Detecting Early Glaucoma by Assessment of Retinal Nerve Fiber Layer Thickness and Visual Function

Christopher Bowd,[1] *Linda M. Zangwill,*[1] *Charles C. Berry,*[2] *Eytan Z. Blumenthal,*[1] *Cristiana Vasile,*[1] *Cesar Sanchez-Galeana,*[1] *Charles F. Bosworth,*[1] *Pamela A. Sample,*[1] *and Robert N. Weinreb*[1]

Purpose. To compare the abilities of scanning laser polarimetry (SLP), optical coherence tomography (OCT), short-wavelength automated perimetry (SWAP), and frequency-doubling technology (FDT) perimetry to discriminate between healthy eyes and those with early glaucoma, classified based on standard automated perimetry (SAP) and optic disc appearance. To determine the agreement among instruments for classifying eyes as glaucomatous.

Methods. One eye of each of 94 subjects was included. Healthy eyes (n = 38) had both normal-appearing optic discs and normal SAP results. Glaucoma by SAP (n = 42) required a repeatable abnormal result (glaucoma hemifield test [GHT] or corrected pattern standard deviation [CPSD] outside normal limits). Glaucoma by disc appearance (n = 51) was based on masked stereoscopic photograph evaluation. Receiver operating characteristic (ROC) curve areas, sensitivities, and specificities were calculated for each instrument separately for each diagnosis.

Results. The largest area under the ROC curve was found for OCT inferior quadrant thickness (0.91 for diagnosis based on SAP, 0.89 for diagnosis based on disc appearance), followed by the FDT number of total deviation plot points of ≤5% (0.88 and 0.87, respectively), SLP linear discriminant function (0.79 and 0.81, respectively), and SWAP PSD (0.78 and 0.76, respectively). For diagnosis based on SAP, the ROC curve area was significantly larger for OCT than for SLP and SWAP. For diagnosis based on disc appearance, the ROC curve area was significantly larger for OCT than for SWAP. For both diagnostic criteria, at specificities of ≥90% and ≥70%, the most sensitive OCT parameter was more sensitive than the most sensitive SWAP and SLP parameters. For diagnosis based on SAP, the most sensitive FDT parameter was more sensitive than the most sensitive SLP parameter at specificities of ≥90% and ≥70% and was more sensitive than the most sensitive SWAP parameter at specificity of ≥70%. For diagnosis based on disc appearance at specificity of ≥90%, the most sensitive FDT parameter was more sensitive than the most sensitive SWAP and SLP parameters. At specificity ≥ 90%, agreement among instruments for classifying eyes as glaucomatous was poor.

Conclusions. In general, areas under the ROC curve were largest (although not always significantly so) for OCT parameters, followed by FDT, SLP, and SWAP, regardless of the definition of glaucoma used. The most sensitive OCT and FDT parameters tended to be more sensitive than the most sensitive SWAP and SLP parameters at the specificities investigated, regardless of diagnostic criteria. (*Invest Ophthalmol Vis Sci.* 2001;42:1993–2003)

In some patients with glaucoma, standard (achromatic) automated perimetry (SAP) does not detect visual field defects until approximately 30% to 50% of retinal ganglion cell axons have been lost.[1–3] Further, glaucoma often may not be detected until substantial retinal nerve fiber layer (RNFL) damage has occurred.[1,4–8] Clearly, there is a compelling need for more sensitive glaucoma diagnostic tests.

Several diagnostic techniques to assess the RNFL and visual function have been introduced recently to aid in the early diagnosis and monitoring of patients with diagnosed or suspected glaucoma. The objective of this study was to compare the diagnostic ability of two methods for quantitatively assessing the RNFL (scanning laser polarimetry [SLP] and optical coherence tomography [OCT]), and two retinal ganglion cell–specific methods for testing visual function (short-wavelength automated perimetry [SWAP] and frequency-doubling technology [FDT] perimetry) in glaucomatous and healthy eyes in a single-sample population. The advantage of examining the diagnostic performance of these instruments in a single population is that population-characteristic–based variables are eliminated, thus allowing direct comparison of results obtained with the different instruments.

There may be biases in evaluating the diagnostic ability of RNFL imaging when using optic disc appearance as a gold standard or in evaluating visual function when using SAP as a gold standard. Our goal was therefore to evaluate the diagnostic precision of RNFL assessment for detecting eyes with glaucomatous visual function defects and to evaluate the diagnostic precision of visual function assessment for detecting eyes with a glaucomatous optic disc appearance. To directly compare results from structural and functional tests in the same sample population, we also evaluated RNFL assessment, with optic disc appearance as the diagnostic criterion, and examined visual function tests, with SAP as the diagnostic criterion.

Methods

Subjects

Ninety-four healthy subjects or patients with glaucoma, prospectively enrolled as longitudinal study participants at the Glaucoma Center, University of California, San Diego, were included. Each had all required diagnostic testing and examinations (described later) within 1 year (70 [75%] of 94 had all tests within 3 months, and 76 [80%] of 94 had all tests within 6 months). One eye of each subject was randomly

From the [1]Glaucoma Center, Department of Ophthalmology and [2]Department of Family and Preventive Medicine, University of California San Diego, La Jolla.

Supported by National Institutes of Health Grants EY11008 (LMZ) and EY08208 (PAS), the Glaucoma Research Foundation (PAS), the Research to Prevent Blindness Lew R. Wasserman award (PAS), and the Foundation for Eye Research (EZB, CV).

Submitted for publication September 21, 2000; revised April 9, 2001; accepted April 26, 2001.

Commercial relationships policy: N.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Corresponding author: Linda M. Zangwill, Glaucoma Center and Diagnostic Imaging Laboratory, University of California, San Diego, 9500 Gilman Drive, La Jolla, CA 92093-0946.
zangwill@eyecenter.ucsd.edu

selected for study. Informed consent was obtained from each participant and the University of California San Diego Human Subjects Committee approved all protocols. All methods adhered to the Declaration of Helsinki for research involving human subjects.

All subject eyes had open angles, best corrected acuity of 20/40 or better, sphere within ±5.0 diopters (D), and cylinder within ± 3.0 D at time of testing. Subjects had no history of diabetes or other systemic disease and no reported ophthalmic or neurologic surgery or other diseases affecting visual fields or color vision. All visual function tests were reliable (≤25% false positives, false negatives, fixation losses) and all RNFL images obtained were judged to be of acceptable quality by experienced operators.

Healthy eyes in this study ($n = 38$) had a measured IOP of 22 mm Hg or less with no history of elevated IOP. These eyes had healthy-appearing optic discs, based on masked consensus grading of simultaneous stereoscopic photographs by two expert graders, and SAP results within normal limits. Average ± SD healthy subject age was 58.13 ± 12.18 years.

In determining diagnostic sensitivity for RNFL measurements, glaucoma was diagnosed based on SAP results. The 42 eyes in the group with diagnosis by SAP had repeatable abnormal SAP results (either glaucoma hemifield test [GHT] results or corrected pattern standard deviation [CPSD] outside normal limits). Optic disc appearance and IOP were not used as diagnostic criteria. Eyes in this group had a mean ± SD mean deviation (MD) of −4.0 ± 4.2 dB, indicating primarily early glaucoma. The patients' mean age was 64.4 ± 11.7 years.

When determining diagnostic sensitivity of visual function tests, glaucoma was diagnosed based on optic disc appearance. These 51 eyes had either focal rim notching, rim thinning, excavation of the rim, or RNFL defects. SAP results and IOP were not used as diagnostic criteria. Eyes in this group had a mean ± SD MD of −3.5 ± 4.0 dB indicating primarily early glaucoma. The patients' mean age ± SD was 63.2 ± 11.9 years. Diagnosis in 37 eyes overlapped, and these were included in both the diagnosis-by-disc-appearance and diagnosis-by-SAP result groups. Mean SAP MD and mean age were not significantly different between the two glaucoma diagnosis groups ($P > 0.05$).

There were no significant differences in age (ANOVA, $P > 0.05$) or ethnic origin among the healthy subjects, patients with glaucoma diagnosed by disc appearance, and those with glaucoma diagnosed by SAP results. Ninety-one percent of subjects were white, 3% were Asian, 3% were Hispanic, and 2% were African American.

To directly compare results and assess bias from RNFL-based techniques and visual function techniques, we also evaluated the diagnostic ability of RNFL tests in healthy subjects and patients with glaucoma diagnosed by disc appearance, and we evaluated the diagnostic ability of visual function tests in healthy subjects and patients with glaucoma diagnosed by SAP results.

Instrumentation

Scanning Laser Polarimetry. The scanning laser polarimeter (GDx Nerve Fiber Analyzer; Laser Diagnostic Technologies, San Diego, CA) uses confocal scanning diode technology coupled with an integrated polarization modulator to measure retardation of light that has double passed the birefringent fibers of the RNFL. Retardation measurements have been shown to correlate with RNFL thickness measurements.[9] Details of this instrument and descriptions of parameters have been provided elsewhere.[10–14]

We examined 27 parameters automatically provided by the GDx nerve fiber analyzer software (ver. 2.0.01; Laser Diagnostic Technologies). Parameters investigated were: GDx Number (neural network result), average thickness, ellipse average, ellipse modulation, inferior average thickness, inferior integral, inferior maximum-nasal median ratio, inferior maximum thickness, inferior-nasal integral ratio, inferior-nasal mean ratio, inferior ratio, inferior-temporal integral ratio, inferior-temporal mean ratio, maximum modulation, superior average thickness, superior-inferior integral ratio, superior-inferior mean ratio, superior integral, superior maximum thickness, superior-nasal integral ratio, superior-nasal mean ratio, superior-nasal ratio, superior ratio, superior-temporal integral ratio, superior-temporal mean ratio, symmetry, and total polar integral. Finally, we examined the value of a discriminant analysis model (linear discriminant function [LDF]) proposed by Weinreb et al.[14] [LDF = −4.442655 − (0.156 · average thickness) + (0.935 · ellipse modulation) + (0.183 · ellipse average)]. All parameters investigated are shown in Tables 1 and 2.

Three scans (approximately 15° field of view) centered on the optic disc were obtained for each test eye. A mean retardation map comprising these three scans was created using the nerve fiber analyzer software. The optic disc margin was outlined on the mean retardation image by a trained technician.

Optical Coherence Tomography. The optical coherence tomograph (OCT 2000; Humphrey-Zeiss Instruments, Dublin, CA) uses low-coherence interferometry to assess peripapillary RNFL thickness. This instrument measures RNFL thickness by measuring the difference in temporal delay of back-scattered light from the RNFL and a reference mirror. RNFL is differentiated from other retinal layers using an edge detection algorithm (version A4X1). RNFL thickness is defined as the number of pixels between its anterior and posterior boundaries.

OCT parameters investigated in this study were mean RNFL thickness (360° measure), temporal quadrant thickness (316–45° unit circle), superior quadrant thickness (46–135°), nasal quadrant thickness (136–225°), inferior quadrant thickness (226–315°), and thickness measures at three superior clock hours (11, 12, and 1 o'clock, with 11o'clock located superior temporally) and three inferior clock hours (5, 6, and 7 o'clock, with 7 o'clock located inferior temporally). We also developed a modulation parameter (called max-min) calculated by subtracting the RNFL thickness measurement of the thinnest quadrant from the measurement of the thickest quadrant. This parameter is designed to assess the amplitude of the characteristic double-hump pattern of RNFL thickness and is analogous to SLP modulation parameters.

Three circular scans of 3.4-mm diameter centered on the optic disc were obtained for each test eye. This approximate scan diameter was found to be optimal for RNFL analysis in a prototype instrument.[15] Mean RNFL thickness values for quadrant and clock-hour measurements were determined from the three scans obtained.

Because OCT RNFL measures are taken nearer to the disc in subjects with larger discs because of the set radius of the circular scan, we examined the correlation between disc area and RNFL quadrant thickness measures (Pearson's r) and examined the difference in disc size between glaucomatous and healthy eyes (t-test) and found no significant results (all $P > 0.1$). Disc area was measured by confocal scanning laser tomograph (Heidelberg Retina Tomograph; Heidelberg Engineering, Heidelberg, Germany).

Short-Wavelength Automated Perimetry. SWAP is a modification of SAP (Humphrey Field Analyzer II, program 24-2; Humphrey Instruments, San Leandro, CA) in which a 440-nm narrow-band target is presented on a 100-candela (cd)/m^2 yellow adaptation field to selectively stress short-wavelength cones and small bistratified blue–yellow ganglion cells.[16] The same parameters and stimulus programs are used in SWAP as in SAP and have been discussed in detail elsewhere.[17]

SWAP parameters investigated in this study were MD, PSD, total number of abnormal points in the total deviation and pattern deviation plots (at $P \leq 0.05$ or worse and $P \leq 0.01$ or worse), number of abnormal points in the superior hemifield of the total and pattern deviation plots (at $P \leq 0.05$ or worse and $P \leq 0.01$ or worse), and number of abnormal points in the in the inferior hemifield of the total and pattern deviation plots (at $P \leq 0.05$ or worse and $P \leq 0.01$ or worse). Abnormality for SWAP parameters was determined by comparison to our normative database of 342 eyes. Only 6 of the 38 healthy eyes used in this study contributed to this database.

Frequency-Doubling Testing. FDT perimetry (Humphrey Visual Field Instrument, using Welch Allyn FDT, Skaneateles Falls, NY) is based on the frequency-doubling effect, in which a low-spatial-frequency sine-wave grating, undergoing high-temporal-frequency

TABLE 1. SLP Parameter ROC Curve Areas, Sensitivities, and Specificities for Detecting Glaucoma Based on Optic Disc and Standard Perimetry Criteria

	Glaucoma by Optic Disc Appearance			Glaucoma by Standard Perimetry		
SLP Parameter	ROC Area (± SE)	Sensitivity–Specificity (%) (Specificity ≥90%)	Sensitivity–Specificity (%) (Specificity ≥70%)	ROC Area (± SE)	Sensitivity–Specificity (%) (Specificity ≥90%)	Sensitivity–Specificity (%) (Specificity ≥70%)
GDx LDF	**0.81 ± 0.05**	**33/92**	**73/71**	**0.79 ± 0.05**	**33/92**	**71/71**
GDx Number	**0.79 ± 0.05**	**35/92**	**73/71**	**0.77 ± 0.05**	**33/92**	**71/71**
Inf max/nas median	**0.73 ± 0.05**	29/92	61/71	0.69 ± 0.06	26/92	55/71
Maximum modulation	**0.73 ± 0.06**	22/92	63/71	**0.70 ± 0.06**	21/92	38/71
Sup-nas ratio	**0.71 ± 0.06**	27/92	**67/71**	0.68 ± 0.06	26/92	**62/71**
Sup-temp mean ratio	0.70 ± 0.06	14/92	**69/71**	0.67 ± 0.06	10/92	52/71
Ellipse modulation	0.70 ± 0.06	25/92	47/71	0.65 ± 0.06	29/92	38/71
Sup-temp integ ratio	0.67 ± 0.06	8/92	**67/71**	0.66 ± 0.06	7/92	**62/71**
Sup ratio	0.67 ± 0.06	12/92	53/71	0.66 ± 0.06	10/92	**57/71**
Inf-temp mean ratio	0.66 ± 0.06	18/92	47/71	0.64 ± 0.06	14/92	45/71
Inf-temp integ ratio	0.66 ± 0.06	8/92	**67/71**	0.64 ± 0.06	7/92	45/71
Sup-nas mean ratio	0.66 ± 0.06	31/92	57/71	0.64 ± 0.06	**33/92**	52/71
Sup average	0.65 ± 0.06	**35/92**	55/71	0.65 ± 0.06	**40/92**	55/71
Sup integral	0.65 ± 0.06	23/92	53/71	0.68 ± 0.06	31/92	**57/71**
Sup-nas integ ratio	0.64 ± 0.06	14/92	45/71	0.64 ± 0.06	17/92	48/71
Inf-nas mean ratio	0.64 ± 0.06	10/92	57/71	0.65 ± 0.06	10/92	52/71
Inf ratio	0.62 ± 0.06	12/92	41/71	0.61 ± 0.06	10/92	45/71
Inf-nas integ ratio	0.62 ± 0.06	14/92	45/71	0.63 ± 0.06	17/92	48/71
Sup maximum	0.62 ± 0.06	**39/92**	51/71	0.63 ± 0.06	**36/92**	55/71
Inf maximum	0.61 ± 0.06	**33/92**	49/71	0.62 ± 0.06	**33/92**	48/71
Inf average	0.61 ± 0.06	29/92	47/71	0.61 ± 0.06	29/92	48/71
Ellipse average	0.60 ± 0.06	**35/92**	49/71	0.62 ± 0.06	**33/92**	50/71
Inf integral	0.60 ± 0.06	24/92	43/71	0.65 ± 0.06	26/92	48/71
Total polar integral	0.58 ± 0.06	27/92	49/71	0.63 ± 0.06	**33/92**	**57/71**
Sup-inf integ ratio	0.57 ± 0.06	23/92	41/71	0.54 ± 0.06	21/92	38/71
Sup-inf mean ratio	0.57 ± 0.06	26/92	47/71	0.54 ± 0.07	26/92	40/71
Average thickness	0.55 ± 0.06	29/92	41/74	0.56 ± 0.06	29/92	43/74
Symmetry	0.53 ± 0.06	20/92	35/71	0.53 ± 0.07	24/92	36/71

The three highest ROC areas and sensitivities at specified sensitivities for each parameter are in bold. Parameters are ordered based on highest to lowest ROC area when glaucoma was diagnosed by disc appearance. Inf, inferior; nas, nasal; sup, superior; temp, temporal; integ, integral.

counter-phase flicker, appears to have double its true spatial frequency. This effect has been attributed to processing by a subset of magnocellular ganglion cells with nonlinear response properties,[18] although there is evidence to show that at contrast threshold, all magnocellular cells are likely to be responsive to this type of stimuli.[19,20]

Target stimuli consist of individual 10° square, 0.25 cyc/deg sinusoidal gratings, counter-phasing at 25 Hz. Targets are presented in one of 17 test areas located within the central 20° of the visual field (threshold program C-20; Welch Allyn). This test measures the stimulus contrast detection threshold using a modified binary staircase procedure.

For each stimulus presentation, observers responded with a button-press when the stimulus was detected. Maximum stimulus duration using this instrument is 720 msec. During the first 160 msec, the stimulus contrast is ramped up from 0 to that selected for the presentation. If it is not immediately detected, the stimulus remains at the selected contrast for 400 msec and then is ramped down to 0 during the final 160 msec. The interstimulus interval is randomly selected up to 500 msec and target location is pseudorandomly selected for each presentation.

FDT parameters investigated in this study were the same as for SWAP. Abnormality for FDT parameters was determined by comparison to the manufacturer's internal normative database (Viewfinder version 1.02, program C-20; Welch-Allyn).

Data Analysis

For each measured parameter from each of the instruments, area under the ROC curve for discriminating between glaucomatous and healthy eyes was calculated. Differences between ROC curve areas among instruments were determined by the method of DeLong et al.[21] Sensitivity and specificity for detection of glaucomatous eyes was determined by obtaining the highest sensitivity values with a target specificity set at ≥90% and again with a target specificity set at ≥70%. Depending on the goals of the screening program and characteristics of the target population, either of these target specificities may be useful.[22] Differences between sensitivities at set specificities among instruments were determined using the McNemar test. ROC curve and sensitivity and specificity analyses were performed twice for each instrument, once using disc appearance as a criterion for a glaucoma classification and once using SAP results. For each instrument, parameter comparisons between glaucomatous and healthy eyes were performed using T-tests with Bonferroni-corrected α.

RESULTS

SLP: Glaucoma Diagnosed by SAP

For SLP parameters, areas under the ROC curve ranged from 0.79 to 0.53 when diagnosis was based on SAP (Table 1). The three largest areas under the ROC curves (ROC area, SE) were for GDx LDF (0.79, 0.05), GDx Number (0.77, 0.05), and maximum modulation (0.70, 0.06).

When target specificity was set at ≥90%, the parameters with the three highest sensitivities (sensitivity, specificity; Table 1) were superior average thickness (40%, 92%); superior maximum thickness (36%, 92%); and GDx LDF, GDx Number, superior-nasal mean ratio, inferior maximum thickness, ellipse

TABLE 2. Comparison of SLP Parameter Measures between Normal and Glaucomatous Eyes, Based on Optic Disc and Standard Perimetry Criteria

		Glaucoma by Optic Disc Appearance		Glaucoma by Standard Perimetry	
SLP Parameter	**Normal (n = 38)**	**Glaucoma (n = 51)**	**P***	**Glaucoma (n = 42)**	**P***
GDx LDF	0.375 ± 0.145	−0.580 ± 0.087	*<0.0001*	−0526 ± 0.105	*<0.0001*
GDx Number	19.3 ± 2.9	37.4 ± 2.5	*<0.0001*	35.4 ± 2.6	*<0.0001*
Inf max-nas median	2.02 ± 0.06	1.74 ± 0.04	*0.0001*	1.78 ± 0.05	*0.002*
Maximum modulation	1.50 ± 0.09	1.09 ± 0.04	*<0.0001*	1.13 ± 0.06	*0.0004*
Sup-nas ratio	2.04 ± 0.08	1.71 ± 0.04	*0.0002*	1.75 ± 0.06	*0.002*
Sup-temp mean ratio	1.88 ± 0.06	1.59 ± 0.05	*0.0004*	1.63 ± 0.04	0.003
Ellipse modulation	2.54 ± 0.13	2.00 ± 0.08	*0.0002*	2.09 ± 0.09	0.004
Sup-temp integ ratio	4.57 ± 0.16	3.91 ± 0.14	*0.002*	3.97 ± 0.15	0.009
Sup ratio	2.21 ± 0.10	1.88 ± 0.05	*0.001*	1.91 ± 0.06	0.006
Inf-temp mean ratio	1.91 ± 0.06	1.66 ± 0.05	*0.001*	1.69 ± 0.05	0.007
Inf-temp integ ratio	4.69 ± 0.19	4.11 ± 0.11	0.006	4.15 ± 0.12	0.019
Sup-nas mean ratio	1.57 ± 0.05	1.42 ± 0.03	0.009	1.43 ± 0.04	0.028
Sup average (μm)	78.74 ± 1.76	72.00 ± 2.42	0.037	71.80 ± 2.56	0.031
Sup integral	0.247 ± 0.008	0.223 ± 0.008	0.052	0.217 ± 0.009	0.016
Sup-nas integ ratio	2.75 ± 0.09	2.48 ± 0.07	0.023	2.49 ± 0.08	0.044
Inf-nas mean ratio	1.58 ± 0.04	1.47 ± 0.03	0.022	1.47 ± 0.03	0.030
Inf ratio	2.21 ± 0.09	1.92 ± 0.05	0.003	1.95 ± 0.06	0.015
Inf-nas integ ratio	2.81 ± 0.09	2.61 ± 0.07	0.061	2.60 ± 0.07	0.063
Sup maximum (μm)	90.50 ± 2.33	82.65 ± 2.86	0.046	82.31 ± 2.86	0.031
Inf maximum (μm)	90.57 ± 2.34	83.64 ± 2.51	0.054	83.34 ± 2.34	0.043
Inf average (μm)	80.21 ± 2.08	74.52 ± 2.06	0.060	73.79 ± 2.11	0.033
Ellipse average (μm)	69.51 ± 1.58	65.50 ± 1.92	0.128	64.89 ± 1.98	0.076
Inf integral	0.254 ± 0.009	0.223 ± 0.008	0.054	0.223 ± 0.007	0.008
Total polar integral	0.653 ± 0.022	0.607 ± 0.019	0.123	0.585 ± 0.021	0.028
Sup-inf integ ratio	0.981 ± 0.018	0.957 ± 0.019	0.382	0.966 ± 0.022	0.610
Sup-inf mean ratio	0.990 ± 0.018	0.967 ± 0.019	0.387	0.975 ± 0.022	0.599
Average thickness (μm)	65.9 ± 1.58	64.0 ± 1.90	0.475	63.5 ± 2.0	0.359
Symmetry	1.01 ± 0.02	0.99 ± 0.02	0.573	0.99 ± 0.02	0.693

Data are means ± SE. Significant differences between diagnostic groups (after Bonferroni correction) are italicized. Parameters are ordered as in Table 1. See Table 1 for abbreviations.

* Mean measurements in glaucomatous eyes compared with normal eyes. Bonferroni corrected $\alpha = 0.002$.

average thickness, and total polar integral (all 33%, 92%). When target specificity was set at ≥70%, the parameters with the three highest sensitivities were GDx LDF and the GDx Number (both 71%, 71%); superior-nasal ratio and superior-temporal integral ratio (both 62%, 71%); and superior ratio, superior integral, and total polar integral (all 57%, 71%).

When SLP parameters in eyes classified as glaucomatous based on SAP results were compared with healthy eyes, 5 of 28 parameters were significantly different (after Bonferroni correction, $\alpha = 0.002$) between groups in the predicted directions (lower GDx LDF result, higher GDx Number, with less thickness modulation in glaucomatous eyes; Table 2).

OCT: Glaucoma Diagnosed by SAP

For OCT parameters, areas under the ROC curve for diagnosis based on SAP ranged from 0.91 to 0.66 (Table 3). The three largest areas under the ROC curves (ROC area, SE) were for inferior quadrant thickness (0.91, 0.03), 6 o'clock thickness (0.90, 0.04), and mean thickness (0.89, 0.04).

When target specificity was set at ≥90%, the parameters with the three highest sensitivities (Table 3) were inferior quadrant thickness, thickness at 6 o'clock, and thickness at 7 o'clock (inferior temporal; all 79%, 92%). When target specificity was set at ≥70%, the parameters with the three highest sensitivities were inferior quadrant thickness (88%, 71%), mean thickness and thickness at 6 o'clock (both 86%, 71%), and thickness at 7 o'clock (83%, 71%).

When OCT parameters in eyes classified as glaucomatous based on SAP results were compared with healthy eyes, all RNFL thickness measures were significantly different (after Bonferroni correction, $\alpha = 0.004$) between groups (thinner RNFL measures in glaucomatous eyes) except for the max-min (modulation) parameter ($P = 0.01$; Table 4).

SWAP: Glaucoma Diagnosed by Optic Disc Appearance

Table 5 shows ROC curve areas and sensitivities and specificities for all SWAP parameters investigated. Areas under the ROC curve for diagnosis based on disc appearance ranged from 0.76 to 0.59. The three largest areas under the ROC curves (ROC area, SE) were for PSD (0.76, 0.05); total deviation plot points ≤1%, pattern deviation plot points ≤5%, and superior quadrant pattern deviation plot points ≤5% (all 0.74, 0.05); and pattern deviation plot points ≤1%, total deviation plot points ≤5%, and superior quadrant total deviation plot points ≤5% (all 0.73, 0.05).

When target specificity was set at ≥90%, the parameters with the three highest sensitivities were pattern deviation plot points ≤5% (43%, 92%), PSD and superior quadrant total deviation plot points ≤5% (both 41%, 92%), and total deviation plot points ≤1% (39%, 92%). When target specificity was set at ≥70%, the parameters with the three highest sensitivities were total deviation plot points ≤1%, superior quadrant pattern deviation points ≤5%, and superior quadrant total deviation plot points ≤5% (all 73%, 71%), superior quadrant total deviation plot points ≤1% and superior quadrant pattern deviation plot points ≤5% (both 63%, 74%), and PSD and pattern deviation plot points ≤1% (both 61%, 71%).

TABLE 3. OCT Parameter ROC Curve Areas, Sensitivities, and Specificities for Detecting Glaucoma Based on Optic Disc and Standard Perimetry Criteria

	Glaucoma by Optic Disc Appearance			Glaucoma by Standard Perimetry		
OCT Parameter	ROC Area (± SE)	Sensitivity–Specificity (%) (Specificity ≥90%)	Sensitivity–Specificity (%) (Specificity ≥70%)	ROC Area (± SE)	Sensitivity–Specificity (%) (Specificity ≥90%)	Sensitivity–Specificity (%) (Specificity ≥70%)
Inf thickness	**0.89 ± 0.03**	**69/92**	**88/71**	**0.91 ± 0.03**	**79/92**	**88/71**
6 o'clock thickness	**0.87 ± 0.04**	**75/92**	77/71	**0.90 ± 0.04**	**79/92**	**86/71**
Mean thickness	**0.85 ± 0.04**	**65/92**	**86/71**	**0.89 ± 0.04**	**74/92**	**86/71**
7 o'clock (inf-temp) thickness	0.84 ± 0.04	**65/92**	**79/71**	0.88 ± 0.04	**79/92**	**83/71**
Sup thickness	0.80 ± 0.05	59/92	75/71	0.82 ± 0.05	62/92	79/71
12 o'clock thickness	0.77 ± 0.05	45/92	67/71	0.78 ± 0.05	48/92	69/71
1 o'clock thickness	0.76 ± 0.05	20/92	75/71	0.78 ± 0.06	24/92	79/71
5 o'clock thickness	0.76 ± 0.05	37/92	71/71	0.78 ± 0.05	45/92	74/71
Temp thickness	0.75 ± 0.05	39/92	69/71	0.82 ± 0.05	52/92	81/71
11 o'clock thickness	0.72 ± 0.06	59/92	63/71	0.78 ± 0.05	**67/92**	69/71
Nas thickness	0.69 ± 0.06	28/92	63/71	0.72 ± 0.06	26/92	63/71
Max-min (modulation)	0.67 ± 0.06	39/92	61/71	0.66 ± 0.06	36/92	62/71

The three highest ROC areas and sensitivities are specified sensitivities for each parameter are in bold. Parameters are ordered based on highest to lowest ROC area when glaucoma was diagnosed by disc appearance. See Table 1 for abbreviations.

When SWAP parameters in eyes classified as glaucomatous based on optic disc appearance were compared with healthy eyes, all parameters were significantly different (after Bonferroni correction, $\alpha = 0.004$), except for inferior total deviation plot points ≤5% and ≤1%, and inferior pattern deviation plot points ≤5% and ≤1% (all $P \geq 0.01$). Glaucoma eyes had more negative MDs and larger PSDs and, in general, had more localized defects than healthy eyes (Table 6).

FDT: Glaucoma Diagnosed by Optic Disc Appearance

Table 7 shows ROC curve areas and sensitivities and specificities for all FDT parameters investigated. Areas under the ROC curve for diagnosis based on disc appearance ranged from 0.87 to 0.63. The three largest areas under the ROC curves (ROC area, SE) were for total deviation plot points ≤5% (0.87, 0.04), MD (0.83, 0.04), and both superior hemifield total deviation plot points ≤5% and pattern deviation plot points ≤5% (0.82, 0.05).

When target specificity was set at ≥90%, the parameters with the three highest sensitivities were MD and pattern deviation plot points ≤5% (both 61%, 92%), pattern deviation plot points ≤1% (57%, 92%), and total deviation points ≤5% (51%, 95%). When target specificity was set at ≥70%, the parameters with the three highest sensitivities were MD (80%, 71%), total deviation plot points ≤5% (78%, 84%), and both superior hemifield total deviation plot points ≤5% (78%, 76%) and PSD (78%, 71%).

When FDT parameters in eyes classified as glaucomatous based on optic disc appearance were compared with healthy eyes, all parameters were significantly different (after Bonferroni correction, $\alpha = 0.004$) between groups (more negative MD, larger PSD, more localized defects in the glaucomatous eyes; Table 8).

TABLE 4. Comparison of OCT Parameter Measures between Normal and Gluacomatous Eyes, Based on Optic Disc and Standard Perimetry Criteria

		Glaucoma by Optic Disc Appearance		Glaucoma by Standard Perimetry	
OCT Parameter (μm)	Normal (n = 38)	Glaucoma (n = 51)	P*	Glaucoma (n = 42)	P*
Inf thickness	142.7 ± 3.6	105.6 ± 3.1	*0.0001*	100.9 ± 3.4	*0.0001*
6 o'clock thickness	150.0 ± 3.6	110.1 ± 4.2	*0.0001*	105.5 ± 4.5	*0.0001*
Mean thickness	119.0 ± 3.0	93.7 ± 2.6	*0.0001*	89.8 ± 2.7	*0.0001*
7 o'clock (inf-temp) thickness	152.9 ± 5.2	107.3 ± 4.5	*0.0001*	100.4 ± 4.6	*0.0001*
Sup thickness	141.2 ± 4.1	118.8 ± 3.6	*0.0001*	108.7 ± 3.8	*0.0001*
12 o'clock thickness	140.3 ± 4.7	109.5 ± 4.1	*0.0001*	108.3 ± 4.5	*0.0001*
1 o'clock thickness	132.7 ± 4.9	104.3 ± 4.2	*0.0001*	102.0 ± 4.8	*0.0001*
5 o'clock thickness	119.8 ± 4.2	96.7 ± 3.6	*0.0001*	94.4 ± 4.1	*0.0001*
Temp thickness	98.8 ± 3.4	79.5 ± 2.9	*0.0001*	74.2 ± 3.0	*0.0001*
11 o'clock thickness	150.4 ± 5.2	120.2 ± 4.5	*0.0001*	115.1 ± 4.7	*0.0001*
Nas thickness	94.2 ± 4.0	77.3 ± 3.2	*0.002*	75.5 ± 3.8	*0.001*
Max-min (modulation)	62.4 ± 3.1	50.9 ± 2.7	0.006	51.8 ± 2.9	0.01

Data are mean ± SE. Significant differences between diagnostic groups (after Bonferroni correction) are italicized. Parameters are ordered as in Table 3. See Table 1 for abbreviations.

* Mean measurements in glaucomatous eyes compared with normal eyes. Bonferroni corrected $\alpha = 0.004$.

TABLE 5. SWAP Parameter ROC Curve Areas, Sensitivities, and Specificities for Detecting Glaucoma Based on Optic Disc and Standard Perimetry Criteria

	Glaucoma by Optic Disc Appearance			Glaucoma by Standard Perimetry		
SWAP Parameter	ROC Area (± SE)	Sensitivity–Specificity (%) (Specificity ≥90%)	Sensitivity–Specificity (%) (Specificity ≥70%)	ROC Area (SE)	Sensitivity–Specificity (%) (Specificity ≥90%)	Sensitivity–Specificity (%) (Specificity ≥70%)
PSD	**0.76 ± 0.05**	**41/92**	**61/71**	**0.78 ± 0.05**	**52/92**	**69/71**
Total dev pts ≤1%	**0.74 ± 0.05**	**39/92**	**73/71**	**0.78 ± 0.05**	**48/92**	**76/71**
Pattern dev pts ≤5%	**0.74 ± 0.05**	**43/92**	60/74	**0.76 ± 0.05**	**50/92**	64/74
Sup pattern dev pts ≤5%	**0.74 ± 0.05**	25/97	**73/71**	**0.76 ± 0.05**	36/97	**76/71**
Pattern dev pts ≤1%	**0.73 ± 0.05**	35/95	**61/74**	**0.78 ± 0.05**	45/95	**74/74**
Total dev pts ≤5%	**0.73 ± 0.05**	33/92	57/71	**0.77 ± 0.05**	43/92	64/71
Sup total dev pts ≤5%	**0.73 ± 0.05**	**41/92**	**73/71**	**0.76 ± 0.05**	**52/92**	**76/71**
Sup total dev pts ≤1%	0.71 ± 0.05	32/92	**63/74**	0.74 ± 0.05	43/92	64/74
Sup pattern dev pts ≤5%	0.70 ± 0.05	33/92	**63/74**	0.73 ± 0.05	45/92	64/74
MD	0.70 ± 0.06	27/92	49/71	0.71 ± 0.06	33/92	52/71
Inf total dev pts ≤5%	0.63 ± 0.06	22/92	49/76	0.66 ± 0.06	26/92	52/76
Inf total dev pts ≤1%	0.62 ± 0.05	25/92	45/76	0.65 ± 0.05	29/92	50/76
Inf pattern dev pts ≤5%	0.62 ± 0.06	14/97	43/82	0.63 ± 0.06	14/97	48/82
Inf pattern dev pts ≤1%	0.59 ± 0.05	14/95	43/74	0.64 ± 0.05	17/95	55/74

The 3 highest ROC areas and sensitivities at specified sensitivities for each parameter are in bold. Parameters are ordered based on highest to lowest ROC area when glaucoma was diagnosed by disc appearance. Dev, deviation; pts, points. See Table 1 for remaining abbreviations.

Comparing SLP, OCT, SWAP, and FDT

To directly compare SLP, OCT, SWAP, and FDT parameters, we additionally determined ROC curve areas and sensitivities at specificities of ≥90% and ≥70% for SWAP and FDT, using the diagnosis by SAP (Tables 5, 7), and determined ROC curve areas and sensitivities at specificities of ≥90% and ≥70% for SLP and OCT, using the diagnosis by disc appearance (Tables 1, 3). We then compared all four techniques, according to each diagnostic criterion.

Diagnosis by SAP

Using the best parameter from each instrument, the largest area under the ROC curve was found for OCT, followed by FDT, SLP, and SWAP (Fig. 1). For diagnosis based on SAP, we found significant differences in ROC curve area between OCT inferior quadrant thickness (0.91, 0.03) and both GDx LDF (0.79, 0.05) and SWAP PSD (0.78, 0.05; both $P \leq 0.02$). No other significant differences were found between parameters with the highest ROC curve areas from other instruments.

At target specificity set at ≥90% for the most sensitive parameter from each instrument, OCT inferior thickness, OCT thickness at 6 o'clock, and OCT thickness at 7 o'clock (all 79%, 92%) were significantly more sensitive than SWAP PSD, SWAP superior total deviation plot points ≤5%, SWAP pattern deviation points ≤1% (all 52%, 92%), and SLP superior average thickness (40%, 92%; all $P \leq 0.01$). FDT superior pattern deviation plot points ≤5% (71%, 92%) was more sensitive than SLP inferior average thickness ($P < 0.03$). At target specificity set at ≥70%, OCT inferior quadrant thickness (88%, 71%) was

TABLE 6. Comparison of SWAP Parameter Measures between Normal and Glaucomatous Eyes, Based on Optic Disc and Standard Perimetry Criteria

		Glaucoma by Optic Disc Appearance		Glaucoma by Standard Perimetry	
SWAP Parameter	Normal (n = 38)	Glaucoma (n = 51)	P*	Glaucoma (n = 42)	P*
PSD (dB)	3.24 ± 0.27	5.03 ± 0.31	*0.0003*	5.43 ± 0.35	*<0.0001*
Total dev pts ≤1%	1.4 ± 0.5	7.0 ± 1.3	*0.0007*	8.5 ± 1.5	*<0.0001*
Pattern dev pts ≤5%	4.7 ± 0.8	9.7 ± 1.0	*0.0003*	10.8 ± 1.1	*<0.0001*
Sup pattern dev pts ≤5%	2.9 ± 0.5	6.6 ± 0.8	*0.0006*	7.5 ± 0.9	*<0.0001*
Pattern dev pts ≤1%	2.6 ± 0.5	5.7 ± 0.8	*0.001*	7.1 ± 0.9	*<0.0001*
Total dev pts ≤5%	4.3 ± 1.2	12.7 ± 1.9	*0.001*	14.5 ± 1.8	*0.0001*
Sup total dev pts ≤5%	2.1 ± 0.7	7.6 ± 1.2	*0.0003*	8.9 ± 1.3	*<0.0001*
Sup total dev pts ≤1%	0.8 ± 0.3	4.8 ± 0.8	*0.0008*	5.9 ± 0.8	*<0.0001*
Sup pattern dev pts ≤1%	1.8 ± 0.3	4.5 ± 0.7	*0.004*	5.3 ± 0.8	*0.0006*
MD (dB)	−0.55 ± 0.71	−3.81 ± 0.61	*0.0008*	−4.27 ± 0.69	*0.0004*
Inf total dev pts ≤5%	2.2 ± 0.6	5.1 ± 0.9	0.022	5.6 ± 1.1	0.011
Inf total dev pts ≤1%	0.6 ± 0.2	2.2 ± 0.6	0.022	2.6 ± 0.7	0.009
Inf pattern dev pts ≤5%	1.7 ± 0.4	3.1 ± 0.5	0.056	3.3 ± 0.6	0.036
Inf pattern dev pts ≤1%	0.7 ± 0.2	1.6 ± 0.4	0.094	1.9 ± 0.5	0.035

Data are mean ± SE. Significant differences between diagnostic groups (after Bonferroni correction) are italicized. Parameters are ordered as in Table 5. Dev, deviation; pts, points. See Table 1 for remaining abbreviations.

* Mean measurements in glaucomatous eyes compared with normal eyes. Bonferroni corrected $\alpha = 0.004$.

TABLE 7. FDT Parameter ROC Curve Areas, Sensitivities, and Specificities for Detecting Glaucoma Based on Optic Disc and Standard Perimetry Criteria

	Glaucoma by Optic Disc Appearance			Glaucoma by Standard Perimetry		
FDT Parameter	ROC Area (± SE)	Sensitivity–Specificity (%) (Specificity ≥90%)	Sensitivity–Specificity (%) (Specificity ≥70%)	ROC Area (± SE)	Sensitivity–Specificity (%) (Specificity ≥90%)	Sensitivity–Specificity (%) (Specificity ≥70%)
Total dev pts ≤5%	**0.87 ± 0.04**	**51/95**	**78/84**	**0.88 ± 0.04**	48/95	**86/84**
MD	**0.83 ± 0.04**	**61/92**	**80/71**	0.84 ± 0.04	**57/92**	83/71
Sup total dev pts ≤5%	**0.82 ± 0.04**	49/97	**78/76**	**0.85 ± 0.04**	50/97	83/76
Pattern dev pts ≤5%	**0.82 ± 0.05**	**61/92**	**75/71**	**0.87 ± 0.04**	**71/92**	**86/71**
PSD	0.81 ± 0.05	39/92	**78/71**	**0.85 ± 0.04**	48/92	**88/71**
Total dev pts ≤1%	0.80 ± 0.04	47/92	69/87	0.84 ± 0.04	**57/92**	76/87
Inf total dev pts ≤5%	0.79 ± 0.04	41/97	71/82	0.78 ± 0.05	40/97	69/82
Pattern dev pts ≤1%	0.76 ± 0.04	**57/92**	65/84	**0.85 ± 0.04**	**67/92**	81/84
Sup total dev pts ≤1%	0.76 ± 0.04	47/95	59/90	0.78 ± 0.04	55/95	62/90
Sup pattern dev pts ≤5%	0.76 ± 0.05	41/97	55/79	0.80 ± 0.05	45/97	64/79
Inf pattern dev pts ≤5%	0.70 ± 0.05	26/95	43/90	0.74 ± 0.05	31/95	48/90
Inf total dev pts ≤1%	0.70 ± 0.04	45/92	45/92	0.76 ± 0.04	**57/92**	57/92
Sup pattern dev pts ≤1%	0.70 ± 0.04	35/95	49/87	0.77 ± 0.05	48/95	64/87
Inf pattern dev pts ≤1%	0.63 ± 0.04	33/92	33/92	0.67 ± 0.04	41/92	41/92

The three highest ROC areas and sensitivities at specified sensitivities for each parameter are in bold. Parameters are ordered based on highest to lowest ROC area when glaucoma was diagnosed by disc appearance. Dev, deviation; pts, points. See Table 1 for remaining abbreviations.

more sensitive than SWAP total deviation plot points ≤1%, SWAP superior pattern deviation points ≤5%, SWAP superior total deviation points ≤5% (all 76%, 71%), and GDx LDF and GDx Number (71%, 71%; all $P \leq 0.02$). FDT PSD (88%, 71%) was more sensitive than SWAP total deviation plot points ≤1%, SWAP superior pattern deviation points ≤5%, SWAP superior total deviation points ≤5% (all 76%, 71%), GDx LDF, and GDx Number (71%, 71%; all $P \leq 0.02$).

At ≥90% specificity for each instrument, agreement was poor among pairs of parameters with the highest ROC curve area for classifying eyes as glaucomatous. The κ statistic ranged from −0.32 between OCT inferior quadrant thickness and FDT total deviation plot points ≤5% to 0.17 between OCT inferior quadrant thickness and SWAP PSD. In Figure 2 Venn diagrams are used to show the number of eyes correctly classified as glaucomatous by the four instruments when diagnosis was based on SAP.

Diagnosis by Optic Disc Appearance

We found a significant difference in ROC curve area between OCT inferior quadrant thickness (0.89, 0.03) and SWAP PSD (0.76, 0.05; $P \leq 0.02$). No significant differences were found between parameters with the highest ROC curve areas from other instruments (Fig. 3).

At target specificity set at ≥90% for the most sensitive parameter from each instrument, OCT 6 o'clock thickness was significantly more sensitive (75%, 92%) than SWAP pattern deviation plot points ≤5% (43%, 92%) and SLP superior maximum thickness (39%, 92%; all $P \leq 0.01$). FDT MD and FDT

TABLE 8. Comparison of FDT Parameter Measures between Normal and Glaucomatous Eyes, Based on Optic Disc and Standard Perimetry Criteria

		Glaucoma by Optic Disc Appearance		Glaucoma by Standard Perimetry	
FDT Parameter	Normal (n = 38)	Glaucoma (n = 51)	P*	Glaucoma (n = 42)	P*
Total dev pts ≤5%	1.0 ± 0.6	6.2 ± 0.5	*<0.0001*	6.5 ± 0.6	*<0.0001*
MD (dB)	−0.16 ± 0.50	−3.8 ± 0.4	*<0.0001*	−4.1 ± 0.4	*<0.0001*
Sup total dev pts ≤5%	0.6 ± 0.4	3.7 ± 0.3	*<0.0001*	3.9 ± 0.4	*<0.0001*
Pattern dev pts ≤5%	1.3 ± 0.4	4.3 ± 0.4	*<0.0001*	4.9 ± 0.4	*<0.0001*
PSD (dB)	3.8 ± 0.41	6.5 ± 0.4	*<0.0001*	7.0 ± 0.4	*<0.0001*
Total dev pts ≤1%	0.3 ± 0.5	3.8 ± 0.4	*<0.0001*	4.4 ± 0.5	*<0.0001*
Inf total dev pts ≤5%	0.4 ± 0.3	2.6 ± 0.3	*<0.0001*	2.6 ± 0.3	*<0.0001*
Pattern dev pts ≤1%	0.3 ± 0.3	2.1 ± 0.3	*<0.0001*	2.5 ± 0.4	*<0.0001*
Sup total dev pts ≤1%	0.2 ± 0.3	2.4 ± 0.3	*<0.0001*	2.8 ± 0.3	*<0.0001*
Sup pattern dev pts ≤5%	0.8 ± 0.3	2.8 ± 0.3	*<0.0001*	3.1 ± 0.3	*<0.0001*
Inf pattern dev pts ≤5%	0.5 ± 0.2	1.6 ± 0.2	*0.0008*	1.8 ± 0.2	*<0.0001*
Inf total dev pts ≤1%	0.1 ± 0.3	1.4 ± 0.2	*0.0002*	1.7 ± 0.3	*<0.0001*
Sup pattern dev pts ≤1%	0.2 ± 0.3	1.5 ± 0.2	*0.0004*	2.0 ± 0.3	*<0.0001*
Inf pattern dev pts ≤1%	0.1 ± 0.1	0.5 ± 0.1	*0.003*	0.6 ± 0.1	*0.0009*

Data are mean ± SE. Significant differences between diagnostic groups (after Bonferroni correction) are italicized. Parameters are ordered as in Table 7. Dev, deviation; pts, points. See Table 1 for remaining abbreviations.

* Mean measurements in glaucomatous eyes compared with normal eyes. Bonferroni corrected $\alpha = 0.004$.

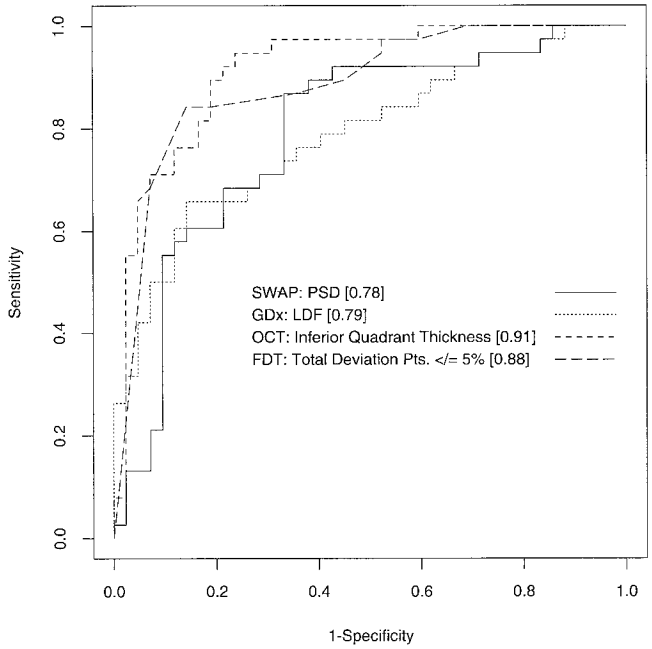

FIGURE 1. ROC curves (and areas) for the best parameters from each instrument when glaucoma was diagnosed by standard perimetry. The ROC curve area for OCT inferior quadrant thickness was significantly greater than the ROC curve area for GDx LDF and SWAP PSD (both $P \leq 0.02$; method of DeLong et al.[21]).

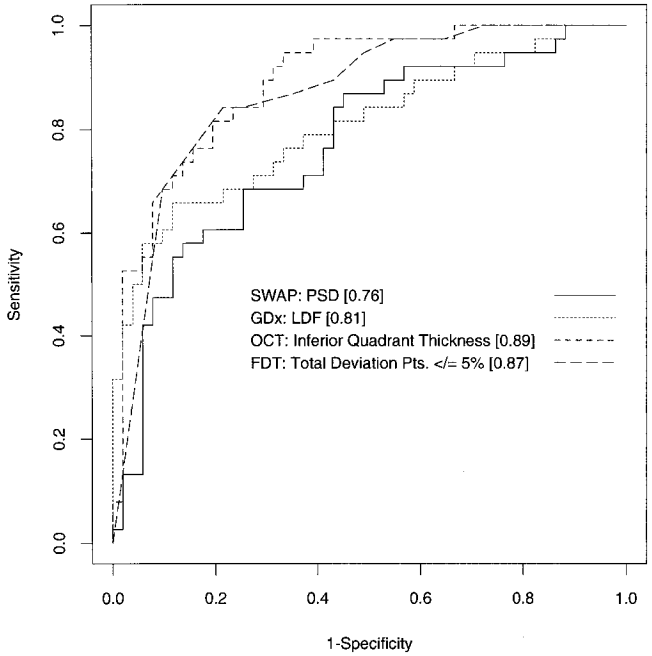

FIGURE 3. ROC curves (and areas) for the best parameters from each instrument when glaucoma was diagnosed by optic disc appearance. The ROC curve area for OCT inferior quadrant thickness was significantly greater than the ROC curve area for SWAP PSD ($P \leq 0.02$; method of DeLong et al.[21]).

pattern deviation plot points ≤5% (both 61%, 92%) were significantly more sensitive than SWAP pattern deviation plot points ≤5% and SLP superior maximum thickness (all $P \leq 0.02$). When target specificity was set at ≥70%, OCT inferior quadrant thickness was significantly more sensitive (88%, 71%) than SWAP total deviation points ≤1%, SWAP superior pattern deviation plots ≤5%, SWAP superior total deviation points ≤5%, GDx LDF, and GDx Number (all 73%, 71%; all $P \leq 0.03$).

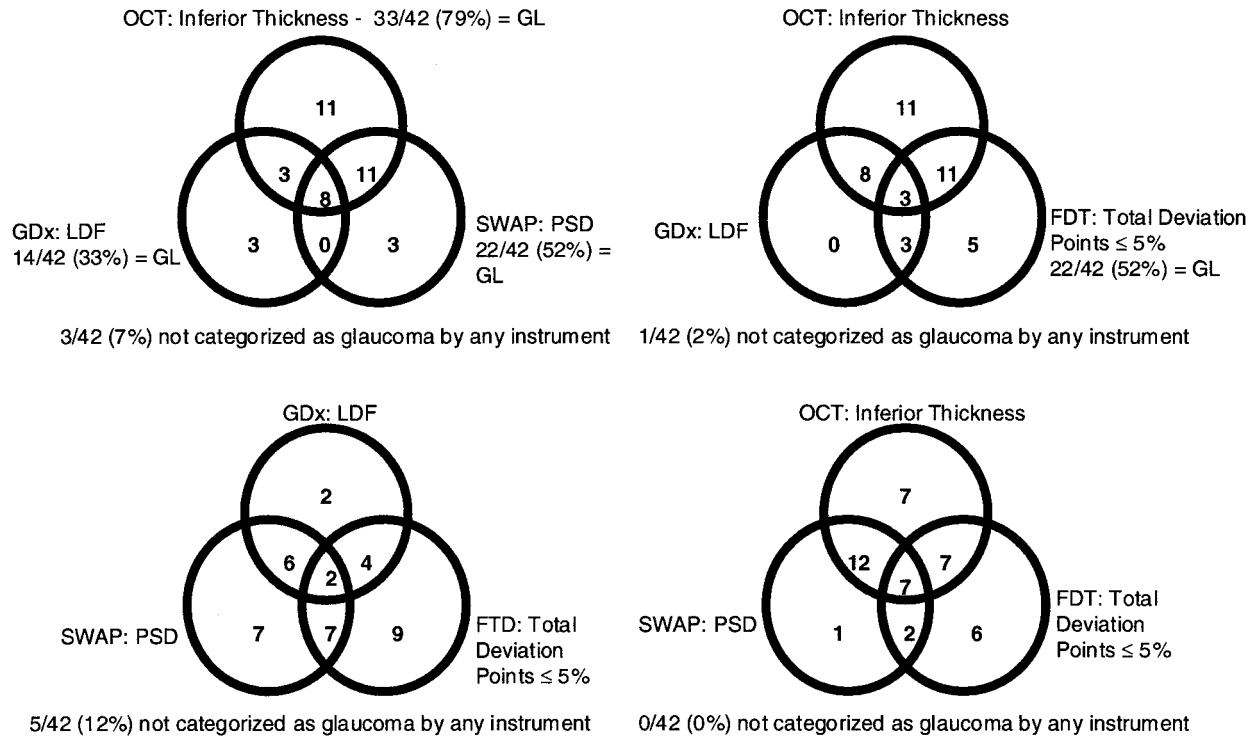

FIGURE 2. Venn diagrams showing the number of eyes correctly classified as glaucomatous by the parameter with the largest ROC curve area from each instrument when specificity was set at ≥90% and glaucoma was diagnosed by standard perimetry results. Agreement was poor among instruments ($\kappa < 0.34$, all pairs).

Similar to diagnosis by SAP comparisons, at ≥90% specificity for each instrument, agreement was poor between pairs of parameters with the largest ROC curve areas. The κ ranged from −0.27 between FDT total deviation plot points ≤5% and SWAP PSD to 0.34 between OCT inferior quadrant thickness and SWAP PSD. Because the number of eyes correctly classified as glaucomatous among all four instruments was similar to when glaucoma was diagnosed by SAP results, the data are not shown.

Discussion

In general, areas under the ROC curve were largest for OCT and FDT parameters, followed by SLP and SWAP parameters, regardless of whether glaucoma was defined based on SAP results or optic disc appearance. Sensitivities of the best SLP and SWAP parameters were significantly lower than those of the best parameters of OCT and FDT when specificity was set at ≥90% when glaucoma was diagnosed based on disc appearance. This general pattern of results was similar when specificity was set at ≥70%. The 90% and 70% specificities were chosen to represent high and moderate specificities, respectively, not in an attempt to find the best cutoff for any particular parameter. ROC curve area figures (Figs. 1, 3) indicate that at some specificities, SWAP and SLP parameter sensitivities were similar to those of OCT and FDT.

The diagnostic criteria used (SAP or disc appearance) had a minimal effect on the performance of SLP, OCT, SWAP, and FDT parameters in discriminating between glaucomatous and healthy eyes. For the most part, areas under the ROC curves and sensitivities were similar. This finding was probably affected by the fact that 37 of the same patients were included in both diagnostic groups. Analysis of independent patient populations for the two diagnostic groups may have provided different results. However, we expect that the percentage of patients with glaucomatous optic discs in an independent standard perimetry diagnosis group would have been similar and comparable to our group, because in our clinic population (as in many other clinics) few patients with glaucoma have repeatable standard visual field defects without observable optic disc damage.

Despite this, the diagnostic ability (based on ROC curve area and sensitivities and specificities) of most OCT, SWAP, and FDT parameters increased slightly (although not significantly) when glaucoma was defined based on functional criteria. This is not surprising, because it is likely that patients with repeatable visual field defects (and thus more advanced disease) have more RNFL damage compared with those who have not yet developed visual field defects. This increased RNFL damage may be detected using OCT, thus improving glaucoma detection. Although when evaluating structure-based tests it is theoretically best to use functional criteria as a diagnostic gold standard, and when evaluating function-based tests, it is best to use structural (optic disc) criteria, our results suggest that in some populations there may be little practical difference. Further research is needed to determine whether this is the case when using independent diagnostic groups.

In the present study, and inherent in all studies evaluating diagnostic procedures in glaucoma, there exists no perfect gold standard for diagnosis. Although stereophotography and standard visual field testing are the current standards used for glaucoma diagnosis in research, it is possible that newly developed instruments are better at detecting glaucoma. Estimates of sensitivity and specificity and ROC curve areas are reliant on the quality of the gold standard applied. The use of an imperfect gold standard can affect ROC areas in two ways, depending on the relationship between errors in the test and the standard.[23] If errors in the test and standard are independent, ROC curve area is underestimated because different "mistakes" are made by both, thus decreasing sensitivity and specificity of test and standard. If errors in the test and the standard are positively dependent, ROC areas are overestimated. The extreme example of this case is when errors in test and standard are perfectly correlated. Because both tests make the same mistakes for each diagnosis, ROC area (and therefore sensitivity and specificity) is 1.0, regardless of the true utility of the test. Therefore, on theoretical grounds, SAP should not be used as the standard for evaluating other function tests, because it is likely that many errors in SAP are correlated with errors in SWAP and FDT. It is less clear whether errors in assessing optic disc appearance using stereophotographs are correlated with errors in SLP or OCT measures.

In general, sensitivities, specificities, and ROC areas in our study are quite similar to those reported by others. For instance, using OCT, we observed ROC curve areas in the 0.85 to 0.90 range and sensitivities (with specificity set at ≥90%) in the 70% to 80% range for the best parameters. Also using OCT, Greaney et al.[24] reported a best ROC curve area of 0.90 and a best sensitivity and specificity of 78% and 90%, respectively. For FDT we observed best ROC curve areas in the 0.80 to 0.90 range and sensitivities (with specificity set at ≥90%) in the 60% to 70% range for the best parameters. Similarly, Cello et al.[25] reported an ROC curve area of 0.93 and a sensitivity and specificity of 85% and 90%, respectively, for patients with early glaucoma defined by SAP. FDT ROC curve areas and sensitivities in our study may be somewhat underestimated, because we used the C-20 program (Welch Allyn), which omits nasal points and, therefore, nasal-step information. Using SWAP, our best ROC curve areas were slightly less than 0.80 and sensitivities (with specificity set at ≥90%) were in the 40% to 52% range for the best parameters. Sample et al.[26] reported a sensitivity of 61% and a specificity of 86% using SWAP test results outside of normal limits as a criterion for diagnostic classification when disc appearance was the gold standard.

In some cases, however, our values were considerably lower than those previously reported. For instance, using SLP, we reported a best ROC curve area of 0.81 and a best sensitivity (with specificity set at ≥90%) of 40%. Others have reported sensitivities and specificities in the 80% to 90% range.[12,27] Recently, some researchers have reported improved SLP results (ROC curve areas of 0.90 for discriminating between glaucomatous and healthy eyes, discrimination of ocular hypertensive eyes from healthy eyes) using nonstandard data analysis methods.[13,28–30] Differences in SLP performance among studies may be due to differences in severity of glaucoma, differences in software, differences in corneal polarization axes,[31] or other methodological differences.

Statistically significant differences in variable measures between glaucoma and healthy eyes found in the present study confirm previous findings for all instruments.[10,12–14,26,27,30,32–39] We found that a limited number of SLP parameters were significantly different between diagnostic groups. We suspected that this may be due in part to the necessary correction of α($\alpha = 0.002$) with such a large number of parameters investigated for this instrument. However, if we use the same α (i.e., 0.002) to define statistically significant differences using OCT, SWAP, and FDT, almost all parameter differences between diagnostic groups remain. Similarly, if we relax α for the SLP comparisons to the level used for SWAP and FDT ($\alpha = 0.004$), few new significant differences appear. Although SAP-measured MD was only 0.5 dB lower in the SAP diagnostic group than in the optic disc diagnostic group, it is interesting that all OCT and SWAP parameters and all but one FDT parameter (inferior total deviation plot points ≤5%) changed in the direction predicted by an increase in glaucoma severity in the SAP diagnostic group (although no changes

were statistically significant). This general trend was not observed for more than half of the SLP parameters.

It also is interesting to note that when specificity was set at ≥90% and diagnostic agreement was compared across techniques, we found poor agreement among the best (largest ROC curve area) parameter from each instrument for determining which eyes were glaucomatous (all between-instrument comparisons, $\kappa < 0.35$; Fig. 2). This finding indicates that different techniques may detect different characteristics of glaucoma. Poor agreement between SWAP and FDT may be partially related to the different subsets of retinal ganglion cells tested. Possibly, early in glaucoma, some eyes lose more of one cell type than another. This may lead to nonoverlapping groups of patients with glaucoma with different patterns of cell loss.[26] Our study did not address this hypothesis.

The primary strength of our study is that instruments were compared in a single population. Therefore, the sensitivity and specificity values per se were less important than their relative values among instruments. Limitations of this study include the small number of subjects. Our inability to find significant differences in ROC curve areas among the best parameters in most cases may be related to sample size. Another limitation is that all tests had to be completed within 1 year. Ideally, this maximum should be shortened to obtain the best cross-sectional comparison of different diagnostic techniques. However, because 75% of patients had tests completed within 3 months and 80% had tests completed within 6 months, it is unlikely that clinically meaningful glaucomatous change occurred within this time frame. It is possible that glaucoma will develop later in some of the healthy eyes included in this study. Therefore, longitudinal study is the only way to truly determine the sensitivity and specificity of these tests.

Another limitation, inherent in any comparable study, is that different diagnostic techniques evaluated in this study are currently at different stages of development. More established techniques (SWAP and SLP) were compared with newer technologies (FDT and OCT). In general, established technologies benefit from robust normative databases and more sophisticated analysis strategies. For instance, a complex parameter such as the GDx Number (a neural network–derived analysis) may be expected to outperform a crude measurement such as OCT-measured RNFL thickness at 6 o'clock (derived from measurements at only eight or nine points). Similarly, a visual field analysis parameter such as PSD is likely to be more thoroughly derived from a standard 24-2 grid (used for SWAP) than from the more crude 16-location FDT grid. In our study, however, the more recently developed technologies generally performed better than the older ones. Because SLP corneal polarization compensation is inadequate in a sizable number of patients,[31] implementation of a method that correctly compensates for corneal polarization axis in individual patients will probably improve the diagnostic precision of the instrument.

In conclusion, the largest ROC curve area for OCT (inferior quadrant thickness) was larger than the largest ROC curve area for SLP (LDF) and SWAP (PSD) when diagnosis was based on SAP, and the largest ROC curve area for OCT (inferior quadrant thickness) was larger than the largest ROC curve area for SWAP (PSD) when diagnosis was based on disc appearance. ROC curve areas among other instruments were not significantly different for either diagnostic criterion. Sensitivities were best (although not always significantly so) for OCT and FDT measurements followed by SWAP and SLP. However, the sensitivity and specificity of even the best parameter of the best instrument are probably not sufficient to warrant use as a sole screening method in the general population. In contrast, for screening in situations in which treatment is at a premium (e.g., developing nations), a sensitivity and specificity of 79% and 92% (for several OCT measures, for example) may be acceptable, assuming that the technique is relatively simple and quick. The poor diagnostic agreement found among instruments suggests that different techniques may identify different characteristics of glaucomatous damage.

References

1. Quigley HA, Addicks EM, Green WR. Optic nerve damage in human glaucoma. III: quantitative correlation of nerve fiber loss and visual field defect in glaucoma, ischemic neuropathy, papilledema, and toxic neuropathy. *Arch Ophthalmol.* 1982;100:135–146.
2. Quigley HA, Dunkelberger GR, Green WR. Retinal ganglion cell atrophy correlated with automated perimetry in human eyes with glaucoma. *Am J Ophthalmol.* 1989;107:453–464.
3. Kerrigan-Baumrind LA, Quigley HA, Pease ME, Kerrigan DF, Mitchell RS. Number of ganglion cells in glaucoma eyes compared with threshold visual field tests in the same persons. *Invest Ophthalmol Vis Sci.* 2000;41:741–748.
4. Sommer A, Miller NR, Pollack I, Maumenee AE, George T. The nerve fiber layer in the diagnosis of glaucoma. *Arch Ophthalmol.* 1977;95:2149–2156.
5. Sommer A, Pollack I, Maumenee AE. Optic disc parameters and onset of glaucomatous field loss. I: methods and progressive changes in disc morphology. *Arch Ophthalmol.* 1979;97:1444–1448.
6. Pederson JE, Anderson DR. The mode of progressive disc cupping in ocular hypertension and glaucoma. *Arch Ophthalmol.* 1980;98:490–495.
7. Sommer A, Katz J, Quigley HA, et al. Clinical detectable nerve fiber atrophy precedes the onset of glaucomatous field loss. *Arch Ophthalmol.* 1991;109:77–83.
8. Zeyen TG, Caprioli J. Progression of disc and field damage in early glaucoma. *Arch Ophthalmol.* 1993;111:62–65.
9. Weinreb RN, Dreher AW, Coleman A, Quigley H, Shaw B, Reiter K. Histopathologic validation of Fourier-ellipsometry measurements of retinal nerve fiber layer thickness. *Arch Ophthalmol.* 1990;108:557–560.
10. Weinreb RN, Shakiba S, Zangwill L. Scanning laser polarimetry to measure the nerve fiber layer of normal and glaucomatous eyes. *Am J Ophthalmol.* 1995;119:627–636.
11. Poinoosawmy D, Fontana L, Wu JX, Fitzke FW, Hitchings RA. Variation of nerve fibre layer thickness measurements with age and ethnicity by scanning laser polarimetry. *Br J Ophthalmol.* 1997;81:350–354.
12. Tjon-Fo-Sang MJ, Lemij HG. The sensitivity and specificity of nerve fiber layer measurements in glaucoma as determined with scanning laser polarimetry. *Am J Ophthalmol.* 1997;123:62–69.
13. Choplin NT, Lundy DC, Dreher AW. Differentiating patients with glaucoma from glaucoma suspects and normal subjects by nerve fiber layer assessment with scanning laser polarimetry. *Ophthalmology.* 1998;105:2068–2076.
14. Weinreb RN, Zangwill L, Berry CC, Bathija R, Sample PA. Detection of glaucoma with scanning laser polarimetry. *Arch Ophthalmol.* 1998;116:1583–589.
15. Schuman JS, Hee MR, Puliafito CA, et al. Quantification of nerve fiber layer thickness in normal and glaucomatous eyes using optical coherence tomography. *Arch Ophthalmol.* 1995;113:586–596.
16. Dacey DM, Lee BB. The "blue-on" opponent pathway in primate retina originates from a distinct bistratified ganglion cell type. *Nature.* 1994;367:731–735.
17. Sample PA, Johnson CA, Haegerstrom-Portnoy G, Adams AJ. Optimum parameters for short-wavelength automated perimetry. *J Glaucoma.* 1996;5:375–383.
18. Maddess T, Hemmi JM, James AC. Evidence for spatial aliasing effects in the Y-like cells of the magnocellular visual pathway. *Vision Res.* 1998;38:1843–1859.
19. Tyler CW. Observations on spatial-frequency doubling. *Perception.* 1974;3:81–86.
20. Tyler CW. Specific deficits of flicker sensitivity in glaucoma and ocular hypertension. *Invest Ophthalmol Vis Sci.* 1981;20:204–212.

21. 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.
22. Quigley HA. Current and future approaches to glaucoma screening. *J Glaucoma.* 1998;7:210–220.
23. Phelps CE, Hutson A. Estimating diagnostic test accuracy using a "fuzzy gold standard." *Med Decis Making*. 1995;15:44–57.
24. Greaney M, Garway-Heath D, Hoffman D, Caprioli J. Comparison of imaging methods to distinguish normal eyes from those with early glaucoma [ARVO abstract]. *Invest Ophthalmol Vis Sci.* 2000; 41(4):**S**121. Abstract nr 623.
25. Cello KE, Nelson-Quigg JM, Johnson CA. Frequency doubling technology perimetry for detection of glaucomatous visual field loss. *Am J Ophthalmol.* 2000;129:314–322.
26. Sample PA, Bosworth CF, Blumenthal EZ, Girkin C, Weinreb RN. Visual function testing of different ganglion cell populations in glaucoma, ocular hypertensive, and normal eyes. *Invest Ophthalmol Vis Sci.* 2000;41:1783–1790.
27. Horn FK, Jonas JB, Martus P, Mardin CY, Budde WM. Polarimetric measurement of retinal nerve fiber layer thickness in glaucoma diagnosis. *J Glaucoma.* 1999;8:353–362.
28. Bryant F, Sinai M, Fechtner R, Essock E. An extension of fourier analysis techniques of nerve fiber layer measurements from scanning laser polarimetry [ARVO abstract]. *Invest Ophthalmol Vis Sci.* 2000;41(4):**S**92. Abstract nr 485.
29. Costa V, Lauande-Pimentel R, Fonesca R, Oliveira H, Goncalves D, Silva L. Discrimination between normal and glaucomatous eyes with functional and structural parameters [ARVO abstract]. *Invest Ophthalmol Vis Sci.* 2000;41(4):**S**90. Abstract nr 470.
30. Kogure S, Iijima H, Tsukahara S. A new parameter for assessing the thickness of the retinal nerve fiber layer for glaucoma diagnosis. *Eur J Ophthalmol.* 1999;9:93–98.
31. Greenfield D, Knighton R, Huang X. The effect of corneal polarization axis upon retinal nerve fiber layer thickness assessment using scanning laser polarimetry. *Am J Ophthalmol.* 2000;129: 715–722.
32. Bosworth CF, Sample PA, Gupta N, Bathija R, Weinreb RN. Motion automated perimetry identifies early glaucomatous field defects. *Arch Ophthalmol.* 1998;116:1153–1158.
33. Bowd C, Weinreb RN, Williams JM, Zangwill LM. The retinal nerve fiber layer thickness in ocular hypertensive, normal, and glaucomatous eyes with optical coherence tomography. *Arch Ophthalmol.* 2000;118:22–26.
34. Brusini P, Busatto P. Frequency doubling perimetry in glaucoma early diagnosis. *Acta Ophthalmol Scand Suppl.* 1998;227:23–24.
35. Hoh ST, Greenfield DS, Mistlberger A, Liebmann JM, Ishikawa H, Ritch R. Optical coherence tomography and scanning laser polarimetry in normal, ocular hypertensive, and glaucomatous eyes. *Am J Ophthalmol.* 2000;129:129–135.
36. Johnson CA. Diagnostic value of short-wavelength automated perimetry. *Curr Opin Ophthalmol.* 1996;7:54–58.
37. Mansberger SL, Sample PA, Zangwill L, Weinreb RN. Achromatic and short-wavelength automated perimetry in patients with glaucomatous large cups. *Arch Ophthalmol.* 1999;117: 1473–1477.
38. Sample PA, Bosworth CF, Weinreb RN. Short-wavelength automated perimetry and motion automated perimetry in patients with glaucoma. *Arch Ophthalmol.* 1997;115:1129–1133.
39. Tjon-Fo-Sang MJ, de Vries J, Lemij HG. Measurement by nerve fiber analyzer of retinal nerve fiber layer thickness in normal subjects and patients with ocular hypertension. *Am J Ophthalmol.* 1996; 122:220–227.