Combining Spectral Domain Optical Coherence Tomography Structural Parameters for the Diagnosis of Glaucoma With Early Visual Field Loss

Jean-Claude Mwanza,[1] Joshua L. Warren,[2] and Donald L. Budenz[1] for the Ganglion Cell Analysis Study Group

[1]Department of Ophthalmology, School of Medicine, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina
[2]Department of Biostatistics, Gillings School of Global Public Health, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina

Correspondence: Donald L. Budenz, Department of Ophthalmology, University of North Carolina at Chapel Hill, 5151 Bioinformatics Building, CB #7040, 130 Mason Farm Road, Chapel Hill, NC 27599; donald_budenz@med.unc.edu.

See the appendix for the members of the Ganglion Cell Analysis Study Group.

Submitted: July 5, 2013
Accepted: November 16, 2013

Citation: Mwanza J-C, Warren JL, Budenz DL for the Ganglion Cell Analysis Study Group. Combining spectral domain optical coherence tomography structural parameters for the diagnosis of glaucoma with early visual field loss. *Invest Ophthalmol Vis Sci*. 2013;54:8393–8400. DOI: 10.1167/iovs.13-12749

Purpose. To create a multivariable predictive model for glaucoma with early visual field loss using a combination of spectral-domain optical coherence tomography (SD-OCT) parameters, and to compare the results with single variable models.

Methods. Two hundred fifty-three subjects (149 healthy controls and 104 with early glaucoma) underwent optic disc and macular scanning using SD-OCT in one randomly selected eye per subject. Sixteen parameters (rim area, cup-to-disc area ratio, vertical cup-to-disc diameter ratio, average and quadrant RNFL thicknesses, average, minimum, and sectoral ganglion cell inner-plexiform layer [GCIPL] thicknesses) were collected and submitted to an exploratory factor analysis (EFA) followed by logistic regression with the backward elimination variable selection technique. Area under the curve (AUC) of the receiver operating characteristic (ROC), sensitivity, specificity, Akaike's information criterion (AIC), predicted probability, prediction interval length (PIL), and classification rates were used to determine the performances of the univariable and multivariable models.

Results. The multivariable model had an AUC of 0.995 with 98.6% sensitivity, 96.0% specificity, and an AIC value of 43.29. Single variable models yielded AUCs of 0.943 to 0.987, sensitivities of 82.6% to 95.7%, specificities of 88.0% to 94.0%, and AICs of 113.16 to 59.64 (smaller is preferred). The EFA logistic regression model correctly classified 91.67% of cases with a median PIL of 0.050 in the validation set. Univariable models correctly classified 80.62% to 90.48% of cases with median PILs 1.9 to 3.0 times larger.

Conclusions. The multivariable model was successful in predicting glaucoma with early visual field loss and outperformed univariable models in terms of AUC, AIC, PILs, and classification rates.

Keywords: glaucoma, diagnosis, optical coherence tomography

Glaucoma is the second leading cause of blindness in the world.[1] Recent estimates based on meta-analysis of population-based studies indicate that in 2010 there were 44.7 million people with open-angle glaucoma (OAG) worldwide and 2.8 million people in the United States.[2] These numbers are projected to reach 58.6 million worldwide[2] and 3.4 million in the United States in 2020.[3] The current standard for the diagnosis of glaucoma is based on the presence of typical structural changes with corresponding functional deficits. However, studies on early changes in glaucoma have shown that structural changes indicated by significant loss of retinal ganglion cells and their axons may precede detection of functional deficit by currently available standard automated perimetry (SAP) methods.[4,5] In addition, SAP is subject to fluctuation even in clinically stable glaucoma. Ophthalmoscopy and optic disc photography traditionally have been used as primary methods for structural assessment in glaucoma. However, substantial interobserver variability and low to medium interobserver agreement in detecting subtle changes makes ophthalmoscopy or fundus photography alone poor methods for detecting changes indicative of early glaucoma or glaucomatous progression.[6,7]

With the advent and continuous improvement of imaging devices such as scanning laser ophthalmoscopy (SLO), scanning laser polarimetry (SLP), and time and spectral-domain optical coherence tomography (SD-OCT) in recent years, a great deal of effort has been invested in identifying quantifiable parameters for objective assessment of structural glaucomatous damage. Of these devices, SD-OCT has rapidly become one of the most widely used technologies in daily clinic due to its high image resolution and measurement precision. The caveat, however, is that the clinician is presented with an array of quantitative information to mentally process as part of the diagnosis process. The multitude of parameters—most of which are highly correlated and are redundant to some extent—oftentimes renders the interpretation process difficult, particularly when structural changes conflict in their results. For instance, optic nerve head (ONH) parameters may appear normal but RNFL measurements may appear abnormal. One way of circumventing this issue would be to use a multivariable model that reduces the

number of structural parameters provided by the OCT output into a set of fewer parameters containing the most useful and relevant information from the original set while also explaining a majority of variability in the original dataset. Earlier studies have combined structural parameters measured by one or several of the devices listed above to assess the glaucoma diagnostic ability using various methods such as machine learning classifiers, linear discriminant functions (LDF), and principal component analysis (PCA).[8–12] In addition, three recent studies used linear discriminant analysis to assess diagnostic ability of combined structural parameters measured SD-OCT. One of these studies combined ONH and peripapillary RNFL parameters,[13] whereas the two others used a combination of ONH, peripapillary RNFL, and ganglion cell of complex (GCC) parameters,[14,15] which is anatomically different from ganglion cell inner-plexiform layer (GCIPL).[16] The purpose of this study was to create a multivariable predictive model for early glaucoma using a combination of ONH, peripapillary RNFL, and macular GCIPL parameters measured with SD-OCT, to determine its sensitivity, and to compare the results with single variable models.

Methods

Study Subjects and Database

The data used in this study were those of subjects who participated in two earlier glaucoma SD-OCT imaging studies[16,17] and one ongoing study, which obtained institutional review board approval and whose procedures adhered to the principles of the Declaration of Helsinki for research involving human subjects. Written informed consent was obtained from all participants. Subjects were 104 patients with definitive early glaucoma and 149 normals. This dataset was randomly divided into a modeling set (2/3 of the original sample) and a validation set (1/3 of the original sample). It was important to have a larger sample size for the modeling set since the introduced method consists of two stages (factor analysis and multiple logistic regression). The modeling set included 69 early-glaucoma patients and 100 normals, while the validation set included 35 early-glaucoma patients and 49 normals. Both healthy subjects and patients with glaucoma underwent an eligibility complete ophthalmologic examination that included measurement of visual acuity, IOP, and refraction, and biomicroscopy of the anterior segment and dilated fundus examination. Glaucomatous patients also underwent visual field testing with a Humphrey Visual Field Analyzer (Carl Zeiss Meditec, Inc., Dublin, CA) using the Swedish Interactive Thresholding Algorithm standard program. Criteria for the diagnosis of glaucoma were typical glaucomatous optic disc changes with corresponding visual field defects. Optic disc change was indicated by cup-to-disc ratio (CDR) greater than 0.5 in either eye, CDR asymmetry greater than or equal to 0.2, or focal thinning of the rim in either eye. A glaucomatous visual field defect was defined as glaucoma hemifield test outside normal limits, pattern SD with a P value less than 5%, or a cluster three points or more in the pattern deviation plot in a single hemifield (superior or inferior) with a P value of less than 5%, one of which having a P value of less than 1%. All participants were excluded based on best-corrected visual acuity worse than 20/40 in either eye, refraction error outside the interval −12 to +8 spherical diopters (D) or worse than 3 cylindrical D, active infection of the anterior or posterior segment of either eye, previous or current vitreoretinal diseases or surgery in the study eye, or evidence of diabetic retinopathy or macular edema on dilated ophthalmoscopic examination or retinal photograph evaluation. In addition, glaucoma patients were excluded if the visual field MD was worse than −6 dB. Only one randomly selected eye was tested for the study from each participant.

OCT Imaging

All subjects underwent Cirrus HD-OCT (Carl Zeiss Meditec, Inc.) macular (Macular Cube 200×200 protocol) and optic disc (Optic Disc Cube 200×200 protocol) scans. All scans were visually reviewed for quality control, and only scans with signal strength greater than or equal to 6, without RNFL misalignment or discontinuity, blinking or involuntary saccade artifacts, and an absence of algorithm segmentation failure on careful visual inspection were retained for analysis. The macular scan was used to measure the thickness of the GCIPL, whereas the optic disc scan served for measuring peripapillary RNFL and optic disc topography.[16,17] Peripapillary RNFL variables considered for analysis were average and quadrant (superior, nasal, inferior, and temporal) thicknesses. Ganglion cell inner-plexiform layer thickness variables included average, minimum (minimum of the average GCIPL thickness along a given radial spoke in the elliptical annulus),[18] and sectoral (superior, superonasal, inferonasal, inferior, inferotemporal, and superotemporal) thicknesses. Parameters from the ONH analysis included rim area, CDR, and vertical cup-to-disc diameter ratio (VCDR).

Data Management and Statistical Analysis

Pearson's correlation coefficient was calculated to explore pairwise linear relationships between the 16 (5 RNFL, 3 ONH, and 8 GCIPL) continuous variables used for analysis in this study, as a preparatory step for factor analysis. Factor analysis was subsequently performed on the data set of 16 variables using the method of exploratory factor analysis (EFA) with a promax rotation to identify latent factors accounting for a large proportion of the variability seen in the original set of variables. Use of the oblique promax rotation resulted in an improved interpretation of latent factors, which are not uniquely identified. It was preferred over other orthogonal rotation methods (varimax, equamax, orthomax, quartimax, and parsimax) in this setting, due to its ability to reduce cross-loadings, which lead to improved factor interpretations and its similar performance in the multiple logistic regression models using a varying number of retained factors. During the EFA, standardized scoring coefficients are estimated for each factor and variable combination. A set of coefficients for a chosen factor serves as a weight for each variable and is multiplied by each person's standardized variable response. These weighted values are then summed to create the estimated factor score for a person. This process is repeated for each factor and each person in order to obtain the complete set of estimated factor scores for each individual.

A logistic regression model with the backward elimination variable selection technique was then fitted with early glaucoma as the outcome variable and the estimated latent factor scores as candidate explanatory variables such that the probability of glaucoma for an individual is modeled as

$$\text{logit}[\text{Prob}(\text{Glaucoma})] = \beta_0 + \beta_1{*}\text{F1} + \ldots + \beta_5{*}\text{F5} + \beta_6{*}\text{F1}{*}\text{F2} + \cdots + \beta_{15}{*}\text{F4}{*}\text{F5}, \quad (1)$$

where F1, F2, F3, F4, and F5 are the five estimated factor scores for an individual in the study (see Results section for more information on choice of five factors). From the estimated logistic regression coefficients, predicted probabilities for early glaucoma status along with 95% prediction intervals were calculated and submitted to a receiver operating characteristic (ROC) curve analysis. Cutoff points from the ROC plot were

Investigative Ophthalmology & Visual Science

Investigative Ophthalmology & Visual Science

Table 1. Sample Correlation Matrix for All 16 Variables Included in the Analysis (Modeling Set)

	Average GCIPL	Minimum GCIPL	Superotemporal GCIPL	Superior GCIPL	Superonasal GCIPL	Inferonasal GCIPL	Inferior GCIPL	Inferotemporal GCIPL	Average RNFL	Temporal RNFL	Superior RNFL	Nasal RNFL	Inferonasal RNFL	Rim Area	CDR	VCDR
Average GCIPL	1.000	0.936	0.918	0.913	0.913	0.931	0.923	0.886	0.781	0.604	0.710	0.361	0.696	0.660	−0.496	−0.522
Minimum GCIPL		1.000	0.861	0.827	0.805	0.843	0.903	0.900	0.705	0.533	0.623	0.289	0.668	0.640	−0.531	−0.575
Superotemporal GCIPL			1.000	0.913	0.853	0.758	0.749	0.772	0.727	0.595	0.700	0.338	0.593	0.607	−0.464	−0.491
Superior GCIPL				1.000	0.894	0.779	0.739	0.679	0.710	0.565	0.678	0.342	0.589	0.582	−0.409	−0.426
Superonasal GCIPL					1.000	0.865	0.740	0.669	0.671	0.571	0.614	0.368	0.539	0.502	−0.355	−0.365
Inferonasal GCIPL						1.000	0.906	0.804	0.698	0.560	0.610	0.321	0.634	0.589	−0.440	−0.453
Inferior GCIPL							1.000	0.931	0.723	0.526	0.629	0.294	0.708	0.646	−0.507	−0.540
Inferotemporal GCIPL								1.000	0.748	0.510	0.654	0.321	0.741	0.685	−0.546	−0.588
Average RNFL									1.000	0.654	0.907	0.561	0.909	0.770	−0.583	−0.624
Temporal RNFL										1.000	0.517	0.059	0.493	0.483	−0.474	−0.486
Superior RNFL											1.000	0.449	0.739	0.702	−0.482	−0.524
Nasal RNFL												1.000	0.420	0.369	−0.232	−0.228
Inferior RNFL													1.000	0.738	−0.562	−0.615
Rim area														1.000	−0.741	−0.731
CDR															1.000	0.978
VCDR																1.000

used to classify subjects based on predicted probabilities. Similar logistic regression and ROC curve analyses were carried out for single variable models and compared to the multivariable analysis. The robustness of the models was validated in a separate set including 49 healthy subjects and 35 subjects with early glaucoma. Area under the curve (AUC) of the ROC, sensitivity, specificity, median 95% prediction interval length (PIL) of all prediction intervals, Akaike's information criterion (AIC), and classification rates were used to determine the performances of the models. Statistical analysis was performed with SAS version 9.2 (SAS, Cary, NC), with statistical significance level set at P less than 0.05.

Results

The mean age in the modeling set was 66.0 ± 11.85 years for patients with glaucoma and 62.8 ± 9.47 years for healthy subjects ($P = 0.06$). The mean age in the validation set was 67.9 ± 12.56 years for patients with glaucoma and 61.7 ± 9.56 years for healthy subjects ($P = 0.01$). Since the groups differed statistically in terms of age in the validation set, we controlled for the age of the subjects through use of a categorical age variable in the logistic regression models. The categories included, less than or equal to 57, 58 to 64, 65 to 72, and greater than 72, where approximately 25% of the sample was contained in each category. Controlling for age had little effect on the modeling results. The mean visual field MD of patients with glaucoma was −3.19 ± 1.69 dB. All 16 structural parameters significantly differed between subjects with glaucoma and healthy subjects in both the modeling (all $P \leq 0.001$) and the validation set (all $P < 0.02$).

EFA and Logistic Regression

Table 1 displays the sample correlation matrix produced by the EFA method, which includes all 16 continuous variables in the analysis. The estimated scoring coefficients from the EFA model needed to construct the latent factors are given in Table

Table 2. Scoring Coefficients for the Variables in the Modeling Set

Variables	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5
Average GCIPL	3.31	0.23	−0.32	0.62	0.71
Minimum GCIPL	0.07	−0.02	0.01	−0.04	−0.04
Superotemporal GCIPL	−0.37	−0.05	−0.08	0.06	−0.01
Superior GCIPL	−0.48	−0.07	−0.10	0.15	0.09
Superonasal GCIPL	−0.36	0.04	−0.16	0.09	0.04
Inferonasal GCIPL	−0.42	−0.03	0.11	−0.22	−0.21
Inferior GCIPL	−0.36	−0.02	0.24	−0.42	−0.35
Inferotemporal GCIPL	−0.45	−0.13	0.33	−0.49	−0.38
Average RNFL	0.46	0.06	1.29	3.20	3.81
Temporal RNFL	−0.15	−0.02	−0.27	0.00	−0.95
Superior RNFL	−0.27	0.02	−0.11	−0.85	−1.18
Nasal RNFL	−0.10	−0.01	−0.23	−0.61	0.06
Inferior RNFL	−0.27	−0.04	0.05	−1.41	−1.66
Rim Area	0.00	−0.03	0.00	−0.01	0.00
CDR	0.05	0.61	−0.13	−0.08	−0.12
VCDR	−0.04	0.35	0.19	0.10	0.12

To calculate *factor #j* for an individual, we multiply the individual's original variables (standardized) by the scoring coefficients for *factor #j* and sum each of the components. For example, *factor #1* = average GCIPL*(3.31) + minimum GCIPL*(0.07) + ... + CDR*(0.05) + VCDR*(–0.04). The estimated factor scores are then entered into the multiple logistic regression model and results are presented in Table 3. Note that only the first three factors are retained after implementing the backward elimination variable selection technique.

Table 3. Factor Loadings of the Variables in the Modeling Set

Variables	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5
Average GCIPL	0.892*	−0.006	0.104	0.070	0.021
Minimum GCIPL	0.867*	−0.127	0.086	−0.052	−0.052
Superotemporal GCIPL	0.786*	−0.063	−0.062	0.232	0.097
Superior GCIPL	0.817*	0.007	−0.105	0.256	0.140
Superonasal GCIPL	0.906*	0.077	−0.152	0.212	0.144
Inferonasal GCIPL	0.877*	0.044	0.138	−0.002	−0.039
Inferior GCIPL	0.808*	−0.021	0.328	−0.140	−0.128
Inferotemporal GCIPL	0.682*	−0.085	0.413	−0.160	−0.102
Average RNFL	0.079	−0.034	0.695*	0.294	0.211
Temporal RNFL	0.132	−0.101	0.216	0.824*	−0.352
Superior RNFL	0.062	0.047	0.647*	0.328	0.208
Nasal RNFL	0.044	−0.051	0.154	−0.297	0.963*
Inferior RNFL	0.045	−0.038	0.856*	−0.074	0.055
Rim area	0.122	−0.439*	0.368	0.066	0.103
CDR	0.000	1.004*	0.053	−0.033	−0.031
VCDR	−0.003	0.951*	−0.035	−0.037	−0.002

* Denote loaded factors (>0.43 or <−0.43).

2. Exploratory factor analysis with promax rotation identified five latent factors, which explained a large proportion of the variability seen in the original set of 16 variables (Table 3). We note that using the oblique promax rotation results in factor loadings that are no longer correlation coefficients, but instead represent regression coefficients that are not confined between −1 and 1.[19] The choice to retain five factors was made based on the scree plot of eigenvalues and analysis of accounted for variance. The scree plot is shown in Figures 1 and 2, the proportion of accounted for variance for each factor is displayed along with the cumulative proportions. The factor loadings, which reflect on the relative weights of the variable in the component, indicated that factor #1 was dominated by all GCIPL variables, factor #2 by all three ONH parameters, factor #3 by average, superior, and inferior quadrant RNFL, factor #4 by temporal RNFL, and factor #5 by nasal RNFL. Factors #1 and #3 through #5 each weighted positively on their respective variables while factor #2 had positive loadings for CDR and VCDR but a negative loading on rim area. This indicates that factor #2 represents the contrast between CDR/VCDR and rim area. The cutoff point of 0.43 used in Table 3 is based on simulation studies carried out by Hair et al.,[20] which

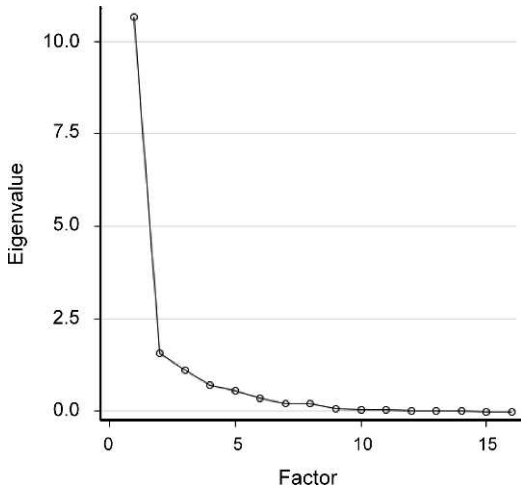

Figure 1. Scree plot of eigenvalues from the exploratory factor analysis.

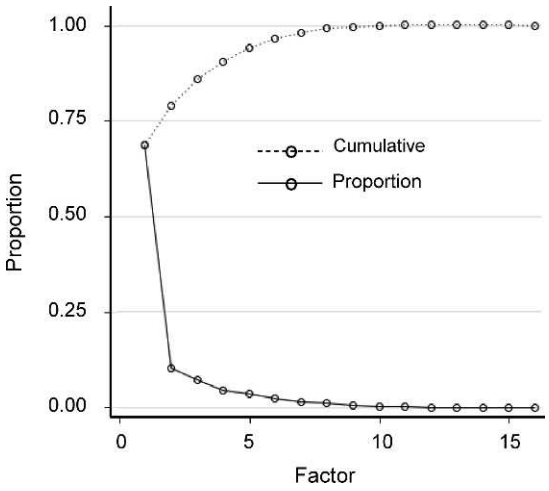

Figure 2. Proportion of variance explained by each identified factor with cumulative proportions.

considered sample size, a power level of 0.80, and level of significance of 0.05. These five factors combined accounted for 94.18% of the variability seen in the original set of 16 variables, greatly reducing the number of variables in the multiple logistic regression analysis while maintaining a majority of the available information. The results of logistic regression with backward elimination in the modeling set that included early glaucoma status as the dependent variable and these five estimated factor scores as predictors (along with their first order interactions), calculated for each subject, identified factor #1, factor #2, and factor #3 (Table 4) as statistically significant predictors of glaucoma, with an AIC of 43.29 (Table 5). The results in Table 3 indicate that an increase in an individual's score for factor #2 leads to an increase in the probability of glaucoma, implying that as the difference between of CDR/VCDR and rim area measurements increases, the probability of glaucoma increases. Also, the results confirm that as the GCIPL and RNFL factor scores increase, the probability of glaucoma decreases.

Table 5 shows that the AUC and sensitivity at fixed levels of specificity of the multivariable model was slightly higher than corresponding values of the best three single variables, while the AIC significantly improved from 113.16 to 59.64 to 43.29 in the univariable and multivariable models, respectively. Differences of four or more in AIC indicate that the lower valued model is preferred.[21] The AUC value for the multivariable model is also shown to be significantly larger than each of the single variable models at the 0.10 level of significance. The results are based on a nonparametric statistical test, which compares the areas under correlated ROC curves and is available in SAS version 9.2.[22] Figure 3 displays the ROC curves for all of the considered models in the analysis of the modeling set.

Table 4. Variable Selection by Logistic Regression With Backward Elimination for the Modeling Set*

Variables	Estimate	*P*
Intercept	−2.410	0.0075
Factor 1: all GCIPL	−2.685	0.0116
Factor 2: all ONH	7.307	0.0036
Factor 3: RNFL subset	−3.344	0.0044

* Performed after controlling for age of the subjects.

Table 5. Discriminating Ability of the Multivariable and Single Variable Analyses in Modeling Set and Proportion of Subjects Correctly Classified (%CC) and Median Length of Prediction Intervals (MPLI) in the Validation Set (Controlling for Age of the Subjects)

Variables	AUC (95% CI)	Sensitivity (90% Spec)	Sensitivity (95% Spec)	AIC	%CC	MLPI
Modeling set						
EFA GLM	0.995 (0.989, 1.000)	100.0	98.6 (87.8, 99.8)	43.29	-	-
VCDR	0.987 (0.976, 0.998)*	97.1 (87.0, 99.4)	89.9 (74.2, 96.5)	59.64	-	-
Inferior RNFL	0.943 (0.908, 0.979)†	82.6 (66.7, 91.8)	79.7 (60.8, 90.9)	108.37	-	-
Minimum GCIPL	0.958 (0.930, 0.987)†	85.5 (70.4, 93.6)	78.3 (59.0, 90.0)	113.16	-	-
Validation set						
EFA GLM	0.974 (0.948, 1.000)	91.4 (70.0, 98.0)	88.6 (62.5, 97.3)	43.43	91.67	0.050
VCDR	0.969 (0.938, 1.000)	91.4 (70.0, 98.0)	88.6 (62.5, 97.3)	41.90	90.48	0.095
Inferior RNFL	0.929 (0.877, 0.981)*	74.3 (48.8, 89.7)	65.7 (36.8, 86.3)	60.32	85.71	0.130
Minimum GCIPL	0.862 (0.780, 0.945)†	60.0 (35.0, 80.7)	51.4 (25.0, 77.0)	86.02	80.95	0.152

* Indicates that the AUC is significantly different from the EFA GLM AUC at the 0.1 level of significance.
† Indicates that the AUC is significantly different from the EFA GLM AUC at the 0.05 level of significance.

The multivariable model applied in a blinded manner to a separate set of healthy subjects and subjects with early glaucoma (validation set) produced another significant discriminant function that correctly classified 91.67%, compared with 80.95% to 90.48% for single variables (Table 5). The multivariable PILs were 1.9 to 3.0 times narrower on average than those of single variable models, indicating that it would be easier to predict the status of a subject using the multivariable rather than the univariable model by correctly characterizing the uncertainty associated with the predictions.

Discussion

In this paper, a classification method for early glaucoma was developed based on EFA followed by logistic regression of constructed glaucomatous structural parameters measured with SD-OCT. This study was designed on the assumption that a method combining structural parameters from ONH, RNFL, and GCIPL may improve the discrimination between eyes with early glaucoma and healthy eyes. With the availability of various imaging devices that provide quantitative evaluation of structural parameters, the number of parameters has increased significantly, making the interpretation more difficult. The multivariate analysis described herein produced a better predictive model than the univariable analysis. The multivariate analysis was achieved by combining OCT structural parameters in a weighted manner that was dictated by the data rather than an arbitrary combination. This weighting was done through two sequential methods: (1) EFA, which combined the OCT ONH, RNFL, and GCIPL measures into factor scores that represented a variable's performance more parsimoniously and with greater interpretability, and (2) multiple logistic regression modeling, which selected and weighted the factor scores with the greatest power to differentiate patients with early glaucoma from healthy subjects. Our results confirmed that the OCT measures are sensitive to group differences between early-stage glaucoma and normal state.

Significant effort has been invested in reducing the number of parameters through combinations that will increase glaucoma diagnostic accuracy. In this perspective, studies have used combination of parameters obtained with different devices (i.e., SLO and SLP or SLO and OCT)[8,11,12] or parameters derived from one device (i.e., SLO or OCT).[10,14,23–26] Badala et al.[8] used a classification and regression tree analysis (CART) to compare early glaucoma diagnostic performance of individual parameters with time domain OCT, SLP, SLO, and qualitative assessment of optic disc photographs and their combinations. No significant difference was found between individual best parameters from the four methods, but the combination of time domain OCT average RNFL thickness and SLO-based CDR resulted in a good diagnostic precision (93%), sensitivity (91%), and specificity (96%). The ability of SLP and SLO parameters to discriminate between healthy and glaucomatous eyes was also investigated by Magacho et al.[11] using multivariate discriminant formulas for each device individually and for both devices combined. The combination of parameters resulted in higher sensitivity, specificity, and accuracy compared with individual parameters on each device. Similarly, combining parameters of both devices achieved higher performances compared with combination of parameters of either device.

Using the or-Logic combination of criteria from SLO and SLP, Toth et al.[12] reported that combining various SLP criteria improved the accuracy and positive likelihood ratio (PLR) for early glaucoma. No improvement was observed following combination of SLO-based criteria alone; however, combining the best SLP and SLO criteria increased the PLR compared with

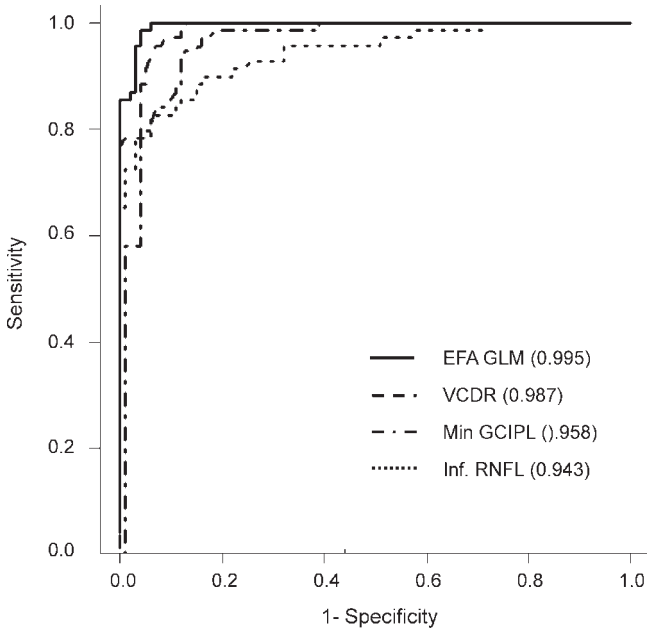

Figure 3. Receiver operating characteristic curves for all considered univariable and multivariable logistic regression models using the modeling set (controlling for age of the subjects).

combinations of SLO criteria alone, but decreased the PLR of SLP combination criteria. The or-Logic combination approach was also investigated by Budenz et al.[23] and in the Advanced Imaging Glaucoma Study (AIGS)[25] using time domain OCT-derived RNFL criteria. In the former study, the combination of one or more abnormal quadrants or clock hours and abnormal average RNFL at less than 5% level provided the best sensitivity for the diagnosis of glaucoma. However, AUCs for those combinations were not provided so that it was unclear what the best combinations were. In the latter study, the combination of average, inferior, and superior quadrant RNFL yielded an AUC of 0.92 compared with best single parameter (0.89). In a subsequent investigation, the AIGS[26] evaluated the performance of the or-Logic combination, support vector machine, relevance vector machine, and LDF methods using time domain OCT ONH, RNFL, and macular variables. The AUCs of the combination of the best three RNFL variables (inferior quadrant, overall average, and superior quadrant RNFL) with the best three ONH variables (horizontal integrated rim width, vertical integrated rim area, and VCDR) were similar (0.943–0.954), but greater than those of the overall best single parameter. However, in both AGISs, the study populations included all severity stages of glaucoma and no separate analysis was performed or provided for early glaucoma. Medeiros et al.[9] also showed that the PCA followed by LDF combining time domain OCT average, RNFL at 7 o'clock, and RNFL at 11 o'clock with VCDR achieved a significantly larger AUC of 0.97 than that 0.91 for best single parameter.

Other studies have assessed the diagnostic performance SLO-based optic disc parameters individually and in combination oftentimes using pre-established LDF or machine learning classifiers. One of these studies compared the diagnostic performances of three LDFs (Bathija's,[27] Mikelberg's,[28] and Burk's[29]) and the Moorfields Regression Analysis. At fixed specificity of 95%, these four methods had comparable low sensitivities that did not exceed 60%.[10] In another study the LDF from stepwise linear regression of SLO optic disc parameters had a sensitivity of 74.2% compared with SLO built-in LDFs (70.4% for Mikelberg's and 67.6% for Burk's) at fixed specificity of 85%.[30] The study by Iester et al.[31] compared the ability for early detection of glaucomatous optic disc by five formulas: Bathija's,[27] Mikelberg's,[28] Mardin's,[32] the sector-based LDF by Iester et al.,[33] and the CSLO-based cup shape measure. The optic disc sector-based LDF outperformed the other formulas, with single cup shape measure having the lowest performance. The same formulas, except Mardin's, were used in a subsequent study by the same investigators to compare their ability to detect glaucomatous visual field defects.[24] The results confirmed their previous finding that the optic disc-based sector LDF and the cup shape measure were the best and worst formulas, respectively.[31] Because SLO-based LDFs rely on manual outlining of the optic disc contour, these methods are likely to be affected by the interobserver variability.

Unlike earlier studies that used SLO, SPL, and time-domain OCT, two recent studies have reported on the diagnostic performance of combined ONH, peripapillary RNFL, GCC parameters measured with SD-OCT (RTVue-100; Optovue, Fremont, CA). Vertical cup-to-disc diameter ratio, average RNFL thickness, and average GCC thickness area provided the best AUCs for discriminating between healthy subjects and subjects with perimetric glaucoma,[14] or for discriminating between glaucoma suspects and subjects with definite glaucoma.[15] The combination of parameters resulted in increased diagnostic performance in both studies, but the resulting LDFs did not include any of the ganglion cell layer variables. Anatomically, the GCC is made of the RNFL and ganglion cell layer, and the inner plexiform layer. Thus, combining GCC and RNFL parameters is somewhat flawed because the GCC already contains the RNFL, and it is not clear how this may have affected the linear discriminant analysis. Combination of Cirrus HD-OCT (Carl Zeiss Meditec, Inc.) structural parameters was recently shown to enhance significantly discrimination between healthy subjects and subjects with mild glaucoma using both linear discriminant analysis and CART.[13] However, this study did not include the GCIPL in the combination. As we are finalizing this paper, we are not aware of another study that assessed the diagnostic ability of combined ONH, RNFL, and GCIPL parameters.

With regard to study populations, our study only included patients with early glaucoma because it is oftentimes hard to distinguish early stages of the disease and normal state. This may be a limitation of this study. Indeed, using a group of patients at all stages of disease severity in the development of the underlying structure would tend to avoid restricting the range in the test measures and therefore attenuating correlations among variables that can result in falsely low estimates of factor loadings.[34] In contrast to our study, most previous studies[9–12,23–26,30,31] included patients with all spectra of glaucoma severity, but did not provide the diagnostic performance for early glaucoma. Because the performance of glaucoma diagnostic devices is often dependent on disease severity, the glaucoma performances reported in those studies would have been lower if the analysis was restricted to early glaucoma. Our study also differs from two of the earlier studies[9,35] in that EFA rather than PCA was performed prior to logistic regression. Principal component analysis is only a data reduction method and it is computed without regard to any underlying structure influenced by latent variables. In addition, in PCA, components are calculated using all of the variance of the manifest variables, so that the whole variance appears in the solution. Principal component analysis does not discriminate between shared and unique variance.[36] On the contrary, EFA reveals latent variables that possibly cause the manifest variables to covary. During factor extraction the shared variance of a variable is partitioned from its unique variance and error variance to reveal the underlying factor structure; only shared variance appears in the solution. If the factors are uncorrelated and communalities are moderate, it tends to overestimate values of variance accounted for by the components. Since EFA only considers shared variance, it should yield similar results while also avoiding the inflation of estimates of variance accounted for.[37,38] The diagnostic method used in this study, namely sequentially combining EFA and logistic regression modeling, benefited from several advantages. First, the use of EFA reorganizes a large amount of data into a more parsimonious set of component scores. Because each EFA component groups together correlated test measures, the component scores more directly gauge a variable's performance with regard to glaucoma status. Second, because the component structure was created from the data of both glaucoma and control subjects, the component structure reflects the structural differences between the two groups as well as the differences among subjects within each group. Third, the discriminant function weights the components in terms of their contributions to discriminating patients with early glaucoma from controls and then classifies each individual with high accuracy, sensitivity, and specificity. Fourth, our method went beyond simple calculation of AUC, sensitivity and specificity by providing AIC, and most importantly the predicted probability of early glaucoma for an individual along with a 95% prediction interval, which might prove extremely useful and may influence the physician's decision to initiate treatment. Despite only slight increase in AUC, the AIC and PIL of the multivariable model were significantly lower than those of the three best single

parameters (Table 5), indicating that our multivariable predictive model performed better and was more accurate than univariable models both at detecting the disease and differentiating between affected and unaffected individuals. An additional difference to consider is that most prior studies did not validate their discriminant functions in separate set of subjects. Validation is important because it ensures that the proposed model is robust to the subjects included in the analysis and will be useful for analyzing future datasets. Fifth, both the modeling and validation sets were drawn from the same sample selected with same inclusion and exclusion criteria. It is unclear whether this affected the performance of the model. Not having performed visual field in healthy subjects may be another methodological limitation of this study. Indeed, relying on IOP measurement and the ophthalmoscopic appearance of the optic disc as assessed by fellowship-trained glaucoma subspecialists may not have been sufficient to confidently exclude all subjects with glaucoma among normals. This is particularly true in subjects with early stages of glaucoma where functional deficits may precede detectable structural changes. Whether some glaucoma patients were missed among healthy subjects and whether this may have affected our results in a significant manner is unknown. On the other hand, the diagnostic performance of our model may have been inflated to some extent as a result of studying two clinically well-defined populations, namely nonglaucomatous and glaucomatous subjects. Therefore, it will be interesting to evaluate the diagnostic performance of this model in subjects suspected of having glaucoma. Doing so will both comply with the general principle that a diagnostic test is useful if it can decrease or eliminate the uncertainty with respect to the diagnosis (i.e., in glaucoma suspects) or to the disease stage and determine the "true" diagnostic performance of this model.

In conclusion, we have shown that the model based on the sequential partnering of EFA and multiple logistic regression of SD-OCT structural parameters produced good results in the modeling and the validation sets. This model was successful in predicting early glaucoma status and outperformed univariable models in terms AIC, median PILs, and classification rates. The classification method based on EFA of structural parameters developed in this paper is a robust and powerful method for the diagnosis of early glaucoma. It may become a valuable part of the toolbox of discriminatory analysis.

Acknowledgments

Supported by an unrestricted grant from Research to Prevent Blindness, New York, New York.

Disclosure: **J.-C. Mwanza**, Optovue, Inc. (F); **J.L. Warren**, None; **D.L. Budenz**, Optovue, Inc. (F)

References

1. Kingman S. Glaucoma is second leading cause of blindness globally. *Bull World Health Organ*. 2004;82:887–888.
2. Quigley HA, Broman AT. The number of people with glaucoma worldwide in 2010 and 2020. *Br J Ophthalmol*. 2006;90:262–267.
3. Friedman DS, Wolfs RC, O'Colmain BJ, et al. Prevalence of open-angle glaucoma among adults in the United States. *Arch Ophthalmol*. 2004;122:532–538.
4. Mikelberg FS, Yidegiligne HM, Schulzer M. Optic nerve axon count and axon diameter in patients with ocular hypertension and normal visual fields. *Ophthalmology*. 1995;102:342–348.
5. 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.
6. Tielsch JM, Katz J, Quigley HA, Miller NR, Sommer A. Intraobserver and interobserver agreement in measurement of optic disc characteristics. *Ophthalmology*. 1988;95:350–356.
7. Azuara-Blanco A, Katz LJ, Spaeth GL, Vernon SA, Spencer F, Lanzl IM. Clinical agreement among glaucoma experts in the detection of glaucomatous changes of the optic disk using simultaneous stereoscopic photographs. *Am J Ophthalmol*. 2003;136:949–950.
8. Badala F, Nouri-Mahdavi K, Raoof DA, Leeprechanon N, Law SK, Caprioli J. Optic disc and nerve fiber layer imaging to detect glaucoma. *Arch Ophthalmol*. 2007;144:724–732.
9. 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.
10. Ford BA, Artes PH, McCormick TA, Nicolela MT, LeBlanc RP, Chauhan BC. Comparison of data analysis tools for detection of glaucoma with the Heidelberg Retina Tomograph. *Ophthalmology*. 2003;110:1145–1150.
11. 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.
12. Toth M, Kothy P, Hollo G. Accuracy of scanning laser polarimetry, scanning laser tomography, and their combination in a glaucoma screening trial. *J Glaucoma*. 2008;17:639–646.
13. Baskaran M, Ong EL, Li JL, et al. Classification algorithms enhance the discrimination of glaucoma from normal eyes using high-definition optical coherence tomography. *Invest Ophthalmol Vis Sci*. 2012;53:2314–2320.
14. Fang Y, Pan YZ, Li M, Qiao RH, Cai Y. Diagnostic capability of Fourier-Domain optical coherence tomography in early primary open angle glaucoma. *Chin Med J (Engl)*. 2010;123:2045–2050.
15. Huang JY, Pekmezci M, Mesiwala N, Kao A, Lin S. Diagnostic power of optic disc morphology, peripapillary retinal nerve fiber layer thickness, and macular inner retinal layer thickness in glaucoma diagnosis with fourier-domain optical coherence tomography. *J Glaucoma*. 2011;20:87–94.
16. Mwanza JC, Oakley JD, Budenz DL, Chang RT, Knight OJ, Feuer WJ. Macular ganglion cell-inner plexiform layer: automated detection and thickness reproducibility with spectral domain-optical coherence tomography in glaucoma. *Invest Ophthalmol Vis Sci*. 2011;52:8323–8329.
17. Mwanza JC, Oakley JD, Budenz DL, Anderson DR. Ability of cirrus HD-OCT optic nerve head parameters to discriminate normal from glaucomatous eyes. *Ophthalmology*. 2011;118:241–248.
18. Mwanza JC, Durbin MK, Budenz DL, et al. Glaucoma diagnostic accuracy of ganglion cell-inner plexiform layer thickness: comparison with nerve fiber layer and optic nerve head. *Ophthalmology*. 2012;119:1151–1158.
19. Jöreskog KG. How large can a standardized coefficient be? SSI. 1999:1–3. Available at http://www.ssicentral.com/lisrel/techdocs/HowLargeCanaStandardizedCoefficientbe.pdf. Accessed September 13, 2013.
20. Hair JF, Black WC, Babin BJ, Anderson RE. *Multivariate Data Analysis*. 7th ed. Upper Saddle River, NJ: Prentice Hall; 2009.
21. Burhman KP, Anderdon DR. Multimodel inference: understanding AIC and BIC in model selection. *Sociol Methods Res*. 2004;33:261–304.

22. 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.
23. Budenz DL, Michael A, Chang RT, McSoley J, Katz J. Sensitivity and specificity of the StratusOCT for perimetric glaucoma. *Ophthalmology*. 2005;112:3–9.
24. Iester M, Mardin CY, Budde WM, Junemann AG, Hayler JK, Jonas JB. Discriminate analysis formulas of optic nerve head parameters measured by confocal scanning laser tomography. *J Glaucoma*. 2002;11:97–104.
25. Lu AT, Wang M, Varma R, et al. Combining nerve fiber layer parameters to optimize glaucoma diagnosis with optical coherence tomography. *Ophthalmology*. 2008;115:1352–1357.
26. Wang M, Lu AT, Varma R, Schuman JS, Greenfield DS, Huang D. Combining information from 3 anatomic regions in the diagnosis of glaucoma with time-domain optical coherence tomography [published online ahead print June 23, 2012]. *J Glaucoma*. doi:10.1097/IJG.0b013e318264b941
27. Bathija R, Zangwill L, Berry CC, Sample PA, Weinreb RN. Detection of early glaucomatous structural damage with confocal scanning laser tomography. *J Glaucoma*. 1998;7:121–127.
28. Mikelberg FS, Parfitt CM, Swindale NV, Graham SL, Drance SM, Gosine R. Ability of the Heidelberg retina tomograph to detect early glaucomatous visual field loss. *J Glaucoma*. 1995;4:242–247.
29. Burk RO, Noack H, Rohrschneider K, Volcker HE. Prediction of glaucomatous visual field defects by reference plane independent three-dimensional optic nerve head parameters. In: Wall M, Wild JM, eds. *XIIIth International Perimetric Society Meeting*. Gardone Riviera, Italy: Kugler; 1998:463–474.
30. Ferreras A, Pablo E, Larrosa JM, Polo V, Pajarin AB, Honrubia FM. Discriminating between normal and glaucoma-damaged eyes with the Heidelberg Retina Tomograph 3. *Ophthalmology*. 2008;115:775–781.
31. Iester M, Jonas JB, Mardin CY, Budde WM. Discriminant analysis models for early glaucoma detection of glaucomatous optic disc changes. *Br J Ophthalmol*. 2000;84:464–468.
32. Mardin CY, Horn FK, Jonas JB, Budde WM. Preperimetric glaucoma diagnosis by confocal scanning laser tomography of the optic disc. *Br J Ophthalmol*. 1999;83:299–304.
33. Iester M, Swindale NV, Mikelberg FS. Sector-based analysis of optic nerve head shape parameters and visual field indices in healthy and glaucomatous eyes. *J Glaucoma*. 1997;6:370–376.
34. Fabrigar LR, Wegener DT, MacCallum RC, Strahan EJ. Evaluating the use of explanatory factor analysis in psychological research. *Psychol Methods*. 1999;4:272–299.
35. Huang ML, Chen HY. Development and comparison of automated classifiers for glaucoma diagnosis using Stratus optical coherence tomography. *Invest Ophthalmol Vis Sci*. 2005;46:4121–4129.
36. Gorsuch RL. Common factor-analysis versus component analysis—some well and little known facts. *Mutivariate Behav Res*. 1990;25:33–39.
37. McArdle JJ. Principles versus pincipals of structural factor-analyses. *Multivariate Behav Res*. 1990;25:81–87.
38. Gorsuch RL. Exploratory factor analysis: its role in item analysis. *J Pers Assess*. 1997;68:532–560.

Appendix

The Ganglion Cell Analysis Study Group

Donald L. Budenz, MD, MPH (Principal Investigator) and Jean-Claude Mwanza, MD, MPH, PhD, Department of Ophthalmology, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina;

Robert T. Chang, MD (Principal Investigator), Department of Ophthalmology, Stanford University, Palo Alto, California;

Arvind Neelakantan, MD (Principal Investigator) and David G. Godfrey, MD, Glaucoma Associates of Texas, Dallas, Texas; and

Alan S. Crandall, MD (Principal Investigator) and Randy Carter, OD, Eye Institute of Utah, Salt Lake City, Utah.

Investigative Ophthalmology & Visual Science