Optic Disk and Nerve Fiber Layer Imaging to Detect Glaucoma

FEDERICO BADALÀ, KOUROS NOURI-MAHDAVI, DUNA A. RAOOF, NARAKORN LEEPRECHANON, SIMON K. LAW, AND JOSEPH CAPRIOLI

• PURPOSE: To compare optic disk and retinal nerve fiber layer (RNFL) imaging methods to discriminate eyes with early glaucoma from normal eyes.

• DESIGN: Retrospective, cross-sectional study.

• METHODS: In a tertiary care academic glaucoma center, 92 eyes of 92 subjects (46 with early perimetric open-angle glaucoma and 46 controls) were studied. Diagnostic performance of optical coherence tomography (StratusOCT; Carl Zeiss Meditec, Dublin, California, USA), scanning laser polarimetry (GDx VCC; Laser Diagnostic Technologies, San Diego, California, USA), confocal laser ophthalmoscopy (Heidelberg Retinal Tomograph [HRT] III; Heidelberg Engineering GmbH, Heidelberg, Germany), and qualitative assessment of stereoscopic optic disk photographs were compared. Outcome measures were areas under receiver operator characteristic curves (AUCs) and sensitivities at fixed specificities. Classification and regression tree (CART) analysis was used to evaluate combinations of quantitative parameters.

• RESULTS: The average (± standard deviation) visual field mean deviation for glaucomatous eyes was −4.0 ± 2.5 dB (decibels). Parameters with largest AUCs (± standard error) were: average RNFL thickness for StratusOCT (0.96 ± 0.02), nerve fiber indicator for GDx VCC (0.92 ± 0.03), Frederick S. Mikelberg (FSM) discriminant function for HRT III (0.91 ± 0.03), and 0.97 ± 0.02 for disk photograph evaluation. At 95% specificity, sensitivity of disk photograph evaluation (90%) was greater than GDx VCC ($P = .05$) and HRT III ($P = .002$) results, but not significantly different than those of StratusOCT ($P > .05$). The combination of StratusOCT average RNFL thickness and HRT III cup-to-disk area with CART produced a sensitivity of 91% and specificity of 96%.

• CONCLUSIONS: StratusOCT, GDx VCC, and HRT III performed as well as, but not better than, qualitative evaluation of optic disk stereophotographs for detection of early perimetric glaucoma. The combination of StratusOCT average RNFL thickness and HRT III cup-to-disk area ratio provided a high diagnostic precision. (Am J Ophthalmol 2007;144:724–732. © 2007 by Elsevier Inc. All rights reserved.)

Accepted for publication Jul 9, 2007.

From the Glaucoma Division, Jules Stein Eye Institute, David Geffen School of Medicine, University of California Los Angeles, Los Angeles, California.

Inquiries to Joseph Caprioli, Glaucoma Division, Jules Stein Eye Institute, 100 Stein Plaza, Los Angeles, CA 90095; e-mail: caprioli@jsei.ucla.edu

STRUCTURAL CHANGES OF THE OPTIC DISK AND RETinal nerve fiber layer (RNFL) often precede visual field defects measured with standard achromatic perimetry in early glaucoma.[1–6] Recognizing these morphologic abnormalities is important clinically for the early diagnosis of the disease. In recent years, optic disk and RNFL imaging techniques increasingly have been used in clinical practice. Studies with previous versions of optical coherence tomography, confocal scanning laser ophthalmoscopy, and scanning laser polarimetry showed no significant difference between these technologies and evaluation of the optic nerve head by expert observers.[7,8] More advanced versions of the same instruments and their performance relative to clinical assessment of the optic disk was examined in this study. This issue is important both from clinical and economic standpoints. Use of the new quantitative instruments without adequate validation may lead to improper management decisions. The purpose of the present study was to compare the ability of four currently used optic disk and RNFL imaging methods to discriminate eyes with early, reproducible glaucomatous visual field loss from healthy eyes.

METHODS

THE AUTHORS REVIEWED THE CLINICAL DATABASE OF THE Glaucoma Division at Jules Stein Eye Institute (University of California, Los Angeles) for patients who underwent visual field testing and optic disk imaging with optical coherence tomography, scanning laser ophthalmoscopy, scanning laser polarimetry, and stereoscopic optic disk photographs at the same visit between April 1, 2003 and April 1, 2006. Subjects with poor-quality imaging or unreliable visual fields were excluded. One eye each of 46 patients with open-angle glaucoma and 46 normal controls older than 40 years of age were enrolled in this cross-sectional observational study. All eyes were required to have visual acuity of 20/40 or better and ametropia of 5 diopters (D) or less (spherical equivalent). Glaucomatous eyes had open angles, confirmed early defects on white-on-white automated perimetry, and no history of other ocular diseases. One eye of each patient was selected randomly when both eyes fulfilled the inclusion criteria. Normal subjects were recruited among staff, patients' spouses, and volunteers; they had normal eye examination results, including a normal optic disk, intraocular pressure

  0002-9394/07/$32.00
doi:10.1016/j.ajo.2007.07.010

TABLE 1. Demographic Characteristics of Glaucoma Cases and Normal Controls

	Normal Group	Glaucoma Group	*P* value
No. of eyes	46 (50%)	46 (50%)	
Age (mean ± SD), yrs	58.9 ± 6.8	61.8 ± 9.7	.1*
Gender (female/male)	26 (57%)/20 (43%)	29 (63%)/17 (37%)	.5†
Ethnicity			.08†
White	25 (54%)	31 (67%)	
Black	1 (2%)	5 (11%)	
Hispanic	9 (20%)	4 (9%)	
Asian	11 (24%)	6 (13%)	
IOP (mean ± SD), mm Hg	14.0 ± 2.9	14.7 ± 4.3	.4*
MD (mean ± SD), dB	0.1 ± 1.2	−4.0 ± 2.5	< .001‡
PSD (mean ± SD), dB	1.4 ± 0.2	5.5 ± 2.5	< .001‡
Spherical equivalent (D)	−0.4 ± 1.7	−0.9 ± 2.1	.3‡

D = diopters; IOP = intraocular pressure; MD = mean deviation; PSD = pattern standard deviation; SD = standard deviation.

The study population consisted of one eye each from a group of glaucomatous patients with confirmed early to moderately advanced visual field defects and one eye each from a group of normal subjects.

*Independent sample *t* test.

†Chi-square test.

‡Mann–Whitney *U* test.

of less than 21 mm Hg, no history of ocular surgery or trauma, and a normal achromatic visual field. Normal subjects were matched with glaucomatous patients for age, gender, and ethnicity.

• **VISUAL FIELD TESTING:** Visual field testing was performed with the Humphrey Field Analyzer 750 (Allergan Humphrey, San Leandro, California, USA). Achromatic, standard 24-2 Swedish interactive threshold algorithm visual fields were obtained. Only patients with reliable fields (fixation loss rate, ≤ 33%; false-positive and false-negative rates, ≤ 20%) were included. Glaucoma patients included in the study had a mean deviation (MD) of more than −8 decibels (dB), glaucoma hemifield test (GHT) results outside normal limits, and a pattern standard deviation (PSD) with $P < .05$, all confirmed on two consecutive visual fields. A normal visual field was defined as one having a GHT within normal limits and a PSD with a $P > .05$ on two consecutive examinations.

• **IMAGING METHODS:** The StratusOCT Fast Retinal Nerve Fiber Layer Thickness algorithm (Carl Zeiss Meditec, Dublin, California, USA) was used to evaluate peripapillary RNFL thickness. Three images were obtained automatically for each eye. Each image consisted of 256 A scans along a circular ring (3.4 mm in diameter) around the optic disk. Only well-centered and focused images with more than 90% good quality A scans and a signal-to-noise ratio of more than 25 dB were included (Zeiss StratusOCT Model 3000 User Manual, PN 55153-1-Rev. A 12/02, pages 6–13). Data were exported to a personal computer and the three measurements obtained for each eye were averaged; left eye data were converted automatically into right eye format during the export procedure. The parameters calculated by the StratusOCT software (version 4.0) and evaluated in this study were: average RNFL thickness; RNFL thickness in each quadrant (superior, inferior, nasal, and temporal); RNFL thickness in each of the 12 clock-hour sectors; maximum RNFL thickness in the superior and inferior quadrants (S_{max} and I_{max}, respectively); the difference between the thickest and thinnest measured points (maximum − minimum); and the ratios I_{max}/S_{max}, S_{max}/I_{max}, S_{max}/temporal average RNFL thickness (T_{avg}), I_{max}/T_{avg}, S_{max}/nasal average RNFL thickness (N_{avg}). The signal strength, a measure of image quality, also was recorded.

Scanning laser polarimetry with variable corneal compensation, GDx VCC (Laser Diagnostic Technologies, San Diego, California, USA; software version 5.2.3), was used to evaluate the peripapillary RNFL. The RNFL thickness was measured along a calculation circle (0.4-mm in width between the 1.6-mm outer radius and 1.24-mm inner radius) centered on the optic nerve head. This circle was divided into superior (120 degrees), inferior (120 degrees), nasal (50 degrees), and temporal (70 degrees) quadrants. The parameters calculated by the GDx VCC software and evaluated in this study were: temporal, superior, nasal, inferior, and temporal again (TSNIT; completing a circle around the optic disk) average; superior average; inferior average; and TSNIT standard deviation (TSNIT SD). TSNIT average is the average RNFL thickness along the calculation circle. Superior and infe-

TABLE 2. StratusOCT Parameters in Glaucomatous and Normal Eyes, Areas Under the Receiver Operator Characteristic (ROC) Curves, and Sensitivities at Fixed Specificities

RNFL Thickness (μm)	Glaucoma Group (mean ± SD)	Normal Group (mean ± SD)	*P* value	ROC Curve ± SE	Sensitivity at 80% Specificity ± SE	Sensitivity at 95% Specificity ± SE
Average thickness	69.7 ± 12.2	98.0 ± 11.2	< .001*	0.96 ± 0.02	96% ± 3%	89% ± 5%
Inferior quadrant	80.1 ± 20.7	128.0 ± 17.5	< .001*	0.95 ± 0.02	91% ± 4%	87% ± 5%
Sector 7	78.9 ± 32.3	141.9 ± 20.9	< .001†	0.93 ± 0.03	89% ± 5%	78% ± 6%
Inferior maximum	108.0 ± 29.9	163.9 ± 21.0	< .001*	0.93 ± 0.03	89% ± 5%	78% ± 6%
Sector 6	86.0 ± 26.2	137.5 ± 24.7	< .001†	0.92 ± 0.03	85% ± 5%	74% ± 7%
Superior quadrant	85.6 ± 20.1	122.1 ± 18.6	< .001*	0.92 ± 0.03	91% ± 4%	54% ± 7%
Superior maximum	113.8 ± 25.6	153.9 ± 19.0	< .001†	0.91 ± 0.03	85% ± 5%	57% ± 7%
Inferior average	80.1 ± 20.7	116.1 ± 17.7	< .001*	0.90 ± 0.03	87% ± 5%	72% ± 7%
Maximum thickness − minimum thickness	126.5 ± 16.9	92.3 ± 22.7	< .001*	0.88 ± 0.03	78% ± 6%	65% ± 7%
Sector 11	90.2 ± 26.7	128.8 ± 25.1	< .001*	0.86 ± 0.04	87% ± 5%	35% ± 7%
Sector 1	78.6 ± 20.5	112.0 ± 23.7	< .001*	0.85 ± 0.04	76% ± 6%	28% ± 7%
Sector 12	88.1 ± 25.9	125.4 ± 23.4	< .001*	0.85 ± 0.04	72% ± 7%	46% ± 7%
Superior average	85.6 ± 20.1	112.7 ± 18.9	< .001*	0.85 ± 0.04	72% ± 7%	50% ± 7%
Sector 5	75.3 ± 20.3	104.4 ± 20.6	< .001*	0.83 ± 0.04	70% ± 7%	57% ± 7%
Sector 8	54.3 ± 16.9	77.0 ± 17.1	< .001*	0.83 ± 0.04	63% ± 7%	35% ± 7%
Temporal quadrant	55.3 ± 14.4	71.2 ± 13.5	< .001*	0.79 ± 0.05	72% ± 7%	30% ± 7%
Sector 2	67.2 ± 18.1	89.4 ± 21.3	< .001*	0.78 ± 0.05	57% ± 7%	37% ± 7%
Nasal quadrant	57.9 ± 13.2	70.8 ± 16.1	< .001†	0.74 ± 0.05	52% ± 7%	35% ± 7%
Sector 10	64.6 ± 20.9	81.7 ± 18.4	< .001*	0.73 ± 0.05	65% ± 7%	33% ± 7%
Sector 9	47.1 ± 14.7	55.2 ± 11.0	.004*	0.69 ± 0.06	54% ± 7%	33% ± 7%
Sector 4	57.0 ± 15.6	66.5 ± 18.1	.007†	0.66 ± 0.06	44% ± 7%	22% ± 6%
Sector 3	49.4 ± 13.1	56.8 ± 14.4	.008†	0.66 ± 0.06	35% ± 7%	20% ± 6%
S_{max}/I_{max}	1.12 ± 0.36	0.95 ± 0.13	.008†	0.66 ± 0.06	50% ± 7%	35% ± 7%
I_{max}/S_{max}	0.99 ± 0.37	1.08 ± 0.16	.008†	0.66 ± 0.06	50% ± 7%	35% ± 7%
I_{max}/T_{avg}	2.05 ± 0.65	2.38 ± 0.53	.010*	0.64 ± 0.06	54% ± 7%	41% ± 7%
S_{max}/N_{avg}	2.02 ± 0.48	2.26 ± 0.48	.020*	0.64 ± 0.06	44% ± 7%	26% ± 7%
S_{max}/T_{avg}	2.15 ± 0.57	2.22 ± 0.42	.494*	0.54 ± 0.06	33% ± 7%	20% ± 6%

I_{max} = inferior quadrant maximum thickness; N_{avg} = nasal quadrant average thickness; RNFL = retinal nerve fiber layer; ROC = receiver operator characteristic; SD = standard deviation; SE = standard error; S_{max} = superior quadrant maximum thickness; T_{avg} = temporal quadrant average thickness.

*Independent sample *t* test.

†Mann–Whitney *U* test.

rior averages are the average RNFL thicknesses in the superior and inferior quadrants of the calculation circle. TSNIT SD is the standard deviation of measurements along the calculation circle. An index called the nerve fiber indicator (NFI) was calculated with a neural network algorithm. It varied between zero and 100, with 100 representing cases with the most severe glaucoma. GDx VCC images with good alignment and fixation and a quality score of eight or more for both the corneal and RNFL images were included. The image quality scores were averaged and reported. Raw data were exported to a personal computer and used for further analysis.

Confocal scanning laser ophthalmoscopy was performed with the Heidelberg Retina Tomograph (HRT) II and data were analyzed with HRT III software (Heidelberg Engineering GmbH, Heidelberg, Germany). Three 15-degree topographic images, obtained at the same sitting, were aligned and averaged to obtain the mean topography. Only mean topographic images of good quality (standard deviation, ≤ 50 μm) were included. The following parameters, as calculated with HRT III software, were examined: disk area, cup area, rim area, cup-to-disk area ratio, rim-to-disk area ratio, cup volume, rim volume, mean cup depth, maximum cup depth, height variation contour, cup shape measure, mean RNFL thickness, RNFL cross-sectional area, horizontal and vertical cup-to-disk ratios, maximum contour elevation and depression, temporal-inferior contour line modulation (CLM), temporal-superior CLM, reference height, Frederick S. Mikelberg (FSM) discriminant function (Mikelberg and associates[9]), and Reinhard O. W. Burk (RB) discriminant function. Although the HRT III has the same scanning specifications as the HRT II, the HRT III uses an expanded normative database with ethnic-specific data.

TABLE 3. Scanning Laser Polarimetry with Variable Corneal Compensation (GDx VCC) Parameters in Glaucomatous and Normal Eyes, Areas under the Receiver Operator Characteristic (ROC) Curves, and Sensitivities at Fixed Specificities

Parameter	Glaucoma Group (mean ± SD)	Normal Group (mean ± SD)	*P* value	ROC Curve ± SE	Sensitivity at 80% Specificity ± SE	Sensitivity at 95% Specificity ± SE
NFI	51.6 ± 20.9	19.2 ± 8.2	< .001†	0.92 ± 0.03	89% ± 5%	78% ± 6%
Superior average	50.2 ± 10.4	65.9 ± 7.9	< .001*	0.88 ± 0.04	85% ± 5%	54% ± 7%
TSNIT SD	14.9 ± 4.0	21.3 ± 4.7	< .001*	0.85 ± 0.04	67% ± 7%	54% ± 7%
Inferior average	49.5 ± 13.0	64.8 ± 8.7	< .001*	0.84 ± 0.05	76% ± 6%	59% ± 7%
TSNIT average	44.6 ± 9.7	55.1 ± 5.9	< .001*	0.83 ± 0.05	80% ± 6%	63% ± 7%

NFI = nerve fiber indicator; ROC = receiver operator characteristic; SD = standard deviation; SE = standard error; TSNIT = temporal-superior-nasal-inferior-temporal.

Retinal nerve fiber layer thicknesses are in micrometers.

*Independent sample *t* test.

†Mann–Whitney *U* test.

Stereoscopic optic disk photographs (ODPs) were obtained sequentially with a fundus camera (Fundus Flash III; Carl Zeiss, Oberkochen, Germany) by an experienced ophthalmic photographer. The ODPs were reviewed by three experienced observers (J.C., K.N.M., S.L.). The reviewers were masked to each patient's identity, diagnosis, each other's scores, and all other clinical data. Each ODP was graded as 1 = normal, 2 = undetermined, or 3 = glaucoma, and a cumulative score was calculated by adding the scores assigned by the three observers. Ninety-two additional ODPs belonging to subjects not included in the study (normals, subjects suspected to have glaucoma on the basis of the optic diskappearance, and individuals with early to moderate disk damage) were added to the ODP pool to minimize evaluation bias. Image quality scores were assigned to each ODP pair with regard to image clarity (1 = acceptable, 2 = good, 3 = excellent) and stereoscopic quality (1 = flat, 2 = moderate depth effect, 3 = excellent stereopsis). Only pictures with good centration and quality were included in the study.

• **STATISTICAL ANALYSES:** The normality of the distribution of numeric variables was evaluated with the Wilk-Shapiro test. Normally distributed variables were compared with the independent sample *t* test. Numeric variables that were not normally distributed were compared with the Mann–Whitney *U* test. Receiver operator characteristic (ROC) curves were used to evaluate the performance of each imaging method for discriminating glaucomatous from normal eyes. Areas under the ROC curves (AUCs) were compared with the method described by DeLong and associates.[10] Specificity cut-offs of 80% and 95% were used to compare sensitivities of the best parameter (the one with the highest AUC) for each technique. Parameters with the highest sensitivities at 80% and 95% specificities were compared with McNemar test. The image quality for each of the four techniques also was assessed.

Classification and regression tree (CART) analysis (SPSS Answer Tree version 3.1; SPSS Inc, Chicago, Illinois, USA) was used to identify the best combination of quantitative parameters from the StratusOCT, GDx VCC, and HRT III methods. Maximum tree depth was specified at five. The minimum number of cases in parent and child nodes was set to 10 and five, respectively. To obtain the best diagnostic precision, the Chi-square automatic interaction detection growing method with 25-fold cross-validation and an α equal to 0.05 for splitting and merging of tree branches was used.

RESULTS

THE AVERAGE VISUAL FIELD MD (± SD) IN THE NORMAL group was 0.1 ± 1.2 dB and was −4.0 ± 2.5 dB in the glaucoma group ($P < .001$, *t* test). The demographic characteristics of the study sample are summarized in Table 1.

• **OPTICAL COHERENCE TOMOGRAPHY (STRATUSOCT):** The mean (± SD) and median (range) of the signal strengths in normal eyes were 8.8 (± 0.73) and eight (eight to nine) and in glaucomatous eyes were 8.5 (± 0.55) and 8 (eight to ten; $P = 0.09$, *t* test). StratusOCT parameters in normal and glaucomatous eyes, AUCs, and sensitivities at fixed specificities are presented in Table 2. The three parameters with the largest AUCs (± standard error [SE]) were: average RNFL thickness (0.96 ± 0.02), inferior quadrant RNFL thickness (0.95 ± 0.02), and RNFL thickness at inferotemporal sector 7 (0.93 ± 0.03).

• **SCANNING LASER POLARIMETER (GDx VCC):** The mean (± SD) and median (range) of image quality scores for GDx VCC in normal eyes were 8.8 (± 0.7) and nine (eight to nine) and in glaucomatous eyes were 8.5 (± 0.5) and eight (eight to 10; $P = .09$, *t* test). GDx VCC

TABLE 4. Heidelberg Retinal Tomograph III Parameters in Glaucomatous and Normal Eyes, Areas under the Receiver Operator Characteristic (ROC) Curves, and Sensitivities at Fixed Specificities

Parameter	Glaucoma Group (mean ± SD)	Normal Group (mean ± SD)	*P* Value	ROC Curve ± SE	Sensitivity at 80% Specificity ± SE	Sensitivity at 95% Specificity ± SE
FSM discriminant function	−1.69 ± 1.9	1.77 ± 1.7	< .001†	0.91 ± 0.03	87% ± 16%	70% ± 13%
Cup-to-disk area ratio	0.48 ± 0.2	0.2 ± 0.1	< .001†	0.91 ± 0.03	83% ± 15%	67% ± 12%
Rim-to-disk area ratio	0.52 ± 0.2	0.8 ± 0.1	< .001†	0.91 ± 0.03	83% ± 15%	67% ± 12%
Vertical cup-to-disk ratio	0.67 ± 0.2	0.31 ± 0.2	< .001*	0.90 ± 0.03	89% ± 18%	67% ± 13%
Rim area (mm^2)	0.98 ± 0.3	1.47 ± 0.3	< .001†	0.89 ± 0.04	83% ± 11%	50% ± 7%
Cup shape measure	−0.09 ± 0.1	−0.2 ± 0.1	< .001†	0.87 ± 0.04	83% ± 7%	33% ± 6%
Cup area (mm^2)	0.98 ± 0.5	0.39 ± 0.3	< .001†	0.87 ± 0.04	80% ± 10%	50% ± 8%
RB discriminant function	0.09 ± 1.1	1.51 ± 0.7	< .001*	0.86 ± 0.04	85% ± 16%	70% ± 9%
Cup volume (mm^3)	0.29 ± 0.3	0.09 ± 0.1	< .001*	0.85 ± 0.04	80% ± 11%	50% ± 10%
Rim volume (mm^3)	0.23 ± 0.1	0.41 ± 0.1	< .001†	0.84 ± 0.05	83% ± 13%	61% ± 8%
Mean RNFL thickness (mm)	0.18 ± 0.1	0.27 ± 0.1	< .001*	0.83 ± 0.05	80% ± 12%	50% ± 8%
Mean cup depth (mm)	0.31 ± 0.1	0.19 ± 0.1	< .001*	0.82 ± 0.04	74% ± 6%	33% ± 5%
Horizontal cup-to-disk ratio	0.68 ± 0.2	0.42 ± 0.2	< .001†	0.82 ± 0.04	70% ± 9%	38% ± 8%
Maximum contour elevation (mm)	0.01 ± 0.1	−0.09 ± 0.1	< .001†	0.81 ± 0.05	65% ± 8%	39% ± 8%
RNFL cross-sectional area (mm^2)	0.9 ± 0.5	1.29 ± 0.3	< .001†	0.8 ± 0.05	76% ± 8%	37% ± 9%
CLM temporal-inferior (mm)	0.07 ± 0.2	0.2 ± 0.1	< .001†	0.79 ± 0.05	63% ± 13%	43% ± 13%
Maximum cup depth (mm)	0.72 ± 0.2	0.54 ± 0.2	< .001†	0.72 ± 0.05	48% ± 6%	28% ± 4%
Maximum contour depression (mm)	0.44 ± 0.2	0.32 ± 0.1	< .001†	0.7 ± 0.05	48% ± 5%	15% ± 4%
CLM temporal-superior (mm)	0.16 ± 0.1	0.2 ± 0.1	.03†	0.63 ± 0.06	43% ± 4%	17% ± 5%
Reference height (mm)	0.39 ± 0.1	0.35 ± 0.1	.30†	0.56 ± 0.06	24% ± 2%	4% ± 2%
Disk area (mm^2)	1.95 ± 0.5	1.86 ± 0.4	.60†	0.54 ± 0.06	22% ± 2%	9% ± 2%
Height variation contour (mm)	0.43 ± 0.2	0.41 ± 0.1	.52*	0.5 ± 0.06	41% ± 4%	26% ± 3%

CLM = contour line modulation; FSM = Frederick S. Mikelberg; RB = Reinhard O. W. Burk; RNFL = retinal nerve fiber layer; ROC = receiver operator characteristic; SD = standard deviation; SE = standard error.

*Independent sample *t* test.

†Mann–Whitney *U* test.

parameters in normal and glaucomatous eyes, AUCs, and sensitivities at fixed specificities are presented in Table 3. All parameters were significantly different in the two subject groups ($P < .001$). The three parameters with the largest AUCs (± SE) were: NFI (0.92 ± 0.03), superior average RNFL thickness (0.88 ± 0.04), and TSNIT SD (0.85 ± 0.04).

• **SCANNING LASER OPHTHALMOSCOPE (HRT III):** The mean (± SD) and median (range) of standard deviation of image measurements (a quality measure for HRT III) in normal eyes were 19.2 μm (± 12.3 μm) and 16.5 μm (3 to 10 μm) and in glaucomatous eyes were 14.6 μm (± 4.6 μm) and 13.5 μm (9 to 29; $P < .001$, *t* test). Table 4 shows the HRT III parameters and the FSM discriminant function in normal and glaucomatous eyes, AUCs, and sensitivities at fixed specificities. All parameters, except temporal superior CLM, disk area, height variation contour, and reference height, were statistically different in the two groups ($P < .001$). The three parameters with the largest AUCs (± SE) were: FSM discriminant function (0.93 ± 0.03), cup-to-disk area ratio (0.91 ± 0.03), and rim-to-disk area ratio (0.91 ± 0.03).

• **OPTIC DISK STEREOPHOTOGRAPHS:** The mean (± SD) scores for image clarity in normal and glaucomatous eyes were 2.13 ± 0.78 and 2.69 ± 0.62 ($P < .001$), respectively. The mean (± SD) scores for stereopsis in normal and glaucomatous eyes were 1.87 ± 0.82 and 1.91 ± 0.74 ($P = .81$). The AUCs (± SE) for the cumulative score and for observers 1, 2, and 3 were 0.97 ± 0.02, 0.98 ± 0.02, 0.94 ± 0.03, and 0.89 ± 0.04, respectively. The agreement among observers as measured with the κ statistic was moderate to almost perfect on the scale developed by Landis and Koch, where $\kappa > 0.8$ is almost perfect, $\kappa = 0.6$ to 0.8 is substantial, $\kappa = 0.4$ to 0.6 is moderate, $\kappa = 0.2$ to 0.4 is fair, and $\kappa < 0.2$ is slight.[11] Observers 1 and 2 agreed in 76% of the cases ($\kappa = 0.61$; 95% confidence interval [CI], 0.47 to 0.75). Agreement between observers 2 and 3 was found in 89% of the cases ($\kappa = 0.82$; 95% CI, 0.71 to 0.92), whereas observers 1 and 3 agreed 74% of the time ($\kappa = 0.58$; 95% CI, 0.44 to 0.73).

• **COMPARISON OF IMAGING METHODS:** No significant difference was found among AUCs for the best parameters of the three devices (StratusOCT average RNFL thickness, GDx VCC nerve fiber indicator, and HRT III FSM

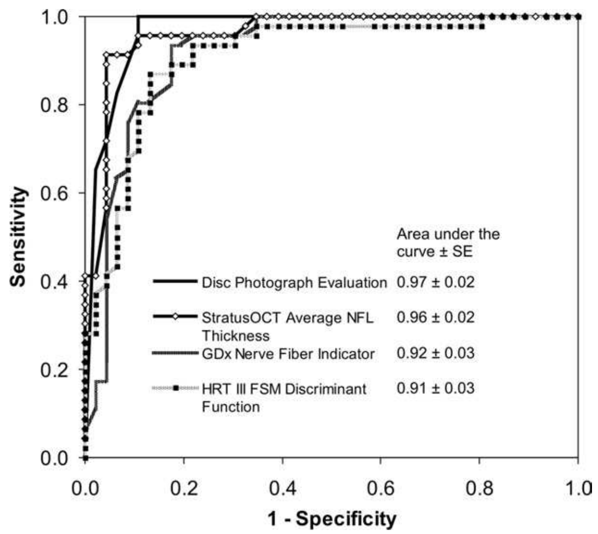

FIGURE 1. Graph showing a comparison of the areas under receiver operator characteristic curves (± standard error [SE]) of the best parameters from StratusOCT; average retinal nerve fiber layer thickness [RNFL], 0.96 ± 0.02), GDx VCC; nerve fiber indicator, 0.92 ± 0.03), Heidelberg Retina Tomograph (HRT) III; (Frederick S. Mikelberg [FSM] discriminant function, 0.91 ± 0.03), and the cumulative score of the three observers for disk photograph evaluation (0.97 ± 0.02).

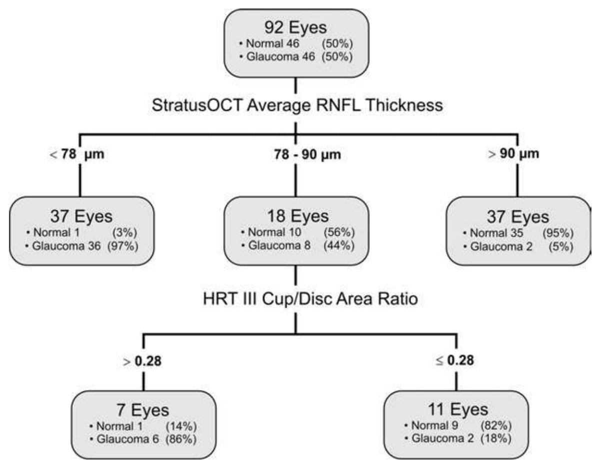

FIGURE 2. Flow chart showing results of classification and regression tree (CART) analysis for the combination of the best parameters from optical coherence tomography StratusOCT; scanning laser polarimetry GDx VCC; and scanning laser ophthalmoscopy HRT III. The three best parameters from each instrument were entered into the final CART analysis. With appropriate cut-off points for the StratusOCT's average RNFL thickness and HRT III's cup-to-disk area ratio, their combination provides a sensitivity of 91% (42 of 46 glaucomatous eyes identified as such) and a specificity of 96% (44 of 46 normal eyes classified as glaucomatous).

discriminant function) compared with the observers' qualitative assessment of optic diskphotographs (all P values $>$ .05, DeLong and associates test[10]). The corresponding ROC curves are shown in Figure 1. When parameters with the highest sensitivities at 80% specificities were compared, no significant differences were observed ($P > .05$ for all comparisons, McNemar test) among the four imaging methods. At 95% specificity, sensitivities were: StratusOCT average RNFL thickness, 89%; GDx VCC nerve fiber indicator, 78% HRT III FSM discriminant function, 70%; and the average of three observers' qualitative evaluation of disk photographs, 90%. The sensitivity at this specificity was higher for disk photographs than GDx VCC and HRT III parameters ($P = .05$ and $P = .002$, respectively, McNemar test) and was higher for StratusOCT than for HRT III ($P = .004$, McNemar test). There was no significant difference between StratusOCT and qualitative evaluation of disk photographs ($P = .98$, McNemar test).

• **OPTIMAL COMBINATION OF PARAMETERS:** CART analysis showed StratusOCT average RNFL thickness and HRT III cup-to-disk area to be the best combination of quantitative parameters to distinguish between normal and glaucomatous eyes. The combination of the two parameters provided a sensitivity of 91%, a specificity of 96%, and a diagnostic precision (± SE) of 93% ± 3%. These results are presented in Figure 2.

DISCUSSION

THIS STUDY COMPARED THE DIAGNOSTIC PERFORMANCE of the latest versions of optical coherence tomography (StratusOCT), scanning laser polarimetry (GDx VCC), confocal scanning laser ophthalmoscopy (HRT III), and qualitative assessment of stereoscopic optic disk photographs in eyes with early perimetric glaucoma. The three quantitative imaging instruments performed as well as, but not better than, the evaluation of optic disk photographs by glaucoma specialists. Studies comparing previous versions of the same imaging technologies with clinical evaluation of the optic nerve did not detect significant differences in performance between the automated imaging systems and experienced clinicians.[7,8] In the present investigation, the diagnostic precision of a large number of parameters was evaluated with AUCs and sensitivities at fixed specificities.

For StratusOCT, we found average RNFL thickness and thickness in the inferior quadrant to have the largest AUC and the greatest sensitivities at 80% and at 95% specificities. Our results are consistent with a recent report by Medeiros and associates comparing the ability of various StratusOCT algorithms to differentiate between normal and glaucomatous eyes.[12] In this study, among RNFL thickness, optic nerve head evaluation, and macular thickness parameters, those with the largest AUCs were average and inferior RNFL thickness (both with an AUC value of 0.91). Budenz and associates also reported that the RFNL thickness in the inferior quadrant and average RNFL thickness were the best parameters (AUCs, 0.97 and 0.96, respectively) for discriminating healthy eyes from those with moderate glaucoma with an average MD of −8.4.[13] The RNFL thickness in the inferior quadrant was found to have the best ROC curve both with the StratusOCT and with OCT 2000 in a recent study by Bourne and associates, who compared the two instruments.[14] Previous studies with the OCT 2000 also reported that inferior quadrant RNFL thickness and average RNFL thickness best differentiated between normal and glaucomatous eyes.[15–17]

Rim area and cup shape measurement were found to have the largest AUCs in two studies investigating diagnostic ability of the HRT I (Iester and associates[18] and Greaney and associates[7]). Cup-to-disk area ratio and rim volume were the HRT parameters with the highest diagnostic precision in a study by Schuman and associates comparing OCT with HRT I.[19] It should be noted that none of the above studies evaluated the performance of the FSM discriminant function. Although the FSM discriminant function is not provided in the HRT printout, it can be obtained from the stereometric parameters displayed on the monitor. Many investigators have shown that multivariate discriminant analysis of combinations of parameters provides better diagnostic precision compared with any single variable.[20–23] With the HRT III software, we found the FSM discriminant function, cup-to-disk area ratio, and rim-to-disk area ratio to have the largest AUCs (0.91 for all). The FSM discriminant function had the highest sensitivity (70%) at 95% specificity.

Among GDx VCC parameters, NFI and superior average RNFL thickness had the largest AUCs and the highest sensitivities at 80% and 95% specificities. These results are consistent with those reported by Reus and Lemij.[24] They evaluated the ability of GDx VCC to discriminate normal eyes from those with moderately advanced glaucomatous visual field defects (average MD, −8.5). The NFI and superior average RNFL thickness had the largest AUC in their study (0.98 and 0.94, respectively). The higher diagnostic accuracy reported in the above study may be attributed to the more advanced stage of glaucoma. Our results are also in line with those obtained by Colen and associates, who described the NFI as the most accurate of all the GDx parameters (AUC, 0.90),[25] and with those of Medeiros and associates, who used the GDx VCC to discriminate normal eyes from those with progressive optic disk change.[26] The NFI was the best parameter from the GDx VCC for discrimination of healthy eyes from eyes that demonstrated progressive optic disk change in that study (AUC, 0.92).[27,28] Another study comparing the GDx VCC and the older version of the GDx with fixed corneal compensation reported similar findings and showed that GDx performance has improved after the introduction of variable corneal compensation.[29]

Medeiros and associates[30] compared the ability of StratusOCT, GDx VCC, and HRT II to discriminate between normal and glaucomatous eyes. They found these parameters to have the largest AUC: inferior and average RNFL thickness for StratusOCT (AUCs, 0.92 and 0.91, respectively); the RB and FSM discriminant functions and linear cup-to-disk ratio for HRT II (AUCs, 0.86, 0.83, and 0.83, respectively); and NFI, inferior normalized area, and TSNIT average for GDx VCC (AUCs, 0.91, 0.86, and 0.85, respectively). These results are similar to ours, and no significant difference among the largest AUCs was found among the different imaging methods in both investigations. Statistically significant differences in sensitivities at 80% specificities were reported by Medeiros and associates, where OCT (inferior RNFL thickness) and GDx (NFI) performed significantly better than HRT (FSM discriminant function).

A recent study of StratusOCT, GDx VCC, HRT II, and disk stereophotograph grading by DeLeön and associates showed qualitative assessment of the disk to be the parameter with the largest AUC (0.90) and with the highest sensitivity (77%) at 80% specificity.[31] We compared StratusOCT, GDx VCC, and HRT III with clinical evaluation of stereoscopic optic disk photographs to distinguish between healthy eyes and eyes with early to moderately advanced perimetric glaucoma. Qualitative evaluation of optic disk photographs by experienced observers performed as well as any of the three quantitative imaging instruments. The diagnostic accuracy of qualitative methods depends to a large extent on the observer's experience; therefore, the results may be different in other clinical settings and with other observers.

Various possible combinations of quantitative parameters were explored with CART analysis. The combination of StratusOCT average RNFL thickness and HRT III cup-to-disk area ratio was found to perform best among the quantitative parameters with a cross-validated diagnostic precision of 93%. The classification tree in Figure 2 shows that appropriate cut-offs for StratusOCT's average RNFL thickness and HRT III's cup-to-disk area ratio provided a sensitivity of 91% (42 of 46 glaucomatous eyes identified as such) and a specificity of 96% (44 of 46 normal eyes classified as normal). If the above sensitivity and specificity values are reproduced in subsequent studies, a reasonable method for glaucoma screening could be developed based on such a combination of parameters. However, further studies are necessary to validate the effectiveness of this approach. A similar approach was used in the past by Magacho and associates,[32] who reported similar findings using multivariate discriminant formulas to combine quantitative parameters from GDx and the Topographic Scanning System (formerly manufactured by Diagnostic Laser Technologies). In view of the high performance of the techniques investigated, very large study samples may be necessary to provide the power to detect small but statistically significant differences among the techniques. Such small differences, however, may not be relevant clinically.

In conclusion, we studied the diagnostic performance of qualitative evaluation of stereoscopic optic disk photographs by glaucoma specialists and contemporary versions of three quantitative imaging techniques (StratusOCT GDx VCC, and HRT III) in patients with early to moderate perimetric glaucoma. Each of the quantitative imaging techniques independently performed as well as, but not better than, evaluation of stereophotographs by experienced clinicians. However, clinical assessment of optic disk photos is influenced largely by the examiner's experience, and results may vary in different clinical settings. Among the quantitative techniques, StratusOCT was more sensitive for detection of glaucoma at high specificity (95%) than was HRT III. The best categorization of patients was obtained with a combination of the quantitative parameters StratusOCT, average RNFL thickness, and HRT III cup-to-disk area ratio, where the values for sensitivity (91%) and specificity (96%) approach those required for successful screening of early perimetric glaucoma.

THIS STUDY WAS SUPPORTED IN PART BY A PHYSICIAN SCIENTIST AWARD TO J.C. FROM RESEARCH TO PREVENT BLINDNESS, Inc, New York, New York, and Grant no. R01-EY12738 from the National Institutes of Health, Bethesda, Maryland. The authors indicate no financial conflict of interest. Involved in design and conduct of study (F.B., K.N.-M., J.C.); data collection and management (F.B., D.R., N.L.); data analysis and interpretation (F.B., K.N.-M., S.K.L., J.C.); and manuscript preparation and review (F.B., D.R., K.N.-M., J.C.). This study was approved by the Institutional Review Board of the University of California Los Angeles and adhered to the tenets of the Declaration of Helsinki. Each patient provided written informed consent for inclusion in the study.

The authors would like to thank Jeffrey A. Gornbein, Department of Biomathematics, UCLA School of Medicine, Los Angeles, California, for his valuable assistance with statistical computing.

REFERENCES

1. Sommer A, Katz J, Quigley HA, et al. Clinically detectable nerve fiber atrophy precedes the onset of glaucomatous field loss. Arch Ophthalmol 1991;109:77–83.
2. Johnson CA, Sample PA, Zangwill LM, et al. Structure and function evaluation (SAFE): II. Comparison of optic disk and visual field characteristics. Am J Ophthalmol 2003;135: 148–154.
3. Tuulonen A, Airaksinen PJ. Initial glaucomatous optic disk and retinal nerve fiber layer abnormalities and their progression. Am J Ophthalmol 1991;111:485–490.
4. Airaksinen PJ, Drance SM, Douglas GR, Schulzer M, Wijsman K. Visual field and retinal nerve fiber layer comparisons in glaucoma. Arch Ophthalmol 1985;103:205–207.
5. Kass MA, Heuer DK, Higginbotham EJ, et al. The Ocular Hypertension Treatment Study: a randomized trial determines that topical ocular hypotensive medication delays or prevents the onset of primary open-angle glaucoma. Arch Ophthalmol 2002;120:701–713; discussion 829–830.
6. Harwerth RS, Carter-Dawson L, Shen F, Smith EL III, Crawford ML. Ganglion cell losses underlying visual field defects from experimental glaucoma. Invest Ophthalmol Vis Sci 1999;40:2242–2250.
7. Greaney MJ, Hoffman DC, Garway-Heath DF, Nakla M, Coleman AL, Caprioli J. Comparison of optic nerve imaging methods to distinguish normal eyes from those with glaucoma. Invest Ophthalmol Vis Sci 2002;43:140–145.
8. 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.
9. Mikelberg FS, Parfitt CM, Swindale NV, Iester M. Ability of the Heidelberg retina tomograph to detect early glaucomatous visual field loss. J Glaucoma 1995;4:242–247.
10. 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.
11. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics 1977;33:159–174.
12. Medeiros FA, Zangwill LM, Bowd C, Vessani RM, Susanna R Jr, Weinreb RN. 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.
13. Budenz DL, Michael A, Chang RT, McSoley J, Katz J. Sensitivity and specificity of the StratusOCT for perimetric glaucoma. Ophthalmology 2005;112:3–9.
14. Bourne RR, Medeiros FA, Bowd C, Jahanbakhsh K, Zangwill LM, Weinreb RN. Comparability of retinal nerve fiber layer thickness measurements of optical coherence tomography instruments. Invest Ophthalmol Vis Sci 2005;46:1280–1285.
15. Mok KH, Lee VW, So KF. Retinal nerve fiber layer measurement by optical coherence tomography in glaucoma suspects with short-wavelength perimetry abnormalities. J Glaucoma 2003;12:45–49.
16. Kanamori A, Nakamura M, Escano MF, Seya R, Maeda H, Negi A. Evaluation of the glaucomatous damage on retinal nerve fiber layer thickness measured by optical coherence tomography. Am J Ophthalmol 2003;135:513–520.
17. Nouri-Mahdavi K, Hoffman D, Tannenbaum DP, Law SK, Caprioli J. Identifying early glaucoma with optical coherence tomography. Am J Ophthalmol 2004;137:228–235.
18. Iester M, Mikelberg FS, Swindale NV, Drance SM. ROC analysis of Heidelberg retina tomograph optic disc shape measures in glaucoma. Can J Ophthalmol 1997;32:382–388.
19. Schuman JS, Wollstein G, Farra T, et al. Comparison of optic nerve head measurements obtained by optical coherence tomography and confocal scanning laser ophthalmoscopy. Am J Ophthalmol 2003;135:504–512.
20. Iester M, Jonas JB, Mardin CY, Budde WM. Discriminant analysis models for early detection of glaucomatous optic disc changes. Br J Ophthalmol 2000;84:464–468.
21. Uchida H, Brigatti L, Caprioli J. Detection of structural damage from glaucoma with confocal laser image analysis. Invest Ophthalmol Vis Sci 1996;37:2393–2401.

22. Iester M, Mardin CY, Budde WM, Junemann AG, Hayler JK, Jonas JB. Discriminant analysis formulas of optic nerve head parameters measured by confocal scanning laser tomography. J Glaucoma 2002;11:97–104.
23. Iester M, Mikelberg FS, Drance SM. The effect of optic disc size on diagnostic precision with the Heidelberg retina tomograph. Ophthalmology 1997;104:545–548.
24. Reus NJ, Lemij HG. Diagnostic accuracy of the GDx VCC for glaucoma. Ophthalmology 2004;111:1860–1865.
25. Colen TP, Tang NE, Mulder PG, Lemij HG. Sensitivity and specificity of new GDx parameters. J Glaucoma 2004;13:28–33.
26. Medeiros FA, Zangwill LM, Bowd C, Sample PA, Weinreb RN. Use of progressive glaucomatous optic disk change as the reference standard for evaluation of diagnostic tests in glaucoma. Am J Ophthalmol 2005;139:1010–1018.
27. Henderson PA, Medeiros FA, Zangwill LM, Weinreb RN. Relationship between central corneal thickness and retinal nerve fiber layer thickness in ocular hypertensive patients. Ophthalmology 2005;112:251–256.
28. Mohammadi K, Bowd C, Weinreb RN, Medeiros FA, Sample PA, Zangwill LM. Retinal nerve fiber layer thickness measurements with scanning laser polarimetry predict glaucomatous visual field loss. Am J Ophthalmol 2004;138:592–601.
29. Weinreb RN, Bowd C, Zangwill LM. Glaucoma detection using scanning laser polarimetry with variable corneal polarization compensation. Arch Ophthalmol 2003;121:218–224.
30. Medeiros FA, Zangwill LM, Bowd C, Weinreb RN. Comparison of the GDx VCC scanning laser polarimeter, HRT II confocal scanning laser ophthalmoscope, and StratusOCT optical coherence tomograph for the detection of glaucoma. Arch Ophthalmol 2004;122:827–837.
31. DeLeön JE, Arthur SN, McGwin G Jr, Xie A, Monheit BE, Girkin CA. Discrimination between glaucomatous and nonglaucomatous eyes using quantitative imaging devices and subjective optic nerve head assessment. Invest Ophthalmol Vis Sci 2006;47:3374–3380.
32. Magacho L, Marcondes AM, Costa VP. Discrimination between normal and glaucomatous eyes with scanning laser polarimetry and optic disc topography: a preliminary report. Eur J Ophthalmol 2005;15:353–359.