Published in final edited form as:
Invest Ophthalmol Vis Sci. 2006 August ; 47(8): . doi:10.1167/iovs.05-1239.

Discrimination between Glaucomatous and Nonglaucomatous Eyes Using Quantitative Imaging Devices and Subjective Optic Nerve Head Assessment

Julio E. DeLeón-Ortega[1], **Stella N. Arthur**[1], **Gerald McGwin Jr**[1,2], **Aiyuan Xie**[1], **Blythe E. Monheit**[1], and **Christopher A. Girkin**[1]

[1]Department of Ophthalmology, School of Medicine, University of Alabama at Birmingham, Birmingham, Alabama

[2]Department of Epidemiology and International Health, School of Public Health, University of Alabama at Birmingham, Birmingham, Alabama

Abstract

Purpose—To compare the diagnostic ability of the confocal scanning laser ophthalmoscope (HRT-II; Heidelberg Engineering, Heidelberg, Germany), scanning laser polarimeter (GDx-VCC; Carl Zeiss Meditec, Inc., Dublin, CA), and optical coherence tomographer (StratusOCT, Carl Zeiss Meditec, Inc.) with subjective assessment of optic nerve head (ONH) stereophotographs in discriminating glaucomatous from nonglaucomatous eyes.

Methods—Data from 79 glaucomatous and 149 normal eyes of 228 subjects were included in the analysis. Three independent graders evaluated ONH stereophotographs. Receiver operating characteristic curves were constructed for each technique and sensitivity was estimated at 80% of specificity. Comparisons of areas under these curves (aROC) and agreement (κ) were determined between stereophoto grading and best parameter from each technique.

Results—Stereophotograph grading had the largest aROC and sensitivity (0.903, 77.22%) in comparison with the best parameter from each technique: HRT-II global cup-to-disc area ratio (0.861, 75.95%); GDx-VCC Nerve Fiber Indicator (NFI; 0.836, 68.35%); and StratusOCT retinal nerve fiber layer (RNFL) thickness (0.844, 69.62%), ONH vertical integrated rim area (VIRA; 0.854, 73.42%), and macular thickness (0.815, 67.09%). The κ between photograph grading and imaging parameters was 0.71 for StratusOCT-VIRA, 0.57 for HRT-II cup-to-disc area ratio, 0.51 for GDX-VCC NFI, 0.33 for StratusOCT RNFL, and 0.28 for StratusOCT macular thickness.

Conclusions—Similar diagnostic ability was found for all imaging techniques, but none demonstrated superiority to subjective assessment of the ONH. Agreement between disease classification with subjective assessment of ONH and imaging techniques was greater for techniques that evaluate ONH topography than with techniques that evaluate RNFL parameters. A combination of subjective ONH evaluation with RNFL parameters provides additive information, may have clinical impact, and deserves to be considered in the design of future studies comparing objective techniques with subjective evaluation by general eye care providers.

Corresponding author: Christopher A. Girkin, UAB Department of Ophthalmology, 700 South 18th Street, Suite 406, Birmingham, AL 35233; cgirkin@uab.edu.

Disclosure: **J.E. DeLeón-Ortega**, None; **S.N. Arthur**, None; **G. Mc- Gwin, Jr**, None; **A. Xie**, None; **B.E. Monheit**, None; **C.A. Girkin**, None

In an effort to detect and document changes indicative of glaucoma at earlier stages, a variety of techniques have evolved to provide quantitative estimates of optic nerve head (ONH) topography, retinal nerve fiber layer (RNFL) thickness, and macular thickness. Several studies have demonstrated that in prior versions each of some of these techniques discriminate between glaucomatous and normal populations with a high degree of sensitivity and specificity.[1–5] However, few studies have compared the diagnostic performance of these instruments in the same study population and in comparison to subjective assessment of the ONH.[1,3] These prior studies used older versions of the instruments and demonstrated that objective optic nerve imaging modalities were equivalent to subjective assessment performed by masked expert stereophotograph graders.

Since the earlier studies, significant modifications have occurred with each of these quantitative imaging techniques that have improved their ability to detect glaucomatous damage. With scanning laser polarimetry (GDx-VCC; Carl Zeiss Meditec, Inc., Dublin, CA), a conversion to variable corneal compensation that provides individualized adjustment of anterior segment birefringence has improved the sensitivity and specificity of this technique.[6,7] The current generation of optical coherence tomography (StratusOCT; Carl Zeiss Meditec, Inc.) has increased scan rate and scan resolution and can also be used to obtain macular and ONH measurements.[8] The most recent version of the confocal scanning laser ophthalmoscope (HRT II; Heidelberg Engineering, Heidelberg, Germany) has been modified significantly with automation of the examination procedure focused on optic disc topography, which has improved the reproducibility and efficacy of this instrument in the detection of glaucoma.[2,4,9,10]

A recent study by Medeiros et al.[11] using the current version of these instruments has demonstrated that each performs with similar efficacy in the diagnosis of glaucoma. However, there was no comparison between these new versions of objective imaging methods with subjective ONH evaluation or objective evaluation of the optic nerve head with ONH analysis and macular thickness from StratusOCT. The purpose of this study was to compare the diagnostic ability of the confocal scanning laser ophthalmoscopy with the HRT II, scanning laser polarimetry with the GDX-VCC, and retinal nerve fiber layer thickness, ONH analysis and macular thickness measurements with the StratusOCT with subjective masked expert assessment of stereophotographs in the same study population.

Methods

Data were obtained from the University of Alabama at Birmingham (UAB) Optic Nerve Imaging Center database, consisting of 124 subjects with glaucoma and 149 normal subjects who had undergone optic disc imaging and visual functional testing between January 2003 and February 2005 as part of ongoing longitudinal glaucoma studies. Patients were obtained from the UAB glaucoma service, and control subjects were obtained primarily from referrals and UAB employees. Informed consent was obtained from all participants, and the University of Alabama at Birmingham Human Subjects Committee approved the methodology. All aspects of the protocol adhered to the tenets of the Declaration of Helsinki. The study eye was randomly selected. All subjects had a complete ophthalmic examination, including slit lamp biomicroscopy, intraocular pressure (IOP) measurement, standard visual field testing, stereoscopic fundus examination, simultaneous stereoscopic photographs of the optic discs, and HRT II, GDx VCC, and Stratus OCT imaging. All testing was completed within 1 to 8 weeks (4.5 ± 3.5 weeks; mean ± 1 SD).

Normal participants were included if they had bilateral highest documented IOP of ≤22 mm Hg; bilateral normal eye examination findings, including dilated fundus examination; and bilateral normal visual field results defined as pattern standard deviation (PSD) within the

95% normal limits and a glaucoma hemifield test (GHT) result within 99% limits. Glaucomatous visual field loss was defined as PSD outside 95% normal limits or GHT outside 99% normal limits, confirmed with a second visual field test. Patients with a mean defect of ≤ −15 dB were excluded. In addition, subjects were excluded if they had best corrected visual acuity worse than 20/40, spherical refraction outside ±5.0 D and cylinder refraction outside ± 3.0 D, or unreliable visual fields (fixation losses and false-positive and -negative responses exceeding 30%) or were using medications known to affect visual sensitivity at the time of visual field testing, those with comorbid ophthalmic or neurologic surgery or disease and those with inadequate imaging–photograph quality were also excluded.

Simultaneous ONH stereophotographs were obtained after dilation of the pupil with a fundus camera (3-Dx; Nidek Technology America, Inc., Greensboro NC). Three masked experienced observers (CAG, BEM, JED-O) independently graded the ONH photographs according to a 5-point scale similar to the method described by Greaney et al.[3] and Girkin et al.[12] The grade increases with the clinical impression of glaucoma (1, definitely normal; 2, probably normal; 3, unsure; 4, probably glaucomatous; and 5, definitely glaucomatous). Criteria for classification was on the basis of the observing typical ONH characteristics consistent with glaucomatous optic neuropathy including the presence of neuroretinal rim thinning, notching, or undermining, nerve fiber layer defects, and optic disc hemorrhages. An overall photograph grade score was developed by the summation of scores to produce a 15-point scale. Furthermore, to compare this grading scale with more commonly used forced-choice grading; the stereophotographs were also graded in a dichotomous manner (glaucoma or normal) by two of the graders (JED-O, BEM), while the third grader (CAG) adjudicated cases of disagreement. The quality of the stereophotograph was evaluated at the time of masked grading. One participant had suboptimal quality on the stereophotograph and stereophotography was repeated, improving its quality.

ONH topography was obtained with the confocal scanning laser ophthalmoscope (HRT II). Details of the HRT II operation have been described before.[13–15] Experienced operators (JED-O, SNA) evaluated image quality and outlined the disc margin while viewing stereoscopic photographs of the optic disc. The HRT II software-determined parameters of RNFL thickness, RNFL cross-sectional area, rim area and volume, mean height contour, cup area and volume, cup shape, mean cup depth, maximum cup depth, optic disc area, and cup-to-disc area ratio and results from discriminant analysis formulas developed by Bathija et al.,[16] Mardin et al.,[17] and Mikelberg et al.[18] were included in the analysis. Topographies with acquisition sensitivity above 90%, SD greater than 40 μm, ONH not centered, excessive movement during acquisition, floaters over or adjacent to the ONH, poor clarity of image or framing were excluded (21 subjects). Most were excluded because of inadequate scan sensitivity and poor clarity (possibly because of reduced media clarity). The remaining had vitreous floaters adjacent to the disc or inadequate framing.

RNFL thickness was evaluated with both the StratusOCT and the GDx VCC (both Carl Zeiss Meditec Inc.). Details of StratusOCT operation have been described elsewhere.[1,8,13,19,20] RNFL thickness measurements consisted of four quadrants and 12 sectors around the ONH (in clock hours). On left eyes (and transposed data from right eyes), these sectors corresponded to superior (11–1 o'clock), temporal (2–4 o'clock), inferior (5–7 o'clock) and nasal (8–10 o'clock) quadrants. The landmark option was used, and images were excluded if the video image had a poor quality, if the scan beam was not centered on the ONH, and if the subject was unable to maintain stable fixation. Twelve subjects were excluded.

Details of the GDX operation have been described previously,[20,6,21] and parameters included were TSNIT (temporal, superior, nasal, inferior, and temporal) average, superior average, inferior average, TSNIT SD, intereye symmetry, superior and inferior average, superior and inferior ratio, superior nasal, maximum, inferior and ellipse modulation, superior and inferior normalized area, and nerve fiber indicator (NFI). NFI is a support vector machine derived parameter indicating the likelihood that an eye has glaucoma.[22] The mean of three 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 artifacts on the retardance image. Ten subjects were excluded because of poor image quality.

Automated ONH measurements were obtained with the StratusOCT, using the ONH protocol of six 4-mm radial line scans centered on the ONH. The mean from three images was used for the analysis, and the parameters included were vertical integrated rim area, horizontal integrated rim width, cup and rim area, cup-to-disc area ratio, and vertical and horizontal cup-to-disc ratio. Images were excluded if the video image was of inadequate quality, the ONH was not properly centered, and fixation losses were present. Thirteen subjects were excluded because of poor image quality.

Macular thickness was obtained with the StratusOCT, using six 6-mm radial-line scans centered on the fovea. Three measurements were obtained, and the mean was determined for each of nine locations (fovea, temporal inner and outer macula, superior inner and outer macula, nasal inner and outer macula, and inferior inner and outer macula). Similar quality-control criteria as used in the ONH and RNFL protocols were implemented, and 12 subjects were excluded.

Including control subjects with previously normal eye examination results is likely to bias the results in favor of subjective assessment of the optic disc. Although classification of control eyes using visual field criterion alone and ignoring clinical examination results would be optimal in comparing stereophotos with quantitative imaging, this strategy may create some degree of misclassification bias with the inclusion of preperimetric glaucoma. To investigate, we explored the comparison in two separate analyses. One using control subjects defined by normal visual fields and a normal eye examination and a secondary analysis using control subjects defined by visual fields alone. For this secondary analysis, we added 29 additional eyes from 29 participants enrolled as control subjects during the same period. These subjects were added to the control group because they had reliable normal visual fields, but we ignored any information from the initial dilated fundus examination. During the screening interview, the subjects self-reported as having no ocular disease, as did the 149 control subjects originally enrolled.

Two-tailed *t*-tests were used to compare glaucomatous and normal eyes with respect to continuous variables with normal distribution based on histograms plots. Nonparametric (Wilcoxon) tests were used for continuous variables not distributed normally. Similar group comparisons were conducted for categorical variables by χ^2 tests.

To compare the relative ability of each imaging technique in discriminating glaucomatous from normal eyes, the area under the receiver operating characteristic curve (aROC) adjusted for age was calculated for each of the quantitative imaging parameters and for the overall stereophotograph grade. The aROC provides an evaluation of the ability to discriminate between those who experience the outcome of interest and those who do not. A logistic regression model containing the stereophotograph grade was compared to a model contained the most efficient (i.e., largest aROC) imaging parameters for each instrument. Sensitivity was assessed fixed at 80% of specificity. Statistical comparisons of the aROC were performed using previously described methods,[23,24] adjusting for multiple comparisons with Bonferroni correction.

To compare the agreement in disease classification between the techniques, upper and lower limits from the control group for each best parameter were defined as two standard deviations from the control population mean, if the data had a normal distribution as determined with histograms plots or within the 2.5 to 97.5 percentiles otherwise. Eyes that fell outside these normal limits were classified as glaucomatous and agreement with dichotomous classification from stereophotograph grading was evaluated with linear κ statistics. The strength of agreement was interpreted as follows: 0.0, no agreement; <0.40, fair agreement; 0.40 to 0.59, moderate agreement; 0.60 to 0.75, good agreement; >0.75 to 0.99, excellent agreement; and 1.0, perfect agreement.[25] Statistical analyses were performed on computer (SAS and JMP; SAS Institute, Inc., Cary, NC).

Results

In our main analysis, data from 79 glaucomatous and 149 normal subjects met the inclusion criteria. Patients with glaucoma were older than control subjects (mean ± 1 SD = 56.0 ± 13.9 years in the glaucoma group and 40.3 ± 11.3 years in the control group; $P < 0.001$). For this reason, age-adjusted aROCs were used in comparing the imaging techniques and subjective assessment. In addition, patients and control subjects with African ancestry and female gender predominated in both study groups (African-American in the glaucoma group, 46 [58.2%], and in the control group, 82 [55.0%]); women in the glaucoma group, 54 [68.4%], and in the control group, 111 [74.5%]). However, univariate analysis showed no race or gender differences between the control and glaucoma groups ($P = 0.67$ and 0.35, respectively). There was a difference in spherical refraction between the two study groups (−0.4 ± 1.9 D in the glaucoma group, −1.1 ± 1.5 D in the control group; $P = 0.0074$). The difference in cylindrical refraction was of borderline significance (0.6 ± 0.7 D, glaucoma group; 0.4 ± 0.6 D, control group; $P = 0.040$). However, regression analysis showed no evidence that refraction state (sphere, cylinder, or spherical equivalent) was significantly associated with any of the imaging parameters. Refraction was not included in any further analysis. No differences were found in optic disc area between glaucoma and control group (disc area with HRT II for glaucoma group = 2.5 ± 0.5 mm^2, for control group = 2.2 ± 0.5 mm^2, $P = 0.082$; disc area with StratusOCT for glaucoma group = 2.5 ± 0.5 mm^2, for control group 2.4 ± 0.4 mm^2, $P = 0.12$).

The glaucoma group had an average mean deviation (MD ± SD) of −3.8 ± 3.6 dB and control group had an average MD of 0.2 ± 1.0 dB. Furthermore, 44 (55.7%) eyes had an early defect, 31 (39.2%) eyes had a moderate defect, and 4 (5.1%) eyes had a severe defect, according to the classification of severity of field loss by Hodapp et al..[26] Thus, the cohort with glaucomatous eyes predominantly had early to moderate visual field defects.

The stereophotograph grading obtained from the 15-point likelihood score scale fell within a range of 3 to 10 for the control group (mean 5.2 ± 1.9 SD) and 75% of control eyes received a score from 3 to 6. In the glaucoma group, the range of score was from 3 to 15 (mean, 10.6 ± 3.7 SD) and 70% of glaucomatous eyes received a score from 9 to 15. The mean of scores significantly differed between the two study groups ($P < 0.0001$). The 15-point likelihood score for stereophotograph grading had an overall sensitivity of 77.2% fixed at ≥80.0% specificity (aROC = 0.903; SE = 0.03), adjusted for age.

Significant differences were found between control and glaucomatous eyes in GDx VCC, HRT II, and StratusOCT measurements (Table 1 GDx VCC, Table 2 global HRT II, and Table 3 StratusOCT). Also shown are the aROC and sensitivities at fixed specificity of at least 80%. Data for sectoral HRT II parameters are not shown because of space limitations. To compare diagnostic methods, we selected the parameter with the largest aROC from each technique. For GDx, the NFI showed the largest aROC (0.836). For the HRT II, the best global parameter was cup-to-disc area ratio, with an aROC of 0.861 and the best sectoral HRT II parameter was temporal inferior cup volume with a comparable aROC (0.854). For the StratusOCT, the ONH analysis parameter with the largest aROC was the vertical integrated rim area (0.854). Similarly, the RNFL thickness at the inferior quadrant (6:00 sector) had the largest aROC (0.844). Last, for the macular thickness, the largest aROC was obtained from thickness at the superior outer macular location (0.815).

Figure 1 shows the age-adjusted ROC curves for each best parameter compared with stereophotograph grading. Compared with all measures, the stereophotograph grading showed the largest aROC, in both age-adjusted and crude analysis. Table 4 shows pair-wise comparisons of aROC between stereophotograph grading and each imaging technique. After Bonferroni correction, a borderline significance was found between stereophotograph grading and HRT global cup-to-disc area ratio, StratusOCT ONH (vertical integrated rim area), and StratusOCT RNFL thickness in the inferior quadrant, whereas macular thickness and GDx VCC NFI had significantly lower aROCs ($P = 0.001$ and 0.007, respectively). No significant differences in aROC were found between each imaging technique, suggesting HRT II (global and sectoral), StratusOCT (ONH, RNFL, and macular thickness), and GDx VCC were equivalent in their diagnostic ability (data not shown because of space restrictions).

Agreement of dichotomous classification between stereophotographs and classification from the best parameter of each imaging technique was good between stereophotograph grading and StratusOCT vertical integrated rim area, moderate between stereophotograph grading and HRT II global cup-to-disc area ratio and GDx-VCC NFI, and fair between stereophotograph grading and StratusOCT RNFL and macular thickness. Venn diagrams (Fig. 2) illustrate the results of disease classification by each technique. Classification based on subjective assessment of the optic disc (stereophotos) was compared against imaging techniques measuring optic disc topography (Fig. 2A) and separately against techniques that assessed RNFL integrity (Fig. 2B). Subjective assessment correctly identified more patients with glaucoma than did objective methods. As demonstrated by the levels of agreement in Table 5, techniques that evaluated optic disc characteristics had higher levels of agreement with subjective assessment of the optic disc than did techniques that evaluated RNFL parameters. Overall, the combination of subjective assessment of the optic disc with RNFL associated parameters (Fig. 2A) correctly identified more subjects (70/79 eyes) than did the combination of subjective disc assessment with optic disc parameters (Fig. 2B; 65/79 eyes).

Similar results were found in the secondary analysis using control subjects defined only by visual field tests, ignoring the information from the initial dilated fundus examination. Results are presented as age-adjusted aROC (SE, percentage of sensitivity at 80% specificity): for stereophotograph grading 0.879 (0.03, 76.25), HRT II cup-to-disc ratio 0.849 (0.03, 75.00), GDx VCC NFI 0.843 (0.03, 70.00), StratusOCT ONH VIRA 0.848 (0.03, 69.23), StratusOCT RNFL 0.836 (0.03, 69.23), and StratusOCT macular thickness 0.826 (0.03, 67.95). Stereophoto grading still provided significantly greater diagnostic efficacy that the other modalities.

Discussion

The present study compared the diagnostic performance of quantitative estimates of ONH, RNFL thickness and macular thickness with subjective evaluation of the optic nerve, using the same population. In agreement with previous studies,[1,3,11] similar diagnostic performance was observed for all three instruments (HRT II, GDx VCC, and StratusOCT) when using the best parameter from each instrument on the basis of largest aROC. Furthermore, the best performance was with subjective stereophotograph grading and confocal laser scanning ophthalmoscopy. Although StratusOCT (both ONH and RNFL parameters) performed less well than did subjective stereophotograph grading, these differences in aROC were of borderline significance after correction for multiple comparisons. GDx VCC and macular thickness performed significantly less well than subjective stereophotograph grading.

Compared with the previous study using the newer generation of these instruments,[11] consistent results were found on some of the best parameters from each instrument. For example, the largest aROC found for GDx VCC was NFI and the largest aROC found for StratusOCT was RNFL thickness at inferior sector. For the HRT II, however, global parameters such as cup-to-disc area ratio gave the largest aROC. This is in contrast to the high aROC, sensitivity, and specificity reported for linear discriminant functions (Bathija et al.,[16] Mikelberg et al.,[18] and Mardin et al.[17]). These discrepancies may reflect differences between study populations in the present study and that in others and demonstrates the need to evaluate these discriminant functions in study populations independent from the study populations from which they were developed.

The aROCs, sensitivities, and specificities reported in the present study are lower than the ones reported by Greaney et al.[3] These lower values may reflect differences in enrollment criteria for control subjects. The present study obtained control subjects from those seeking eye care from referring practices or from employees who obtain eye care services at our facility, which may better approximate the source population for the cases compared with other studies, which have demonstrated greater diagnostic performance. If so, this study would more accurately reflect the performance of these techniques in the screening setting.

Although expert consensus assessment of masked stereophotographs demonstrated the highest performance in discriminating glaucomatous and nonglaucomatous eyes, this result may not reflect optic nerve head assessment in clinical practice. Level of training has been shown to affect stereophotograph grading, where glaucoma experts performed better than optometrists, residents, and general ophthalmologists.[27] In addition, the optic nerve head evaluation in clinical practice is often performed in the busy clinical setting with less time for thorough assessment. Thus, this study may overestimate the diagnostic performance of subjective optic nerve head assessment as reflected in the primary care setting. Furthermore, subjects recruited as part of the control group are likely to have nonglaucomatous optic disc appearance because of the low prevalence of glaucoma in the normal population,[28] which was the source population for the control subjects.

Macular thickness had a significantly worse discriminatory ability than stereophotograph grading and the other imaging techniques. This finding is consistent with a study by Medeiros et al.[29] A probable explanation is that most of our glaucoma eyes had early to moderate visual field loss, and more advanced cases may produce greater changes at the macular thickness than the changes observed in our study population.

A substantial agreement was found between stereophotograph grading and results from StratusOCT ONH analysis, suggesting these two techniques detect similar characteristics of the optic nerve head and further supports the usefulness of StratusOCT technique. Furthermore, a greater number of eyes were correctly classified by combining RNFL parameters (GDx VCC NFI and StratusOCT RNFL at the temporal inferior region) with subjective disc assessment than by combining ONH parameters (StratusOCT ONH analysis and HRT II global cup-to-disc area ratio) with subjective disc assessment. This finding probably reflects the higher level of agreement between subjective disc evaluation and ONH imaging methods. Thus, much of the information obtained from ONH imaging may be assessed subjectively, whereas the RNFL measurements may provide a greater degree of additive information when combined with subjective assessment. This further information obtained from the combination of subjective evaluation of ONH and objective quantitative evaluation of RNLF may have a clinical impact, and such combination should be considered in the design of future studies, particularly when the subjective evaluation is perform by general eye care providers.

Ideally, classification of control eyes using visual field criterion alone would be optimal in comparing the results of stereophotos with quantitative imaging results. However, this strategy may create some degree of misclassification bias with the inclusion of preperimetric glaucoma. Alternatively, including control subjects with previously normal eye examination results is likely to bias the results in favor of subjective assessment of the optic disc. To investigate, we explored the comparison using both control subjects defined by normal visual fields and a normal eye examination and those defined by normal visual fields alone, adding the subjects who were eliminated based on abnormal-appearing optic discs. Although the diagnostic performance was predictably worse with all techniques when the field alone defined control subjects, we found that subjective assessment of the optic disc performed with significantly greater efficacy than did quantitative optic disc imaging.

In conclusion, subjective assessment of the ONH provided the best diagnostic efficacy in the detection of glaucoma defined by visual field defects alone. After correcting for multiple comparisons, differences between subjective ONH assessment and StratusOCT (ONH analysis and RNFL thickness) and HRT II were of borderline significance, whereas GDx VCC and macular thickness performed significantly less well than subjective ONH assessment. However, the optimal parameters from these newer generations of quantitative imaging techniques to detect glaucoma did not differ significantly in their discriminatory ability when compared with each other. In addition, combining RNFL measurements with subjective optic nerve head assessment correctly diagnosed more subjects with glaucoma than did the combination of optic nerve head topography with subjective disc assessment. Studies comparing objective techniques with subjective evaluation by general eye care providers will help to determine the usefulness of these techniques in clinical settings.

Acknowledgments

Supported by National Eye Institute Grants K23-EY 13959-0 (CAG), R21-EY14071 (AX), and 1U10-EY 14267-01 (JED-O, CAG), and the Eyesight Foundation of Alabama (JED-O, SNA).

References

1. Zangwill LM, Bowd C, Berry CC, et al. Discriminating between normal and glaucomatous eyes using the Heidelberg Retina Tomograph, GDx Nerve Fiber Analyzer, and Optical Coherence Tomograph. Arch Ophthalmol. 2001; 119:985–993. [PubMed: 11448320]
2. Wollstein G, Garway-Heath DF, Fontana L, Hitchings RA. Identifying early glaucomatous changes: comparison between expert clinical assessment of optic disc photographs and confocal scanning ophthalmoscopy. Ophthalmology. 2000; 107:2272–2277. [PubMed: 11097609]
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; 43:140–145. [PubMed: 11773024]
4. Miglior S, Guareschi M, Albe E, et al. Detection of glaucomatous visual field changes using the Moorfields regression analysis of the Heidelberg retina tomograph. Am J Ophthalmol. 2003; 136:26–33. [PubMed: 12834666]

5. Kanamori A, Nagai-Kusuhara A, Escano MF, et al. Comparison of confocal scanning laser ophthalmoscopy, scanning laser polarimetry and optical coherence tomography to discriminate ocular hypertension and glaucoma at an early stage. Graefes Arch Clin Exp Ophthalmol. 2006; 244:58–68. [PubMed: 16044326]
6. Weinreb RN, Bowd C, Greenfield DS, Zangwill LM. Measurement of the magnitude and axis of corneal polarization with scanning laser polarimetry. Arch Ophthalmol. 2002; 120:901–906. [PubMed: 12096960]
7. Zhou Q, Weinreb RN. Individualized compensation of anterior segment birefringence during scanning laser polarimetry. Invest Ophthalmol Vis Sci. 2002; 43:2221–2228. [PubMed: 12091420]
8. Jaffe GJ, Caprioli J. Optical coherence tomography to detect and manage retinal disease and glaucoma. Am J Ophthalmol. 2004; 137:156–169. [PubMed: 14700659]
9. Miglior S, Casula M, Guareschi M, et al. Clinical ability of Heidelberg retinal tomograph examination to detect glaucomatous visual field changes. Ophthalmology. 2001; 108:1621–1627. [PubMed: 11535460]
10. Wollstein G, Garway-Heath DF, Hitchings RA. Identification of early glaucoma cases with the scanning laser ophthalmoscope. Ophthalmology. 1998; 105:1557–1563. [PubMed: 9709774]
11. 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. Arch Ophthalmol. 2004; 122:827–837. [PubMed: 15197057]
12. 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; 45:2272–2278. [PubMed: 15223805]
13. Stein DM, Wollstein G, Schuman JS. Imaging in glaucoma. Ophthalmol Clin North Am. 2004; 17:33–52. [PubMed: 15102512]
14. Chauhan DS, Marshall J. The interpretation of optical coherence tomography images of the retina. Invest Ophthalmol Vis Sci. 1999; 40:2332–2342. [PubMed: 10476800]
15. Girkin CA, McGwin G Jr, McNeal SF, DeLeon-Ortega J. Racial differences in the association between optic disc topography and early glaucoma. Invest Ophthalmol Vis Sci. 2003; 44:3382–3387. [PubMed: 12882785]
16. Bathija R, Zangwill L, Berry CC, et al. Detection of early glaucomatous structural damage with confocal scanning laser tomography. J Glaucoma. 1998; 7:121–127. [PubMed: 9559499]
17. Mardin CY, Horn FK, Jonas JB, Budde WM. Preperimetric glaucoma diagnosis by confocal scanning laser tomography of the optic disc. Br J Ophthalmol. 1999; 83:299–304. [PubMed: 10365037]
18. Mikelberg FS, Parfitt CM, Swindale NV, et al. Ability of the Heidelberg Retina Tomograph to detect early glaucomatous visual field loss. J Glaucoma. 1995; 4:242–247. [PubMed: 19920681]
19. Blumenthal EZ, Williams JM, Weinreb RN, et al. Reproducibility of nerve fiber layer thickness measurements by use of optical coherence tomography. Ophthalmology. 2000; 107:2278–2282. [PubMed: 11097610]
20. Zangwill LM, Chang CF, Williams JM, Weinreb RN. New technologies for diagnosing and monitoring glaucomatous optic neuropathy. Optom Vis Sci. 1999; 76:526–536. [PubMed: 10472958]
21. Weinreb RN, Bowd C, Zangwill LM. Scanning laser polarimetry in monkey eyes using variable corneal polarization compensation. J Glaucoma. 2002; 11:378–384. [PubMed: 12362075]
22. Reus NJ, Lemij HG. Diagnostic accuracy of the GDx VCC for glaucoma. Ophthalmology. 2004; 111:1860–1865. [PubMed: 15465547]
23. DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. Biometrics. 1988; 44:837–845. [PubMed: 3203132]
24. Hanley JA, McNeil BJ. A method of comparing the areas under receiver operator characteristic curve derived from the same cases. Radiology. 1983:839–843. [PubMed: 6878708]
25. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics. 1977; 33:159–174. [PubMed: 843571]

26. Hodapp, E.; Parrish, RK., II; Anderson, DR. Clinical Decision in Glaucoma. St. Louis: Mosby-Year Book; 1993.
27. Abrams L, Scott IU, Spaeth GL, et al. Agreement among optometrists, ophthalmologist, and residents in evaluating the optic disc for glaucoma. Ophthalmology. 1994; 101:1662–1667. [PubMed: 7936564]
28. Quigley HA. Number of people with glaucoma worldwide. Br J Ophthalmol. 1996; 80:389–393. [PubMed: 8695555]
29. Medeiros FA, Zangwill LM, Bowd C, et al. Evaluation of retinal nerve fiber layer, optic nerve head, and macular thickness measurements for glaucoma detection using optical coherence tomography. Am J Ophthalmol. 2005; 139:44–55. [PubMed: 15652827]

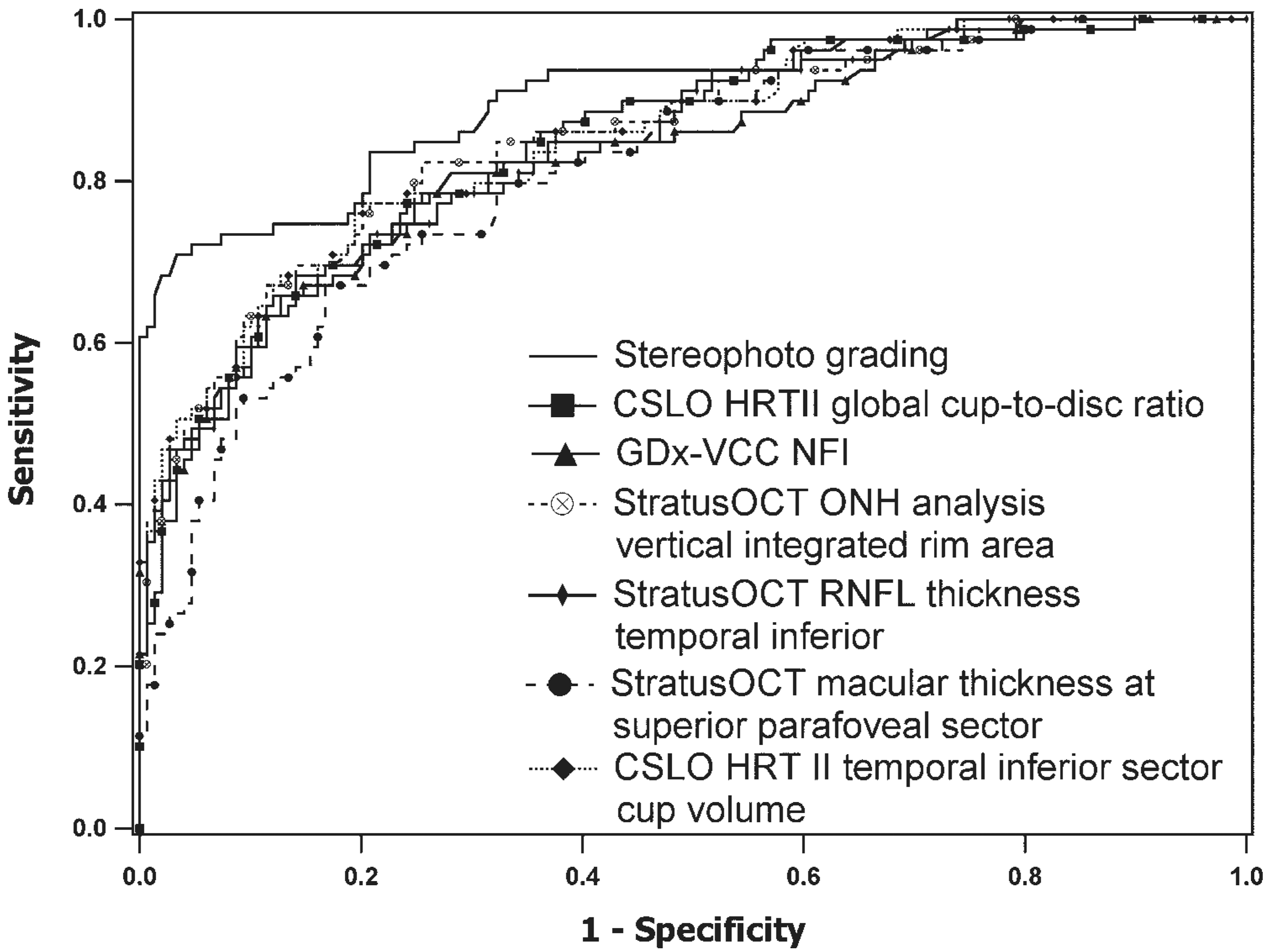

Figure 1.
ROC curves were constructed for stereophotograph grading and for the best parameter from each imaging technique.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

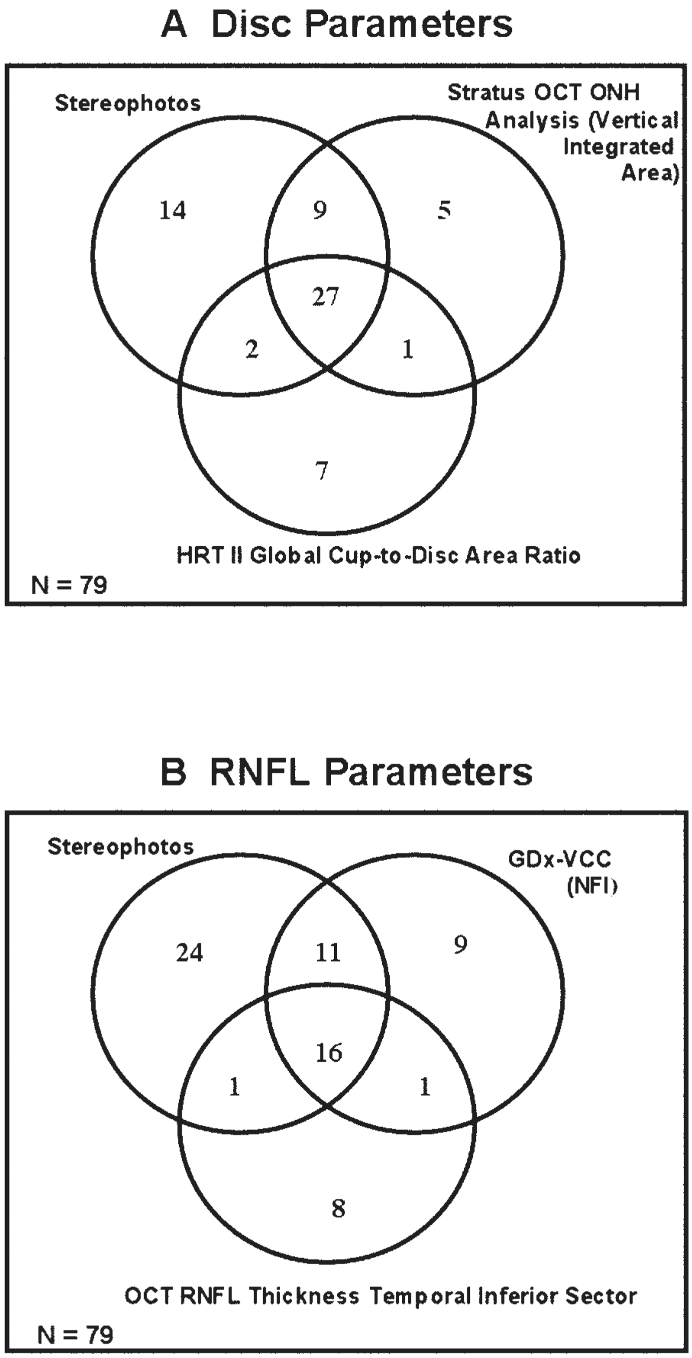

Figure 2.

Venn diagrams showing the number of eyes classified as glaucomatous by stereophotograph grading and by the best parameter from each imaging technique. Imaging data were divided into (**A**) disc and (**B**) retinal nerve fiber layer (RNFL) parameters. *n* = 79 glaucomatous eyes, on the basis of standard perimetry alone. None of the techniques correctly identified all glaucomatous eyes. A combination of stereophotograph grading+GDx-VCC NFI +StratusOCT RNFL identified most glaucomatous eyes (70/79).

Table 1

GDx-VCC Data

	Glaucoma*	Normal*	*P*	aROC (SE)	Sensitivity at ≥80% Specificity
NFI	33.9 ± 21.7	16.7 ± 7.3	<0.0001	0.836 (0.03)	68.35
TSNIT average	48.7 ± 8.8	56.0 ± 5.4	<0.0001	0.831 (0.03)	72.15
Inferior normalized area	0.11 ± 0.04	0.14 ± 0.03	<0.0001	0.831 (0.03)	69.8
Inferior average	55.3 ± 12.2	65.4 ± 8.1	<0.0001	0.829 (0.03)	65.82
Inferior maximum	70.0 ± 15.0	82.0 ± 11.0	<0.0001	0.826 (0.03)	64.56
Superior average	58.8 ± 12.5	68.5 ± 7.8	<0.0001	0.818 (0.03)	70.89
Superior normalized area	0.11 ± 0.04	0.14 ± 0.02	<0.0001	0.818 (0.03)	69.62
TSNIT SD	19.4 ± 5.9	24.0 ± 4.3	<0.0001	0.815 (0.03)	63.29
Superior maximum	70.0 ± 15.5	81.0 ± 10.5	<0.0001	0.814 (0.03)	65.82
Maximum modulation	2.4 ± 1.2	2.9 ± 1.2	0.0092	0.812 (0.04)	64.56
Ellipse modulation	3.8 ± 1.7	4.3 ± 1.8	0.0302	0.812 (0.04)	64.26
Superior ratio	3.0 ± 1.2	3.6 ± 1.2	0.0015	0.810 (0.04)	63.29
Inferior ratio	3.0 ± 1.2	3.6 ± 1.1	0.0008	0.810 (0.04)	63.29

*Data are the mean ± 1 SD.

NIH-PA Author Manuscript
NIH-PA Author Manuscript

Table 2

CSLO (HRT II) Global Data

	Glaucoma*	Normal*	*P*	aROC (SE)	Sensitivity at ≥80% Specificity
Cup-disc area ratio	0.4 ± 0.2	0.2 ± 0.1	<0.0001	0.861 (0.03)	75.95
Rim-disc area ratio	0.6 ± 0.2	0.8 ± 0.1	<0.0001	0.861 (0.03)	75.95
Linear cup-to-disc ratio	0.6 ± 0.2	0.4 ± 0.1	<0.0001	0.858 (0.03)	77.22
Cup volume (mm^3)	0.3 ± 0.3	0.2 ± 0.1	<0.0001	0.857 (0.03)	74.68
Maximum contour elevation	−0.0 ± 0.1	−0.1 ± 0.1	<0.0001	0.857 (0.03)	72.22
Mean cup-depth (mm)	0.3 ± 0.1	0.2 ± 0.1	<0.0001	0.856 (0.04)	78.4
Cup area (mm^2)	0.9 ± 0.6	0.5 ± 0.3	<0.0001	0.854 (0.04)	74.68
Vertical cup-disk ratio	1.0 ± 0.2	0.3 ± 0.2	<0.0001	0.854 (0.03)	75.95
Horizontal cup-disk ratio	1.0 ± 0.2	0.4 ± 0.2	<0.0001	0.851 (0.04)	74.68
Mardin LDF†	−1.4 ± 2.4	0.8 ± 1.5	<0.0001	0.851 (0.03)	74.68
Maximum cup depth (mm)	1.0 ± 0.2	1.0 ± 0.2	<0.0001	0.844 (0.04)	74.68
Bathija et al. LDF	0.1 ± 1.8	1.4 ± 1.0	<0.0001	0.844 (0.03)	73.42
Rim area (mm^2)	1.5 ± 0.5	1.7 ± 0.4	<0.0001	0.84 (0.03)	75.69
Maximum contour depression	0.3 ± 0.1	0.3 ± 0.1	<0.0001	0.836 (0.04)	73.42
Rim volume (mm^3)	0.4 ± 0.2	0.5 ± 0.2	<0.0001	0.833 (0.03)	67.09
Reference height (mm)	0.4 ± 0.1	0.3 ± 0.1	0.0019	0.832 (0.04)	74.68
Cup shape measure	−0.2 ± 0.1	−0.2 ± 0.1	<0.0001	0.826 (0.04)	70.89
RB discriminant function value	1.0 ± 1.2	2.0 ± 1.0	<0.0001	0.826 (0.03)	66.67
Mikelberg et al. LDF	5.9 ± 4.7	8.6 ± 3.9	<0.0001	0.823 (0.04)	62.03
FSM discriminant function value	0.1 ± 2.4	2.3 ± 2.1	<0.0001	0.82 (0.03)	65.85
CLM temporal superior (mm)	0.2 ± 0.1	0.2 ± 0.1	0.0015	0.813 (0.04)	64.56
Mean RNFL thickness (mm)	0.2 ± 0.1	0.3 ± 0.1	0.0003	0.807 (0.04)	64.56
RNFL cross-sectional area	1.2 ± 1.0	1.4 ± 0.3	0.002	0.807 (0.04)	64.56

* Data is displayed as mean ± 1 SD.

† LDF, Linear discriminant function.

Table 3

StratusOCT Data: Optic Nerve Head, Macular Thickness, and RNFL Analysis

	Glaucoma*	Normal*	*P*	aROC (SE)	Sensitivity at ≥80% Specificity
Optic Nerve Head Analysis					
Vertical rim area	0.3 ± 0.2	0.6 ± 0.3	<0.0001	0.854 (0.03)	73.42
Horizontal rim width	1.5 ± 0.3	2.0 ± 0.3	<0.0001	0.850 (0.03)	72.12
Cup area	1.1 ± 1.0	1.0 ± 0.4	<0.0001	0.844 (0.04)	72.15
Cup-to-disc area ratio	0.5 ± 0.2	0.3 ± 0.2	<0.0001	0.844 (0.03)	74.68
Cup-to-disc horizontal ratio	1.0 ± 0.2	1.0 ± 0.2	<0.0001	0.841 (0.04)	69.62
Cup-to-disc vertical ratio	0.6 ± 0.2	0.5 ± 0.2	<0.0001	0.841 (0.03)	68.35
Rim area	1.3 ± 0.5	2.0 ± 1.0	<0.0001	0.822 (0.03)	70.89
Macular Thickness					
Superior outer macula	221.0 ± 18.4	234.0 ± 15.0	<0.0001	0.815 (0.03)	67.09
Nasal outer macula	236.0 ± 23.0	252.0 ± 17.0	<0.0001	0.814 (0.03)	68.35
Inferior outer macula	217.0 ± 20.2	228.4 ± 14.4	<0.0001	0.813 (0.04)	64.56
Superior inner macula	258.1 ± 23.0	269.2 ± 23.0	0.0001	0.811 (0.04)	63.29
Inferior inner macula	250.4 ± 25.5	264.0 ± 15.0	<0.0001	0.811 (0.04)	68.35
Temporal inner macula	247.0 ± 24.0	254.3 ± 14.5	0.0076	0.81 (0.04)	60.76
Temporal outer macula	213.1 ± 19.0	218.0 ± 13.4	0.0385	0.81 (0.04)	64.56
Nasal inner macula	256.0 ± 25.0	268.0 ± 16.0	0.0002	0.809 (0.04)	62.03
RNFL					
RNFL at 6 o’clock sector	121.6 ± 33.6	152.0 ± 25.1	<0.0001	0.844 (0.03)	69.62
Inferior quadrant	113.0 ± 29.0	137.0 ± 18.0	<0.0001	0.842 (0.03)	70.89
Nasal quadrant	70.0 ± 17.4	85.1 ± 15.0	<0.0001	0.826 (0.03)	69.62
Averaged RNFL	102.3 ± 26.1	123.2 ± 26.1	<0.0001	0.826 (0.03)	70.89
Superior quadrant	103.1 ± 27.0	124.4 ± 17.0	<0.0001	0.816 (0.03)	64.56
Temporal quadrant	64.0 ± 16.0	70.0 ± 14.0	0.0035	0.809 (0.04)	63.29

*Data are the mean ± 1 SD.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Table 4

Comparisons of aROC between Stereophoto Grading and Best Parameter from Quantitative Imaging Measures

	aROC	Sensitivity at ≥80% Specificity	*P*
Stereo photo grading	0.903* (0.884)†	77.22 (75.95)‡	NA
CSLO HRT II: global cup-to-disc area ratio	0.861 (0.795)	75.95 (64.56)	0.013
GDx VCC: NFI	0.836 (0.773)	68.35 (56.96)	0.007
StratusOCT optic nerve head analysis: vertical integrated rim area	0.854 (0.822)	73.42 (64.56)	0.011
StratusOCT macular thickness: superior outer sector	0.815 (0.705)	67.09 (51.90)	0.001
StratusOCT RNFL thickness: temporal inferior (6 o'clock)	0.844 (0.759)	69.62 (56.96)	0.015

Data in parenthesis correspond to unadjusted area under ROC curve (†) and sensitivity (‡) to same level of specificity as in the age-adjusted model. Bonferroni correction, $\alpha = 0.01$, Five comparisons

*Adjusted for age.

Table 5

Agreement between Stereophoto Grading and Cut-Off Point from Best Parameter of Each Quantitative Imaging Technique in Classifying Eyes Either as Glaucomatous or Normal

	StratusOCT ONH Analysis: Vertical Integrated Rim Area	HRT II Global Cup-To-Disc Area Ratio	GDx-VCC NFI	StratusOCT RNFL Temporal Inferior Sector	StratusOCT Macular Thickness at Parafoveal Superior Location
Photo grading	0.71 (0.07)	0.57 (0.07)	0.51 (0.07)	0.33 (0.03)	0.28 (0.07)
StratusOCT ONH analysis: vertical integrated rim area		0.65 (0.07)	0.57 (0.07)	0.35 (0.08)	0.32 (0.08)
HRT II global cup-to-disc area ratio			0.48 (0.08)	0.35 (0.09)	0.21 (0.08)
GDx-VCC NFI				0.47 (0.08)	0.40 (0.09)
StratusOCT RNFL temporal inferior sector					0.07 (0.07)

Data are kappa statistic and standard error in parentheses.