Comparison of HRT-3 Glaucoma Probability Score and Subjective Stereophotograph Assessment for Prediction of Progression in Glaucoma

Luciana M. Alencar,[1,2] *Christopher Bowd,*[1] *Robert N. Weinreb,*[1] *Linda M. Zangwill,*[1] *Pamela A. Sample,*[1] *and Felipe A. Medeiros*[1,2]

PURPOSE. To assess whether baseline Glaucoma Probability Score (GPS; HRT-3; Heidelberg Engineering, Dossenheim, Germany) results are predictive of progression in patients with suspected glaucoma. The GPS is a new feature of the confocal scanning laser ophthalmoscope that generates an operator-independent, three-dimensional model of the optic nerve head and gives a score for the probability that this model is consistent with glaucomatous damage.

METHODS. The study included 223 patients with suspected glaucoma during an average follow-up of 63.3 months. Included subjects had a suspect optic disc appearance and/or elevated intraocular pressure, but normal visual fields. Conversion was defined as development of either repeatable abnormal visual fields or glaucomatous deterioration in the appearance of the optic disc during the study period. The association between baseline GPS and conversion was investigated by Cox regression models.

RESULTS. Fifty-four (24.2%) eyes converted. In multivariate models, both higher values of GPS global and subjective stereophotograph assessment (larger cup–disc ratio and glaucomatous grading) were predictive of conversion: adjusted hazard ratios (95% CI): 1.31 (1.15–1.50) per 0.1 higher global GPS, 1.34 (1.12–1.62) per 0.1 higher CDR, and 2.34 (1.22–4.47) for abnormal grading, respectively. No significant differences ($P > 0.05$ for all comparisons) were found between the c-index values (equivalent to area under ROC curve) for the multivariate models (0.732, 0.705, and 0.699, respectively).

CONCLUSIONS. GPS values were predictive of conversion in our population of patients with suspected glaucoma. Further, they performed as well as subjective assessment of the optic disc. These results suggest that GPS could potentially replace stereophotograph as a tool for estimating the likelihood of conversion to glaucoma. (*Invest Ophthalmol Vis Sci.* 2008;49: 1898–1906) DOI:10.1167/iovs.07-0111

S everal studies have shown that the appearance of the optic nerve and retinal nerve fiber layer (RNFL) can precede visual field loss in glaucoma.[1,2] However, optic disc evaluation is subjective, and stereoscopic photographs are not readily obtained in clinical practice. In fact, several research groups have found a low rate of conformance with recommended optic disc examination and documentation practice patterns for glaucoma.[3–5] In a retrospective review of 395 medical charts, Fremont et al.[4] found that almost half of the patients with glaucoma did not have a photograph or drawing of the optic nerve head made at the time of their initial evaluation.

Recently, several imaging technologies have become available for structural evaluation of the optic disc and RNFL and to assist in the diagnosis of glaucoma. One of these technologies, confocal scanning laser ophthalmoscopy (CSLO) provides objective measurements of optic disc topography that have been shown to predict progression in those with suspected glaucoma.[6,7] As part of the ocular hypertension treatment study, Zangwill et al.[6] showed that CSLO stereometric parameters and the Moorfields regression analysis (MRA) had a significant predictive value in discriminating those ocular hypertensive patients who converted to glaucoma from those who did not.

Previous CSLO measures used for evaluation of optic disc topography have been limited by the need for an examiner to approximate the optic disc margin with a contour line to calculate stereometric parameters and the MRA. This requirement added an undesirable subjectivity to the examination and may have resulted in significant differences in the topographic parameter values obtained by different examiners.[8] Also, accurate tracing of the contour line optimally requires simultaneous visualization of optic disc photographs,[8] which obviously decreases the value of CSLO as a replacement method for stereophotographs in the assessment and documentation of the optic disc. To overcome this limitation, a new version of CSLO (HRT-3; Heidelberg Engineering GmbH, Dossenheim, Germany) includes the Glaucoma Probability Score (GPS), an index that is independent of the contour line traced by the examiner. The GPS is based on a three-dimensional model of the entire topographical image and has been shown to discriminate glaucomatous from normal subjects in cross-sectional studies.[9,10]

The purpose of the present study was to evaluate the ability of the contour-line–independent parameter GPS to predict the development of visual field loss or optic disc deterioration in persons with suspected glaucoma. Further, the predictive ability of this parameter was compared to that of subjective stereophotograph assessment.

METHODS

This was an observational cohort study. Patients in this study participated in a prospective longitudinal study designed to evaluate optic

nerve structure and visual function in glaucoma (DIGS; Diagnostic Innovations in Glaucoma Study) conducted at the Hamilton Glaucoma Center (University of California, San Diego; UCSD). All patients from the DIGS who met the inclusion criteria described below were enrolled in the present study. Informed consent was obtained from all participants. The UCSD Human Subjects Committee approved all protocols and the methods described adhered to the tenets of the Declaration of Helsinki.

Subjects were defined as having suspected glaucoma according to the clinical examination by two glaucoma specialists (FAM, RNW). Included were those with suspect optic disc appearance (as determined by subjective assessment) and/or elevated intraocular pressure (>21 mm Hg). All subjects had normal and reliable standard automated perimetry (SAP) visual fields at baseline, as defined later in the paper. Eligible subjects were required to have had a visual field examination and optic disc stereophotograph taken close in time to a baseline HRT scan used for evaluation. Baseline was set at the first occurrence of this matching, and the HRT date was used as the baseline date. The average time interval between examinations was 1.4 months (median: 0.6 months, first quartile: 0.2 months, third quartile: 1.7 months). Only subjects with open angles on gonioscopy were included. Subjects were excluded if they presented best-corrected visual acuity less than 20/40, spherical refraction outside ± 5.0 D and/or cylinder correction outside 3.0 D, or any other ocular or systemic disease that could affect the optic nerve or the visual field. One eye of each patient was randomly selected for analysis.

For each eye, central corneal thickness (CCT) was calculated as the average of three measurements obtained during the same visit using an ultrasound pachymeter (Pachette GDH 500; DGH Technology, Inc, Philadelphia, PA).

Standard Automated Perimetry

Only patients with normal and reliable visual fields on the baseline were included. Standard automated perimetry (SAP) visual fields were obtained using either 24-2 Full Threshold or Swedish Interactive Thresholding Algorithm (SITA) (Humphrey Field Analyzer; Carl Zeiss Meditec, Dublin, CA) strategies. Only tests with reliable results ($\leq$33% fixation losses, false positives, and false negatives) were included.

Glaucomatous conversion by visual field was defined as the development of three consecutive abnormal examinations during follow-up, or two consecutive when these were the last examination results available during follow-up. An abnormal result followed by a normal one was not considered conversion. An abnormal visual field was defined as a pattern standard deviation (PSD) with $P < 0.05$ and/or a glaucoma hemifield test (GHT; Humphrey Perimeter; Carl Zeiss Meditec, Inc., Oberkochen, Germany) with results outside normal limits. Two experienced glaucoma specialists verified that the visual field defects were consistent with glaucoma.

Stereophotograph Grading

Simultaneous stereoscopic optic disc photographs (TRC-SS; Topcon Instrument Corp of America, Paramus, NJ) were reviewed with a stereoscopic viewer (Pentax Stereo Viewer II; Asahi Optical Co., Tokyo, Japan). Baseline stereophotographs were evaluated by two masked, experienced graders and classified as glaucomatous or normal. Glaucomatous optic disc appearance was defined based on the presence of neuroretinal rim thinning, excavation, notching, or characteristic retinal nerve fiber layer defects. Vertical cup-to-disc ratio (CDR) was assessed by visually estimating the CDR based on the contour of the cup. The average value between examiners was calculated and used for analysis. For progression assessment, each patient's most recent stereophotograph was compared with the baseline one. Each grader was masked to the temporal sequence of the photographs. Definition of change was based on focal or diffuse thinning of the neuroretinal rim, increased excavation, and the appearance or enlargement of RNFL defects.

Discrepancies between the two graders were resolved either by consensus or by adjudication of a third experienced grader. Only photographs with adequate quality were included. From an initial group of 310 patients who fulfilled inclusion criteria, 5 (2%) subjects had poor-quality photographs at baseline, and 29 (9%) did not have follow-up stereophotographs to assess progression and were excluded from further analysis.

Confocal Scanning Laser Ophthalmoscopy

CSLO images were acquired using either the HRT-I or –II (Heidelberg Engineering, GmbH) and analyzed on each respective machine, using HRT-3 software. Further details on these instruments have been described previously.[11-13] Only 15° images were used. For each patient, three topographical images were obtained and then combined and automatically aligned to make a single mean topography used for analysis. Magnification errors were corrected using patients' corneal curvature measurements. Good-quality images required a focused reflectance image with a standard deviation not greater than 50 μm and centered GPS analysis. From an initial group of 310 patients who fulfilled inclusion criteria, 15 (5%) were excluded because the 15° HRT baseline image could not be retrieved, 26 (8%) were excluded after quality control of the HRT mean image, 4 (1%) were excluded because the HRT was not able to run the GPS analysis, and 11 (4%) were excluded as a result of highly off centered analysis of the GPS algorithm.

The GPS is obtained using a new automated analysis independent of either contour line tracing or a reference plane. The software analyzes the optic disc and parapapillary retina topography and builds a three-dimensional (3-D) model using five shape-based measures: cup size, cup depth, and rim steepness (referring to the optic disc) and vertical (superior to inferior) and horizontal (nasal to temporal) parapapillary nerve fiber layer curvatures. The values of the parameters are then fed into a machine-learning classifier analysis, a relevance vector machine (RVM), which compares the patient's results to previously defined healthy and glaucomatous models. Glaucomatous eyes usually present with flatter RNFL curvature and increased cup size, depth, and slope (rim steepness). The final GPS is the probability that the model has structural differences from the normal model that are compatible with glaucomatous damage. The higher the GPS, the more similar it is to the glaucoma model.

GPS results were obtained for the global region, as well as for six predefined sectors (with 0° as temporal): temporal superior (45°–90°), nasal superior (91°–135°), nasal (136°–225°), nasal inferior (226°–270°), temporal inferior (271°–315°), and temporal (316°–44°).

Follow-up and Definition of Study Endpoints

Conversion to glaucoma in this study was defined by either visual field test results or optic disc stereophotograph evaluation, whichever came first. Eyes in which a confirmed visual field defect developed or optic disc deterioration on stereophotographs were referred to as converters.

For converters, follow-up time was defined as the time between the HRT baseline visit and the date of the first abnormal visual field result or the first optic disc stereophotograph showing deterioration (the study endpoint). For nonconverters in both visual field and optic disc evaluation, follow-up time was defined as the time between the HRT baseline visit and date of last available visual field test result or stereophotograph on DIGS, whichever came first. During follow-up time, each patient was treated at the discretion of the attending ophthalmologist.

Statistical Analysis

The primary purpose of the study was to determine whether the baseline GPS is predictive of progression. Other variables analyzed as potential risk factors were age, baseline IOP, CCT, and the baseline SAP visual field index PSD. Hazard ratios (HRs) for the association between GPS parameters and the development of a documented progression were obtained by Cox proportional hazards models. We report HRs from univariate models, which do not adjust for the presence of other factors, as well as adjusted HRs from multivariate Cox proportional

hazards models. For the multivariate models, we report hazard ratios after adjustment for age, baseline IOP, CCT, and SAP PSD. These variables have been reported to be significantly associated with the risk of development of glaucomatous visual field loss or optic disc deterioration among patients with ocular hypertension or suspected glaucoma.[14–16]

We also evaluated the ability of subjective stereophotograph evaluation (grading and vertical cup/disc ratio) in predicting the development of progression. Univariate hazard ratios were reported for stereophotograph grading (glaucoma versus normal) as well as for vertical CDR. Adjusted HRs were also reported for these variables after adjustment for age, baseline IOP, CCT, and SAP PSD.

As the magnitude of a hazard ratio for a particular variable depends on its unit of measurement, a direct comparison of HRs would be an inappropriate way of comparing the predictive abilities of GPS and stereophotograph assessment. For this purpose, we used the c-index, as suggested by Harrell.[17] The c-index is similar to the area under the receiver operating characteristic (ROC) curve and is frequently used to evaluate the discriminating ability of predictive models in survival data. It is calculated as the proportion of all usable subject pairs in which the predictions and outcomes are concordant. If the predicted survival time is larger for the subject who actually survived longer, the predictions of the pair are concordant with the outcomes. In predicting the time to an event, c is calculated by including all possible pairs of subjects, at least one of whom has experienced the event (viz., progression). Two subjects' survival times cannot be ordered if both subjects are censored or if one has failed and the follow-up time of the other is less than the failure time of the first.[18] A c-index of 0.5 indicates random predictions, whereas 1.0 indicates perfect prediction. The c-index was calculated for multivariate models, including GPS, and adjusting for age, baseline IOP, CCT, and SAP PSD, as well as for multivariate models including stereophotograph parameters and adjusting for the same variables. Therefore, each multivariate model contained the combination of an optic disc parameter (objective versus subjective) plus other variables previously identified as significantly associated with the risk of the development of glaucoma. To test for the significance of the difference in discrimination between two models, we used the rcorrp.cens function from Harrell's Hmisc/Design library.[17] This computes U statistics for testing whether the predictions of one model are more concordant with actual observations than those of another model.

To adjust for potentially confounding effects of treatment, these analyses were also performed incorporating treatment as a time-dependent covariate.

Statistical analyses were performed with commercial software packages (SPSS, ver. 15.0; SPSS, Chicago, IL; Stata, ver. 9.0; StataCorp, College Station, TX; and S-PLUS ver. 6.0; Insightful Corp., Seattle, WA). The α level (type I error) was set at 0.05.

RESULTS

Two-hundred and twenty-three eyes of 223 patients with suspected glaucoma were included. Table 1 shows baseline demographic and clinical characteristics of the patients included in the study. Fifty-four eyes (24.2%) showed conversion during the follow-up. Of these, 13 (24%) showed progression based first on development of optic disc changes, 32 (59.3%) based first on progressive visual field loss, and 9 (16.7%) based on both optic disc and visual field changes concomitantly. For the 41 patients whose endpoint was determined by VF change, 33 had 3 consecutive abnormal examination results, and 8 had only two repeatable abnormal visual fields, which were the last available examinations during follow-up. None of these eight cases showed progression by stereophotographs during the study period. Exclusion of these eight converters did not change the results of the study. Therefore, we report only analyses including the full group. Mean follow-up time until conversion was 53.1 ± 31.0 months (median 47.7 months; range, 11.4–132.9 months). Mean follow-up time for nonconverters was 66.6 ± 38.7 months (median 59.0 months; range, 6.5–156.1 months). Figure 1 illustrates the Kaplan-Meier estimated cumulative probability of development of visual field loss or optic disc changes during the study. At 5 years of follow-up, the overall cumulative probability of conversion was 21%.

Table 2 shows univariate HRs with 95% CI for each putative predictive factor for conversion. In univariate analyses, both global and sectoral GPS results were significantly associated with progression, with HRs ranging from 1.29 (temporal superior and temporal inferior) to 1.31 (global, temporal, and nasal inferior). A 0.1 larger global GPS was associated with a 31% increase in the risk of progression (HR: 1.31 per 0.1 higher; 95% CI: 1.16–1.47; $P < 0.001$).

HRs were also assessed for GPS parameters classified according to the manufacturer's suggested cutoffs for global and sectoral analyses, and for the final classification provided on the HRT 3.0 printout (Table 3). According to these cutoffs, GPS results between 0 and 0.27 are classified as within normal limits, between 0.28 and 0.64 as borderline, and between 0.65 and 1.0 as outside normal limits. The final classification provided is outside normal limits if any sector or if the global

TABLE 1. Baseline Demographic and Clinical Characteristics of Study Patients

	All ($n = 223$)	Converters ($n = 54$)	Nonconverters ($n = 169$)
Age (y)	59.0 ± 12.7	63.7 ± 11.6	57.5 ± 12.7
IOP (mm Hg)	22.5 ± 5.7	21.7 ± 5.6	22.8 ± 5.7
CCT (μm)	565 ± 38	548 ± 39	570 ± 36
PSD (dB)	1.94 ± 0.68	2.32 ± 0.84	1.82 ± 0.56
Vertical cup/disc ratio	0.59 ± 0.19	0.68 ± 0.17	0.56 ± 0.18
Glaucomatous grading, n (%)	113 (50)	40 (73)	73 (43)
History of high IOP, n (%)*	191 (85)	49 (89)	142 (84)
GPS global	0.34 ± 0.23	0.43 ± 0.24	0.31 ± 0.21
GPS temporal	0.34 ± 0.23	0.44 ± 0.25	0.31 ± 0.21
GPS temporal superior	0.31 ± 0.22	0.41 ± 0.24	0.28 ± 0.21
GPS temporal inferior	0.34 ± 0.23	0.45 ± 0.24	0.31 ± 0.22
GPS nasal	0.33 ± 0.23	0.43 ± 0.25	0.30 ± 0.21
GPS nasal superior	0.31 ± 0.22	0.40 ± 0.24	0.28 ± 0.20
GPS nasal inferior	0.32 ± 0.22	0.42 ± 0.24	0.29 ± 0.21

Converters are patients who experienced development of either visual field loss or optic disc deterioration, and nonconverters are those who did not. Data are the mean ± SD, unless otherwise stated.
* IOP ≥ 21 mm Hg.
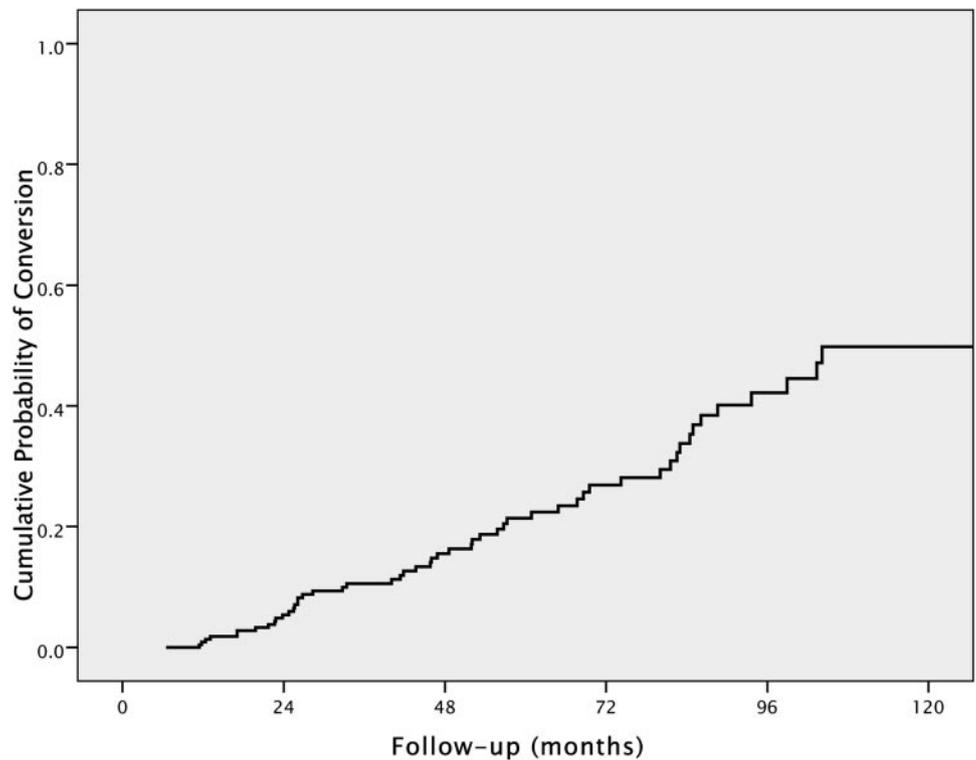
FIGURE 1. Cumulative probability (Kaplan-Meier survival curve) of development of visual field loss or optic disc deterioration during the study.

assessment is flagged as outside normal limits. Using the within normal limits result as the reference category, an outside normal limits result on the GPS final classification had a univariate HR of 4.70 (95% CI: 2.27–9.75; *P* < 0.001), and a borderline classification had a univariate HR of 2.00 (95% CI: 1.02–3.94; *P* = 0.044). Figure 2 shows the cumulative probability (Kaplan-Meier survival curves) for development of visual field loss or optic disc changes according to the categorization suggested by the manufacturer for the final classification (*P* < 0.001, log rank test).

A glaucomatous grading on the baseline stereophotograph was also significantly associated with progression (HR: 3.08; 95% CI: 1.69–5.61; *P* < 0.001). Figure 3 shows the cumulative

probability (Kaplan-Meier survival curves) for development of visual field loss or optic disc changes according to the categorization used for stereophotograph grading. For vertical cup/disc ratio, the univariate HR was 1.41 per 0.1 larger (95% CI: 1.19–1.67; *P* < 0.001).

When multivariate models were constructed adjusting for age, baseline IOP, CCT and SAP PSD, the GPS global, and sectoral values were still significantly associated with increased risk of progression to glaucoma, with adjusted HRs ranging from 1.29–1.32 for continuous variables (Table 2). Each 0.1 larger value of Global GPS was associated with a 31% increase in the risk of converting to glaucoma (adjusted HR: 1.31; 95% CI: 1.15–1.50; *P* < 0.001). The adjusted HRs for outside normal

TABLE 2. Univariate and Multivariate HRs with 95% CIs for the Development of Visual Field Loss or Optic Disc Deterioration

	Univariate		Multivariate		
	HR	95% CI	HR	95% CI	*c*-Index
Age (per decade)	1.53*	1.19–1.98	—	—	—
CCT (per 40 μm thinner)	1.73*	1.29–2.33	—	—	—
PSD (per 0.2 larger)	1.08*	1.03–1.13	—	—	—
IOP (per 1 mm Hg higher)	0.98	0.93–1.03	—	—	—
CDR (per 0.1 larger)	1.41*	1.19–1.67	1.34*	1.12–1.62	0.705
Photograph grading (glaucomatous)	3.08*	1.69–5.61	2.34*	1.22–4.47	0.699
GPS (per 0.1 higher)					
Global	1.31*	1.16–1.47	1.31*	1.15–1.50	0.732
Temporal	1.31*	1.17–1.47	1.32*	1.15–1.50	0.732
Temporal superior	1.29*	1.16–1.44	1.29*	1.14–1.46	0.739
Temporal inferior	1.29*	1.15–1.44	1.29*	1.13–1.46	0.731
Nasal	1.30*	1.16–1.46	1.30*	1.14–1.48	0.733
Nasal superior	1.30*	1.16–1.46	1.30*	1.14–1.49	0.733
Nasal inferior	1.31*	1.17–1.48	1.32*	1.15–1.51	0.727

The *c*-index is given for each multivariate model, adjusted for age, baseline IOP, CCT, and PSD, CDR, cup-to-disc ratio.

* Significant at *P* < 0.05.

TABLE 3. Univariate and Multivariate HR with 95% CI for the Development of Visual Field Loss or Optic Disc Deterioration

	Univariate		Multivariate	
	HR	95% CI	HR	95% CI
GPS (outside vs. within normal limits)				
Final	4.70*	2.27–9.75	4.90*	2.21–10.87
Global	6.89*	3.18–14.95	7.00*	3.09–15.88
Temporal	5.19*	2.49–10.81	5.43*	2.50–11.80
Temporal superior	4.18*	1.93–9.06	4.08*	1.77–9.41
Temporal inferior	5.00*	2.29–10.88	4.80*	2.06–11.19
Nasal	4.88*	2.31–10.31	5.03*	2.28–11.13
Nasal superior	5.04*	2.35–10.79	4.79*	2.09–10.98
Nasal inferior	4.91*	2.24–10.75	4.83*	2.03–11.52

Data are GPS parameters, categorized according to manufacturer-suggested cutoffs. Multivariate models adjust for age, baseline IOP, CCT, and PSD.

* Significant at $P < 0.05$.

limits results (using manufacturer suggested cutoffs) are presented in Table 3. An outside normal limits result on GPS final classification had an adjusted HR of 4.90 (95% CI: 2.21–10.87; $P < 0.001$). In multivariate models with adjustment for the same variables, adjusted HRs for subjective stereophotograph assessment were 2.34 (95% CI: 1.22–4.47; $P = 0.010$) for a grade indicating glaucoma and 1.34 (95% CI: 1.12–1.62; $P = 0.002$) for a 0.1 higher vertical cup–disc ratio.

The c-index was used to evaluate the predictive abilities of the multivariate models (Table 2). Models containing global or sectoral GPS results had very similar c-index values. The c-index for the multivariate model containing GPS global was 0.732. For the multivariate model containing stereophotograph grading (glaucoma versus normal), the c-index was 0.699, whereas it was 0.705 for the multivariate model containing vertical cup–disc ratio. There were no statistically significant differ-ences in the predictive abilities of multivariate models containing GPS global results, stereophotograph CDR, or stereophotograph grading ($P > 0.05$ for all comparisons). Similar results were found when analyses were repeated using treatment as a time-dependent covariate.

Spatial agreement between visual field loss and GPS abnormalities was evaluated for those who developed a repeatable abnormal visual field. Twenty-four subjects presented an inferior visual field defect at conversion. Of these, there were 6 with an outside normal limits classification for GPS global analysis and 6 for temporal, 5 for nasal, 10 for inferior (inferior temporal and inferior nasal), and 10 for superior (superior temporal and superior nasal) visual field defects. Thirty-two subjects presented a superior visual field defect at conversion (10 had loss in both hemispheres). Of these, there were 8 with an outside normal limits classification for GPS global analysis and 9 for temporal, 9 for nasal, 14 for inferior (inferior temporal and inferior nasal), and 12 for superior (superior temporal and superior nasal) visual field defects. In several cases, there was more than one sector outside normal limits, and there was also an overlap of subjects with superior and inferior visual field defects.

Table 4 shows the HRs for the association of each variable and progression as assessed by stereophotograph evaluation and by visual fields separately. When stereophotographs were used as the only endpoint, only CCT and stereophotograph assessment at baseline were significantly associated with the risk of progression. It is important to note, however, that the low number of patients who converted by stereophotograph during follow-up did not allow precise estimation of HRs in this analysis, and the CIs were considerably larger than in the complete analysis. When only consecutive abnormal visual fields were considered as the endpoint, results were similar to those of the complete analysis.

Figure 4 shows Venn diagrams illustrating the agreement between categorical GPS final classification and stereophotograph grading in converters and nonconverters.
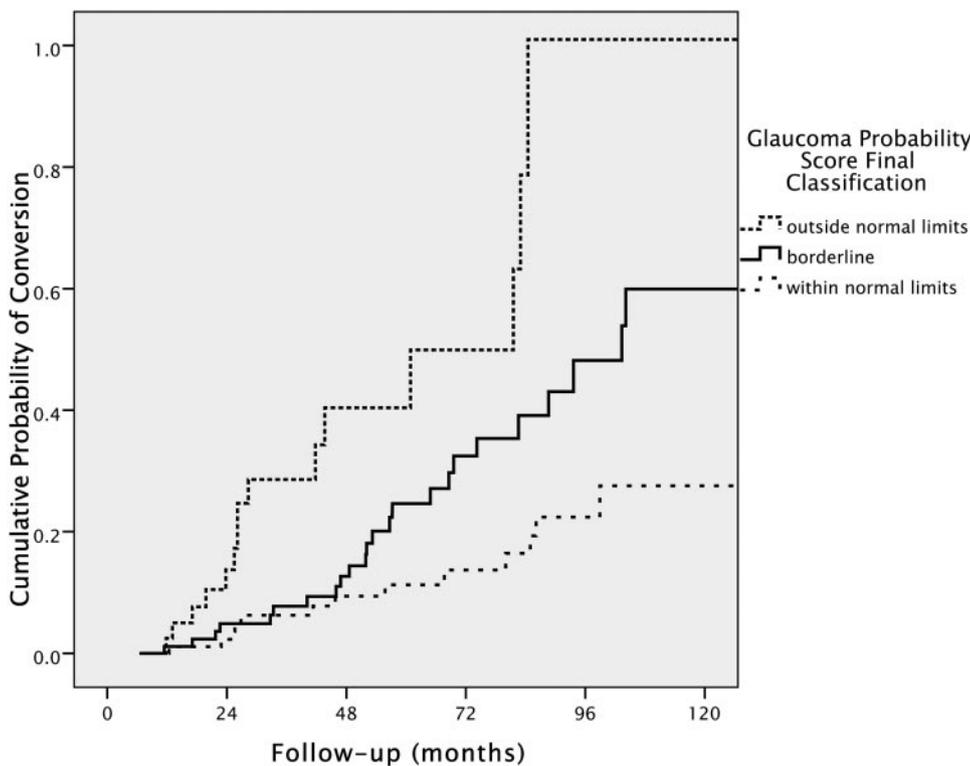
FIGURE 2. Cumulative probability (Kaplan-Meier survival curve) of development of visual field loss or optic disc deterioration in patients with suspected glaucoma according to the baseline final classification of the GPS.
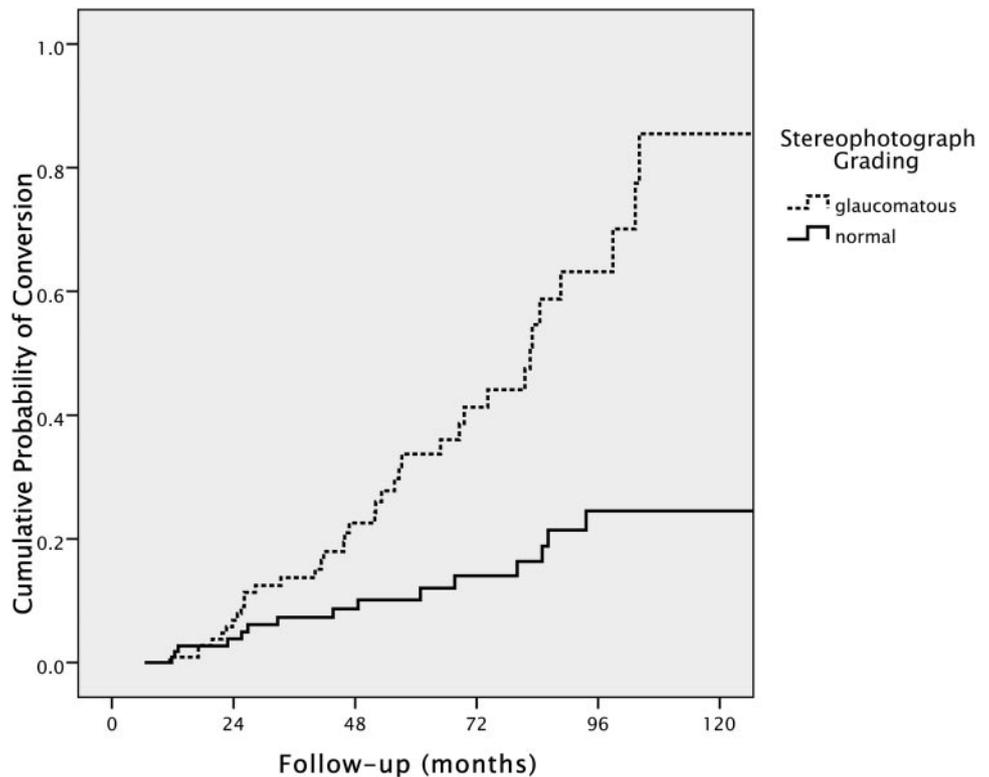
FIGURE 3. Cumulative probability (Kaplan-Meier survival curve) of development of visual field loss or optic disc deterioration in patients with suspected glaucoma according to the baseline grading of optic disc stereophotographs.

DISCUSSION

In this study, objective optic nerve head assessment with the HRT parameter GPS was predictive of future progression in patients with suspected glaucoma. Subjects with higher scores at baseline had a greater risk of disease progression during the follow-up period. In addition, GPS results performed similarly

TABLE 4. HRs for the Association of Each Explanatory Variable and Glaucoma Progression

	Stereophotographs		Visual Fields	
	HR	95% CI	HR	95% CI
Univariate				
Age (per decade)	0.95	0.68–1.34	1.69*	1.26–2.27
CCT (per 40 μm thinner)	1.95*	1.22–3.12	1.89*	1.37–2.62
PSD (per 0.2 larger)	1.02	0.92–1.12	1.11*	1.05–1.16
IOP (per 1 mm Hg higher)	1.00	0.93–1.09	0.97	0.92–1.02
CDR (per 0.1 higher)	1.70*	1.26–2.29	1.38*	1.15–1.65
Photograph grading (glaucomatous)	10.76*	2.51–46.09	2.93*	1.55–5.52
GPS global (per 0.1 higher)	1.12	0.93–1.35	1.34*	1.18–1.51
Multivariate				
CDR (per 0.1 higher)	1.76*	1.26–2.44	1.24*	1.02–1.52
Photograph grading (glaucomatous)	11.45*	2.56–51.16	1.88	0.93–3.79
GPS global (per 0.1 higher)	1.14	0.93–1.41	1.31*	1.13–1.51

Variables were assessed separately by changes on the stereophotographs and development of abnormal visual fields.

* Statistically significant ($P < 0.05$).

to expert stereophotograph evaluation in predicting which patients would have glaucoma develop during the follow-up. These findings suggest that objective optic disc analysis using the GPS could replace subjective stereophotograph evaluation in risk assessment of patients with suspected glaucoma.

After adjustment for other potential risk factors, each 0.1 increase in the GPS global score was associated with a 31% higher risk of reaching the study endpoint. No significant improvement in predictive ability was noted when GPS results in the six predefined sectors were considered. When manufacturer's suggested cutoffs were used, an outside normal limits result on the GPS final classification was associated with an approximately fivefold higher risk of glaucoma, compared to a within normal limits result. For example, at 5 years of follow-up, the probability of conversion was 11% for those with within normal limits baseline GPS final classification, compared with 25% for those with borderline GPS and 45% for those with results outside normal limits.

Several investigators have previously evaluated the ability of HRT parameters to predict development of visual field defects or optic disc deterioration in patients with suspected glaucoma. Bowd et al.[7] demonstrated that HRT II MRA results and machine learning classifiers developed using stereometric parameters were able to detect abnormalities in optic disc topography before the development of visual field loss. Zangwill et al.[6] showed that several baseline HRT topographic measurements, alone or in combination with baseline clinical and demographic factors, were significantly associated with the development of glaucoma among participants with ocular hypertension in the Ocular Hypertension Study (OHTS), suggesting that this instrument could be a useful predictive tool in this population. All HRT parameters included in these studies, however, required tracing of a contour line outlining the optic disc margin for their calculation. Although the HRT has been developed to provide objective measurements of optic disc topography, the contour line requirement still imposes limitations on the objectivity and accuracy of this test. In fact,
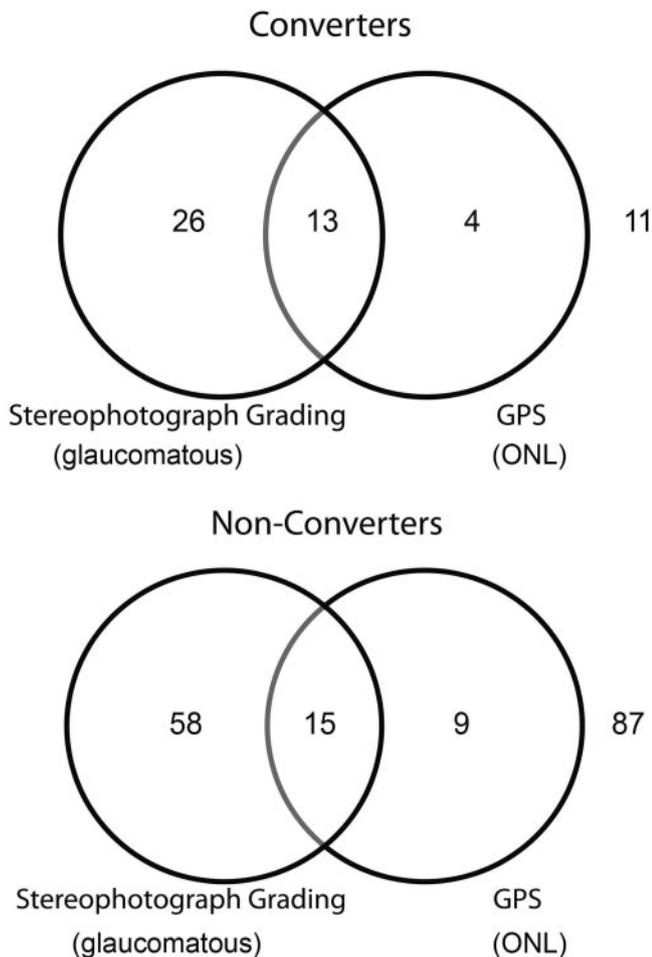
Converters

Non-Converters

FIGURE 4. Venn diagram showing agreement between the GPS final classification and the stereophotograph grading at baseline for converters and nonconverters.

improper contour line drawing is recognized as a frequent source of error and misinterpretation of HRT exams.[8] With the development of the GPS analysis, these limitations have been largely overcome. GPS calculations are independent of contour-line tracing and do not rely on a reference plane. Therefore, use of the GPS parameter removes the subjectivity that would be added by relying on contour line tracing performed by an examiner. To our knowledge, the present study is the first to demonstrate that the GPS parameters are also significantly predictive of conversion in patients with suspected glaucoma.

We also compared the performance of the GPS to that of subjective expert stereophotograph assessment for predicting which patients would have glaucoma develop during follow-up. Both methods had similar ability to discriminate converters from nonconverters, as assessed by the c-index. For the multivariate model containing GPS global, age, baseline IOP, CCT, and SAP PSD, the c-index was 0.732. A c-index of 0.73 indicates that, in approximately 73% of the cases, the model allocated a higher predicted probability for a subject who actually converted than for a subject who did not. For multivariate models containing stereophotograph grading and vertical cup–disc ratio, corresponding values were 0.699 and 0.705. It is interesting to note that the c-indexes for the multivariate models reported in our study are similar to those reported in other studies conducted to evaluate predictive models to estimate risk of glaucoma. Medeiros et al.[19] found c-indexes from 0.68 to

0.73 when predictive models derived from the OHTS were applied to estimate risk of conversion to glaucoma in a group of 126 ocular hypertensive patients followed for approximately 8 years. Similarly, a predictive model developed from combined results of the OHTS and the European Glaucoma Prevention Study (EGPS) had a c-index of 0.74 for discriminating ocular hypertensive subjects who converted to glaucoma from those who did not.[20]

Several cross-sectional studies have previously compared the performances of objective structural assessment by imaging instruments and subjective stereophotograph evaluation. Girkin et al.[21] and DeLeón-Ortega et al.[22] showed that subjective stereophotograph assessment outperformed HRT parameters in discriminating patients with glaucomatous visual field loss from normal subjects. Both studies, however, are limited by their cross-sectional design and lack of a reference standard that was completely independent of the tests being evaluated. Although visual fields were reported to be the main test used to classify patients in both studies, it is very likely that clinical examination of the optic disc was used at some point to classify participants, due to the cross-sectional design and glaucoma clinic-based samples in the studies. The appearance of the optic disc on clinical examination is more likely to be related to its appearance on stereophotographs than to results of imaging instruments, and this tendency could introduce a bias in favor of stereophotograph assessment. Also, although diagnostic accuracy measures obtained from the studies just mentioned are useful in providing an initial evaluation of the ability of these tests to detect glaucomatous damage, it is clear that in clinical practice a clinician does not need an imaging test just to help differentiate a patient with repeatable glaucomatous visual field loss from a healthy subject without suspect findings. In fact, clinicians are most interested in the ability of the test to diagnose or predict damage in patients with suspected disease who do not show any clear abnormality such as repeatable visual field loss.[23] The longitudinal design used in our study largely overcomes these limitations when imaging instruments and stereophotographs are compared.

Our study has limitations. Patients were not randomized for treatment or no treatment, and the decision as to whether to initiate treatment might have been based on disc assessment results and other risk factors. It is possible that patients with more risk factors or a more suspect appearance of the optic disc were the ones who received treatment during follow-up. It might be argued that these patients would be less likely to progress—a situation that would underestimate the true predictive value of baseline optic disc evaluation for conversion to glaucoma. However, this probably would affect the predictive ability of both methods of optic disc assessment evaluated in our study. Further, when treatment was included as a time-dependent covariate, no significant differences were seen on the results. It must be emphasized that stereophotographs in our study were obtained by specialized personnel and evaluated by highly trained individuals from a reading center, blinded to chronological order. Such expertise is unlikely to be available to most eye care providers, and thus the performance of subjective assessment of optic disc morphology was probably overstated.

It is known that, in a considerable number of patients, damage to the optic nerve head occurs despite any evidence of functional loss.[1,2] Therefore, structural and functional assessments were used to define conversion in our study. When we performed separate analyses using stereophotographs or visual fields as endpoints, significant differences were observed on the predictive abilities of the baseline variables. It is noteworthy that baseline stereophotograph assessment seemed to perform better than GPS when stereophotographs were used as the only endpoint. This, however, may reflect a bias of using

the same method as the predictive variable and for determination of the endpoint, which likely resulted in overestimation of the predictive ability of stereophotograph assessment. Also, the relatively small number of patients in whom conversion was identified by stereophotograph resulted in large CIs and less precise estimates of the predictive abilities of the baseline variables. Therefore, it is likely that a more fair comparison of the predictive abilities of GPS and stereophotographs was obtained in the analysis that used visual fields as the endpoint measure.

It also might be argued that some patients included in our study already had glaucomatous optic neuropathy (GON) at baseline and therefore would not truly have suspected disease. However, the primary purpose of the study was to assess the predictive performance of an objective method of structural assessment, the HRT GPS parameter, without relying on subjective assessment of optic disc morphology. To accomplish this, it was necessary to define the inclusion criteria based solely on visual fields. Although the inclusion of subjects with GON at baseline could overestimate the predictive ability of GPS, the design of our study replicates the clinical situation in which this technology is used to replace subjective assessment of the optic disc for evaluation of the risk of development of glaucoma. It is also important to emphasize that, although an outside normal limits result on the GPS was associated with increased likelihood of progression to visual field loss or optic disc damage, most of the subjects with GPS outside normal limits did not reach the endpoint during the study period. For example, for final GPS classification, 41 (18%) of all patients had an outside normal limits result at baseline. Of these, 17 (42%) had conversion to glaucoma and 24 (58%) did not. Further follow-up will be necessary to ascertain whether these patients will have visual field damage in the future or whether these were false-positive results.

Although stereophotograph assessment and GPS were found to have similar predictive abilities in our study, their agreement was only fair. As these two methods use different techniques to measure different aspects of the optic disc and RNFL, the low agreement is not surprising. Stereophotograph evaluation relies on subjective assessment by graders, which takes into account several features of the optic nerve that would not be evaluated by the GPS, such as presence of hemorrhages, parapapillary atrophy, and localized RNFL defects. On the other hand, GPS analysis is an objective method, and its classification relies on cutoffs to achieve predetermined specificity levels according to a comparison to a normative database. Further, the GPS was originally designed as a global parameter, and modifications to evaluate sectoral optic nerve damage included division of the optic nerve head by sectors. However, only two of the original five parameters are provided in sectors—rim steepness and cup size—whereas cup depth and vertical and horizontal nerve fiber layer curvatures are only given as global measurements. Therefore, the sectoral parameters do not incorporate information about localized damage to the RNFL or localized changes in cup depth. The inability to measure local damage may also explain, at least in part, the weak relationship between the abnormal GPS sector and the subsequent area of VF loss found in our study. In fact, the performance of sectoral parameters has not been demonstrated to be superior to that of GPS global for detection of glaucomatous damage.[24]

In conclusion, the results of our study demonstrate that the HRT parameter GPS is able to predict development of glaucomatous field loss or optic disc deterioration in patients with suspected glaucoma. Further, they demonstrate that GPS analysis has predictive ability similar to that of subjective stereophotograph assessment by glaucoma experts. These findings suggest that objective assessment of the optic disc topography using the HRT GPS technology has the potential to replace subjective optic disc evaluation as a predictive tool for estimating the likelihood of conversion in a patient with suspected glaucoma.

References

1. Tuulonen A, Airaksinen PJ. Initial glaucomatous optic disk and retinal nerve fiber layer abnormalities and their progression. *Am J Ophthalmol.* 1991;111:485–490.

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. Hertzog LH, Albrecht KG, LaBree L, Lee PP. Glaucoma care and conformance with preferred practice patterns: examination of the private, community-based ophthalmologist. *Ophthalmology.* 1996;103:1009–1013.

4. Fremont AM, Lee PP, Mangione CM, et al. Patterns of care for open-angle glaucoma in managed care. *Arch Ophthalmol.* 2003; 121:777–783.

5. Shingleton BJ, Crandall A, Johnstone M, Robin A, Brown R. Medical treatment patterns of ASCRS members for primary open-angle glaucoma: 1998 survey. *J Cataract Refract Surg.* 1999;25:118–127.

6. Zangwill LM, Weinreb RN, Beiser JA, et al. Baseline topographic optic disc measurements are associated with the development of primary open-angle glaucoma: the Confocal Scanning Laser Ophthalmoscopy Ancillary Study to the Ocular Hypertension Treatment Study. *Arch Ophthalmol.* 2005;123:1188–1197.

7. Bowd C, Zangwill LM, Medeiros FA, et al. Confocal scanning laser ophthalmoscopy classifiers and stereophotograph evaluation for prediction of visual field abnormalities in glaucoma-suspect eyes. *Invest Ophthalmol Vis Sci.* 2004;45:2255–2262.

8. Iester M, Mikelberg FS, Courtright P, et al. Interobserver variability of optic disk variables measured by confocal scanning laser tomography. *Am J Ophthalmol.* 2001;132:57–62.

9. Bursansky-Eliash Z, Wollstein G, Bilonick RA, Ishikawa H, Kagemann L, Schuman JS. Glaucoma detection with the Heidelberg retina tomograph 3. *Ophthalmology.* 2007;114:466–471.

10. Coops A, Henson DB, Kwartz AJ, Artes PH. Automated analysis of heidelberg retina tomograph optic disc images by glaucoma probability score. *Invest Ophthalmol Vis Sci.* 2006;47:5348–5355.

11. Cioffi GA, Robin AL, Eastman RD, Perell HF, Sarfarazi FA, Kelman SE. Confocal laser scanning ophthalmoscope: reproducibility of optic nerve head topographic measurements with the confocal laser scanning ophthalmoscope. *Ophthalmology.* 1993;100:57–62.

12. Weinreb RN. Laser scanning tomography to diagnose and monitor glaucoma. *Curr Opin Ophthalmol.* 1993;4:3–6.

13. Wollstein G, Garway-Heath DF, Hitchings RA. Identification of early glaucoma cases with the scanning laser ophthalmoscope. *Ophthalmology.* 1998;105:1557–1563.

14. Medeiros FA, Sample PA, Zangwill LM, Bowd C, Aihara M, Weinreb RN. Corneal thickness as a risk factor for visual field loss in patients with preperimetric glaucomatous optic neuropathy. *Am J Ophthalmol.* 2003;136:805–813.

15. Gordon MO, Beiser JA, Brandt JD, et al. The Ocular Hypertension Treatment Study: baseline factors that predict the onset of primary open-angle glaucoma. *Arch Ophthalmol.* 2002;120:714–720, discussion 829–30.

16. Miglior S, Pfeiffer N, Torri V, Zeyen T, Cunha-Vaz J, Adamsons I. Predictive factors for open-angle glaucoma among patients with ocular hypertension in the European Glaucoma Prevention Study. *Ophthalmology.* 2007;114:3–9.

17. Harrell FE Jr. *Regression Modelling Strategies with Applications to Linear Models, Logistic Regression and Survival Analysis.* New York: Springer, 2001.

18. D'Agostino RB, Nam BH. Evaluation of the performance of survival analysis models: discrimination and calibration measures. In: Balakrishnan N, Rao CR, eds. *Handbook of Statistics. Vol.* 23. Amsterdan: Elsevier, 2004.

19. Medeiros FA, Weinreb RN, Sample PA, et al. Validation of a predictive model to estimate the risk of conversion from ocular hypertension to glaucoma. *Arch Ophthalmol.* 2005;123:1351–1360.

20. Gordon MO, Torri V, Miglior S, et al. Validated prediction model for the development of primary open-angle glaucoma in individuals with ocular hypertension. *Ophthalmology.* 2007;114:10–19.

21. Girkin CA, DeLeón-Ortega JE, Xie A, McGwin G, Arthur SN, Monheit BE. Comparison of the Moorfields classification using confocal scanning laser ophthalmoscopy and subjective optic disc classification in detecting glaucoma in blacks and whites. *Ophthalmology.* 2006;113:2144–2149.

22. DeLeón-Ortega JE, Arthur SN, McGwin G Jr, Xie A, Monheit BE, Girkin CA. Discrimination between glaucomatous and nonglaucomatous eyes using quantitative imaging devices and subjective optic nerve head assessment. *Invest Ophthalmol Vis Sci.* 2006;47: 3374–3380.

23. Medeiros FA, Ng D, Zangwill LM, Sample PA, Bowd C, Weinreb RN. The effects of study design and spectrum bias on the evaluation of diagnostic accuracy of confocal scanning laser ophthalmoscopy in glaucoma. *Invest Ophthalmol Vis Sci.* 2007;48:214–222.

24. Swindale NV, Stjepanovic G, Chin A, Mikelberg FS. Automated analysis of normal and glaucomatous optic nerve head topography images. *Invest Ophthalmol Vis Sci.* 2000;41:1730–1742.