tvst

https://doi.org/10.1167/tvst.7.2.16

Article

Validation of the UNC OCT Index for the Diagnosis of Early Glaucoma

Jean-Claude Mwanza[1], Gary Lee[2], Donald L. Budenz[1], Joshua L. Warren[3], Michael Wall[4], Paul H. Artes[5], Thomas M. Callan[2], and John G. Flanagan[6]

[1] Department of Ophthalmology, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA
[2] Research and Development, Carl Zeiss Meditec, Inc., Dublin, CA, USA
[3] Department of Biostatistics, Yale University, New Haven, CT, USA
[4] Department of Ophthalmology and Visual Sciences, The University of Iowa, Iowa City, IA, USA
[5] Eye and Vision Research Group, Institute of Health and Community, Plymouth University, UK
[6] School of Optometry, University of California Berkeley, Berkeley, CA, USA

Correspondence: Jean-Claude Mwanza, Department of Ophthalmology, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599, USA. e-mail: jean-claude_mwanza@med.unc.edu

Received: 22 September 2017
Accepted: 21 February 2018
Published: 3 April 2018

Keywords: early glaucoma; optical coherence tomography; UNC OCT Index

Citation: Mwanza JC, Lee G, Budenz DL, Warren JL, Wall M, Artes PH, Callan TM, Flanagan JG. Validation of the UNC OCT Index for the diagnosis of early glaucoma. Trans Vis Sci Tech. 2018;7(2):16, https://doi.org/10.1167/tvst.7.2.16

Purpose: To independently validate the performance of the University of North Carolina Optical Coherence Tomography (UNC OCT) Index in diagnosing and predicting early glaucoma.

Methods: Data of 118 normal subjects (118 eyes) and 96 subjects (96 eyes) with early glaucoma defined as visual field mean deviation (MD) greater than −4 decibels (dB), aged 40 to 80 years, and who were enrolled in the Full-Threshold Testing Size III, V, VI comparison study were used in this study. CIRRUS OCT average and quadrants' retinal nerve fiber layer (RNFL); optic disc vertical cup-to-disc ratio (VCDR), cup-to-disc area ratio, and rim area; and average, minimum, and six sectoral ganglion cell-inner plexiform layer (GCIPL) measurements were run through the UNC OCT Index algorithm. Area under the receiver operating characteristic curve (AUC) and sensitivities at 95% and 99% specificity were calculated and compared between single parameters and the UNC OCT Index.

Results: Mean age was 60.1 ± 11.0 years for normal subjects and 66.5 ± 8.1 years for glaucoma patients ($P < 0.001$). MD was 0.29 ± 1.04 dB and −1.30 ± 1.35 dB in normal and glaucomatous eyes ($P < 0.001$), respectively. The AUC of the UNC OCT Index was 0.96. The best single metrics when compared to the UNC OCT Index were VCDR (0.93, $P = 0.054$), average RNFL (0.92, $P = 0.014$), and minimum GCIPL (0.91, $P = 0.009$). The sensitivities at 95% and 99% specificity were 85.4% and 76.0% (UNC OCT Index), 71.9% and 62.5% (VCDR, all $P < 0.001$), 64.6% and 53.1% (average RNFL, all $P < 0.001$), and 66.7% and 58.3% (minimum GCIPL, all $P < 0.001$), respectively.

Conclusions: The findings confirm that the UNC OCT Index may provide improved diagnostic perforce over that of single OCT parameters and may be a good tool for detection of early glaucoma.

Translational Relevance: The UNC OCT Index algorithm may be incorporated easily into routine clinical practice and be useful for detecting early glaucoma.

Introduction

Glaucoma is a chronic and slowly progressive optic neuropathy that causes irreversible structural changes to the inner retinal layers and the optic nerve head (ONH), as well as damage to the visual field. Early disease detection and treatment are important for the prevention of visual impairment. Detection is routinely achieved by assessing structural changes using imaging devices and visual field damage using standard automated perimetry (SAP). Ocular imaging techniques for glaucoma assessment have improved considerably in recent years, particularly with the introduction of technologies such as spectral-domain optical coherence tomography (SD-OCT). SD-OCT

can reliably perform in vivo qualitative and quantitative analyses of both superficial and deep structures of the ONH, retinal nerve fiber layer (RNFL), and ganglion cell layer with high reproducibility. Evidence has accumulated that OCT helps in detecting glaucoma in early stages both in the clinical setting[1–4] and in the population[5–7] and can often identify progressive disease before SAP.[8,9] However, despite their good performance in diagnosing early glaucoma, SD-OCT devices can still falsely diagnose healthy eyes as having glaucoma or miss the diagnosis of early glaucoma in substantial proportions of subjects.[10–13] This is particularly true when the diagnosis is based on single parameters such as average RNFL thickness.

For a long time, RNFL thickness was the OCT parameter most frequently used for objective and quantitative assessment of glaucoma. Subsequent developments have resulted in the addition of macula and both superficial and deep ONH parameters. There is now a large number of SD-OCT parameters available for glaucoma assessment. As a consequence, clinicians face the challenging task of mentally processing the information from a whole range of parameters to determine whether an individual patient has early glaucoma or not. This situation is particularly challenging when the OCT results of ONH, RNFL, and ganglion cell analyses do not agree with each other. In order to enhance the performance of OCT to predict and diagnose glaucoma, various statistical methods (i.e., logistic regression, classification trees, linear discriminant analysis, and machine-learning tools) have been proposed to combine single metrics into a single composite. With few exceptions, such combined metrics have generally provided better diagnostic performance than have single metrics.[14–25] We previously developed the University of North Carolina Optical Coherence Tomography (UNC OCT) Index based on exploratory factor analysis (EFA) and multivariable logistic regression.

Compared to single parameters, the UNC OCT Index improved the prediction and detection of early glaucoma based on various performance measures.[26] The current study was designed to validate the performance of the UNC OCT Index in an independent population of patients with very early disease.

Methods

Patients

Data for this investigation were drawn from 118 normal subjects and 96 patients previously diagnosed with glaucoma and under treatment, 40 to 80 years old, and who participated in the study to compare Goldmann stimulus sizes III, V, and VI for automated perimetry (Flanagan JG, et al. *IOVS* 2016;57: ARVO E-Abstract 3417). That study was designed (1) to determine the diagnostic performance of larger stimulus sizes for the detection of visual field damage, including damage close to fixation, in patients with early visual field loss due to glaucoma and (2) to establish normative reference limits for the different stimulus sizes. The study received Institutional Review Board approval from the University of Waterloo (Waterloo, ON, Canada), the University of Iowa (Iowa City, IA), and the Dalhousie University (Halifax, NS, Canada). Written informed consent was obtained from all participants, and the study adhered to the tenets of the Declaration of Helsinki.

All participants underwent an ophthalmological examination that included visual acuity measurement (Early Treatment of Diabetic Retinopathy Study, or EDTRS, acuity chart); refraction (phoropter or trial lens) if necessary; intraocular pressure measurement (Goldmann applanation tonometry); visual field testing (Humphrey 24-2 SITA and Full-Threshold) using stimulus size III, V, and VI, nondilated CIRRUS HD-OCT (Carl Zeiss, Meditec, Inc, Dublin, CA) imaging of the ONH (Optic Disc 200×200 Cube Scan) and the macula (200×200 Cube Scan); and dilated stereoscopic fundus photography centered on the optic discs. Only the results of the visual fields obtained with SITA Standard (stimulus size III) were used in the present study. All OCT data are from scans with signal strength of 6 or more and en face images without sign of saccade or blinking artifacts.

The clinical diagnosis of glaucoma was based on structural defect on CIRRUS HD-OCT defined as at least one yellow quadrant or two red clock hours and/or the rim area or vertical cup-to-disc ratio (VCDR) in the yellow range. Only patients with early glaucoma, defined as visual field mean deviation (MD) of −4 decibels (dB) or better on Humphrey Field Analyzer (HAF IIi Model 750; Carl Zeiss Meditec, Inc) with SITA Standard 24-2 test, were retained for the study. Controls and glaucoma subjects were excluded if their best corrected visual acuity was worse than 20/30 (6/9) in a potential study eye; refractive error was greater than 5 diopters spherical equivalent or 2.5 diopters cylinder; they had a diagnosis or a history of ocular or systemic disease; or they had treatment that may have affected the visual field. Participants with unreliable visual field

test results defined as false-positive rate of 15% or greater and/or false-negative rate of 20% or more were also excluded.

Validation of the UNC OCT Index Performance

CIRRUS HD-OCT data of normal and glaucoma participants were used. For each eye, the UNC OCT Index was derived as previously described[26] from average, superior, inferior, inferior and superior quadrant RNFL; optic disc rim area, VCDR, and cup-to-disc ratio (CDR); and ganglion cell-inner plexiform layer (GCIPL) average, minimum, superior, superonasal, inferonasal, inferior, inferotemporal, and superotemporal measurements. In brief, this algorithm is based on EFA with oblique promax rotation to identify a small set of latent factors that explain the large proportion of the variability observed in the 16 original OCT parameters. The EFA is then followed by a multivariable logistic regression analysis as extensively detailed elsewhere.[26] The EFA identified five latent factors (factor 1, all eight GCIPL parameters; factor 2, all three optic disc parameters; factor 3, average, superior, and inferior RNFL; factor 4, temporal RNFL; and factor 5, nasal RNFL) accounting for a large proportion of the original variability. These factors were then submitted to a multivariable logistic regression analysis using the backward variable selection elimination technique, with glaucoma status as the dependent variable and these five factors (and first-order interactions) as predictors. The logistic regression identified only three factors (factor 1, all GCIPL parameters; factor 2, all ONH parameters; factor 3, average, superior, and inferior RNFL) as significant predictors of early glaucoma. The logistic regression model also controlled for the age category of each patient. Figure 1 provides a diagram of the algorithm.

Statistical Analysis

Statistical analyses were performed with SAS software (version 9.2; SAS, Cary, NC) and in MATLAB (R2017a and R2017b; MathWorks, Natick, MA), with statistical significance level set at $P < 0.05$. Mean values of demographic, visual field, and OCT variables were compared between groups using the *t*-test for independent samples. The diagnostic performance was assessed using the area under the receiver operator characteristic curve (AUC) and sensitivity at specificities fixed at 95% and 99%. AUCs and sensitivities were determined for the whole sample (visual field MD better than −4 dB) and then separately in the subset of patients with visual field MD better than −2 dB) to determine whether there was a change in diagnostic performance. The Delong method[27] was used to compare single parameter AUCs to that of the UNC OCT Index. Sensitivities were compared using McNemar's test.

For predicted early glaucoma probability calculations, we used formulas that differ slightly from those presented in the original publication.[26] There, the full dataset was divided into two thirds of the original size (modeling dataset) and one third of the original size (validation dataset) to evaluate the model validation. In the updated version of the model presented here,[28] both datasets were merged to increase the sample size, leading to improved robustness and stability of the estimation of the unknown model parameters. Stepwise variable selection was used on the full set of factors and first-order interactions and led to the inclusion of an additional significant predictor of early glaucoma: the interaction between factor 1 and factor 3. Other than the inclusion of this additional predictor, the method remains consistent across both studies. The model outputs a predictive probability that ranges from 0.0 to 1.0. We calculated the optimal cutoff as 0.34 by finding the point on the receiver operating characteristic (ROC) curve that minimized the distance between the ROC curve and the (0, 1) point.[29] Predicted probabilities larger than 0.34 indicate a greater likelihood that the structural changes represent glaucomatous damage, whereas values lower than 0.34 indicate that the observed structural changes are likely nonglaucomatous.

Results

Demographic and Ocular Features of the Study Population

The demographic and clinical characteristics of the study population are displayed in Table 1. There were 100 men and 114 women. Glaucoma patients were slightly older and had significantly worse visual field MD, thinner RNFL and GCIPL, smaller rim area, and greater VCDR and CDR than normal subjects (all $P < 0.001$). Glaucoma patients with a visual field MD better than −2 dB ($n = 71$) were also significantly older (66.2 ± 8.0 years), with worse visual fields (−0.71 ± 0.99 dB) and worse structural damage than normal subjects (all $P < 0.001$).

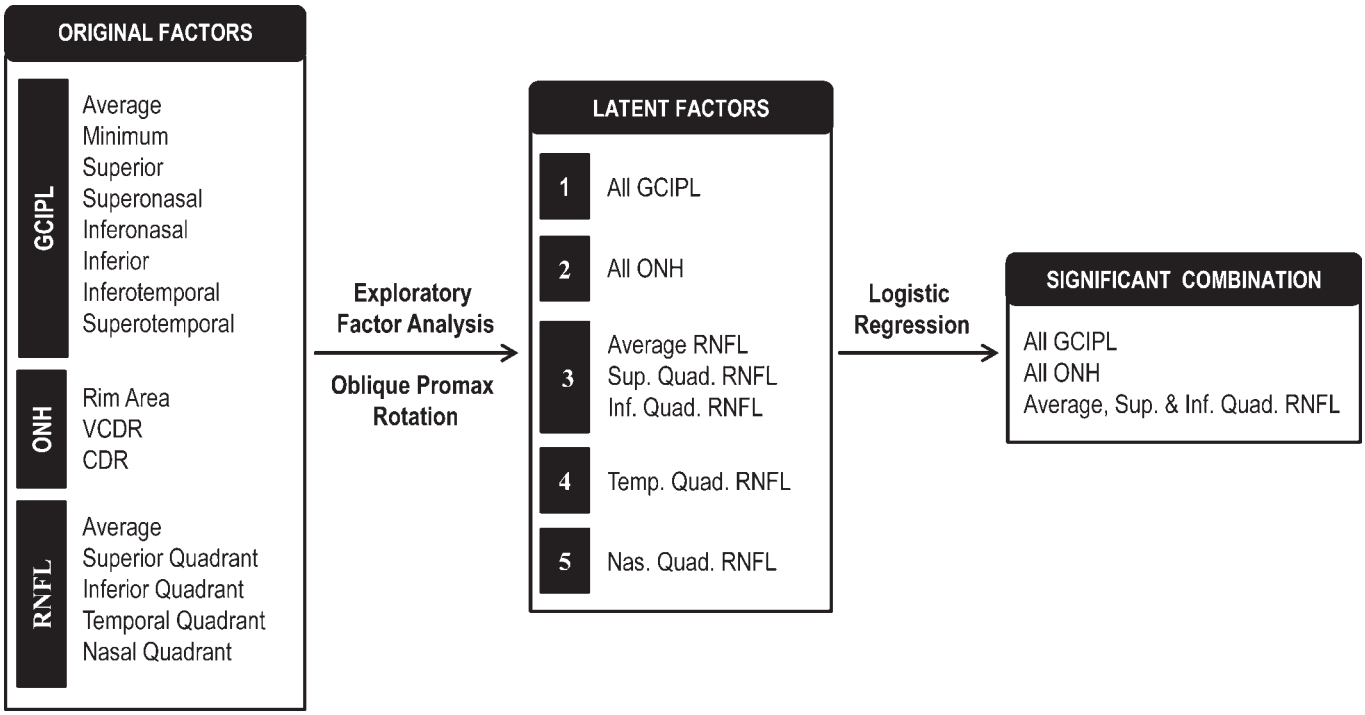

Figure 1. Diagram of steps of the UNC OCT Index algorithm showing all OCT parameters used in the model (*left*), the latent factors obtained after submitting the original parameters to EFA followed by oblique promax rotation (*middle*), and the final composite resulting from the multivariable logistic regression (*right*).

Table 1. Comparison of OCT Parameters in Normal and Glaucomatous Subjects

Parameter	Normal	All Glaucoma	Glaucoma With MD $\geq$ 2	P_1	P_2
Age	60.1 ± 11.0	66.5 ± 8.1	66.2 ± 8.1	<0.001	<0.001
Visual field MD	0.3 ± 1.0	−1.3 ± 1.3	−0.7 ± 1.0	<0.001	<0.001
Average GCIPL	79.9 ± 6.2	69.3 ± 8.0	70.0 ± 7.5	<0.001	<0.001
Minimum GCIPL	78.1 ± 7.9	62.0 ± 10.5	62.8 ± 10.5	<0.001	<0.001
Superotemporal GCIPL	79.2 ± 6.1	69.8 ± 10.3	70.9 ± 9.5	<0.001	<0.001
Superior GCIPL	80.3 ± 6.5	70.4 ± 10.4	71.0 ± 10.3	<0.001	<0.001
Superonasal GCIPL	81.4 ± 6.7	72.8 ± 9.4	73.3 ± 9.0	<0.001	<0.001
Inferonasal GCIPL	79.7 ± 7.1	70.4 ± 8.6	70.8 ± 8.2	<0.001	<0.001
Inferior GCIPL	78.3 ± 6.7	66.3 ± 9.1	66.8 ± 9.1	<0.001	<0.001
Inferotemporal GCIPL	80.5 ± 6.4	66.4 ± 10.8	67.1 ± 10.5	<0.001	<0.001
Average RNFL	92.2 ± 9.2	73.7 ± 9.0	75.1 ± 8.7	<0.001	<0.001
Superior quadrant RNFL	64.0 ± 10.9	54.1 ± 9.6	54.4 ± 9.9	<0.001	<0.001
Superior quadrant RFNL	113.9 ± 15.4	88.7 ± 14.4	90.7 ± 14.3	<0.001	<0.001
Nasal quadrant RNFL	72.3 ± 10.7	65.8 ± 9.2	66.3 ± 9.7	<0.001	<0.001
Inferior quadrant RNFL	118.8 ± 14.8	86.0 ± 19.0	89.1 ± 17.9	<0.001	<0.001
Rim area	1.3 ± 0.3	0.9 ± 0.2	0.9 ± 0.2	<0.001	<0.001
Average CDR	0.5 ± 0.2	0.7 ± 0.1	0.7 ± 0.1	<0.001	<0.001
Vertical CDR	0.5 ± 0.2	0.7 ± 0.1	0.7 ± 0.1	<0.001	<0.001
UNC OCT index	0.0 ± 0.2	0.8 ± 0.3	0.8 ± 0.4	<0.001	<0.001

P_1, significance of the difference between normal and all glaucoma subjects; P_2, significance of the difference between normal and glaucoma subjects with MD $\geq$2 dB; CDR, cup-to-disc ratio.

Table 2. Diagnostic Performance of UNC OCT Index and Single GCIPL, RNFL, and ONH Parameters in Early Glaucoma With Visual Field MD ≥4 dB

Parameter	AUC (95% CI)	P[a]	Sensitivity, %	
			95% Specificity (P)[b]	99% Specificity (P)[b]
UNC index	0.96 (0.93–0.98)	–	85.4	77.1
Average GCIPL	0.84 (0.80–0.90)	0.000	52.1 (<0.001)	33.3 (<0.001)
Minimum GCIPL	0.91 (0.87–0.95)	0.010	66.7 (<0.001)	58.3 (<0.001)
Superotemporal GCIPL	0.81 (0.76–0.87)	0.000	51.0 (<0.001)	31.3 (<0.001)
Superior GCIPL	0.79 (0.73–0.86)	0.000	41.7 (<0.001)	35.4 (<0.001)
Superonasal GCIPL	0.76 (0.70–0.83)	0.000	32.3 (<0.001)	20.8 (<0.001)
Inferonasal GCIPL	0.79 (0.73–0.85)	0.000	37.5 (<0.001)	20.8 (<0.001)
Inferior GCIPL	0.85 (0.80–0.90)	0.000	51.0 (<0.001)	45.8 (<0.001)
Inferotemporal GCIPL	0.86 (0.81–0.91)	0.000	67.7 (<0.001)	56.3 (<0.001)
Average RNFL	0.92 (0.89–0.96)	0.014	64.6 (<0.001)	53.1 (<0.001)
Temporal quadrant RNFL	0.75 (0.69–0.82)	0.000	21.9 (<0.001)	7.3 (<0.001)
Superior quadrant RFNL	0.89 (0.84–0.93)	0.003	47.9 (<0.001)	31.3 (<0.001)
Nasal quadrant RNFL	0.67 (0.61–0.75)	0.000	8.3 (<0.001)	4.2 (<0.001)
Inferior quadrant RNFL	0.91 (0.87–0.95)	0.005	69.8 (<0.001)	49.0 (<0.001)
Rim area	0.92 (0.89–0.96)	0.044	76.0 (0.07)	64.6 (0.019)
Average CDR	0.89 (0.86–0.94)	0.000	52.1 (<0.001)	37.5 (<0.001)
Vertical CDR	0.93 (0.90–0.97)	0.054	71.9 (0.003)	62.5 (<0.001)
Visual field MD	0.83 (0.78–0.89)	0.000	42.7 (<0.001)	18.8 (<0.001)

[a] Significance of the difference compared to the UNC Index AUC.
[b] Significance of the difference compared to the UNC Index sensitivity.

Diagnostic Performance of UNC OCT Index Versus Single OCT Parameters

The AUC of the UNC OCT Index was 0.96 (95% confidence interval [CI]: 0.93–0.98). The best single metrics when compared to UNC OCT Index were VCDR (0.93, CI: 0.90–0.97, $P = 0.053$), average RNFL (0.92, CI: 0.89–0.96, $P = 0.011$), and minimum GCIPL (0.91, CI: 0.87–0.95, $P = 0.009$), as shown in Table 2 and Figure 2 (left panel). The discriminating abilities of the other single metrics are also listed in Table 2. The sensitivities at 95% and 99% fixed specificity were 85.4% and 76.0% (UNC OCT Index); those of the best single metric from each of the three anatomical areas were 71.9% and 62.5% (VCDR), 64.6% and 53.1% (average RNFL), and 66.7% and 58.3% (minimum GCIPL), respectively (Table 2). The sensitivity of the UNC Index was significantly better than sensitivities of individual parameters (all $P < 0.05$), except for rim area at 95% specificity ($P = 0.07$). The results of a separate analysis (Table 3) including only 71 glaucomatous eyes with visual field MD greater than −2 dB (mean: −0.71 ± 0.99 dB) yielded an AUC of 0.95 (CI: 0.92–0.98) for the UNC OCT Index, compared to 0.92 (CI: 0.88–0.97, $P = 0.064$) for VCDR, 0.91 (CI: 0.87–0.95, $P = 0.030$) for average RNFL, and 0.90 (CI: 0.85–0.94, $P = 0.026$) for minimum GCIPL, as also shown in Figure 2 (right panel). The sensitivities in this subgroup of glaucomatous patients at the same specificity levels were 81.7% and 71.8% for UNC OCT Index, 67.6% and 60.6% for VCDR, 57.7% and 45.1% for average RNFL, and 63.4% and 53.4% for minimum GCIPL. Those of the other single variables are also listed in Table 3. Again, the UNC Index was significantly more sensitive than all individual parameters (all $P < 0.05$), with exception of rim area at both 95% and 99% specificity ($P = 0.39$ and 0.09, respectively). The differences in AUCs and sensitivities of the UNC OCT Index, VCDR, average RNFL, and minimum GCIPL in glaucomatous eyes with visual field MD greater than −4 dB were not statistically significant from those in eyes with visual field MD greater than −2 dB ($P = 0.79$–0.84 for AUCs, 0.14–0.52 for sensitivities at 95% specificity, and 0.14–0.72 for sensitivities at 99% specificity).

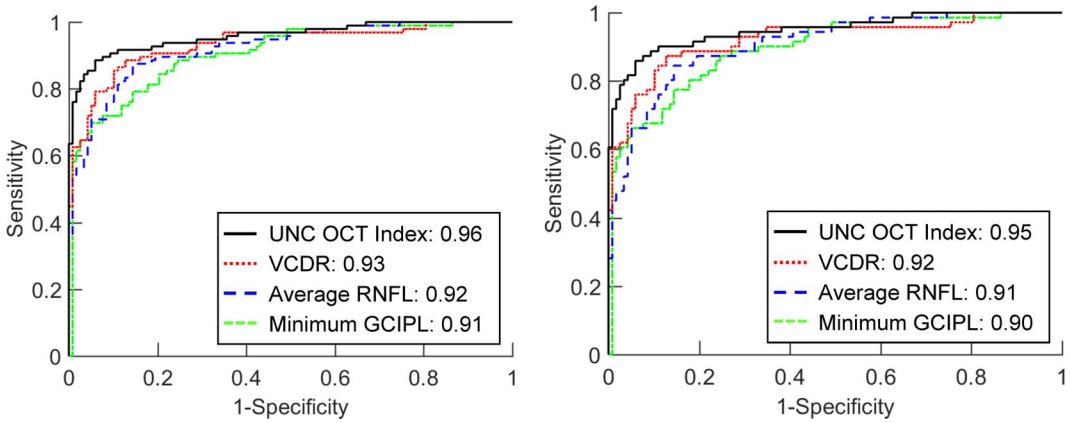

Figure 2. ROC curves of the UNC OCT Index and the best single OCT parameters for each anatomical area in glaucoma patients with visual field MD better than −4 dB (*left*) and those with MD better than −2 dB (*right*).

Predicted Probabilities in Representative Cases

The logistic regression parameter estimates (Table 4) were used to compute the predicted probabilities of early glaucoma (Table 5) for the left eye of a 73-year-old female subject and the right eye of a different 73-year-old female subject, both suspected of having early glaucoma. The OCT outputs of these eyes are shown in Figure 3. Except for inferior quadrant and clock-hour 5 RNFL thickness and the VCDR that were borderline, all other RNFL, ONH, and GCIPL measurements were within normal range in the left eye (Fig. 3A). The predicted probability in this case

Table 3. Diagnostic Performance of UNC OCT Index and Single GCIPL, RNFL, and ONH Parameters in Early Glaucoma With Visual Field MD ≥2 dB

			Sensitivity, %	
Parameter	AUC (95% CI)	*P**	95% Specificity (*P***)	99% Specificity (*P***)
UNC index	0.95 (0.92−0.98)	–	81.7	73.3
Average GCIPL	0.83 (0.78−0.90)	0.000	47.9 (<0.001)	32.4 (<0.001)
Minimum GCIPL	0.90 (0.85−0.94)	0.026	63.4 (0.007)	53.5 (0.003)
Superotemporal GCIPL	0.79 (0.73−0.86)	0.000	46.5 (<0.001)	23.9 (<0.001)
Superior GCIPL	0.78 (0.71−0.86)	0.000	38.0 (<0.001)	33.8 (<0.001)
Superonasal GCIPL	0.75 (0.69−0.83)	0.000	35.2 (<0.001)	19.7 (<0.001)
Inferonasal GCIPL	0.78 (0.72−0.86)	0.000	35.2 (<0.001)	19.7 (<0.001)
Inferior GCIPL	0.84 (0.78−0.90)	0.000	47.9 (<0.001)	43.7 (<0.001)
Inferotemporal GCIPL	0.85 (0.79−0.91)	0.000	64.8 (0.004)	53.5 (0.001)
Average RNFL	0.91 (0.87−0.95)	0.030	57.7 (<0.001)	45.1 (<0.001)
Temporal quadrant RNFL	0.74 (0.68−0.82)	0.000	22.5 (<0.001)	7.0 (<0.001)
Superior quadrant RFNL	0.87 (0.82−0.93)	0.006	42.3 (<0.001)	23.9 (<0.001)
Nasal quadrant RNFL	0.66 (0.58−0.74)	0.000	8.5 (<0.001)	4.2 (<0.001)
Inferior quadrant RNFL	0.90 (0.85−0.95)	0.016	63.4 (0.002)	38.0 (<0.001)
Rim area	0.92 (0.88−0.96)	0.145	74.6 (0.39)	63.4 (0.09)
Average CDR	0.89 (0.85−0.94)	0.001	49.3 (<0.001)	36.6 (<0.001)
Vertical CDR	0.92 (0.88−0.97)	0.064	67.6 (0.008)	60.6 (0.022)
Visual field MD	0.78 (0.71−0.85)	0.000	22.5 (<0.001)	0.0 (<0.001)

P *, significance of the difference relative to UNC OCT Index AUC; *P***, significance of the difference relative to UNC OCT Index sensitivity.

Table 4. Logistic Regression Parameter Estimates

Parameter	Estimate	Odds Ratio	*P* Value
Intercept (β_0)	−1.6385 (−2.5583, −0.7187)	0.1943	0.0005
Factor1 (β_1)	−1.6135 (−2.7050, −0.5220)	0.1992	0.0038
Factor2 (β_2)	5.7379 (3.1443, 8.3315)	310.4119	<0.0001
Factor3 (β_3)	−3.5384 (−5.3572, −1.7196)	0.0291	0.0001
Factor1*Factor3 (β_4)	−1.3297 (−2.2056, −0.4538)	0.2646	0.0029
Age category 1 (β_5)	1.5451 (−0.0180, 3.1082)	4.6884	0.0527
Age category 2 (β_6)	−0.2444 (−1.4435, 0.9547)	0.7832	0.6895
Age category 3 (β_7)	−1.0383 (−2.3432, 0.2666)	0.3541	0.1189

was 0.90 (CI: 0.68–0.97), highly indicative of glaucoma. In the other case (Fig. 3B), the superior quadrant and clock-hour 1 and 6 thicknesses were borderline; the remainder of the measurements were within normal range. The predicted probability was 0.17 (CI: 0.04–0.49), consistent with a healthy eye.

Discussion

The correct identification of glaucoma in its early stages is essential in order to initiate early treatment and to preserve visual function. SD-OCT is a useful imaging modality for the diagnosis and long-term monitoring of glaucoma. However, its limitations in the early stage of the disease have been well documented. The present study was designed to validate the performance of the UNC OCT Index, which was created previously as a combinatorial model for the diagnosis of early glaucoma using CIRRUS HD-OCT structural ONH, RNFL, and GCIPL parameters. We demonstrated that combining single parameters into a single UNC OCT Index improved the diagnostic performance over that obtained with single parameters of each anatomical area.

Although SD-OCT has been available for almost a decade, we have located only four previous papers that have reported on the value of combining ONH, RNFL, and ganglion cell layer parameters for glaucoma diagnosis. Our current findings not only confirm our own previous results,[26] but they also agree with other previous reports, demonstrating that reducing the number of OCT parameters into a composite parameter improves the performance of SD-OCT in diagnosing early glaucoma.[2,19–21] Huang et al.[19] compared the ability of individual ONH, RNFL, and ganglion cell complex (GCC) parameters obtained with RTVue OCT and their combinations to differentiate 74 normal subjects from 146 perimetric glaucoma patients. The combination of individual parameters was achieved using a linear discriminant function (LDF). The AUC of the LDF

Table 5. Predictive Probabilities of Two Representative Subjects

Parameter	Subject BA15, OS	Subject RM9, OD
Intercept	−1.6385	−1.6385
Factor1	−1.6135 × −0.62104 = 1.00204804	−1.6135 × −0.05817 = 0.093857295
Factor2	5.7379 × 0.40731 = 2.337104	5.7379 × 0.21444 = 1.230435276
Factor3	−3.5384 × −0.26266 = 0.929396144	−3.5384 × 0.28954 = −1.024508336
Factor1*Factor3	−1.3297 × (−0.62104 × −0.26266) = −0.2137604795	−1.3297 × (−0.05817 × 0.28954) = 0.022395527
Age category 1	1.5451 × −1 = −1.5451	1.5451 × −1 = −1.5451
Age category 2	−0.2444 × −1 = 0.2444	−0.2444 × −1 = 0.2444
Age category 3	−1.0383 × −1 = 1.0383	−1.0383 × −1 = 1.0383
Sum, logit of predicted probability	2.1538877045	−1.578720238
Predicted probability	1/(1 + exp{−(2.1538877)}) = **0.90**	1/(1 + exp{−(−1.5787202)}) = **0.17**

OS, left eye; OD, right eye.

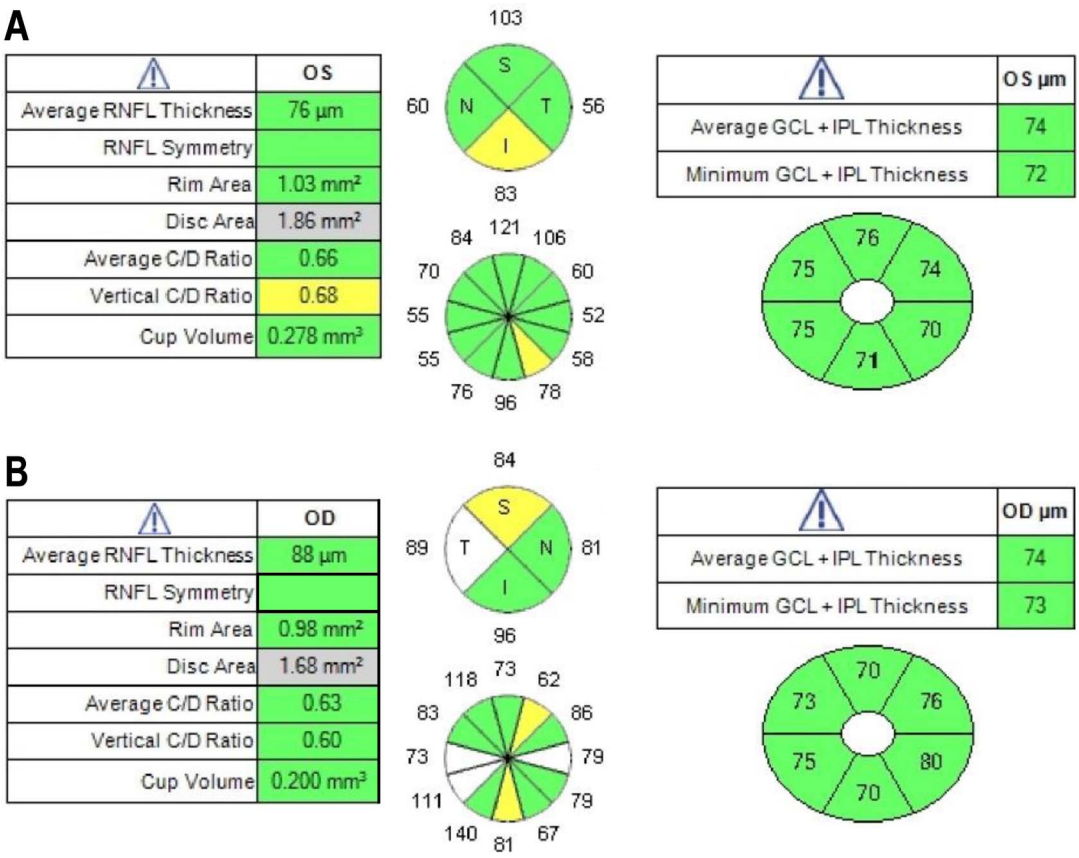

Figure 3. SD-OCT peripapillary RNFL, ONH, and GCIPL profile of two subjects diagnosed as glaucoma suspects. In the first subject (A), OCT data run through the UNC OCT Index yielded a predicted probability for early glaucoma of 0.90, suggesting that this eye was glaucomatous. In the second subject (B), the predicted probability was 0.17, suggesting that the eye was not glaucomatous.

(0.970) was significantly better than those of best single variables, namely average RNFL (0.919), superior hemisphere GCC (0.871), and VCDR (0.854). In another investigation with RTVue OCT, Loewen et al.[21] used a multivariable logistic regression model to generate the glaucoma structural diagnostic index, an indicator ranging from 0 to 1 derived from the combination of ONH, RNFL, and GCC parameters. The GDSI, which included the composites' overall thickness and focal loss volume for RNFL and GCC, and the VCDR had a better diagnostic performance than global loss volume GCC in terms of AUC (0.922 vs. 0.896), sensitivity at 95% and 99% specificity (80.2% and 69% vs. 61.4% and 52.5%), and specificity at 95% and 99% specificity cutoff (60% and 80% vs. 9.3% and 11.6%). It is important to note that this performance was based on glaucoma patients at all severity stages classified according to the enhanced glaucoma severity system.[30] Surprisingly, when only patients with early glaucoma were considered, the AUC of the GDSI (0.840–0.874) no longer outperformed those of best individual parameters. CIRRUS OCT ONH, RNFL, and GCIPL parameters were used in another study to generate three predictive models (1, quantitative; 2, qualitative; and 3, quantitative + qualitative parameters) using multivariable logistic regression that combined the best OCT parameters.[20] Model 3 was the most robust at all glaucoma severity stages with similar AUCs in the training (0.937) and validation groups (0.932) and sensitivities of 77.8% and 78.1% at 95% specificity, respectively. The authors did not provide the diagnostic performance of their combinatorial models in early glaucoma, which would have allowed direct comparison with our findings. Since diagnostic performance increases with disease severity, it is compelling to assume that their model 3 would have had a somewhat lower performance than that of the OCT UNC Index in eyes with early glaucoma. Our current findings support the hypothesis that the binary OR-logic combination of CIRRUS OCT minimum GCIPL and average RNFL or rim area provides better diagnostic performances than those of AND-

logic combinations or best single GCIPL, RNFL, or ONH parameters.[2] In contrast to this, Fang et al.[18] reported no improvement of the ability of the logistic regression-based combination of the best three individual parameters (VCDR, average RNFL, and rim area) over that of the best single parameter (VCDR) measured with Cirrus OCT to discriminate between normal subjects and patients with early glaucoma. Adding more parameters to the model made no further difference. The model of the UNC OCT Index evaluated here differs from all other previous models in that its composite contains 16 factors (Fig. 1) and therefore more information contributing to the diagnostic process compared to three to eight factors in models proposed by others.[18–21]

Validation is an essential step in developing new tools, particularly if they involve complex mathematical functions such as those that the UNC OCT Index is based on.[26] It ensures a good quality and reliable end product. It generates outputs that are compared to those of the developer for accuracy of the tool as initially described. Most glaucoma diagnostic studies assessing the performance of combinatorial tools perform the training and testing set of the algorithm on the same groups of study subjects. This approach is prone to an overly optimistic assessment of performance since the tool is biased to the specific properties of each dataset (overfitting). Although resampling procedures such as cross validation help maximize use of the available data and control the bias that results from testing the algorithm on the same group of subjects from which it was trained, an external validation on a completely independent sample of individuals is a much better approach to estimate the diagnostic performance and the generalizability of the algorithm. Thus, an important difference between the UNC OCT Index and other similar models is that the former has undergone an internal validation at the time of development and an external validation that is reported here. External validation requires a completely independent entity to assess every aspect of the work of a given product. It is an effective way to identify and eventually correct issues early, during the development, rather than later, after product delivery. Independent validation of the UNC OCT Index involved the use of data from an independent source and application of the data to the model by another source, which finally delivered the output that is presented in this report. Our present study demonstrates that the UNC OCT Index generalized well and achieved similar levels of performance as reported from its training and interval validation.[26] Our model is flexible, and other parameters (from OCT or other imaging devices) or other clinical data could be incorporated without sacrificing either the accuracy or the speed of the automated computation. Glaucoma is a multifactorial disease, and factors such as intraocular pressure, demographics (age, race), and a family history of the disease all contribute to the risk of developing the disease.

While the mathematical basis of the UNC OCT Index may appear complex, this is not an impediment to using the index in the consulting room for real-time clinical decision making. Structural changes in the early stages of the disease are often so minimal that it is very challenging to detect them clinically and therefore to differentiate a glaucomatous from a normal eye. OCT may produce measurements that are within normal range in an eye that is glaucomatous, a situation that has been labeled "green disease."[31] Conversely, OCT may also produce values that fall outside normal limits in an eye that is definitely normal, a situation labeled "red disease."[32] The UNC OCT Index not only helps avoid complex mental arithmetic to interpret outputs from individual parameters, but it also minimizes misclassification rates.

Similar to the performance of OCT parameters in glaucoma that is dependent on the disease severity, clinical ascertainment of the presence of the disease is more challenging in early stages than in advanced. For this reason, this investigation was limited to patients with early glaucoma, despite the fact that this may be considered a study limitation because correlation between variables may be inflated, resulting in low factor loadings. It is unlikely that the model was affected since the use of EFA prevents this from happening by considering only shared variance, which ultimately allows avoiding the inflation of estimates of variance accounted for.[33] With regard to interpretation of the predicted probability, we acknowledge that it may alternatively be interpreted in a more continuous way. However, the optimal cut point (which is based on the ROC curve analysis) is given in order to provide clinicians with a method of making a binary diagnosis decision for a patient, but it is also highly recommended that the magnitude of the probability be investigated. For example, a probability of 0.99 provides much stronger evidence of early glaucoma than does a probability of 0.55, even though both would lead to the same diagnosis. We also acknowledge that establishing the diagnosis of glaucoma in the patients included in this study on the

basis of OCT findings is a limitation. However, the final classification of individuals by the UNC OCT Index may still be different from that of clinical categorization. Specifically, an individual classified clinically (including individual OCT data) as having early glaucoma may be categorized as normal through the index. The opposite is also true, meaning a patient classified clinically as not having early glaucoma may be categorized as having the disease by the model. Therefore, the combinatorial nature of the model makes some difference. A study is underway to evaluate the performance of the index in patients diagnosed as having early glaucoma based on clinical ground and visual field only.

In conclusion, glaucoma diagnosis relies on various sources of information, such as structural assessment of the ONH, RNFL and ganglion cell layer, and functional assessment. For structural assessment, the successful validation of the performance of the UNC OCT Index suggests that this model is a promising direction for achieving improved diagnostic information from the ONH, RNFL, and GCIPL and that it could become a useful decision support tool for more accurate diagnosis of early glaucoma.

Acknowledgments

Presented at the Annual Meeting of the Association for Research in Vision and Ophthalmology, May 6–11, 2017, Baltimore, MD, USA.

Supported by Research to Prevent Blindness, New York, New York, USA, and by CTSA Grant UL1 TR001863 from the National Center for Advancing Translational Science.

Disclosure: **J.-C. Mwanza**, P; **G. Lee**, E; **D.L. Budenz**, P; **J.L. Warren**, P; **M. Wall**, None; **P.H. Artes**, None; **T.M. Callan**, E; **J.G. Flanagan**; None

References

1. Lisboa R, Leite MT, Zangwill LM, Tafreshi A, Weinreb RN, Medeiros FA. Diagnosing preperimetric glaucoma with spectral domain optical coherence tomography. *Ophthalmology*. 2012; 119:2261–2269.
2. Mwanza JC, Budenz DL, Godfrey DG, et al. Diagnostic performance of optical coherence tomography ganglion cell–inner plexiform layer thickness measurements in early glaucoma. *Ophthalmology*. 2014;121:849–854.
3. 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.
4. Sung MS, Yoon JH, Park SW. Diagnostic validity of macular ganglion cell-inner plexiform layer thickness deviation map algorithm using cirrus HD-OCT in preperimetric and early glaucoma. *J Glaucoma*. 2014;23:e144–151.
5. Bengtsson B, Andersson S, Heijl A. Performance of time-domain and spectral-domain optical coherence tomography for glaucoma screening. *Acta Ophthalmol*. 2012;90:310–315.
6. Li G, Fansi AK, Boivin JF, Joseph L, Harasymowycz P. Screening for glaucoma in high-risk populations using optical coherence tomography. *Ophthalmology*. 2010;117:453–461.
7. Nakano T, Hayashi T, Nakagawa T, et al. Applicability of automatic spectral domain optical coherence tomography for glaucoma mass screening. *Clin Ophthalmol*. 2017;11:97–103.
8. Yu M, Lin C, Weinreb RN, Lai G, Chiu V, Leung CK. Risk of visual field progression in glaucoma patients with progressive retinal nerve fiber layer thinning: a 5-year prospective study. *Ophthalmology*. 2016;123:1201–1210.
9. Zhang X, Loewen N, Tan O, et al. Predicting development of glaucomatous visual field conversion using baseline Fourier-domain optical coherence tomography. *Am J Ophthalmol*. 2016; 163:29–37.
10. Kim KE, Jeoung JW, Park KH, Kim DM, Kim SH. Diagnostic classification of macular ganglion cell and retinal nerve fiber layer analysis: differentiation of false-positives from glaucoma. *Ophthalmology*. 2015;122:502–510.
11. Leal-Fonseca M, Rebolleda G, Oblanca N, Moreno-Montanes J, Munoz-Negrete FJ. A comparison of false positives in retinal nerve fiber layer, optic nerve head and macular ganglion cell-inner plexiform layer from two spectral-domain optical coherence tomography devices. *Graefes Arch Clin Exp Ophthalmol*. 2014; 252:321–330.
12. Hwang YH, Kim YY, Kim HK, Sohn YH. Ability of cirrus high-definition spectral-domain optical coherence tomography clock-hour, deviation, and thickness maps in detecting photo-

graphic retinal nerve fiber layer abnormalities. *Ophthalmology*. 2013;120:1380–1387.
13. Hwang YH, Jeong YC, Kim HK, Sohn YH. Macular ganglion cell analysis for early detection of glaucoma. *Ophthalmology*. 2014;121:1508–1515.
14. Barella KA, Costa VP, Goncalves Vidotti V, Silva FR, Dias M, Gomi ES. Glaucoma diagnostic accuracy of machine learning classifiers using retinal nerve fiber layer and optic nerve data from SD-OCT. *J Ophthalmol*. 2013;2013:789129.
15. 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.
16. Burgansky-Eliash Z, Wollstein G, Chu T, et al. Optical coherence tomography machine learning classifiers for glaucoma detection: a preliminary study. *Invest Ophthalmol Vis Sci*. 2005;46:4147–4152.
17. Chen HY, Huang ML, Hung PT. Logistic regression analysis for glaucoma diagnosis using Stratus optical coherence tomography. *Optom Vis Sci*. 2006;83:527–534.
18. 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.
19. 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.
20. Larrosa JM, Moreno-Montanes J, Martinez-de-la-Casa JM, et al. A diagnostic calculator for detecting glaucoma on the basis of retinal nerve fiber layer, optic disc, and retinal ganglion cell analysis by optical coherence tomography. *Invest Ophthalmol Vis Sci*. 2015;56:6788–6795.
21. Loewen NA, Zhang X, Tan O, et al. Combining measurements from three anatomical areas for glaucoma diagnosis using Fourier-domain optical coherence tomography. *Br J Ophthalmol*. 2015;99:1224–1229.
22. Pablo LE, Ferreras A, Pajarin AB, Fogagnolo P. Diagnostic ability of a linear discriminant function for optic nerve head parameters measured with optical coherence tomography for perimetric glaucoma. *Eye (Lond)*. 2010;24:1051–1057.
23. Sugimoto K, Murata H, Hirasawa H, Aihara M, Mayama C, Asaoka R. Cross-sectional study: does combining optical coherence tomography measurements using the "Random Forest" decision tree classifier improve the prediction of the presence of perimetric deterioration in glaucoma suspects? *BMJ Open*. 2013;3:e003114.
24. 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. *J Glaucoma*. 2014;23:129–135.
25. Yoshida T, Iwase A, Hirasawa H, et al. Discriminating between glaucoma and normal eyes using optical coherence tomography and the "Random Forests" classifier. *PLoS One*. 2014;9:e106117.
26. Mwanza JC, Warren JL, Budenz DL. 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.
27. 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.
28. Mwanza JC, Budenz DL, Warren JL, inventors; The University of North Carolina at Chapel Hill, assignee. Methods, systems, and computer readable media for predicting early onset glaucoma. US patent 9554755 B2, January 31, 2017.
29. Kumar R, Indrayan A. Receiver operating characteristic (ROC) curve for medical researchers. *Indian Pediatr*. 2011;48:277–287.
30. Brusini P, Filacorda S. Enhanced Glaucoma Staging System (GSS 2) for classifying functional damage in glaucoma. *J Glaucoma*. 2006;15:40–46.
31. Sayed MS, Margolis M, Lee RK. Green disease in optical coherence tomography diagnosis of glaucoma. *Curr Opin Ophthalmol*. 2017;28:139–153.
32. Chong GT, Lee RK. Glaucoma versus red disease: imaging and glaucoma diagnosis. *Curr Opin Ophthalmol*. 2012;23:79–88.
33. Gorsuch RL. Exploratory factor analysis: its role in item analysis. *J Pers Assess*. 1997;68:532–560.