EXTENDED REPORT

Accuracy of GDx VCC, HRT I, and clinical assessment of stereoscopic optic nerve head photographs for diagnosing glaucoma

Nicolaas J Reus, Maartje de Graaf, Hans G Lemij

Br J Ophthalmol 2007;**91**:313–318. doi: 10.1136/bjo.2006.096586

See end of article for authors' affiliations

Correspondence to:
N J Reus, The Rotterdam Eye Hospital, P.O. Box 70030, NL-3000 LM Rotterdam, The Netherlands;
reus@oogziekenhuis.nl

Accepted 3 October 2006
Published Online First 11 October 2006

Aims: To determine and compare the accuracy and reproducibility of GDx variable cornea compensation (VCC) scanning laser polarimetry (SLP) with VCC, Heidelberg retina tomograph (HRT) I confocal scanning laser ophthalmoscopy (CSLO), and clinical assessment of stereoscopic optic nerve head (ONH) photographs for diagnosing glaucoma.

Methods: One eye each of 40 healthy subjects, 48 glaucoma patients, and six patients with ocular hypertension were measured with SLP-VCC and CSLO. Simultaneous stereoscopic ONH photographs were also obtained. Sixteen photographs of healthy and glaucomatous eyes were duplicated for assessing intraobserver agreement. Four glaucoma specialists, four general ophthalmologists, four residents in ophthalmology, and four optometrists classified the ONH photographs as normal or glaucomatous. For SLP-VCC, the nerve fiber indicator (NFI) was evaluated. For CSLO, the Moorfields regression analysis (MRA) and the Bathija linear discriminant function (LDF) were used. Sensitivity, specificity, percentage of correctly classified eyes, and intra- and interobserver agreement, expressed as kappa (κ) were calculated.

Results: SLP-VCC had the highest diagnostic accuracy, with a sensitivity, specificity, and overall correct classification of 91.7%, 95.0% and 93.2%, respectively. CSLO, expressed as Bathija LDF and MRA, had a diagnostic accuracy comparable to glaucoma specialists and general ophthalmologists with an overall accuracy of 89.8%, 86.4%, 86.7% and 85.2%, respectively. Residents classified the fewest eyes correctly. Intraobserver agreement for classifying the ONH photographs ranged between 0.48 (within residents) and 0.78 (within glaucoma specialists). The interobserver agreement ranged between 0.45 (between residents) and 0.74 (between glaucoma specialists). The agreement between observers and CSLO MRA (κ, 0.68) was statistically significantly higher ($p<0.001$; paired *t*-test) than between observers and SLP-VCC NFI (κ, 0.60) and CSLO Bathija LDF (κ, 0.62).

Conclusion: Automated analysis of measurements with GDx VCC and HRT had a similar diagnostic accuracy for glaucoma as classification of stereoscopic ONH photographs by glaucoma specialists, thus bringing all eye-care professionals to this desirable level. The intra- and interobserver agreement for ONH analysis was only moderate to good. We think these imaging techniques may assist clinicians in diagnosing glaucoma.

The ability of eye-care professionals to discriminate between healthy and glaucomatous eyes, based solely on the subjective evaluation of the appearance of the optic nerve head (ONH), is limited.[1–5] Nonetheless, assessment of the ONH's appearance remains a cornerstone for diagnosing glaucoma.

Various objective imaging techniques, such as scanning laser polarimetry (SLP) and confocal scanning laser ophthalmoscopy (CSLO), have been developed to assess glaucomatous damage to the retinal nerve fiber layer (RNFL) and ONH, respectively. SLP, featured in the commercially available GDx variable cornea compensation (VCC), (Carl Zeiss Meditec AG, Jena, Germany), estimates the thickness of the RNFL in the peripapillary area. CSLO, commercially available in the Heidelberg retina tomograph (HRT),(Heidelberg Engineering GmbH, Dossenheim, Germany), creates a topographic image of the ONH.

The purpose of the present study was to determine whether these imaging techniques may assist clinicians in the management of glaucoma. To this end, we compared the accuracy and reproducibility of SLP-VCC and CSLO for diagnosing glaucoma to that of clinical assessment of stereoscopic ONH photographs by various eye-care professionals.

METHODS

Subjects

Forty healthy subjects, 48 glaucoma patients, and six patients with ocular hypertension (OHT) were selected from a cohort of patients and controls who had been originally recruited for an ongoing longitudinal glaucoma study at the Rotterdam Eye Hospital (REH). Healthy subjects had been recruited from spouses and friends of patients and from employees of the REH and their spouses and friends. All subjects were of white ethnic origin and had a best-corrected visual acuity of 20/40 or better. None had a significant history of coexisting ocular diseases, a history of intraocular surgery (except for any uncomplicated cataract surgery or, if applicable, glaucoma surgery), or systemic diseases with possible ocular involvement, such as diabetes mellitus. All subjects had reproducible and reliable

Abbreviations: AUC, area under curves; CSLO, confocal scanning laser ophthalmoscopy; GHT, glaucoma hemifield test; HAP, Hodapp-Anderson-Parrish; HRT, Heidelberg retina tomograph; IOP, intraocular pressure; LDF, linear discriminant function; MRA, Moorfields regression analysis; NFI, nerve fiber indicator; OHT, ocular hypertension; ONH, optic nerve head; REH, Rotterdam Eye Hospital; RNFL, retinal nerve fiber layer; ROC, receiver operating characteristic; SLP, scanning laser polarimetry; VCC, variable cornea compensation visual fields with standard automated perimetry (SAP, Humphrey Field Analyzer II, Carl Zeiss Meditec AG, Jena, Germany). One eye was randomly selected if both eyes were eligible.

All healthy subjects had normal visual fields, that is, a glaucoma hemifield test (GHT) within normal limits and no nerve fiber bundle visual field defects in the total and/or pattern deviation probability plots. In addition, none had relatives in the first and/or second degree with glaucoma. Furthermore, slit-lamp examination was unremarkable; they had healthy-looking ONHs and an intraocular pressure (IOP) $\leqslant$21 mmHg in both eyes.

Glaucoma was diagnosed by one of four glaucoma specialists of the REH based on a glaucomatous appearance of at least their included eyes' ONHs (with notching or thinning of the neuroretinal rim, possibly with disc haemorrhages) and a reproducible corresponding nerve fiber bundle visual field defect with SAP (author, please expand and as this is only used once there can be no abreviation).

Subjects with OHT had an IOP of $\geqslant$22 mmHg and $\leqslant$32 mmHg and normal visual fields on at least two separate occasions. The appearance of the ONH was not a selection criterion.

The research adhered to the tenets of the Declaration of Helsinki. The institutional human experimentation committee had approved the research. Informed consent was obtained from the subjects after explanation of the nature and possible consequences of the study.

GDx VCC scanning laser polarimetry

SLP-VCC measurements were performed with the GDx VCC (software *V*. 5.4.0). This technique has been described in detail elsewhere.[6] For each subject, images were obtained from both eyes with the subject's pupils undilated and the room lights left on. During all measurements, the operator saw to it that subjects had their heads as vertical as possible to maintain the same orientation of the slow axes of the birefringent structures in the eye to that of the instrument's compensator. Only high quality scans, that is with a centered ONH, well-focussed, evenly and justly illuminated throughout the image, and without any motion artifacts, were accepted. Further details on the scanning technique have been described in detail elsewhere.[6] We did not exclude any atypical scans from the analysis.

HRT I confocal scanning laser ophthalmoscopy

CSLO measurements were performed with the HRT I (software *V*. 1.4.0.0). This technique has been described in detail elsewhere.[7] Three high-quality images at a 15°×15° scanning angle were recorded for each eye per subject. A mean image was used for subsequent analysis. All mean images had a mean SD of the height measurements $\leqslant$50 μm. The optic disc margin was manually marked at the inner edge of Elschnig's ring by one of the authors (NJR), with the use of stereoscopic ONH photographs, if required. The standard reference plane was used for calculations of optic disc topography with the relative and tilted coordinate system turned on.

Optic nerve head stereo photography

The ONH in both eyes of all subjects were photographed with a simultaneous stereoscopic camera (TRC-SS2, Topcon Medical Systems, Inc., Paramus, NJ, USA) at a 15° field of view. Only high quality images were used. Nine photographs of glaucoma patients and seven photographs of healthy subjects were randomly selected and duplicated for assessing intraobserver agreement, totaling 110 stereophotographs. The sequence of the photographs was randomised.

Clinical assessment of optic nerve head photographs

Stereoscopic ONH photographs were evaluated by 16 observers: four glaucoma specialists, four general ophthalmologists, four residents in ophthalmology, and four optometrists. Participating glaucoma specialists worked at the REH. The general ophthalmologists worked in two other hospitals. The participating residents had been in training for 1–2 years at the REH. Optometrists were employed by the REH. In The Netherlands, optometry is a relatively new profession. Many optometrists work at optician's stores where they may provide primary eye care. With increasing frequency, however, they work in hospitals, supervised by ophthalmologists.

The graders individually classified the ONHs as either normal or glaucomatous. When in doubt, they were forced to make a classification. All observers classified the set of slides in the same, randomised, order. To better estimate ONH size, reference slides of healthy, small (5th percentile), medium (50th percentile), and large (95th percentile) ONHs were provided.

Data analysis

The sensitivity and specificity were calculated for each observer and imaging technique. We also calculated the sensitivity and specificity for each group of observers by averaging the individual sensitivities and specificities. For the GDx VCC, we determined the sensitivity and specificity of the nerve fiber indicator (NFI) parameter, which has been shown to be its best discriminating parameter.[6] In the present study, we used a cut-off value of 35. For GDx VCC software *V*. 5.3.1 and higher, we have found the best discriminating cut-off value to be 35 (unpublished data) instead of the previously published value of 40.[6] For comparison with other studies, we also determined the diagnostic accuracy of the NFI for cut-off values of 30 and 40. For the HRT, we calculated the sensitivity and specificity for both the Moorfields regression analysis (MRA)[8] and the Bathija linear discriminant function (LDF).[9] The Bathija LDF has been shown to be the best discriminating parameter of the HRT.[10] The MRA is often used in clinical practice. We analysed results of the MRA with both the classification of "Borderline" grouped with "Outside normal limits" (MRA 1) and the classification of "Borderline" grouped with "Within normal limits" (MRA 2). For the continuous parameters NFI and Bathija LDF, we constructed receiver operating characteristic (ROC) curves. The 95% confidence intervals (CIs) for the areas under the ROC curves, (AUC) were calculated as the point estimate of the AUC $\pm$ 1.96 *standard error (SE).

To explore the agreement in classification between the various observers, SLP-VCC and CSLO, we determined the intraobserver and interobserver agreement, expressed as κ. The values for κ were classified as follows: $\leqslant$0.2, poor; 0.21 to 0.40, fair; 0.41 to 0.60, moderate; 0.61 to 0.80, good; and $\geqslant$0.81, very good.[11]

Data of OHT eyes were not analysed in the present study as they had merely been added to present a continuum of glaucomatous damage.

We classified the glaucoma patients by the severity of their visual field damage based on their mean deviation (MD; mild, MD $\geqslant-6$ dB; moderate, -12 dB $\leqslant$ MD <-6 dB; severe, MD <-12 dB) and with the Hodapp-Anderson-Parrish (HAP) grading scale.[12]

We used paired and unpaired *t* tests to evaluate differences in measurements within and between groups, respectively. For differences in dichotomous variables between groups, we used the Pearson χ^2 test. Differences in AUC between NFI and Bathija-LDF were tested for statistical significance with the paired-samples test described by DeLong *et al*.[13] Differences in the number of correctly classified eyes between the various

Table 1 Demographics of Healthy Subjects (n = 40) and Glaucoma Patients (n = 48)

	Unit	Healthy Subjects	Glaucoma Patients	*p*
Age (SD; range)	Years	59 (12)	61 (11)	0.31
Gender (male) (%)	NA	19 (48)	26 (54)	0.53
Randomised eye (right) (%)	NA	21 (53)	18 (38)	0.16
Disc area (SD; range)	mm^2	1.89 (0.33; 1.29 to 2.97)	2.03 (0.37; 1.47 to 2.90)	0.055
MD (SD; range)	Decibels	0.10 (0.93; −1.55 to 2.04)	−6.56 (6.32; −22.26 to 1.25)	<0.001
PSD (SD; range)	Decibels	1.64 (0.38; 0.91 to 3.01)	7.71 (4.03; 2.65 to 15.92)	<0.001

Differences between the groups were tested for statistical significance with an independent samples *t*-test for continuous variables and with the Pearson Chi-Square test for dichotomous variables.
MD, mean deviation; NA, not applicable; PSD, pattern standard deviation; SD, standard deviation

techniques and observers were tested for statistical significance with the McNemar test. In the present study, a p value of less than 0.05 was considered statistically significant. For paired comparisons, the α was adjusted to the number of comparisons within each analysis with the Bonferroni correction to allow for multiple testing. Statistical analyses were performed with SPSS for Windows (release 12.0.1, 2003, SPSS Inc., Chicago, IL, USA).

The demographics of the healthy subjects and glaucoma patients have been presented in table 1.

RESULTS

Diagnostic accuracy

The sensitivity and specificity for discriminating between healthy and glaucomatous eyes by the two imaging technologies, and by the four groups of observers who judged ONH stereo slides, have been presented in fig 1 and in table 2. Overall, the SLP-VCC NFI parameter appeared to have the highest diagnostic accuracy (sensitivity, 91.7%; specificity, 95.0%; overall accuracy, 93.2%). The accuracy of CSLO Bathija LDF was slightly lower (table 2). The AUC of the SLP-VCC NFI parameter (0.98; 95% CI, 0.96–1.00) and CSLO Bathija LDF (0.94; 95% CI, 0.89–0.99) were not statistically significantly different (p = 0.087). Both SLP-VCC NFI and CSLO Bathija LDF had a statistically significantly higher overall accuracy than CSLO MRA 2 and ophthalmology residents (table 3). With regard to classifying the ONH stereo slides, glaucoma specialists had the highest diagnostic accuracy (sensitivity, 80.2%; specificity, 94.4%; overall accuracy, 86.7%), followed subsequently by general ophthalmologists, optometrists, and ophthalmology residents (table 2). The overall accuracy of glaucoma specialists and general ophthalmologists was statistically significantly higher than that of ophthalmology residents (table 3). The CSLO MRA 1 had a diagnostic accuracy that was similar to that of glaucoma specialists judging the ONH stereo slides (table 2). The diagnostic accuracy of CSLO MRA 2 was lower (table 2).

The sensitivity for correctly classifying glaucomatous eyes stratified for mild, moderate, and severe visual field loss has been presented in table 4. Eyes with mild glaucomatous damage were most difficult to correctly classify, for both the imaging techniques and the clinical observers.

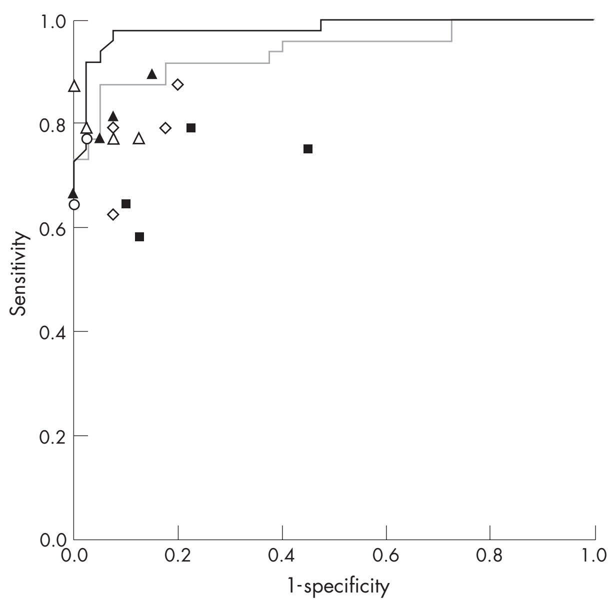

Figure 1 Receiver operating characteristic curves of the scanning laser polarimetry with variable cornea compensator parameter NFI (solid black line) and the confocal scanning laser ophthalmoscopy (CSLO) linear discriminant function by Bathija (solid grey line) for diagnosing glaucoma in 40 healthy and 48 glaucomatous eyes. In addition, data on evaluation of stereoscopic optic nerve head photographs by four glaucoma specialists (open triangles), four general ophthalmologists (closed triangles), four optometrists (open diamonds), and four junior ophthalmology residents (closed squares) as well as the CSLO Moorfields regression analysis (open circles) have been presented.

Intraobserver agreement

The intraobserver agreement (κ, mean; range) in ophthalmology residents for classifying ONH stereo slides was moderate (0.48; 0.13–0.74). The intraobserver agreement in optometrists (0.63; 0.53–0.73), general ophthalmologists (0.71; 0.52–1.00), and glaucoma specialists (0.78; 0.46–1.00) was good.

Interobserver agreement

The diagnostic accuracy of correctly classifying ONH stereo slides varied considerably between observers (fig 1). The average interobserver agreement between glaucoma specialists, general ophthalmologists, optometrists, and ophthalmology residents was 0.74, 0.72, 0.67, and 0.45, respectively (table 5).

The classification of stereoscopic ONH photographs by the observers agreed better with the classification by the CSLO MRA 1 (mean κ, 0.68) than with either the classification by the SLP-VCC NFI parameter (mean κ, 0.60) (p<0.001, paired *t* test) or the classification made by the Bathija LDF (mean κ, 0.62) (p<0.001, paired *t* test). The agreement between observers and SLP-VCC NFI was not statistically significantly different from the agreement between observers and the Bathija LDF (p = 0.090, paired *t* test).

DISCUSSION

In the present study, automated analysis of SLP-VCC and CSLO measurements discriminated similarly between healthy and glaucomatous eyes as clinical assessment of stereoscopic ONH photographs by glaucoma specialists, thus bringing all eye-care professionals to this desirable level. Of all subjective graders of stereoscopic ONH photographs, the glaucoma specialists showed the highest overall diagnostic accuracy. For any

Table 2 Sensitivity and specificity for discriminating between healthy and glaucoma eyes and the percentage of correctly classified eyes by scanning laser polarimetry with variable corneal compensation (SLP-VCC), confocal scanning laser ophthalmoscopy (CSLO), and stereoscopic ONH photography evaluated by four glaucoma specialists, four general ophthalmologists, four optometrists, and four junior ophthalmology residents

	Sensitivity (%)	Specificity (%)	Correctly classified (%)
SLP-VCC			
NFI, cut-off 30	95.8	92.5	94.3
NFI, cut-off 35	91.7	95.0	93.2
NFI, cut-off 40	83.3	97.5	89.8
CSLO			
Bathija LDF	85.4	95.0	89.8
MRA 1 (WNL vs B + ONL)	77.1	97.5	86.4
MRA 2 (WNL + B vs ONL)	64.6	100.0	80.7
Stereoscopic ONH photographs			
Glaucoma specialists (SD)	80.2 (4.9)	94.4 (5.5)	86.7 (5.0)
General ophthalmologists (SD)	78.7 (9.5)	93.1 (6.3)	85.2 (2.5)
Optometrists (SD)	77.1 (10.5)	86.9 (6.6)	81.5 (4.1)
Ophthalmology residents (SD)	69.3 (9.5)	77.5 (15.9)	73.0 (5.5)

B, Borderline; Bathija LDF, Linear Discriminant Function by Bathija *et al*;[9] MRA, Moorfields Regression Analysis; NFI, Nerve Fiber Indicator; ONH, optic nerve head; ONL, outside normal limits; SD, standard deviation; WNL, within normal limits.

technique or grader, eyes with mild glaucomatous loss were more difficult to classify correctly than eyes with more damage.

The diagnostic accuracy of the GDx VCC that we found presently was similar to what we reported earlier,[6] which may be due to our including some of the same eyes in both studies. In a different sample, Medeiros *et al*[10] reported the area under the ROC curve for the GDx VCC NFI parameter to be 0.91,[10] which is slightly smaller than what we presently found (that is 0.98). This difference may be due to the larger proportion of more advanced visual field loss in our study than in theirs. Similarly, we found the ability of the HRT to discriminate between healthy and glaucomatous eyes by the Bathija LDF and the MRA to be slightly higher than reported by both Medeiros *et al*[10] (Bathija LDF) and Ford *et al* (Bathija LDF and MRA).[14] This difference may again be due to their including relatively milder glaucomatous damage than we did. The difference in the diagnostic accuracy of the GDx VCC appeared to be higher than that of the HRT, which was also reported by Medeiros *et al*.[10] However, in neither study was this difference statistically significant. A larger sample size might yield a statistically significant difference.

We found that more experienced graders (that is, glaucoma specialists and general ophthalmologists vs. ophthalmology residents and optometrists) classified ONHs more accurately. These findings correspond well with findings by Varma *et al*[1] and Abrams *et al*[2] in a smaller analysis of 18 glaucomatous discs and 15 healthy ones.

The diagnostic accuracy of glaucoma specialists that we presently found was similar to that found by Greaney *et al*.[3] In contrast, both Wollstein *et al*[4] and Girkin *et al*[5] found a slightly lower sensitivity and specificity for glaucoma specialists than we did, which might be due to their including more glaucoma eyes with milder damage.

In our study, optometrists better classified (although not statistically significantly) ONHs than did junior ophthalmology residents. Abrams *et al*[2] had found the opposite. In their study, however, the graders were senior residents, who were more likely to correctly classify ONHs than the junior residents in our study.

Subjective grading of the ONH plays a key role in the clinical evaluation of a patient for glaucoma, together with all other available clinical information, for example, presence of visual field defects and level of IOP. We have presently shown that automated analysis of measurements with SLP-VCC and CSLO may better discriminate between healthy and glaucomatous eyes than general ophthalmologists and less experienced eye-care professionals who grade the appearance of ONHs. Therefore, we think that the clinical diagnosis of glaucoma may improve when they add imaging techniques to their diagnostic armamentarium. We also think that these techniques may assist them in learning how to correctly judge ONHs.

We found substantial variability in classifying the ONH as normal or glaucomatous both within and between graders. The agreement within an observer and between observers was only

Table 3 Level of statistical significance for differences in correctly classifying healthy and glaucoma eyes between scanning laser polarimetry with variable corneal compensation (SLP-VCC), confocal scanning laser ophthalmoscopy (CSLO), and stereoscopic ONH photography evaluated by four glaucoma specialists, four general ophthalmologists, four optometrists, and four junior ophthalmology residents

	SLP-VCC NFI	CSLO MRA 1	CSLO MRA 2	CSLO Bathija	Glaucoma specialists	General ophthalmologists	Optometrists	Residents
SLP-VCC, NFI	N/A	0.15	**0.019**	0.55	0.55	0.12	0.064	**0.004**
CSLO, MRA 1		N/A	0.13	0.51	0.38	1.00	0.58	0.092
CSLO, MRA 2			N/A	0.057	**0.008**	0.29	0.75	0.77
CSLO, Bathija				N/A	1.00	0.39	0.21	**0.021**
Glaucoma specialists					N/A	0.22	0.070	**0.006**
General ophthalmologists						N/A	0.69	**0.031**
Optometrists							N/A	0.29
Residents								N/A

Bathija LDF, Linear Discriminant Function by Bathija *et al*;[9] MRA, Moorfields Regression Analysis (1, within normal limits (WNL) versus borderline (B) + outside normal limits (ONL); 2, WNL + B vs ONL); NFI, Nerve Fiber Indicator.

Table 4 Sensitivity and specificity for discriminating between healthy eyes and eyes with mild, moderate, and severe glaucomatous damage (classified with the Mean Deviation [MD] and the Hodapp-Anderson-Parrish [HAP] grading scales) by scanning laser polarimetry with variable corneal compensation (SLP-VCC), confocal scanning laser ophthalmoscopy (CSLO), and stereoscopic ONH photography evaluated by four glaucoma specialists, four general ophthalmologists, four optometrists, and four junior ophthalmology residents

		Sensitivity (%)					
		MD			HAP		
	Specificity (%)	Mild (n = 27)	Moderate (n = 13)	Severe (n = 8)	Mild (n = 17)	Moderate (n = 10)	Severe (n = 21)
SLP-VCC							
NFI, cut-off 30	92.5	92.6	100.0	100.0	88.2	100.0	100.0
NFI, cut-off 35	95.0	85.2	100.0	100.0	76.5	100.0	100.0
NFI, cut-off 40	97.5	77.8	92.3	87.5	64.7	100.0	90.5
CSLO							
MRA 1 (WNL vs B + ONL)	97.5	66.7	84.6	100.0	64.7	60.0	95.2
MRA 2 (WNL + B vs ONL)	100.0	51.9	76.9	87.5	52.9	40.0	85.7
Bathija LDF	95.0	77.8	92.3	100.0	76.5	80.0	95.2
Stereoscopic ONH photographs							
Glaucoma specialists (SD)	94.4 (5.5)	65.8 (8.2)	98.1 (3.8)	100.0 (0.0)	60.3 (11.2)	75.0 (5.8)	98.8 (2.4)
General ophthalmologists (SD)	93.1 (6.3)	63.0 (15.4)	98.1 (3.8)	100.0 (0.0)	58.8 (14.4)	75.0 (12.9)	96.4 (4.6)
Optometrists (SD)	86.9 (6.6)	64.8 (12.3)	90.4 (9.7)	96.9 (6.3)	66.2 (10.1)	65.0 (23.8)	91.7 (6.0)
Ophthalmology residents (SD)	77.5 (15.9)	52.8 (12.6)	86.5 (18.2)	96.9 (6.3)	54.4 (13.0)	52.5 (15.0)	89.3 (6.0)

B, borderline; Bathija LDF, Linear Discriminant Function by Bathija *et al*;[9] MRA, Moorfields Regression Analysis; NFI, Nerve Fiber Indicator; ONH, optic nerve head; ONL, outside normal limits; SD, standard deviation; WNL, within normal limits.

moderate to good, which corresponds with previously reports.[1–4] Both intra- and interobserver agreement increased with the expected level of experience in judging ONHs.

Subjective grading of the ONH photographs agreed better with the Moorfields regression analysis than with the other automated analyses, which may not be very surprising because both relate strongly to ONH morphology. This also suggests that the MRA may be of little added diagnostic value to the clinician. Conversely, the GDx VCC NFI parameter relates to only RNFL morphology instead of ONH morphology and the HRT Bathija LDF relates to both, which may perhaps explain the lower agreement between the subjective ONH grading and these

Table 5 Agreement (κ) between classification of healthy and glaucoma eyes by evaluation of stereoscopic ONH photographs by four glaucoma specialists, four general ophthalmologists, four optometrists, and four junior residents in ophthalmology and scanning laser polarimetry with variable corneal compensation (SLP-VCC), confocal scanning laser ophthalmoscopy (CSLO) within various groups

	A	B	C	D	SLP-VCC, NFI (35)	CSLO, MRA 1	CSLO, Bathija LDF
Glaucoma specialists							
A	–	0.66	0.84	0.70	0.71	0.79	0.73
B	–	–	0.77	0.68	0.59	0.68	0.61
C	–	–	–	0.77	0.82	0.86	0.84
D	–	–	–	–	0.68	0.77	0.66
General ophthalmologists							
A	–	0.69	0.79	0.84	0.66	0.75	0.68
B	–	–	0.75	0.58	0.75	0.75	0.68
C	–	–	–	0.68	0.68	0.82	0.75
D	–	–	–	–	0.64	0.76	0.66
Optometrists							
A	–	0.66	0.58	0.62	0.63	0.64	0.66
B	–	–	0.68	0.73	0.57	0.66	0.55
C	–	–	–	0.72	0.44	0.55	0.50
D	–	–	–	–	0.66	0.75	0.68
Ophthalmology residents							
A	–	0.29	0.33	0.47	0.36	0.34	0.39
B	–	–	0.52	0.51	0.44	0.55	0.45
C	–	–	–	0.60	0.48	0.65	0.59
D	–	–	–	–	0.57	0.57	0.55

Bathija LDF, Linear Discriminant Function by Bathija *et al*;[9] MRA, Moorfields Regression Analysis; NFI, Nerve Fiber Indicator.

instrument based classifiers. Therefore, their classifiers may yield clinicians more additional information on the patient's eyes than does the MRA.

To our knowledge, the present study is the first to compare the diagnostic accuracy for glaucoma of subjective grading of stereoscopic ONH photographs by various eye-care professionals to that of automated analysis by the GDx VCC and the HRT. More research is needed to determine the effect of implementing imaging techniques in routine clinical practice.

In conclusion, we have shown that automated analysis of commercially available imaging techniques, such as the GDx VCC and the HRT, discriminate well between glaucomatous and healthy eyes compared with eye-care professionals classifying stereoscopic ONH photographs, suggesting that these objective techniques may assist clinicians in glaucoma management. The use of imaging devices has the added benefit of objectively documenting the ONH and/or RNFL for follow-up.

ACKNOWLEDGEMENTS
The authors would like to thank the glaucoma specialists, general ophthalmologists, ophthalmology residents, and optometrists who participated in the present study.

Authors' affiliations
Nicolaas J Reus, Maartje de Graaf, Hans G Lemij, Glaucoma Service, The Rotterdam Eye Hospital, Rotterdam, The Netherlands

Funding: The Rotterdam Eye Hospital Research Foundation, Rotterdam, The Netherlands; Stichting Glaucoomfonds, Leiden, The Netherlands. The funding sources did not have any involvement in study design; in the collection, analysis, and interpretation of data; in the writing of the report; nor in the decision to submit the paper for publication.

Competing interests: NJR was reimbursed for travelling expenses by Laser Diagnostic Technologies, Inc. (now: Carl Zeiss Meditec, Inc.) for speaking at several symposia. MdG has no competing interests. HGL was reimbursed for travelling expenses by Laser Diagnostic Technologies, Inc. (now: Carl Zeiss Meditec, Inc.) for speaking at several symposia and had a research fellow paid for, in part, by Laser Diagnostic Technologies, Inc. (now: Carl Zeiss Meditec, Inc.). NJR and HGL used a GDx VCC that was on loan by Laser Diagnostic Technologies, Inc. (now: Carl Zeiss Meditec, Inc.).

Meeting presentation: Presented as a poster at the 2005 annual meeting of the Association for Research in Vision and Ophthalmology.

REFERENCES
1 **Varma R**, Steinmann WC, Scott IU. Expert agreement in evaluating the optic disc for glaucoma. *Ophthalmology* 1992;**99**:215–21.
2 **Abrams LS**, Scott IU, Spaeth GL, *et al.* Agreement among optometrists, ophthalmologists, and residents in evaluating the optic disc for glaucoma. *Ophthalmology* 1994;**101**:1662–7.
3 **Greaney MJ**, Hoffman DC, Garway-Heath DF, *et al.* Comparison of optic nerve imaging methods to distinguish normal eyes from those with glaucoma. *Invest Ophthalmol Vis Sci* 2002;**43**:140–5.
4 **Wollstein G**, Garway-Heath DF, Fontana L, *et al.* Identifying early glaucomatous changes. Comparison between expert clinical assessment of optic disc photographs and confocal scanning ophthalmoscopy. *Ophthalmology* 2000;**107**:2272–7.
5 **Girkin CA**, McGwin G Jr, Long C, *et al.* Subjective and objective optic nerve assessment in African Americans and whites. *Invest Ophthalmol Vis Sci* 2004;**45**:2272–8.
6 **Reus NJ**, Lemij HG. Diagnostic accuracy of the GDx VCC for glaucoma. *Ophthalmology* 2004;**111**:1860–5.
7 **Weinreb RN**, Dreher AW, Bille JF. Quantitative assessment of the optic nerve head with the laser tomographic scanner. *Int Ophthalmol* 1989;**13**:25–9.
8 **Wollstein G**, Garway-Heath DF, Hitchings RA. Identification of early glaucoma cases with the scanning laser ophthalmoscope. *Ophthalmology* 1998;**105**:1557–63.
9 **Bathija R**, Zangwill L, Berry CC, *et al.* Detection of early glaucomatous structural damage with confocal scanning laser tomography. *J Glaucoma* 1998;**7**:121–7.
10 **Medeiros FA**, Zangwill LM, Bowd C, *et al.* 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. *Arch Ophthalmol* 2004;**122**:827–37.
11 **Altman DG**. Practical Statistics for Medical Research, 1st ed. London: Chapman and Hall, 1991.
12 **Hodapp E**, Parrish RK, Anderson DR. Clinical Decisions in Glaucoma. St. Louis, MO: Mosby, 1993.
13 **DeLong ER**, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. *Biometrics* 1988;**44**:837–45.
14 **Ford BA**, Artes PH, McCormick TA, *et al.* Comparison of data analysis tools for detection of glaucoma with the Heidelberg Retina Tomograph. *Ophthalmology* 2003;**110**:1145–50.