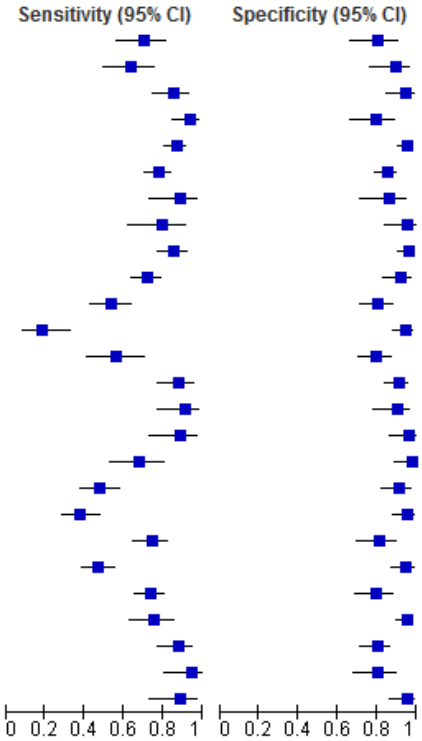
Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Fang 2010	6	2	28	40	0.18 [0.07, 0.35]	0.95 [0.84, 0.99]
Fang 2010	24	6	10	36	0.71 [0.53, 0.85]	0.86 [0.71, 0.95]
Huang 2011	104	13	42	61	0.71 [0.63, 0.78]	0.82 [0.72, 0.90]
Lisboa 2013	10	5	38	89	0.21 [0.10, 0.35]	0.95 [0.88, 0.98]
Lisboa 2013	20	19	28	75	0.42 [0.28, 0.57]	0.80 [0.70, 0.87]
Na 2013a	76	12	29	56	0.72 [0.63, 0.81]	0.82 [0.71, 0.91]
Na 2013a	64	6	41	62	0.61 [0.51, 0.70]	0.91 [0.82, 0.97]
Na 2013a	48	3	57	65	0.46 [0.36, 0.56]	0.96 [0.88, 0.99]
Rao 2010b	49	4	91	70	0.35 [0.27, 0.44]	0.95 [0.87, 0.99]
Rao 2010b	79	15	61	59	0.56 [0.48, 0.65]	0.80 [0.69, 0.88]
Rao 2012a	55	24	10	95	0.85 [0.74, 0.92]	0.80 [0.71, 0.87]
Rao 2012a	39	6	26	113	0.60 [0.47, 0.72]	0.95 [0.89, 0.98]



Test 41. OCT: ONH vertical C/D ratio

OCT: ONH vertical C/D ratio

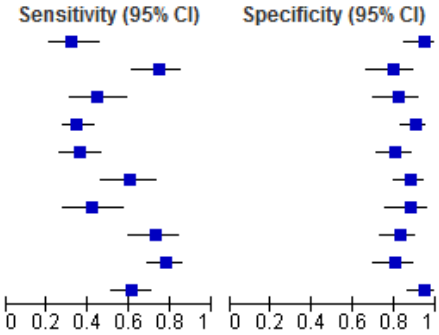
Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Barella 2013	40	9	17	37	0.70 [0.57, 0.82]	0.80 [0.66, 0.91]
Barella 2013	36	5	21	41	0.63 [0.49, 0.76]	0.89 [0.76, 0.96]
Begum 2014a	53	3	9	50	0.85 [0.74, 0.93]	0.94 [0.84, 0.99]
Begum 2014a	58	11	4	42	0.94 [0.84, 0.98]	0.79 [0.66, 0.89]
Calvo 2014	135	9	21	173	0.87 [0.80, 0.91]	0.95 [0.91, 0.98]
Calvo 2014	121	27	35	155	0.78 [0.70, 0.84]	0.85 [0.79, 0.90]
Fang 2010	30	6	4	36	0.88 [0.73, 0.97]	0.86 [0.71, 0.95]
Fang 2010	27	2	7	40	0.79 [0.62, 0.91]	0.95 [0.84, 0.99]
Gonzales de la Rosa 2013	89	4	15	98	0.86 [0.77, 0.92]	0.96 [0.90, 0.99]
Huang 2011	105	6	41	68	0.72 [0.64, 0.79]	0.92 [0.83, 0.97]
Kim 2014a	49	18	43	74	0.53 [0.43, 0.64]	0.80 [0.71, 0.88]
Lisboa 2013	9	5	39	89	0.19 [0.09, 0.33]	0.95 [0.88, 0.98]
Lisboa 2013	27	19	21	75	0.56 [0.41, 0.71]	0.80 [0.70, 0.87]
Mwanza 2012	51	9	7	90	0.88 [0.77, 0.95]	0.91 [0.83, 0.96]
Mwanza 2013	32	5	3	44	0.91 [0.77, 0.98]	0.90 [0.78, 0.97]
Mwanza 2013	31	2	4	47	0.89 [0.73, 0.97]	0.96 [0.86, 1.00]
Mwanza 2014	34	1	16	48	0.68 [0.53, 0.80]	0.98 [0.89, 1.00]
Na 2013a	50	6	55	62	0.48 [0.38, 0.58]	0.91 [0.82, 0.97]
Na 2013a	40	3	65	65	0.38 [0.29, 0.48]	0.96 [0.88, 0.99]
Na 2013a	78	13	27	55	0.74 [0.65, 0.82]	0.81 [0.70, 0.89]
Rao 2010b	66	4	74	70	0.47 [0.39, 0.56]	0.95 [0.87, 0.99]
Rao 2010b	103	15	37	59	0.74 [0.65, 0.81]	0.80 [0.69, 0.88]
Rao 2012a	49	6	16	113	0.75 [0.63, 0.85]	0.95 [0.89, 0.98]
Rao 2012a	57	24	8	95	0.88 [0.77, 0.95]	0.80 [0.71, 0.87]
Rao 2013	32	12	2	48	0.94 [0.80, 0.99]	0.80 [0.68, 0.89]
Rao 2013	30	3	4	57	0.88 [0.73, 0.97]	0.95 [0.86, 0.99]



Test 42. OCT: GCIPL Cirrus Inferonasal quadrant

OCT: GCIPL Cirrus Inferonasal quadrant

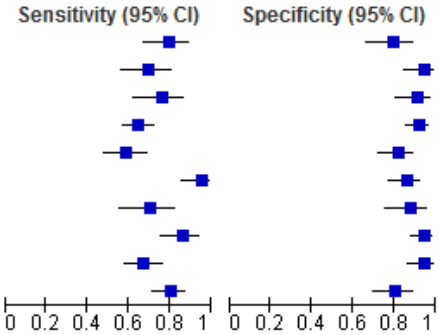
Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Begum 2014a	20	3	42	50	0.32 [0.21, 0.45]	0.94 [0.84, 0.99]
Begum 2014a	46	11	16	42	0.74 [0.62, 0.84]	0.79 [0.66, 0.89]
Choi 2013	24	10	30	46	0.44 [0.31, 0.59]	0.82 [0.70, 0.91]
Jeoung 2013	57	12	107	107	0.35 [0.28, 0.43]	0.90 [0.83, 0.95]
Kim 2014a	33	18	59	74	0.36 [0.26, 0.47]	0.80 [0.71, 0.88]
Mwanza 2012	35	12	23	87	0.60 [0.47, 0.73]	0.88 [0.80, 0.94]
Mwanza 2014	21	6	29	43	0.42 [0.28, 0.57]	0.88 [0.75, 0.95]
Nouri-Mahdavi 2013	43	16	16	75	0.73 [0.60, 0.84]	0.82 [0.73, 0.90]
Sung 2013	83	14	24	58	0.78 [0.68, 0.85]	0.81 [0.70, 0.89]
Sung 2013	65	4	42	68	0.61 [0.51, 0.70]	0.94 [0.86, 0.98]



Test 43. OCT: GCIPL Cirrus Inferotemporal quadrant

OCT: GCIPL Cirrus Inferotemporal quadrant

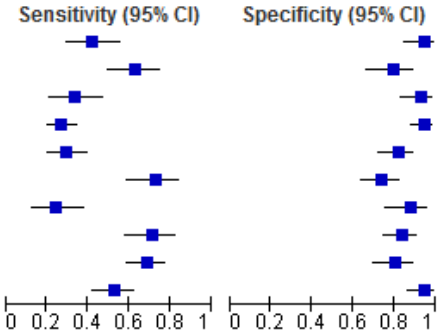
Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Begum 2014a	49	11	13	42	0.79 [0.67, 0.88]	0.79 [0.66, 0.89]
Begum 2014a	43	3	19	50	0.69 [0.56, 0.80]	0.94 [0.84, 0.99]
Choi 2013	41	5	13	51	0.76 [0.62, 0.87]	0.91 [0.80, 0.97]
Jeoung 2013	106	10	58	109	0.65 [0.57, 0.72]	0.92 [0.85, 0.96]
Kim 2014a	54	17	38	75	0.59 [0.48, 0.69]	0.82 [0.72, 0.89]
Mwanza 2012	55	14	3	85	0.95 [0.86, 0.99]	0.86 [0.77, 0.92]
Mwanza 2014	35	6	15	43	0.70 [0.55, 0.82]	0.88 [0.75, 0.95]
Nouri-Mahdavi 2013	51	5	8	86	0.86 [0.75, 0.94]	0.95 [0.88, 0.98]
Sung 2013	72	4	35	68	0.67 [0.58, 0.76]	0.94 [0.86, 0.98]
Sung 2013	86	14	21	58	0.80 [0.72, 0.87]	0.81 [0.70, 0.89]



Test 44. OCT: GCIPL Cirrus Superonasal quadrant

OCT: GCIPL Cirrus Superonasal quadrant

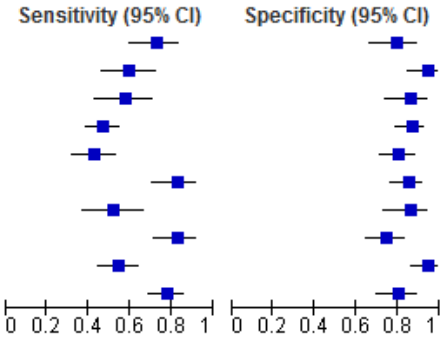
Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Begum 2014a	26	3	36	50	0.42 [0.30, 0.55]	0.94 [0.84, 0.99]
Begum 2014a	39	11	23	42	0.63 [0.50, 0.75]	0.79 [0.66, 0.89]
Choi 2013	18	4	36	52	0.33 [0.21, 0.47]	0.93 [0.83, 0.98]
Jeoung 2013	44	7	120	112	0.27 [0.20, 0.34]	0.94 [0.88, 0.98]
Kim 2014a	27	17	65	75	0.29 [0.20, 0.40]	0.82 [0.72, 0.89]
Mwanza 2012	42	26	16	73	0.72 [0.59, 0.83]	0.74 [0.64, 0.82]
Mwanza 2014	12	6	38	43	0.24 [0.13, 0.38]	0.88 [0.75, 0.95]
Nouri-Mahdavi 2013	42	15	17	76	0.71 [0.58, 0.82]	0.84 [0.74, 0.90]
Sung 2013	73	14	34	58	0.68 [0.59, 0.77]	0.81 [0.70, 0.89]
Sung 2013	56	4	51	68	0.52 [0.42, 0.62]	0.94 [0.86, 0.98]



Test 45. OCT: GCIPL Cirrus Superotemporal quadrant

OCT: GCIPL Cirrus Superotemporal quadrant

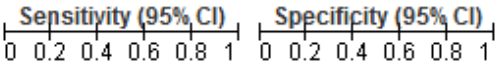
Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Begum 2014a	45	11	17	42	0.73 [0.60, 0.83]	0.79 [0.66, 0.89]
Begum 2014a	37	3	25	50	0.60 [0.46, 0.72]	0.94 [0.84, 0.99]
Choi 2013	31	8	23	48	0.57 [0.43, 0.71]	0.86 [0.74, 0.94]
Jeoung 2013	77	16	87	103	0.47 [0.39, 0.55]	0.87 [0.79, 0.92]
Kim 2014a	39	18	53	74	0.42 [0.32, 0.53]	0.80 [0.71, 0.88]
Mwanza 2012	48	15	10	84	0.83 [0.71, 0.91]	0.85 [0.76, 0.91]
Mwanza 2014	26	7	24	42	0.52 [0.37, 0.66]	0.86 [0.73, 0.94]
Nouri-Mahdavi 2013	49	23	10	68	0.83 [0.71, 0.92]	0.75 [0.65, 0.83]
Sung 2013	58	4	49	68	0.54 [0.44, 0.64]	0.94 [0.86, 0.98]
Sung 2013	83	14	24	58	0.78 [0.68, 0.85]	0.81 [0.70, 0.89]



Test 46. OCT: GCC Spectralis average thickness

OCT: GCC Spectralis average thickness

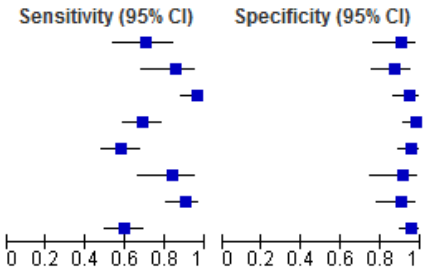
Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)



Test 47. Direct comparison: GDx NFI

Direct comparison: GDx NFI

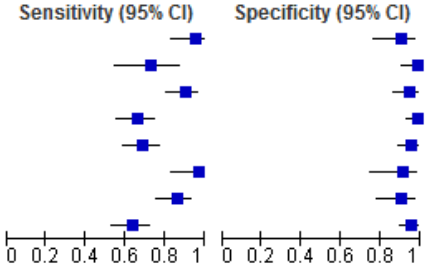
Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Aptel 2010	28	4	12	36	0.70 [0.53, 0.83]	0.90 [0.76, 0.97]
Benitez-del-Castillo 2011	28	7	5	48	0.85 [0.68, 0.95]	0.87 [0.76, 0.95]
Bertuzzi 2014	67	4	3	67	0.96 [0.88, 0.99]	0.94 [0.86, 0.98]
Garas 2012	68	2	31	76	0.69 [0.59, 0.78]	0.97 [0.91, 1.00]
Gonzales de la Rosa 2013	60	5	44	97	0.58 [0.48, 0.67]	0.95 [0.89, 0.98]
Hoesl 2013	26	3	5	29	0.84 [0.66, 0.95]	0.91 [0.75, 0.98]
Oddone 2011	63	5	7	45	0.90 [0.80, 0.96]	0.90 [0.78, 0.97]
Rao 2014	63	5	43	104	0.59 [0.49, 0.69]	0.95 [0.90, 0.98]



Test 48. Direct comparison: OCT RNFL average

Direct comparison: OCT RNFL average

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Aptel 2010	38	4	2	36	0.95 [0.83, 0.99]	0.90 [0.76, 0.97]
Benitez-del-Castillo 2011	24	1	9	54	0.73 [0.54, 0.87]	0.98 [0.90, 1.00]
Bertuzzi 2014	63	4	7	67	0.90 [0.80, 0.96]	0.94 [0.86, 0.98]
Garas 2012	65	1	34	77	0.66 [0.55, 0.75]	0.99 [0.93, 1.00]
Gonzales de la Rosa 2013	71	5	33	97	0.68 [0.58, 0.77]	0.95 [0.89, 0.98]
Hoesl 2013	30	3	1	29	0.97 [0.83, 1.00]	0.91 [0.75, 0.98]
Oddone 2011	60	5	10	45	0.86 [0.75, 0.93]	0.90 [0.78, 0.97]
Rao 2014	67	5	39	104	0.63 [0.53, 0.72]	0.95 [0.90, 0.98]



ADDITIONAL TABLES

Table 1. Accuracy of all parameters for each test

Test (parameter)	**Number of studies** (**Number of patients**)	**Sensitivity**[1]	**Specificity**[1]
GDx			
Inferior sector	30 (4199)	0.61 (0.55 to 0.66)	0.92 (0.90 to 0.94)
Nerve fibre indicator (NFI)	**35 (4958)**	**0.76 (0.70 to 0.81)**	**0.92 (0.90 to 0.94)**
Superior sector	30 (4199)	0.61 (0.56 to 0.66)	0.93 (0.91 to 0.94)
Temporal superior nasal inferior temporal (TSNIT) average	30 (4104)	0.61 (0.57 to 0.66)	0.93 (0.92 to 0.95)
HRT			
Cup disc area ratio	9 (1959)	0.57 (0.46 to 0.68)	0.95 (0.93 to 0.96)
Cup area	7 (1447)	0.43 (0.31 to 0.56)	0.94 (0.92 to 0.96)
Cup shape measure	6 (1343)	0.41 (0.31 to 0.52)	0.94 (0.91 to 0.95)
Cup volume	9 (1959)	0.32 (0.23 to 0.43)	0.95 (0.93 to 0.96)
Frederick S. Mikelberg (FSM) discriminant function	6 (1215)	0.58 (0.36 to 0.77)	0.94 (0.90 to 0.96)
Moorfields regression analysis (MRA)	**8 (1271)**	**0.69 (0.56 to 0.79)**	**0.89 (0.84 to 0.93)**
Reinhard O.W. Burk (RB) discriminant function	6 (1207)	0.53 (0.42 to 0.63)	0.95 (0.93 to 0.96)
Rim volume	6 (1207)	0.53 (0.42 to 0.63)	0.95 (0.93 to 0.96)
Rim area	9 (1958)	0.45 (0.34 to 0.56)	0.95 (0.93 to 0.96)
Vertical cup/disc ratio	8 (1849)	0.67 (0.55 to 0.77)	0.94 (0.92 to 0.95)
OCT ONH			
Cup/disc area ratio	17 (2863)	0.64 (0.54 to 0.73)	0.93 (0.90 to 0.95)
Horizontal cup/disc ratio	6 (1009)	0.41 (0.26 to 0.58)	0.94 (0.90 to 0.96)
Vertical cup/disc ratio	**15 (2389)**	**0.72 (0.60 to 0.81)**	**0.94 (0.92 to 0.95)**
Cup area	9 (1600)	0.45 (0.26 to 0.67)	0.92 (0.87 to 0.95)
Cup volume	9 (1582)	0.30 (0.16 to 0.49)	0.94 (0.92 to 0.96)
Disc area	7 (1032)	0.16 (0.09 to 0.27)	0.93 (0.88 to 0.96)

Table 1. Accuracy of all parameters for each test *(Continued)*

Nerve head volume	4 (749)	0.44 (0.28 to 0.62)	0.93 (0.87 to 0.96)
Rim area	17 (2863)	0.63 (0.54 to 0.70)	0.93 (0.91 to 0.95)
Rim volume	6 (947)	0.49 (0.35 to 0.62)	0.95 (0.92 to 0.96)
OCT RNFL			
Average	57 (8223)	0.69 (0.63 to 0.73)	0.94 (0.93 to 0.95)
Inferior sector	**45 (6542)**	**0.72 (0.65 to 0.77)**	**0.93 (0.92 to 0.95)**
Nasal sector	30 (4395)	0.29 (0.23 to 0.37)	0.93 (0.91 to 0.95)
Superior sector	43 (6395)	0.59 (0.51 to 0.66)	0.94 (0.92 to 0.95)
Temporal sector	30 (4395)	0.30 (0.22 to 0.39)	0.93 (0.91 to 0.95)

[1]Summary sensitivity and specificity pairs of all parameters of each test. Parameters with the highest sensitivity are presented in bold character.
ONH: optic nerve head
RNFL: retinal nerve fibre layer

Table 2. Relative accuracy of all parameters for each test

Test (parameter)	**Sensitivity**	**Specificity**	**Relative DOR**	**P value**
GDx				
Inferior sector	0.62 (0.57 to 0.67)	0.92 (0.90 to 0.94)	0.57 (0.440.74)	< 0.0001
Nerve fibre indica-r (NFI)	**0.74 (0.69 to 0.78)**	**0.92 (0.91 to 0.94)**	**Reference 34.21 (26.50 to 44.15)**	**Reference**
Superior sector	0.63 (0.57 to 0.68)	0.93 (0.91 to 0.95)	0.66 (0.51 to 0.86)	0.0022
Temporal superior nasal inferior temporal (TSNIT) average	0.63 (0.57 to 0.68)	0.94 (0.92 to 0.95)	0.73 (0.56 to 0.95)	0.0213
HRT				
Cup/disc area ratio	0.56 (0.46 to 0.66)	0.95 (0.93 to 0.96)	0.84 (0.55 to 1.30)	0.4326
Cup area	0.44 (0.35 to 0.55)	0.94 (0.93 to 0.96)	0.50 (0.32 to 0.79)	**0.0032**
Cup shape measure	0.37 (0.28 to 0.47)	0.94 (0.92 to 0.96)	0.37 (0.23 to 0.60)	**< 0.0001**
Cup volume	0.31 (0.23 to 0.41)	0.94 (0.92 to 0.96)	0.30 (0.20 to 0.47)	**< 0.0001**
Frederick S. Mikelberg (FSM) discriminant function	0.54 (0.44 to 0.64)	0.94 (0.91 to 0.96)	0.67 (0.41 to 1.10)	0.1092

Table 2. Relative accuracy of all parameters for each test *(Continued)*

Moorfields regression analysis (MRA)	0.74 (0.64 to 0.81)	0.88 (0.84 to 0.91)	0.77 (0.45 to 1.33)	0.3476
Reinhard O.W. Burk (RB) discriminant function	0.52 (0.41 to 0.62)	0.95 (0.92 to 0.97)	0.70 (0.41 to 1.17)	0.1722
Rim volume	0.48 (0.37 to 0.58)	0.94 (0.92 to 0.96)	0.57 (0.36 to 0.90)	**0.0164**
Rim area	0.45 (0.35 to 0.55)	0.95 (0.93 to 0.96)	0.53 (0.34 to 0.81)	**0.0038**
Vertical cup/disc ratio	**0.60 (0.50 to 0.69)**	**0.95 (0.93 to 0.96)**	**Reference 26.81 (17.41 to 41.28)**	**Reference**
OCT ONH				
Cup/disc area ratio	0.66 (0.56 to 0.74)	0.93 (0.90 to 0.95)	0.82 (0.57 to 1.19)	0.2963
Horizontal cup/disc ratio	0.56 (0.45 to 0.66)	0.93 (0.88 to 0.95)	0.49 (0.29 to 0.82)	0.0062
Vertical cup/disc ratio	**0.68 (0.58 to 0.76)**	**0.94 (0.91 to 0.96)**	**Reference 31.63 (18.90 to 52.93)**	**Reference**
Cup area	0.57 (0.46 to 0.67)	0.93 (0.90 to 0.95)	0.57 (0.37 to 0.88)	0.0116
Cup volume	0.44 (0.34 to 0.55)	0.93 (0.90 to 0.96)	0.35 (0.22 to 0.56)	< 0.0001
Disc area	0.31 (0.22 to 0.41)	0.92 (0.87 to 0.95)	0.15 (0.09 to 0.25)	< 0.0001
Nerve head volume	0.59 (0.48 to 0.69)	0.92 (0.88 to 0.96)	0.55 (0.31 to 0.98)	0.0415
Rim area	0.65 (0.55 to 0.73)	0.94 (0.91 to 0.96)	0.90 (0.62 to 1.30)	0.5759
Rim volume	0.57 (0.46 to 0.68)	0.94 (0.91 to 0.97)	0.73 (0.41 to 1.27)	0.2647
OCT RNFL				
Average	**0.69 (0.64 to 0.73)**	**0.95 (0.93 to 0.95)**	**Reference 37.84 (29.66 to 48.29)**	**Reference**
Inferior sector	0.70 (0.66 to 0.75)	0.93 (0.92 to 0.95)	0.90 (0.73 to 1.13)	0.3734
Nasal sector	0.30 (0.25 to 0.35)	0.93 (0.91 to 0.94)	0.15 (0.12 to 0.19)	< 0.0001
Superior sector	0.59 (0.54 to 0.64)	0.94 (0.92 to 0.95)	0.58 (0.46 to 0.72)	< 0.0001
Temporal sector	0.31 (0.26 to 0.36)	0.93 (0.92 to 0.95)	0.17 (0.13 to 0.21)	< 0.0001

DOR: diagnostic odds ratio
ONH: optic nerve head
RNFL: retinal nerve fibre layer

Table 3. Accuracy of macular parameters

OCT macular parameters (models)	**Number of studies (Number of patients)**	**Sensitivity**	**Specificity**
Average (GCC 3D-Topcon, GCC RTVue, GCIPL Cirrus)	32 (5010)	0.63 (0.57 to 0.70)	0.93 (0.91 to 0.94)
Inferior sector(GCC 3D-Topcon, GCC RTVue, GCIPL Cirrus)	27 (4241)	0.63 (0.56 to 0.70)	0.93 (0.01 to 0.94)
Superior sector (GCC 3D-Topcon, GCC RTVue, GCIPL Cirrus)	27 (4241)	0.49 (0.43 to 0.56)	0.93 (0.91 to 0.95)
Focal loss volume (GCC RTVue)	13 (2143)	0.66 (0.50 to 0.78)	0.93 (0.90 to 0.95)
Global loss volume (GCC RTVue)	12 (1939)	0.64 (0.46 to 0.79)	0.93 (0.89 to 0.96)
Minimum sector (GCIPL Cirrus)	9 (1361)	0.76 (0.65 to 0.84)	0.91 (0.87 to 0.93)
Inferonasal sector (GCIPL Cirrus)	8 (1277)	0.48 (0.38 to 0.58)	0.88 (0.84 to 0.91)
Inferotemporal sector (GCIPL Cirrus)	8 (1277)	0.75 (0.65 to 0.82)	0.90 (0.87 to 0.93)
Superonasal sector (GCIPL Cirrus)	8 (1277)	0.43 (0.31 to 0.57)	0.89 (0.83 to 0.93)
Superotemporal sector (GCIPL Cirrus)	8 (1277)	0.61 (0.49 to 0.71)	0.86 (0.81 to 0.90)

Table 4. Accuracy of all parameters: data extracted at the lowest specificity

Test (parameter)	**Sensitivity[1]**	**Specificity[1]**
GDx		
Inferior sector	0.70 (0.65 to 0.74)	0.85 (0.21 to 0.87)
Nerve fibre indicator (NFI)	**0.81 (0.77 to 0.85)**	**0.87 (0.84 to 0.90)**
Superior sector	0.70 (0.65 to 0.73)	0.86 (0.83 to 0.88)
Temporal superior nasal inferior temporal (TSNIT) average	0.69 (0.64 to 0.74)	0.87 (0.84 to 0.90)
HRT		
Cup Disk area ratio	0.65 (0.55 to 0.74)	0.91 (0.87 to 0.94)
Cup area	0.44 (0.31 to 0.58)	0.93 (0.89 to 0.95)
Cup shape measure	0.47 (0.37 to 0.57)	0.92 (0.88 to 0.95)
Cup volume	0.41 (0.29 to 0.54)	0.91 (0.87 to 0.94)
Frederick S. Mikelberg (FSM) discriminant function	0.60 (0.38 to 0.79)	0.92 (0.87 to 0.95)

Table 4. Accuracy of all parameters: data extracted at the lowest specificity *(Continued)*

Moorfields regression analysis (MRA)	0.68 (0.55 to 0.78)	0.88 (0.83 to 0.92)
Reinhard O.W. Burk (RB) discriminant function	0.56 (0.44 to 0.67)	0.93 (0.89 to 0.96)
Rim Volume	0.49 (0.38 to 0.60)	0.92 (0.89 to 0.95)
Rim area	0.54 (0.43 to 0.65)	0.91 (0.87 to 0.94)
Vertical cup disk ratio	**0.71 (0.56 to 0.82)**	**0.92 (0.88 to 0.94)**
OCT ONH		
Cup disk area ratio	0.74 (0.64 to 0.81)	0.84 (0.81 to 0.87)
Cup disk horizontal ratio	0.67 (0.56 to 0.78)	0.81 (0.77 to 0.84)
Cup disk vertical ratio	**0.80 (0.73 to 0.85)**	**0.86 (0.82 to 0.89)**
Cup area	0.56 (0.32 to 0.77)	0.83 (0.79 to 0.87)
Cup volume	0.57 (0.38 to 0.73)	0.83 (0.80 to 0.86)
Disc area	0.32 (0.22 to 0.43)	0.83 (0.77 to 0.87)
Nerve head volume	0.66 (0.57 to 0.73)	0.81 (0.76 to 0.85)
Rim area	0.76 (0.67 to 0.82)	0.85 (0.81 to 0.88)
Rim volume	0.67 (0.59 to 0.74)	0.85 (0.79 to 0.90)
OCT RNFL		
Average	0.78 (0.74 to 0.82)	0.89 (0.86 to 0.91)
Inferior sector	**0.79 (0.75 to 0.82)**	**0.87 (0.84 to 0.89)**
Nasal sector	0.43 (0.36 to 0.50)	0.86 (0.83 to 0.89)
Superior sector	0.71 (0.66 to 0.75)	0.87 (0.85 to 0.90)
Temporal sector	0.41 (0.33 to 0.50)	0.86 (0.83 to 0.89)

[1]Summary sensitivity and specificity pairs of all parameters of each test. Parameters with the highest sensitivity are presented in bold character.
ONH: optic nerve head
RNFL: retinal nerve fibre layer

Table 5. Relative accuracy of the best parameter of each test

Test (parameter)	**Sensitivity**	**Specificity**	**Relative DOR[1]**	**P value**
GDx: Nerve fibre indicator (NFI)	0.70 (0.65 to 0.74)	0.92 (0.91 to 0.94)	0.70 (0.37 to 1.33)	0.2797

Table 5. Relative accuracy of the best parameter of each test *(Continued)*

HRT: Vertical cup/disc ratio	**0.72 (0.61 to 0.80)**	**0.94 (0.91 to 0.96)**	**Reference 40.24 (22.65 to 71.50)**	**Reference**
OCT ONH: Vertical cup/disc ratio	0.72 (0.66 to 0.75)	0.94 (0.92 to 0.95)	0.98 (0.52 to 1.85)	0.9515
OCT RNFL: Average	0.71 (0.67 to 0.75)	0.94 (0.92 to 0.95)	0.99 (0.54 to 1.82)	0.9910

[1]Relative DORs are obtained from HSROC curves assuming parallelism of summary ROC curves by covariate levels, i.e. assuming curves with the same shape.
ONH: optic nerve head
RNFL: retinal nerve fibre layer

Table 6. Heterogeneity investigation[1]

Covariate	Number of studies (Number of patients)	Sensitivity	Specificity	Relative DOR[2]	P value
Reference Standard					
Visual field (VF) alone	27 (4230)	0.71 (0.64 to 0.78)	0.93 (0.91 to 0.95)	**Reference 34.15 (23.59 to 49.44)**	**Reference**
Optic nerve head (ONH) alone	15 (2508)	0.73 (0.68 to 0.77)	0.94 (0.93 to 0.95)	0.56 (0.29 to 1.09)	0.0888
VF + ONH	73 (10681)	0.55 (0.43 to 0.67)	0.94 (0.91 to 0.96)	1.19 (0.77 to 1.85)	0.4278
Mean deviation (MD)					
MD < -6 (more severe glaucoma)	49 (7598)	0.79 (0.74 to 0.83)	0.94 (0.93 to 0.95)	**Reference 57.11 (43.49 to 74.99)**	**Reference**
MD ≥ -6 (less severe glaucoma)	65 (9720)	0.64 (0.60 to 0.69)	0.93 (0.92 to 0.94)	0.45 (0.31 to 0.64)	< 0.0001
Could the conduct or interpretation of the index test have introduced bias?					
Low	68 (9938)	0.70 (0.65 to 0.75)	0.94 (0.92 to 0.95)	**Reference 33.93 (26.44 to 43.54)**	**Reference**
High	33 (5390)	0.70 (0.63 to 0.77)	0.95 (0.94 to 0.96)	1.29 (0.83 to 2.00)	0.2642
Unclear	14 (2091)	0.74 (0.63 to 0.81)	0.91 (0.88 to 0.94)	0.86 (0.49 to 1.51)	0.6003
Could the patient flow have introduced bias?					
Low risk	12 (2155)	0.67 (0.55 to 0.77)	0.61 (0.55 to 0.66)	**Reference 23.42 (13.76 to 39.86)**	**Reference**
High risk	56 (8532)	0.69 (0.64 to 0.74)	0.94 (0.93 to 0.95)	1.48 (0.81 to 2.69)	0.1893

Table 6. Heterogeneity investigation[1] *(Continued)*

Unclear risk	47 (6732)	0.73 (0.68 to 0.78)	0.92 (0.89 to 0.95)	1.81 (0.99 to 3.34)	0.0553
Could the reference standard, its conduct, or its interpretation have introduced bias?					
Low risk	101 (14897)	0.70 (0.66 to 0.74)	0.94 (0.93 to 0.95)	**Reference 35.06 (28.58 to 43.01)**	**Reference**
High risk	1 (120)	0.43 (0.30 to 0.57)	1.00 (0.94 to 1.00)	∞	0.9879
Unclear risk	13 (2402)	0.76 (0.64 to 0.85)	0.93 (0.93 to 0.94)	1.23 (0.65 to 2.36)	0.5221
Could the selection of patients have introduced bias?					
Low risk	2 (284)	0.45 (0.14 to 0.81)	0.95 (0.84 to 0.98)	**Reference**	**Reference**
High risk	111 (16705)	0.71 (0.67 to 0.75)	0.94 (0.93 to 0.94)	2.43 (0.45 to 13.15)	0.3025
Unclear risk	2 (430)	0.61 (0.24 to 0.89)	0.96 (0.86 to 0.99)	2.29 (0.22 to 24.13)	0.4890

[1]Heterogeneity investigation is obtained including the parameter with the best diagnostic odds ratio (DOR) for each test, as found in primary analyses including all studies.
[2]Relative DORs are obtained from HSROC curves assuming parallelism of summary ROC curves by covariate levels, i.e. assuming curves with the same shape.

APPENDICES

Appendix 1. Protocol's original methodological quality assessment criteria using the QUADAS checklist

Assessment of methodological quality: QUADAS and additional items

	Item definition	Item question	Assessment
	Representative spectrum?	Was the spectrum of patients representative of the patients who will receive the test in practice?	Yes: a diverse spectrum of glaucoma and glaucoma suspects is included in the study Unclear: reporting insufficient to assess this item No: a selected type of glaucoma, such as early or late glaucoma only, are included in the study; or specific types of glaucoma or healthy controls which are selected in fixed proportions by design, such as in case-control studies
	Selection criteria reported?	Were the selection criteria clearly described?	Yes: prospective recruitment of patients referred because they are suspected of having glaucoma based on prior testing,

(Continued)

			i.e. any glaucoma screening test by primary care professionals Unclear: reporting insufficient to assess this item No: selection criteria are not reported
	Acceptable reference standard?	Is the reference standard likely to classify the target condition correctly?	Yes: repeatable visual field defect using validated scoring systems for glaucoma, alone or in combination with fundus stereoscopic photography or fundus biomicroscopy of the optic disc by an ophthalmologist or a trained technician (in case photography is used) Unclear: reporting insufficient to assess this item No: definition of glaucoma not based on validated methods regarding visual field and/or optic disc damage
	Acceptable delay between tests?	Is the time period between reference standard and index test short enough to be reasonably sure that the target condition did not change between the two tests?	Yes: if the interval between ONH/RNFL testing and reference standard assessment is one month or less Unclear: reporting insufficient to assess this item No: interval between index test and reference standard declared to be more than one month
	Partial verification avoided?	Did the whole sample, or a random selection of the sample, receive verification using a reference standard of diagnosis?	Yes: there is no pre-selection of patients potentially includable performed according to index test results, i.e. all includable patients, or a random sample, receive both the index and reference test Unclear: reporting insufficient to assess this item No: there are discrepancies and these depend on a pre-selection based on optic nerve head imaging testing
	Differential verification avoided?	Did patients receive the same reference standard regardless of the index test result?	Yes: the same reference standard was used for all patients Unclear: reporting insufficient to assess this item No: different reference standards were used and this selection is potentially associated with index test results
	Incorporation avoided?	Was the reference standard independent of the index test (i.e. the index	Yes: only functional measures of damage have been considered as a reference standard to define glaucoma

(Continued)

		test did not form part of the reference standard)?	Unclear: reporting insufficient to assess this item No: optic disc appearance was part of the reference standard (see Methods section)
	Index test execution described?	Was the execution of the index test described in sufficient detail to permit replication of the test?	Yes: OCT, HRT and GDx model, execution and diagnostic criteria clearly described Unclear: some reporting but insufficient to assess this item No: the above elements not described
	Reference test execution described?	Was the execution of the reference standard described in sufficient detail to permit its replication?	Yes: visual field analyser model, program, threshold strategy (i.e. HFA mod 750, 24-2 program and SITA Standard strategy) and visual field defect criteria were described in detail (i.e. MD or PSD significance thresholds, or Glaucoma Hemifield Test outcome). If included in the reference standard the description of optic disc defects was clearly defined. Unclear: reporting insufficient to assess this item No: either visual field analyser or visual field defect definition or optic disc defect definition are not described
	Index test results masked	Were the index test results interpreted without knowledge of the results of the reference standard?	Yes: it is stated that the index test was performed masked to the results of the reference standard; or it was performed and results recorded prior to the reference standard Unclear: reporting insufficient to assess this item No: the index standard was performed and assessed with knowledge of the results of the reference standard
	Reference test results masked	Were the reference test results interpreted without knowledge of the results of imaging?	Yes: it is stated that the reference standard was performed masked to the results of imaging; or it was performed and results recorded prior to imaging Unclear: reporting insufficient to assess this item No: the reference standard was performed and assessed with knowledge of the results of imaging
	Clinical data available?	Were the same clinical data available when test results were interpreted as would	Yes: intraocular pressure and other clinical data are available as is common in clinical practice

(Continued)

		be available when the test is used in practice?	Unclear: reporting insufficient to assess this item No: intraocular pressure and other clinical data not available
	Uninterpretable results reported?	Were uninterpretable/intermediate test results reported?	Yes: the number of patients with uninterpretable index test results is reported, and the reasons are explained Unclear: reporting insufficient to assess this item No: uninterpretable optic nerve head imaging results not reported
	Withdrawals explained	Were withdrawals from the study explained?	Yes: the number of drop-outs has been reported and reasons have been explained Unclear: reporting insufficient to assess this item No: the number of withdrawals has not been reported
	Sponsoring precluded?	Was the study sponsored by producers of imaging devices?	Yes: no sponsorship or other than imaging producers Unclear: reporting insufficient to assess this item No: sponsored by imaging producers
	Individuals as unit of analyses?	Were eyes or individuals the unit of analyses?	Yes: only one eye of each individual was included or less than 10% of individuals had both eyes included in the analyses Unclear: reporting insufficient to assess this item No: 10% or more of individuals had both eyes included in the analyses

Appendix 2. MEDLINE (Ovid) search strategy

1 exp glaucoma/
2 glaucoma$.tw.
3 exp ocular hypertension/
4 (OHT or IOP).tw.
5 exp intraocular pressure/
6 (((increas$ or elevat$ or high$ or raise$) adj3 (ocular or intraocular or intra-ocular)) and pressure).tw.
7 optic nerve diseases/
8 (optic adj2 nerve$ adj2 head).tw.
9 ONH.tw.
10 optic disk/
11 optic dis$.tw.
12 retinal ganglion cells/
13 retinal ganglion cell$.tw.
14 (retinal adj2 nerve adj2 fiber adj2 layer).tw.
15 (retinal adj2 nerve adj2 fibre adj2 layer).tw.

16 RNFL.tw.
17 or/1-16
18 ophthalmoscopy/
19 (confocal adj2 scan$ adj2 laser adj2 ophthalm$).tw.
20 (Heidelberg adj2 Retina adj2 Tomograph$).tw.
21 HRT.ti,ab.
22 Lasers/du [Diagnostic Use]
23 (scan$ adj2 laser$ adj2 polarimetry).tw.
24 SLP.tw.
25 GDx.tw.
26 VCC.tw.
27 enhanced corneal compensat$.tw.
28 variable corneal compensat$.tw.
29 tomography, optical coherence/
30 tomography, optical/
31 (optical adj2 coherence adj2 tomograph$).tw.
32 OCT.ti,ab.
33 (optical adj2 coherence adj2 interferomet$).tw.
34 or/18-33
35 17 and 34
36 exp animals/
37 exp humans/
38 36 not (36 and 37)
39 35 not 38
40 case reports.pt.
41 39 not 40

Appendix 3. EMBASE (Ovid) search strategy

1 exp glaucoma/
2 glaucoma$.tw.
3 exp intraocular hypertension/
4 (OHT or IOP).tw.
5 exp intraocular pressure/
6 (((increas$ or elevat$ or high$ or raise$) adj3 (ocular or intraocular or intra-ocular)) and pressure).tw.
7 optic nerve disease/
8 (optic adj2 nerve$ adj2 head).tw.
9 ONH.tw.
10 optic disk/
11 optic dis$.tw.
12 retinal ganglion cell/
13 retinal ganglion cell$.tw.
14 (retinal adj2 nerve adj2 fiber adj2 layer).tw.
15 (retinal adj2 nerve adj2 fibre adj2 layer).tw.
16 RNFL.tw.
17 or/1-16
18 ophthalmoscopy/
19 scanning laser ophthalmoscopy/
20 (confocal adj2 scan$ adj2 laser adj2 ophthalm$).tw.
21 (Heidelberg adj2 Retina adj2 Tomograph$).tw.
22 HRT.ti,ab.
23 polarimetry/
24 (scan$ adj2 laser$ adj2 polarimetry).tw.
25 SLP.tw.
26 GDx.tw.
27 VCC.tw.
28 enhanced corneal compensat$.tw.
29 variable corneal compensat$.tw.
30 optical coherence tomography/
31 optical tomography/
32 (optical adj2 coherence adj2 tomograph$).tw.
33 OCT.ti,ab.

34 (optical adj2 coherence adj2 interferomet$).tw.
35 or/18-34
36 17 and 35
37 exp animals/
38 exp humans/
39 37 not (37 and 38)
40 36 not 39
41 case report/
42 40 not 41

Appendix 4. MEDION search strategy

Database will be searched on ICPC code field, using code "f" for ophthalmology.

Appendix 5. ARIF search strategy

glaucoma

Appendix 6. Guidance for extracting study characteristics

Study ID	**First author, year of publication.**
Clinical features and settings	Spectrum of glaucoma severity, previous testing, clinical setting including country where the study was conducted, specialty of clinicians involved in the assessment
Participants	Sample size, age, sex, ethnicity, country, co-morbidities
Study design	Whether the sample was selected as a single group (consecutive series) or as separate groups with and without the target condition (case-control). Whether participants were consecutively enrolled in the study and were identified retrospectively or prospectively. If studies evaluated more than one imaging test, how were individuals allocated to a certain imaging test and whether all participants underwent all imaging tests
Target condition	Manifest glaucoma, including the prevalence of the target condition in the sample
Reference standard	Type of optic nerve head evaluation (photography or biomicroscopy, scoring system). Type of visual field test and criteria used for diagnosing glaucomatous damage (such as a specific scoring system). If the assessment was performed by more than one observer, how were discrepancies between observers resolved. Reliability of the visual field examination indexes
Index tests	Model, manufacturer and any technical characteristics (software spatial analyses) of the imaging method under investigation. Test parameters or diagnostic algorithms used. Quality imaging scan assessment and conflict of interest reporting
Follow up	Not applicable since we will not include studies in which follow up is needed as reference standard
Notes	Source of funding, any other relevant information

Appendix 7. QUADAS 2 guidance adapted from the original QUADAS guidance in Appendix 6

DOMAIN	yes (high)	no	unclear
PATIENT SELECTION	Describe methods of patient selection: Describe included patients (prior testing, presentation, intended use of index test and setting):		

(Continued)

Was a consecutive or random sample of patients enrolled?	Consecutive sampling or random sampling of patients according to inclusion criteria	Non random sampling or sampling based on volunteering or referral.	Unclear whether consecutive or random sampling used
Was a case-control design avoided?	No selective recruitment of participants with well known disease and a control group of healthy patients or nested case-control designs (systematically and randomly selected from a defined population cohort)	Selection of specific types of glaucoma and healthy controls in a predetermined, nonrandom fashion	Unclear selection mechanism
Did the study avoid inappropriate exclusions?	Exclusions are detailed and felt to be appropriate (e.g. non glaucomatous optic neuropathy or neurologic disease affecting visual field assessment)	Inappropriate exclusions, such as "difficult-to-diagnose patients", are reported	Reporting insufficient to assess this item
Risk of bias: Could the selection of patients have introduced bias?	Overall judgement at reviewers' discretion, with reasons		
Concerns regarding applicability: Are there concerns that the included patients do not match the review question?	Inclusion of adult patients with suspect manifest open angle glaucoma	Inclusion of participants whose features (severity of the target condition, presence of comorbid conditions, age, setting of enrollment and previous clinical history) may not match the review question.	Unclear inclusion criteria
INDEX TEST	Describe the index test and how it was conducted and interpreted; Describe the imaging scan's quality assessment criteria and any conflict of interest		
Were the index test results interpreted without knowledge of the results of the reference standard?	Not considered in this review since outcome measures are objectively measured and no effect of masked vs. unmasked evaluation is expected.		
If a threshold was used, was it pre-specified?	Many included index tests are based on continuous measures (e.g. RNFL thickness, rim area, TSNIT average); the study authors used selected cut-off specified a priori to dichotomise data or calculated sensitivity at different fixed level of specifity	The authors selected and used a test threshold based on their own study data, to optimize sensitivity and/or specificity	Reporting insufficient to assess this item
Was the imaging scan's quality assessed?	Imaging scans' quality was assessed and the relative criteria are clearly reported	Imaging scans' quality assessment not used, i.e. no selection is made based on image quality.	Reporting insufficient to assess this item(e.g. scan's quality assessment is mentioned but no specific used criteria are detailed)
Was any conflict of interest avoided?	No author has conflict of interest or commercial relationship with imaging test producer	One or more authors have financial or commercial relationship with the imaging test producer conflict of interest	Reporting insufficient to assess this item
Risk of bias: Could the conduct or interpreta-	Overall judgement at reviewers' discretion, with reasons		

(Continued)

tion of the index test have introduced bias?			
Concerns regarding applicability: Are there concerns that the index test, its conduct, or interpretation differ from the review question?	Tests used and testing procedure clearly reported and tests executed by personnel with sufficient training.	Tests used are not validated or study personnel is insufficiently trained.	Unclear tests or unclear study personnel profile, background and training.
REFERENCE STANDARD	Describe the reference standard and how it was conducted and interpreted:		
Is the reference standard likely to correctly classify the target condition?	Visual field damage used ty classify manifest glaucoma patients according to international guidelines	Optic nerve damage only, not visual field used to classify manifest glaucoma	Reporting insufficient to assess this item
Were the reference standard results interpreted without knowledge of the results of the index test?	Reference standard performed "blinded" or "independently and without knowledge of" index test results are sufficient and full details of the blinding procedure are not required; or clear temporal pattern to the order of testing that precludes the need for formal blinding.	Reference standard was performed and assessed with knowledge of the results of imaging	Unclear whether results were interpreted independently
Risk of bias: Could the reference standard, its conduct, or its interpretation have introduced bias?	Overall judgement at reviewers' discretion, with reasons		
Concerns regarding applicability: Are there concerns that the target condition as defined by the reference standard does not match the review question?	Both optic disc neuropathy and/or visual field defect used and testing procedure and evaluation performed by personnel with sufficient experience	The criteria used to define target condition differ from the criteria specified in the review question	Unclear study personnel profile, background and training or definition criteria
FLOW AND TIMING	Describe any patients who did not receive the index test(s) and/or reference standard or who were excluded from the 2x2 table (refer to flow diagram): Describe the time interval and any interventions between index test(s) and reference standard		
Was there an appropriate interval between index test(s) and reference standard?	Time interval between index and reference test was one month or less	More than one month between index and reference test execution	Unclear whether tests were executed within one month
Did all patients receive a reference standard?	All patients receiving the index test were verified with the reference standard	The verification rate of index test positive and is different than that of negative patients	Unclear whether all subjects receiving the index test were verified with the reference standard
Did all patients receive the same reference standard?	The same reference standard (optic disc appearance assessment or visual field testing or both) were used for all patients	Not all patients were assessed with the same reference standard (e.g. visual field testing was performed for some participants only)	Unclear whether all participants were verified with the same reference test by trained professionals.

(Continued)

Were all patients included in the analysis?	The number of subjects enrolled in the study does match the number in analyses or less than 10% of the whole sample enrolled, was excluded from the analysis	More than 10% of the whole sample enrolled and included in the study, was excluded from the final analysis. .	Reporting insufficient to assess this item(e.g, some patients' exclusion was mentioned but no specific details were reported)
Risk of bias: Could the patient flow have introduced bias?	Overall judgement at reviewers' discretion, with reasons		

Appendix 8. Cochrane Library search strategy

#1 MeSH descriptor Glaucoma
#2 glaucoma*
#3 MeSH descriptor Ocular Hypertension
#4 OHT or IOP
#5 MeSH descriptor Intraocular Pressure
#6 ((increas* or elevat* or high* or raise*) near/3 (ocular or intraocular or intra-ocular) near/3 (pressure))
#7 MeSH descriptor Optic Nerve Diseases
#8 optic near/2 nerve* near/2 head
#9 ONH
#10 MeSH descriptor Optic Disk
#11 optic dis*
#12 MeSH descriptor Retinal Ganglion Cells
#13 retinal ganglion cell*
#14 retinal nerve fiber layer
#15 retinal nerve fibre layer
#16 RNFL
#17 (#1 OR #2 OR #3 OR #4 OR #5 OR #6 OR #7 OR #8 OR #9 OR #10 OR #11 OR #12 OR #13 OR #14 OR #15 OR #16)
#18 MeSH descriptor Ophthalmoscopy
#19 scan* near/2 laser* near/2 ophthalm*
#20 Heidelberg near/2 Retina near/2 Tomograph*
#21 HRT:ti,ab
#22 MeSH descriptor Lasers explode all trees with qualifier: DU
#23 scan* near/2 laser* near/2 polarimetry
#24 SLP
#25 GDX
#26 VCC
#27 enhanced corneal compensat*
#28 variable corneal compensat*
#29 MeSH descriptor Tomography, Optical Coherence
#30 MeSH descriptor Tomography, Optical
#31 optical near/2 coherence near/2 tomograph*
#32 OCT:ti,ab
#33 optical near/2 coherence near/2 interferomet*
#34 (#18 OR #19 OR #20 OR #21 OR #22 OR #23 OR #24 OR #25 OR #26 OR #27 OR #28 OR #29 OR #30 OR #31 OR #32 OR #33)
#35 (#17 AND #34)

WHAT'S NEW

Date	Event	Description
17 August 2020	Amended	In Summary of findings 1, heading for one of the parameters has been amended from OCT RNFL mean thickness to OCT RNFL inferior sector.

HISTORY

Protocol first published: Issue 11, 2010
Review first published: Issue 11, 2015

CONTRIBUTIONS OF AUTHORS

Conceiving the review: FO, GV, SF, SMN
Designing the review: FI, GV, MP, MB, MM
Co-ordinating the review: GV
Data collection for the review:
- Designing electronic search strategies: Cochrane Eyes and Vision editorial base
- Undertaking manual searches: FO, MP, MM
- Screening search results: EL, FO, MP, MM, SMN, SF
- Organising retrieval of papers: EL, FO, MP, MM
- Screening retrieved papers against inclusion criteria: EL, MM, FO, GV
- Appraising quality of papers: EL, FO, GV, CP, MB, MM, SF, SMN
- Extracting data from papers: EL, FO, MP, MM
- Writing to authors of papers for additional information: not applicable
- Providing additional data about papers: not applicable
- Obtaining and screening data on unpublished studies: not applicable
Data management for the review:
- Entering data into RevMan: EL, FO, GV, MM
Analysis of data: EL, FO, GV, MB
Interpretation of data:
- Providing a methodological perspective: MB, GV
- Providing a clinical perspective: FO, GV, CP, MM
- Providing a policy perspective: MB, GV
Writing the review: EL, FI, GV, MP, MB, MM
Providing general advice on the review: EL, FI, GV, MP, MB, NN
Securing funding for the review: FO, GV
Performing previous work that was the foundation of the current study: FO

DECLARATIONS OF INTEREST

Manuele Michelessi: none known
Ersilia Lucenteforte: none known
Francesco Oddone. none known
Miriam Brazzelli: none known
Mariacristina Parravano: none known
Sara Franchi: none known
Sueko M Ng: none known
Gianni Virgili: none known

SOURCES OF SUPPORT

Internal sources

- Department of Translational Surgery and Medicine, University of Florence, Italy, Italy

  Dr Virgili's work is partially supported by the University of Florence

External sources

- Assessorato per il Diritto alla Sallute, Regione Toscana, Italy

  Dr Virgili was partially funded by the Tuscany Region for supporting the conduction of the review
- Fondazione G.B. Bietti per lo studio e la ricerca in Oftalmologia - IRCCS, Rome, Italy, Other

  The contribution of the IRCCS Fondazione Bietti in this paper was supported by Ministry of Health and Fondazione Roma

- National Institute for Health Research (NIHR), UK
  * Richard Wormald, Co-ordinating Editor for Cochrane Eyes and Vision (CEV) acknowledges financial support for his CEV research sessions from the Department of Health through the award made by the National Institute for Health Research to Moorfields Eye Hospital NHS Foundation Trust and UCL Institute of Ophthalmology for a Specialist Biomedical Research Centre for Ophthalmology.
  * The NIHR also funds the CEV Editorial Base in London.

  The views expressed in this publication are those of the authors and not necessarily those of the NIHR, NHS, or the Department of Health.

DIFFERENCES BETWEEN PROTOCOL AND REVIEW

We adapted the original QUADAS guidance (Whiting 2003) developed in the protocol for this review to the framework of QUADAS 2, as recommended.

Originally, we planned to include in this review both cohort studies and case-control studies. A first draft was submitted to the DTA Editorial Team based on a search conducted until 15 June 2013, which identified a large number of case-control studies. During the revision process of the initial first draft, we decided to update the literature search (15 February 2015). The new search identified further case-control studies that are known to be prone to methodological biases. We considered the addition of further poor-quality case-control studies not to be worthwhile, and that they were unlikely to improve the quality of the body of evidence assessed in this review. Future updates of this review should only consider studies where patients are enrolled consecutively based on the same set of inclusion criteria, such as referable patients identified in primary care.

During the review process, we decided to extract OCT parameters that are not related to RNFL and ONH morphology, but rather to macular cell layers affected by glaucoma, such as ganglion cell complex (GCC) and ganglion cell inner plexiform layer (GCIPL), as these parameters have gained currency in recent years.However, these data were not formally analysed and used to formulate conclusions.

We deviated from the HSROC models using SAS rather than Winbugs, as originally planned. We found little variation in specificity, as sensitivity was extracted at fixed specificity in almost all studies.

INDEX TERMS

Medical Subject Headings (MeSH)

Diagnostic Errors [statistics & numerical data]; Glaucoma [*diagnosis]; Nerve Fibers [*pathology]; Odds Ratio; Ophthalmoscopy [*standards]; Optic Disk [*pathology]; Prospective Studies; Retrospective Studies; Scanning Laser Polarimetry [*standards]; Sensitivity and Specificity; Tomography, Optical Coherence [*standards]; Visual Field Tests

MeSH check words

Humans