

Central Macular Topographic and Volumetric Measures: New Biomarkers for Detection of Glaucoma

Vahid Mohammadzadeh¹, Melodyanne Cheng¹, Sepideh Heydar Zadeh¹, Kiumars Edalati¹, Dariush Yalzadeh¹, Joseph Caprioli¹, Sunil Yadav^{2,3}, Ella M. Kadas^{2,3}, Alexander U. Brandt⁴, and Kouros Nouri-Mahdavi¹

¹ Glaucoma Division, Stein Eye Institute, David Geffen School of Medicine, University of California Los Angeles, Los Angeles, CA, USA

² Experimental and Clinical Research Center, Max Delbrück Center for Molecular Medicine, Berlin, Germany

³ Charité – Universitätsmedizin Berlin, corporate member of Freie Universität Berlin, Humboldt-Universität zu Berlin, and Berlin Institute of Health, Berlin, Germany

⁴ Department of Neurology, University of California Irvine, CA, USA

Correspondence: Kouros Nouri-Mahdavi, 100 Stein Plaza, Los Angeles, CA 90095, USA. e-mail: nouri-mahdavi@jsei.ucla.edu

Received: January 28, 2022

Accepted: June 27, 2022

Published: July 29, 2022

Keywords: optical coherence tomography (OCT); macula; foveal shape; topography; volume; deep learning; gradient boost model

Citation: Mohammadzadeh V, Cheng M, Zadeh SH, Edalati K, Yalzadeh D, Caprioli J, Yadav S, Kadas EM, Brandt AU, Nouri-Mahdavi K. Central macular topographic and volumetric measures: New biomarkers for detection of glaucoma. *Transl Vis Sci Technol.* 2022;11(7):25. <https://doi.org/10.1167/tvst.11.7.25>

Purpose: To test the hypothesis that newly developed shape measures using optical coherence tomography (OCT) macular volume scans can discriminate patients with perimetric glaucoma from healthy subjects.

Methods: OCT structural measures defining macular topography and volume were recently developed based on cubic Bézier curves. We exported macular volume scans from 135 eyes with glaucoma (133 patients) and 155 healthy eyes (85 subjects) and estimated global and quadrant-based measures. The best subset of measures to predict glaucoma was explored with a gradient boost model (GBM) with subsequent logistic regression. Accuracy and area under receiver operating curves (AUC) were the primary metrics. In addition, we separately investigated model performance in 66 eyes with mild glaucoma (mean deviation ≥ -6 dB).

Results: Average (\pm SD) 24-2 mean deviation was $-8.2 (\pm 6.1)$ dB in eyes with glaucoma. The main predictive measures for glaucoma were temporal inferior rim height, nasal inferior pit volume, and temporal inferior pit depth. Lower values for these measures predicted higher risk of glaucoma. Sensitivity, specificity, and AUC for discriminating between healthy and glaucoma eyes were 81.5% (95% CI = 76.6–91.9%), 89.7% (95% CI = 78.7–94.2%), and 0.915 (95% CI = 0.882–0.948), respectively. Corresponding metrics for mild glaucoma were 84.8% (95% CI = 72.1%–95.5%), 85.8% (95% CI = 87.1%–97.4%), and 0.913 (95% CI = 0.867–0.958), respectively.

Conclusions: Novel macular shape biomarkers detect early glaucoma with clinically relevant performance. Such biomarkers do not depend on intraretinal segmentation accuracy and may be helpful in eyes with suboptimal macular segmentation.

Translational Relevance: Macular shape biomarkers provide valuable information for detection of early glaucoma and may provide additional information beyond thickness measurements.

Introduction

Glaucoma is characterized by the loss of retinal ganglion cells (RGCs) and may result in functional loss if left untreated.^{1,2} Timely treatment of glaucoma is of utmost importance as glaucoma deterioration is associated with worsening of vision-specific quality of

life.^{3,4} The RGCs most crucial to central vision are located in the macular region. Macular optical coherence tomography (OCT) is commonly used to assess the health of the RGC-axonal complex.^{5,6} Although current OCT imaging approaches have significantly enhanced detection of early signs of glaucoma, identifying eyes with glaucoma at highest risk of visual disability with high accuracy is a pressing unmet need

in the field of glaucoma diagnostics. Automated detection of eyes with a high-risk of early glaucoma with high specificity could make screening or risk stratification possible. This is especially important in earlier stages of the disease so that further functional damage may be averted.⁷

Macular thickness measurements have been the main outcome of interest in both research and clinical setting with thinning of various layers of the inner retina considered the hallmark of glaucoma.^{8–13} The central macula has a complex configuration and variations in its thickness and topography have been reported in eyes with glaucoma.^{14–18} Recent studies reported on the foveal shape in healthy subjects and subjects with glaucoma and explored its influence on structure–function (SF) relationships.^{14,19–21} Yadav et al. and Motamedi and collaborators recently proposed a novel approach to parametrizing the configuration and topography of the central macular and perifoveal region based on cubic Bézier polynomials along with a least square optimization for 3-dimensional modeling.^{22,23} Motamedi et al. evaluated the performance of such measures in differentiating between neuromyelitis optica (NMO) and multiple sclerosis (MS); the resulting area under receiver operating characteristic curve (AUROC) for discriminating between the 2 diseases was 0.800.²³ Yadav and colleagues also reported the differences in the aforementioned shape and topography measures between healthy subjects and patients with various inflammatory optic neuropathies.²²

The goal of the current study is to test the hypothesis that newly developed macular topographic and volumetric measures (called macular shape measures in this paper) can discriminate eyes with glaucoma from healthy eyes with clinically relevant accuracy. We speculate that these novel macular biomarkers could enhance automated detection of early central damage in patients with glaucoma above and beyond macular thickness measures.

Methods

Study Sample

One hundred thirty-five eyes from 133 patients with glaucoma and 155 eyes of 85 healthy subjects were enrolled in this study. In this dataset, 66 eyes of 66 patients were categorized as having mild perimetric damage defined as visual field mean deviation (MD) ≥ -6 dB. The current study was carried out in accordance with the tenets of the Declaration of Helsinki and the Health Insurance Portability and Accountability Act (HIPAA) and was approved by the University

of California Los Angeles (UCLA) Human Research Protection Program.

Inclusion criteria for eyes with glaucoma were: (1) clinical diagnosis of primary open-angle glaucoma, pseudoexfoliative glaucoma, pigmentary glaucoma, or primary angle-closure glaucoma; (2) age between 40 and 80 years; (3) best corrected visual acuity $\geq 20/50$; and (4) no significant confounding retinal or neurological disease. All study eyes underwent a complete eye examination and macular imaging with Spectralis SD-OCT (Heidelberg Engineering, Heidelberg, Germany). A visual field (VF) defect was considered to be present if both of the following criteria were met: (1) Glaucoma Hemifield Test outside the normal limits or pattern standard deviation with $P < 0.05$; and (2) 4 abnormal points with $P < 0.05$ on the pattern deviation plot, both confirmed at least once.²⁴ VF examinations with false positives rates $>15\%$ were excluded from this study. Eligible healthy eyes were required to have normal VFs and normal optic disc photographs based on the review of two individual glaucoma specialists.

Macular OCT Imaging

We used the Posterior Pole Algorithm of the Spectralis SD-OCT for this study. This algorithm acquires 61 horizontal B-scans spanning a 30 degrees \times 25 degrees wide area with activated eye tracker, parallel to the fovea-Bruch's membrane opening axis; each B-scan consists of 768 A-scans. The B-scans are repeated 9 to 11 times (automatic real-time [ART] averaging) to decrease speckle noise and improve image quality. Scans with a quality factor greater than 15 were considered for the analysis. The quality factor is a quality index provided by the proprietary software of Spectralis OCT and ranges from 0 (worst quality) to mid-40s (best quality) under clinical scenarios. We exported .vol files to a personal computer for further analysis by the morphometry algorithm as described below.

All analyses were carried out with a software suite developed by the co-authors (Nocturne GmbH, Berlin, Germany). Nocturne software is a CE certified software (the product meets high safety, health and environmental protection requirements to be sold in the European Economic Area) as a medical device (DIN EN ISO 13845:2016, class IM measurement software). The local installable web application analyzes OCT images of the retina to measure thinning, swelling, and deformation of tissues. After an OCT volume scan is acquired, the user can easily obtain a report in a few minutes by opening the application on a local computer, uploading the scan, and triggering the required analysis. This deep learning-based software

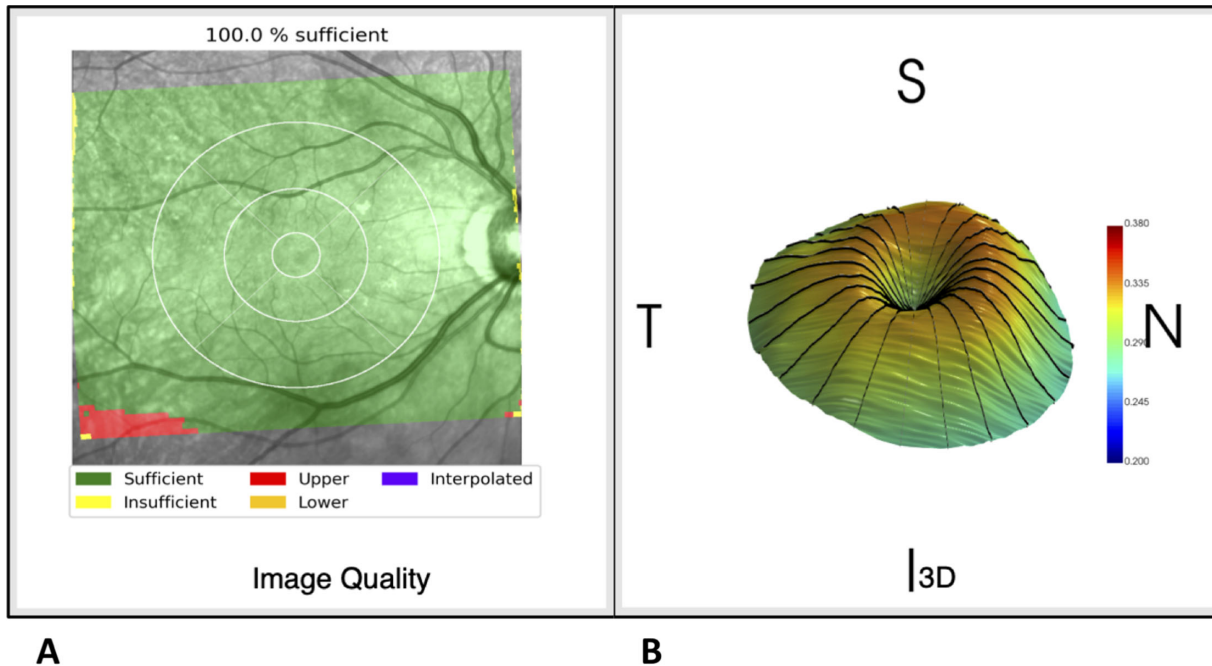


Figure 1. (A) The output from the AQUA algorithm provides information on image quality and segmentation adequacy. AQUA uses a deep learning approach to assess volume scan quality. AQUA accesses each A-scan in the input volume and projects the corresponding quality measure on the SLO image in different colors; for example, sufficient and insufficient signal quality are shown in *green* and *yellow*, respectively. Furthermore, upper and lower image cutoffs are represented in *red* and *orange* colors, respectively. The interpolation feature will generate artificial B-scans if one to three nonconsecutive B-scans have low signal quality. (B) The FOVEAM algorithm utilizes macular OCT volume scans to create a model-driven 3D reconstruction of the central 4 mm of the macula and estimates various shape and volume measures (see text). T, S, N, and I represent the temporal, superior, nasal, and inferior macula, respectively.

suite consists of a quality check algorithm (Automatic Quality Analysis [AQUA]) and a foveal morphometry algorithm (FOVEAM).^{22,25}

Quality Check With the AQUA Algorithm

The OCT quality check and morphometry analyses were carried out with the Nocturne software developed by the co-authors. Only images with a quality factor of >15 were included. In addition, OCT image quality was checked with a customized automated algorithm, AQUA, which is part of the Nocturne software suite and is based on the deep learning modeling.²⁵ This model ensures that OCT volume scan data are of sufficient quality for subsequent model-driven 3D reconstruction of the fovea. The AQUA algorithm provides a fully standardized, grader-independent, and easily applicable real-time method for evaluation of macular OCT image quality (Fig. 1).²⁶ AQUA has been validated against reference data labeled by expert reviewers at Berlin's Charité University Hospital and against the OSCAR-IB quality assessment algorithm, a reference standard for quality check of macular OCT images²⁷; AQUA was able to discriminate between good and poor-quality scans (volume, radial, and

ring patterns) with 97% accuracy. The AQUA method combines three main OSCAR-IB criteria: (1) foveal center detection, (2) signal quality (signal to noise ratio), and (3) image completeness artifacts (scan cuts). These features are evaluated before the foveal morphometry, and only if the input volume is free from missing regions, is well-centered on the fovea, and the signal quality is acceptable, then the morphometric analysis is performed. AQUA predicts the quality of OCT volume scans based on a review of A-scans and if 98% or more of the A-scans have an image quality above the threshold (learned by DCNN), then the input volume will be included for further processing. AQUA uses parabola fitting in the foveal region for accurate center detection. AQUA not only checks image quality but also evaluates the quality of segmented data. In our study, we used the segmentation of the upper (inner limiting membrane [ILM]) and lower boundaries (Bruch's membrane) of the retina as provided in the Spectralis' .vol files; the segmentation of these two boundaries can be more easily done and is of higher quality compared to the inner layers when the quality of OCT images is subpar. Before performing the foveal morphometry,²² we checked the accuracy of these two segmentation lines to avoid including

any data with inappropriate segmentation. Similar to the image quality assessment, AQUA ensures that the segmentation lines are free from missing regions, high jumps, and that the internal limiting membrane and Bruch’s membrane are not intersecting each other.

Macular Shape Analysis

Two-dimensional (2D) and three-dimensional (3D) data from the macular OCT volume scans can assess structural changes of the central macula including the perifoveal area that may correlate with presence of glaucoma or disease progression. Yadav et al. first proposed a method for parametric modeling of the central macular topography and volume from OCT data, which uses a cubic Bézier polynomial and least square optimization alongside landmark features of the macula to create a best-fit parametric model of the foveal topography within the central 4 mm of the macula.²² Specifically, the proposed model uses a least square optimization to fit the cubic Bézier polynomial to interior and exterior regions of the fovea based off stable and reliable features of the foveal pit shape, including the lowest point of the pit and the highest point around the pit region. To generate a high-fidelity parameterization of the fovea, the interior and exterior regions are parametrized using different constraints based on the geometric properties of the macula. The foveal shape has two different intrinsic properties: maximum peaks at rim points and minimum depth at the center of the fovea. We utilized these two properties to force different constraints in different regions. The whole region was split between interior (between maximum and minimum points) and exterior (beyond the maximum points) areas. For parameterization of the interior region, the main constraint is that the end points must have horizontal tangent lines (because of the extreme points). For the exterior region, we used

the same constraints for one side (maxima points) but not for the other side, which is defined by the rim points. Because of the intrinsic shape differences, the constraints are different.

Using the above model-driven foveal shape analysis method developed by Yadav et al., we acquired thickness, volume, and fitting measures to accurately model and reconstruct the 3D profile of the perifoveal region.²² The 3D reconstruction exemplified in Figure 1B represents the macula of a patient’s right eye.

Foveal Morphometry With the FOVEAM Algorithm

The FOVEAM (fovea morphometry) algorithm calculates topographic and volumetric (shape) biomarkers based on a central 4 mm macular area (see Fig. 1B). The FOVEAM algorithm calculates several shape measures from raw macular volume scans (.vol files) as follows (Figs. 2, 3). The macular volume scans sample the central macula using a stack of linear B-scans aligned with the fovea-Bruch’s membrane opening axis. Therefore, radial resampling of the volume scan is needed. The number of radial resampling curves around the fovea is an important variable as too sparse sampling could lead to loss of data and a too dense sampling scheme can result in longer processing time for data extrapolations from the macular volume scans. Yadav et al. estimated the fidelity of the resampling approach with the L2-metric, which assesses the closeness between two geometrical shapes.^{28,29} They found that the resampling process using 12 or more radial lines on the Spectralis volume scans provided adequate resolution when compared to a ground truth of 48 radial scans (average error = 0.2 μm).

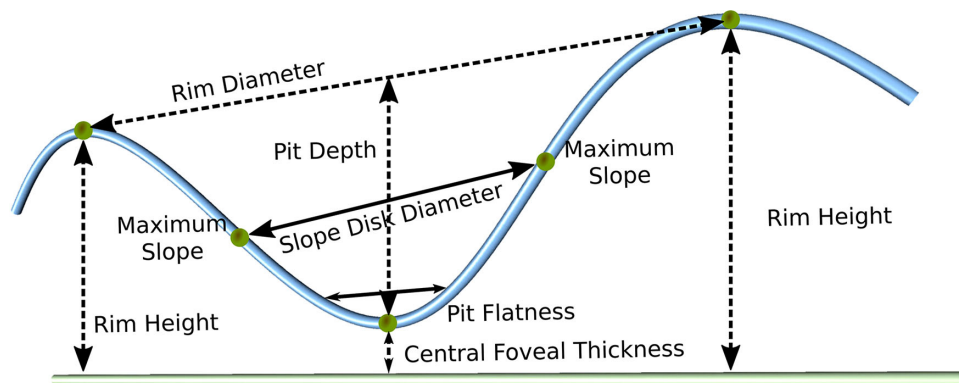


Figure 2. Visualization of 2D measures on the central B-scan passing through the foveal pit.

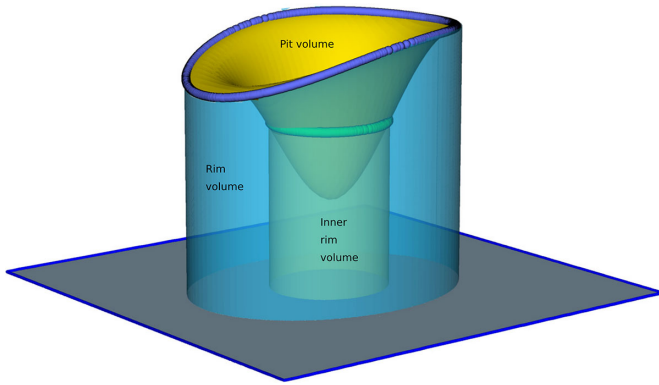


Figure 3. Volumetric measures of interest consist of the rim volume, pit volume, and inner rim volume. The rim volume (represented by the large blue cylinder) describes the total volume of the perifoveal retina delimited internally by the rim disk and externally by the Bruch’s membrane–RPE complex, whereas the inner rim volume (the green cylinder) is defined as the volume of the retina between the Bruch’s membrane–RPE complex and the internal limiting membrane within a 0.5 mm radius of the foveal center. The pit volume is the total volume between the rim disk, the highest points on the inner surface of the macula, and the inner surface of the foveal pit.

Measures of Interest

Table 1 summarizes the topographic and volumetric parameters derived from the FOVEAM algorithm. The current version of the algorithm is able to calculate both global measures and sectoral measures within

four quadrants around the fovea (temporal superior [TS], temporal inferior [TI], nasal superior [NS], and nasal inferior [NI]). Supplementary Figure S1 provides the algorithm pipeline carried out to estimate the shape biomarkers from raw 3D OCT images.

Statistical Analyses

We first explored the performance of each global and sectoral measure one at a time (univariable analyses) for glaucoma risk. Both linear logistic models and restricted cubic spline (RCS) logistic models were fit for each individual measure. Because the shape measures are correlated, we used a Gradient Boost logistic model (GBM) using all potential predictors simultaneously to evaluate the relative influence of each potential predictor and determine the best subset of the predictors to retain. The subset of predictors that accounted for most of the relative influence was retained. A conventional logistic model allowing two-way interactions was subsequently fit using the variables selected by GBM. The model combines the variables into a logit score through the following formula (showing only 3 potential predictors):

$$\text{Logit score} = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_1 x_2 + \beta_5 x_1 x_3$$

where x_1 , x_2 , and x_3 are the selected variables and β_1 , β_2 , and β_3 are their corresponding weights. In

Table 1. Description of Central Macular Topographic and Volumetric Measures (Also See Figs. 2 and 3).

Rim disk area:	Area represented by a spline fit on the maximum perifoveal points on each radial cut (reconstructed radial foveal cross-section) centered on the foveal center.
Rim disk diameter:	Longest distance between opposite maximum perifoveal rim points among radial B-scans.
Rim height:	Average of maximum perifoveal rim height measurements on 12 radial B-scans.
Average pit depth:	Distance between the center of the foveal pit and the average rim height.
Central foveal thickness:	Minimum distance between foveal pit’s center and the underlying retinal pigment epithelium (RPE)-Bruch’s membrane complex.
Slope disk area:	The area derived from a spline fit of all maximum slope points on each reconstructed radial foveal cross-section.
Slope disk diameter:	Longest distance among all the corresponding pairs of maximum slope points on radial cross-sections.
Pit flat disk area:	Quantifies the flatness of the foveal pit surrounding the foveal center.
Pit flat disk diameter:	Distance between two opposite points representing the widest area of the flat portion of foveal pit.
Volume Parameters	
Rim volume:	Total volume of the perifoveal rim delimited internally by the rim disk and externally the BM-RPE complex (large blue cylinder, Fig. 3).
Inner rim volume:	Volume of the fovea between the region of inner limiting membrane within 0.5mm of the foveal center and the BM-RPE complex (green cylinder, Fig. 3).
Pit volume:	Total volume between the rim disk and the inner surface of the foveal pit (yellow valley, Fig. 3).

the equation above, the β_1 and β_2 are the weights of interactions between variables 1 and 2 and variables 1 and 3, respectively. The odds and risk ($P = \text{proportion}$) of glaucoma can be computed from the logit score as: where $\text{odds} = O = \exp(\text{logit score})$ and $\text{risk} = P = O/(1+O)$.

We also included a random subject effect to allow for non-independence between the two eyes of subjects. The Akaike information criterion (AIC) was used to determine which interaction terms were retained in the final logistic model. We report primary outcome measures from the logistic model, including sensitivity, specificity, and AUROC with their corresponding 95% confidence interval (CI) after 5-fold cross-validations. We repeated the same approach to determine the performance of the shape measures for discriminating between mild glaucoma and healthy eyes.

We also investigated the performance of various machine learning (ML) models for discriminating glaucoma eyes from healthy eyes for both the entire dataset and for eyes with early glaucoma. We used a GBM, Naïve Bayes (NB), and a support vector machine (SVM) algorithm for this purpose. We fed the 40 topographic and volumetric parameters of interest to the ML models and carried out parameter tuning with Grid Search Method; a 5-fold cross-validation was done for each ML model. We report the AUROC, sensi-

tivity, specificity, and accuracy along with the corresponding 95% CI for each model.

All the analyses were performed with the software R version 4.0.4 (R Foundation for Statistical Computing, Vienna, Austria).

Results

Table 2 displays the clinical and demographic data for the study eyes. The mean (SD) age was 52.6 (13.6) and 66.7 (9.6) years ($P < 0.001$) and the average (SD) 24-2 visual field MD was -0.73 (1.77) and -8.24 (6.11) dB in the glaucoma and healthy groups, respectively ($P < 0.001$). There were no significant differences between the two groups with regard to gender and ethnicity.

Table 3 describes the performance of each of the global shape measures for detection of glaucoma in the entire patient sample based on linear and RCS. The best performing global measure was the global rim height (AUC = 0.814). The global rim height (AUC = 0.818) was also found to be the best-performing measure based on the RCS logistic model for the entire patient group.

Overall, 40 global and sectoral variables were used in the GBM analysis. Based on the GBM analysis, the

Table 2. Demographic Characteristics of the Healthy Subjects and the Subjects With Glaucoma

	Healthy (N = 155)	Glaucoma (N = 135)	P Value
Age			
Mean (SD)	52.6 (13.6)	66.7 (9.6)	< 0.001
Gender (%)			
Female	55%	60%	0.466
Male	45%	40%	
Race (%)			
Caucasian	32%	54%	0.988
African American	15%	15%	0.988
Hispanic	15%	13%	0.987
Asian	38%	18%	0.988
24–2 MD (dB)			
Mean (SD)	-0.73 (1.77)	-8.24 (6.11)	< 0.001
Axial length (mm)			
Mean (SD)	23.8 (1.2)	24.6 (1.4)	< 0.001
Global rim height (mm)			
Mean (SD)	0.34 (0.01)	0.31 (0.02)	< 0.001
Global pit volume (mm³)			
Mean (SD)	0.25 (0.04)	0.21(0.04)	< 0.001
Global pit depth (mm)			
Mean (SD)	0.11 (0.02)	0.09 (0.02)	< 0.001

Table 3. Performance of Global Topographic and Volumetric Parameters for Detection of Glaucoma Based on Univariable Linear and Restricted Cubic Spline Logistic Models for the Entire Dataset

Global Predictor	Linear Logistic Model				Restricted Cubic Spline Logistic Model			
	Specificity	Sensitivity	Accuracy	AUC	Specificity	Sensitivity	Accuracy	AUC
Rim height	82.1%	68.9%	75.5%	0.814	82.1%	68.9%	75.5%	0.818
Pit depth	62.8%	82.2%	72.5%	0.769	62.8%	82.2%	72.5%	0.769
Rim volume	66.7%	77.0%	71.9%	0.764	66.7%	77.0%	71.9%	0.758
Pit volume	67.9%	74.1%	71.0%	0.739	67.9%	74.1%	71.0%	0.739

AUC, area under the curve.

Table 4. The Coefficients, Standard Errors, and *P* Values for the Predictors in the Final Optimal Logistic Model for Prediction of Glaucoma Based on Topographic and Volumetric Markers for the Entire Glaucoma Group and Separately for Eyes With Mild Glaucoma

Predictor	Entire Glaucoma Eyes			Mild Glaucoma Eyes		
	Estimate	SE	<i>P</i> Value	Estimate	SE	<i>P</i> Value
TI rim height (mm)	84.6	58.8	0.1500	60.4	57.4	0.2927
NI pit volume (mm³)	482.5	247.3	0.0511	334.4	257.0	0.1933
TI pit depth (mm)	200.8	124.2	0.1059	186.3	114.3	0.1030
TI rim height × NI pit volume	-1710.3	776.0	0.0275	-1219.6	807.5	0.1310
TI rim height × TI pit depth	-703.4	389.4	0.0709	-663.0	363.0	0.0678

TI, temporal inferior; NI, nasal inferior; SE, standard error.

• Logit score for entire dataset = $-20.9 + 84.6$ (TI rim height) + 482.5 (NI pit volume) + 200.8 (TI pit depth) - 1710.3 (TI rim height × NI pit volume) - 703.4 (TI rim height × TI pit depth).

• Logit score for eyes with mild glaucoma = $-14.1 + 60.4$ (TI rim height) + 334.4 (NI pit volume) + 186.3 (TI pit depth) - 1219.6 (TI rim height × NI pit volume) - 663.0 (TI rim height × TI pit depth).

3 top variables according to their relative influence were the TI rim height, NI pit volume, and TI pit depth with the 3 of them accounting for 53% of the total influence of all the shape measures (i.e. all other predictors added less than 47%). The final optimal logistic model included the above 3 measures and 2-way interactions between TI rim height and NI pit volume, and TI rim height and TI pit depth. Table 4 provides the estimates, standard errors, and *P* values for the predictors in the final logistic model. The sensitivity, specificity, accuracy, and AUC (95% CI) of the final logistic model for discriminating between healthy eyes and eyes with glaucoma were 81.5% (95% CI = 76.6%, 91.9%), 89.7% (95% CI = 78.7%, 94.2%), 85.6% (95% CI = 84.1%, 87.1%), and 0.915 (95% CI = 0.882, 0.948), respectively. Figure 4 shows the glaucoma risk for a range of values of each predictor with the other 2 variables fixed at their first quartile, mean, and third quartile values. Lower values for TI rim height, NI pit volume, and TI pit depth were associated with a higher risk of glaucoma.

We repeated all the above steps to build a new logistic model based on the healthy eyes and the eyes

with early perimetric glaucoma including only eyes with visual field MD ≥ -6 dB. The GBM selected TS rim height, NI pit volume, TI pit depth, and TS pit depth as the top variables with overall resistive index (RI) of 50%. Table 5 displays the estimates, standard errors, and *P* values for this logistic model using only eyes with early perimetric glaucoma. The logistic model included the interactions between TS pit depth and TS rim height and between TI pit depth and TS rim height. The sensitivity, specificity, accuracy, and AUC (95% CI) were 84.8% (95% CI = 72.1%, 95.5%), 85.8% (95% CI = 87.1%, 97.4%), 85.3% and 0.913 (95% CI = 0.867, 0.958), respectively. When we applied the logistic model derived from the entire glaucoma group, which included TI rim height, NI pit volume, and TI pit depth, and the relevant interactions (see Table 4), to eyes with mild glaucoma and healthy eyes only, the sensitivity, specificity, accuracy, and AUC (95% CI) were 87.9% (95% CI = 68.2%, 95.5%), 76.8% (95% CI = 71.6%, 96.8%), 82.3%, and 0.896 (95% CI = 0.850, 0.942) respectively. Figure 5 represents the receiver operating characteristic (ROC) curves for models used for discriminating normal eyes

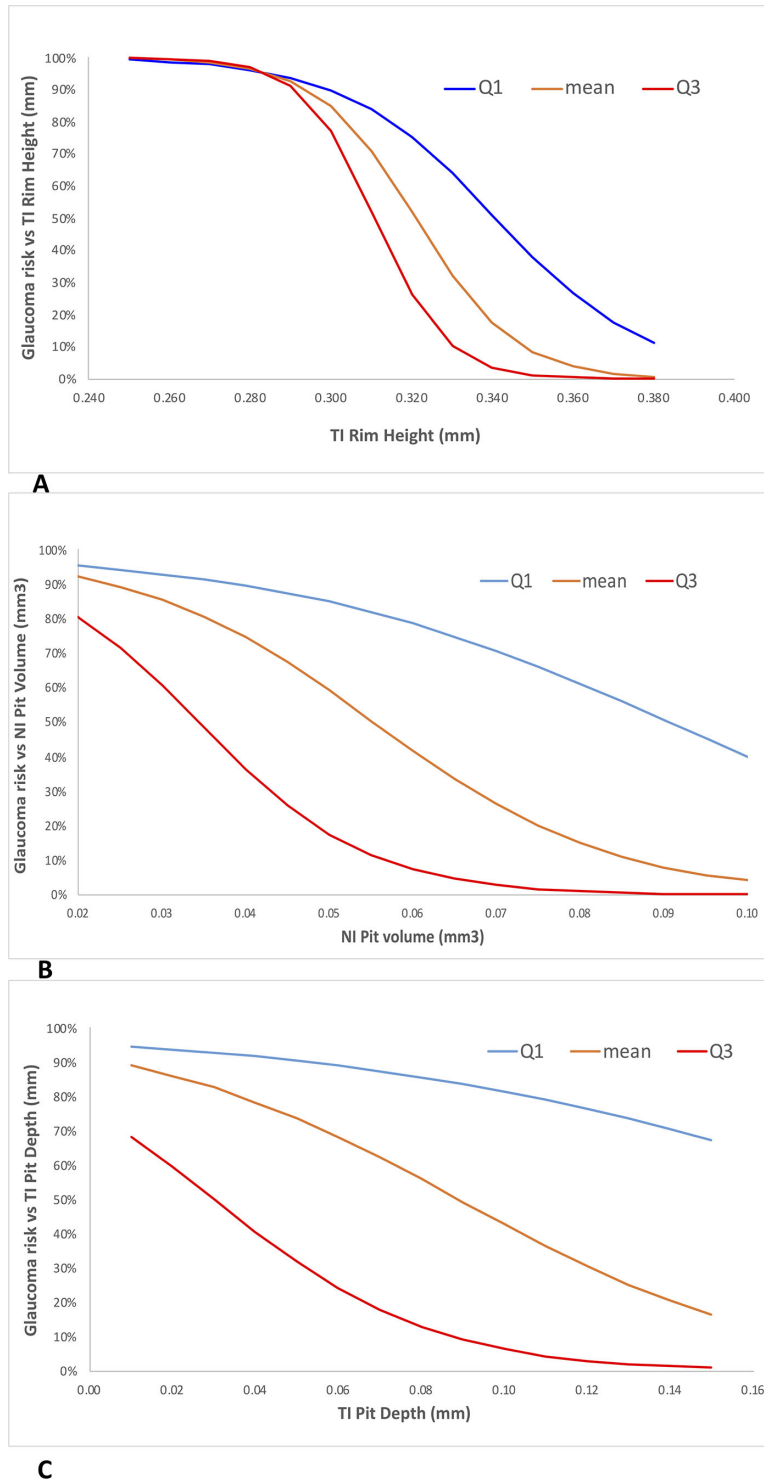


Figure 4. Effect of each of the three shape predictors identified by the gradient boost model (temporal inferior [TI] rim height, nasal inferior [NI] pit volume, and temporal inferior pit depth) on the glaucoma risk, whereas the other two predictors are fixed at their first quartile (Q1), mean, and third quartile (Q3).

from glaucoma eyes and from early stage glaucoma eyes.

The performance of the 3 ML models (GBM, NB, and SVM) are provided in Table 6. For the

entire dataset, GBM demonstrated the highest performance for discriminating eyes with glaucoma from healthy eyes. The AUCs (95% CI) for the classification task were 0.947 (95% CI = 0.921, 0.973), 0.838

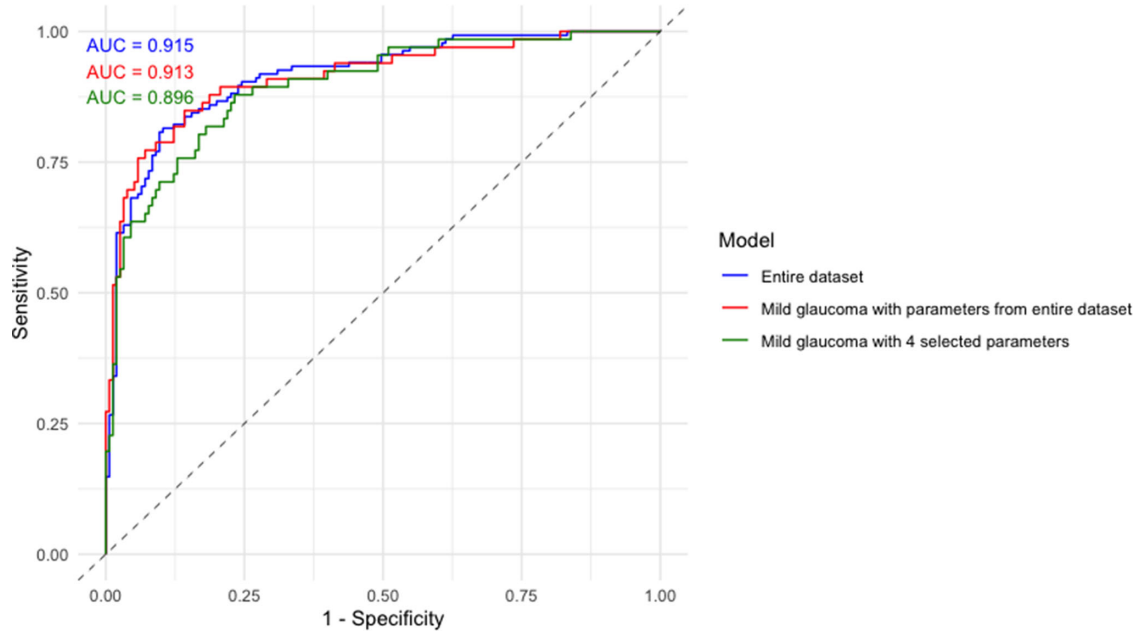


Figure 5. Receiver operating characteristic (ROC) curves for glaucoma prediction models for the entire dataset (red curve, AUC = 0.915), prediction of mild glaucoma using four selected parameters based on repeat analyses (green curve, AUC = 0.913) and based on parameters derived from the entire dataset (blue curve, AUC = 0.896).

Table 5. The Coefficients, Standard Errors, and P Values for the Predictors in the Final Optimal Logistic Model for Prediction of Mild Perimetric Glaucoma (Visual Field Mean Deviation ≥ -6 dB) Based on the Top Four Topographic and Volumetric Markers Selected by the Gradient Boost Model

Predictor	Estimate	SE	P Value
(Intercept)	9.77	10.73	0.3622
TS rim height (mm)	-13.50	34.48	0.6953
NI pit volume (mm ³)	-70.21	17.85	0.0001
TI pit depth (mm)	1356.51	427.07	0.0015
TS pit depth (mm)	-1139.56	402.76	0.0047
TS pit depth \times TS rim height	3758.79	1258.32	0.0028
TI pit depth \times TS rim height	-4502.45	1333.19	0.0007

TI, temporal inferior; NI, nasal inferior; TS, temporal superior; SE, standard error.

Logit score = 9.77 -13.50 (TS rim height) -70.21 (NI pit volume) + 1356.51 (TI pit depth) - 1139.56 (TS pit depth) + 3758.79 (TS pit depth \times TS rim height) -4502.45 (TI pit depth \times TS rim height).

Table 6. Results of Three Machine Learning Approaches for Discriminating all Eyes With Glaucoma and Eyes With Early Glaucoma From Healthy Subjects

	AUC (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)	Accuracy (95% CI)
Entire dataset				
GBM	0.947 (0.921, 0.973)	89.6% (81.4%, 96.3%)	91.0% (82.0%, 97.7%)	90.0% (86.2%, 93.1%)
Naïve Bayes	0.838 (0.782, 0.895)	80.3% (64.4%, 97.7%)	78.0% (57.7%, 90.8%)	78.3% (67.4%, 85.2%)
SVM	0.845 (0.800, 0.890)	83.6% (74.0%, 90.3%)	76.1% (66.4%, 84.5%)	79.3% (74.4%, 83.4%)
Mild glaucoma eyes				
GBM	0.950 (0.918, 0.983)	86.3% (77.2%, 97.7%)	96.1% (81.5%, 99.3%)	92.7% (85.0%, 96.3%)
Naïve Bayes	0.838 (0.782, 0.895)	80.3% (64.4%, 97.7%)	77.4% (57.7%, 90.8%)	78.3% (67.7%, 85.2%)
SVM	0.845 (0.800, 0.890)	83.6% (74.0%, 90.3%)	76.1% (65.1%, 84.5%)	79.3% (74.4%, 83.4%)

GBM, gradient boost model; SVM, support vector machine; CI, confidence interval.

(95% CI = 0.782, 0.895), and 0.845 (95% CI = 0.800, 0.890) for GBM, NB, and SVM, respectively. Similarly, for discriminating eyes with mild glaucoma from the healthy eyes, GBM showed the highest performance among all the ML models, with AUCs of 0.950 (95% CI = 0.918, 0.983), 0.838 (95% CI = 0.782, 0.895), and 0.845 (95% CI = 0.800, 0.890) for GBM, NB, and SVM, respectively.

Discussion

The goal of this study was to assess whether the newly developed macular shape measures, capturing both topographical and volumetric features of the central perifoveal macular region, can discriminate eyes with perimetric glaucoma from healthy eyes. Our results demonstrate that these novel macular biomarkers can detect perimetric glaucoma with clinically relevant performance. Quadrant-based measures performed overall better than global measures. The best-performing model incorporated a combination of a few quadrant-based measures and their interactions. We found an AUC of 0.915 for the detection of glaucoma with the selected measures, namely TI rim height, NI pit volume, and TI pit depth. In addition, we explored 3 ML models (GBM, NB, and SVM) for identifying eyes with glaucoma. All showed good to very good performance for discriminating glaucoma and more specifically, eyes with early glaucoma from healthy eyes with GBM demonstrating the best performance for this purpose (AUC of 0.948 and 0.950 for the entire dataset and mild glaucoma eyes, respectively).

Most prior studies found that macular thickness measurements derived from OCT imaging performed well for identifying glaucoma at different stages of the disease.^{6,8–10,30–36} These studies investigated both full macular thickness and inner retinal layers. However, there are several shortcomings with regard to macular thickness measurements. Various artifacts may affect the segmentation of inner retinal layers, especially in more severe glaucoma or when the quality of the images is subpar.^{37,38} Another caveat is that measurement of perifoveal RGCs is particularly difficult even in eyes with good-quality images as the inner retinal thickness rises from near zero in the center of the fovea to its peak around 700 to 1100 microns from the fovea.³⁹ Thickness measurements may also be misleading in myopic eyes.⁴⁰

The novel proposed biomarkers are not dependent on the segmentation of individual retinal layers and may be especially helpful when accurate segmentation of the inner retinal layers cannot be achieved. We also

hypothesize that such biomarkers may be helpful very early in the disease process when subtle changes in thickness measurements in individual eyes may be difficult to ascertain due to the wide range of thickness measurements observed in OCT normative databases. The software suite developed by the co-authors is equipped with a built-in algorithm (AQUA), which is a deep learning model able to automatically evaluate the quality of the macular volume scans. Hood et al. proposed that two main macular damage phenotypes demonstrating shallow and diffuse versus localized and deep damage may exist.¹³ The proposed novel shape measures would be particularly useful for defining macular phenotypes and our team is exploring deep learning approaches to identify such phenotypes. We also speculate that the macular shape measures could be used to monitor disease progression in various stages of glaucoma.

The deep learning model providing novel 3D volumetric and topographic perifoveal biomarkers with automated image quality check has no precedence in the OCT literature. Vercellin et al. developed an algorithm in MATLAB to provide ganglion cell complex (GCC) volume and reported very good performance in detecting glaucoma.²⁰ However, the scan quality and segmentation accuracy needed to be checked manually. A recent study reported that 3D optic disc neuroretinal rim measures were able to detect glaucoma progression earlier than 2D thickness measurements.²¹ The Nocturne software suite does not rely on segmentation of inner macular layers or thickness measurements. Any artifacts within the macular volume scans can be detected by the AQUA software and depending on the severity may be subsequently removed from the dataset. Asaoka et al. recently demonstrated that a deep learning algorithm using pretraining would perform better than an approach not utilizing pretraining.⁴¹ They reported an AUC of 94% for the new algorithm compared to 77% to 79% without the pretraining. The main difference between the study by Asaoka et al. and ours is that the deep learning on raw images represents a black box with no clue on how the deep learning algorithm is performing this task while our proposed biomarkers are specifically aiming to utilize predefined features of the perifoveal macula. The performance of both approaches seems to be clinically relevant.⁴¹

The Nocturne algorithm provides an array of global and sectoral measures. Due to the collinearity observed between many measures (results not shown), a machine learning approach (GBM) was performed to select the best measures to discriminate between eyes with glaucoma and healthy eyes. After introducing a loss function, GBM selected TI rim height, NI pit volume,

and TI pit depth as the top 3 shape measures with an AUC of 0.915 for discriminating eyes with glaucoma of various stages from healthy eyes. The same model still performed well for detection of mild perimetric glaucoma (AUC = 0.895). In addition, when subset selection was carried out using only eyes with mild glaucoma, the identified subset of predictive measures was quite similar with a higher AUC (0.913). Therefore, the proposed 3D topographic and volumetric measures could be clinically relevant biomarkers for the detection of glaucoma at all stages once their performance is further validated in other external and larger datasets. Two out of three best sectoral measures belonged to the inferotemporal sector, which is well known to be the most common macular region where early signs of the disease tend to manifest.^{12,13,42,43}

Yadav et al. proposed cubic Bézier polynomial modeling and least square optimization of landmark features of the macula for parametric modeling of the perifoveal configuration and volume in a cohort of healthy eyes and eyes from patients with neuroinflammatory central nervous system disorders and optic neuritis.²² Their novel model is based on stable and reliable features of the foveal pit shape, including the lowest point of the pit and the highest retinal point around the foveal pit region. A recent study on foveal morphology showed the importance of accounting for artifacts seen on RGC probability maps in healthy subjects;⁴⁴ the proposed modeling of the perifoveal region by Yadav et al. is an important step toward individualized structure-function maps of the macula as further development of the current deep learning algorithms may help prediction of functional damage in the macular region.^{14,17}

The Nocturne software has an OCT reader module which is capable of reading the image data and other available information in OCT volume scans from different devices (Heidelberg Spectralis, Cirrus, Nidek, and Topcon) and different file types (*.vol*, *.img*, *.dem*, *.xml*, and *.fda*). All these files contain, at the minimum, the raw imaging data (e.g. *.img* format) and potentially the segmentation information (for example, *.vol* format). If the data file has only the raw image information, then the Nocturne segmenter module performs segmentation of the ILM and Bruch's membrane in the foveal region and this segmentation information is utilized for the perifoveal shape analysis module. In this study, we utilized *.vol* files from the Heidelberg Spectralis OCT; these files do contain the segmentation information for the upper (ILM) and lower (Bruch's membrane) retinal boundaries. Therefore, we did not utilize the segmenter module from Nocturne. However, the ILM and Bruch's membrane segmentation quality for each volume scan was checked using the automated

quality analysis module (AQUA) from Nocturne. The foveal shape analysis was performed using the fovea morphometry module utilizing quality checked data. As Cubic Bézier parametrization is a smooth representation of the foveal shape, the effect of noise in segmentation is removed robustly. Furthermore, we have also compared the accuracy of our method with respect to the current state-of-the-art methods and shown that our method is capable of parameterizing the different foveal shape with high fidelity.²² A similar result was shown in a recent paper by Romero-Boscenes et al., who performed a comparison between current state-of-the-art methods for foveal shape modeling in terms of accuracy and robustness.¹⁷

Given the performance of the macular shape parameters, we propose that combining these parameters with macular thickness measures could potentially provide a stronger structural signal for detection of glaucoma, especially in eyes with early-stage glaucoma. Another potential benefit is to explore these novel macular phenotypes for categorizing different stages of glaucoma, which has been an ongoing interest in the glaucoma community.

Our results should be interpreted in light of potential shortcomings. A larger normal database is required to define the factors potentially affecting shape parameters. Specifically, race, gender, and axial length might affect the shape measures in healthy and glaucomatous eyes. The group of eyes with early glaucoma was fairly small and the results need to be confirmed in a larger study and potentially in eyes with preperimetric glaucoma.

In conclusion, the proposed novel topographic and volumetric macular biomarkers were able to detect perimetric glaucoma across the severity spectrum with clinically relevant performance. Incorporating this model in clinical practice, once validated, could help in optimizing detection of glaucoma in early stages especially in eyes with significant early central involvement. The shape biomarkers do not depend on segmentation of intraretinal layers and may prove to be helpful where segmentation of macular layers is suboptimal.

Acknowledgments

The authors thank Jeffrey A. Gornbein, DrPH, for his assistance with regard to the statistical methodology.

Supported by an NIH R01 grant (R01-EY029792), an unrestricted Departmental Grant from Research to Prevent Blindness, and an unrestricted grant from Heidelberg Engineering.

Disclosure: **V. Mohammadzadeh**, None; **M. Cheng**, None; **S.H. Zadeh**, None; **K. Edalati**, None; **D. Yazadeh**, None; **J. Caprioli**, None; **S. Yadav**, None; **E.M. Kadas**, None; **A.U. Brandt**, None; **K. Nouri-Mahdavi**, Heidelberg Engineering

References

1. Garway-Heath DF, Caprioli J, Fitzke FW, Hitchings RA. Scaling the hill of vision: the physiological relationship between light sensitivity and ganglion cell numbers. *Invest Ophthalmol Vis Sci*. 2000;41(7):1774–1782.
2. Nickells RW. The cell and molecular biology of glaucoma: mechanisms of retinal ganglion cell death. *Invest Ophthalmol Vis Sci*. 2012;53(5):2476–2481.
3. Mills RP, Janz NK, Wren PA, Guire KE. Correlation of visual field with quality-of-life measures at diagnosis in the Collaborative Initial Glaucoma Treatment Study (CIGTS). *J Glaucoma*. 2001;10(3):192–198.
4. Ramulu P. Glaucoma and disability: which tasks are affected, and at what stage of disease? *Curr Opin Ophthalmol*. 2009;20(2):92–98.
5. Greenfield DS, Bagga H, Knighton RW. Macular thickness changes in glaucomatous optic neuropathy detected using optical coherence tomography. *Arch Ophthalmol*. 2003;121(1):41–46.
6. Mohammadzadeh V, Fatehi N, Yarmohammadi A, et al. Macular imaging with optical coherence tomography in glaucoma. *Surv Ophthalmol*. 2020;65(6):597–638.
7. Tatham AJ, Weinreb RN, Medeiros FA. Strategies for improving early detection of glaucoma: the combined structure-function index. *Clin Ophthalmol*. 2014;8:611–621.
8. Jeoung JW, Choi YJ, Park KH, Kim DM. Macular ganglion cell imaging study: glaucoma diagnostic accuracy of spectral-domain optical coherence tomography. *Invest Ophthalmol Vis Sci*. 2013;54(7):4422–4429.
9. Yang Z, Tatham AJ, Weinreb RN, Medeiros FA, Liu T, Zangwill LM. Diagnostic ability of macular ganglion cell inner plexiform layer measurements in glaucoma using swept source and spectral domain optical coherence tomography. *PLoS One*. 2015;10(5):e0125957.
10. Zivkovic M, Dayanir V, Zlatanovic M, et al. Ganglion Cell-Inner Plexiform Layer Thickness in Different Glaucoma Stages Measured by Optical Coherence Tomography. *Ophthalmic Res*. 2018;59(3):148–154.
11. Rabiolo A, Mohammadzadeh V, Fatehi N, et al. Comparison of Rates of Progression of Macular OCT Measures in Glaucoma. *Transl Vis Sci Technol*. 2020;9(7):50.
12. Hood DC. Improving our understanding, and detection, of glaucomatous damage: An approach based upon optical coherence tomography (OCT). *Prog Retin Eye Res*. 2017;57:46–75.
13. Hood DC, Raza AS, de Moraes CG, Liebmann JM, Ritch R. Glaucomatous damage of the macula. *Prog Retin Eye Res*. 2013;32:1–21.
14. Sepulveda JA, Turpin A, McKendrick AM. Individual Differences in Foveal Shape: Feasibility of Individual Maps Between Structure and Function Within the Macular Region. *Invest Ophthalmol Vis Sci*. 2016;57(11):4772–4778.
15. Raza AS, Zhang X, De Moraes CG, et al. Improving glaucoma detection using spatially correspondent clusters of damage and by combining standard automated perimetry and optical coherence tomography. *Invest Ophthalmol Vis Sci*. 2014;55(1):612–624.
16. Hood DC, De Moraes CG. Challenges to the Common Clinical Paradigm for Diagnosis of Glaucomatous Damage With OCT and Visual Fields. *Invest Ophthalmol Vis Sci*. 2018;59(2):788–791.
17. Romero-Bascones D, Barrenechea M, Murueta-Goyena A, et al. Foveal Pit Morphology Characterization: A Quantitative Analysis of the Key Methodological Steps. *Entropy (Basel)*. 2021;23(6):699.
18. Maccora KA, Sheth S, Ruddle JB. Optical coherence tomography in paediatric clinical practice. *Clin Exp Optom*. 2019;102(3):300–308.
19. Turpin A, Chen S, Sepulveda JA, McKendrick AM. Customizing Structure-Function Displacements in the Macula for Individual Differences. *Invest Ophthalmol Vis Sci*. 2015;56(10):5984–5989.
20. Verticchio Vercellin AC, Jassim F, Poon LY, et al. Diagnostic Capability of Three-Dimensional Macular Parameters for Glaucoma Using Optical Coherence Tomography Volume Scans. *Invest Ophthalmol Vis Sci*. 2018;59(12):4998–5010.
21. Ratanawongphaibul K, Tsikata E, Zemplenyi M, et al. Earlier Detection of Glaucoma Progression Using High-Density 3-Dimensional Spectral-Domain OCT Optic Nerve Volume Scans. *Ophthalmol Glaucoma*. 2021;4(6):604–616.
22. Yadav SK, Motamedi S, Oberwahrenbrock T, et al. CuBe: parametric modeling of 3D foveal shape using cubic Bezier. *Biomed Opt Express*. 2017;8(9):4181–4199.
23. Motamedi S, Oertel FC, Yadav SK, et al. Altered fovea in AQP4-IgG-seropositive neuromyelitis

- optica spectrum disorders. *Neurol Neuroimmunol Neuroinflamm.* 2020;7(5):e805.
24. Johnson CA, Sample PA, Cioffi GA, Liebmann JR, Weinreb RN. Structure and function evaluation (SAFE): I. criteria for glaucomatous visual field loss using standard automated perimetry (SAP) and short wavelength automated perimetry (SWAP)11Internet Advance publication at ajo.com June 17, 2002. *Am J Ophthalmol.* 2002;134(2):177–185.
 25. Kauer-Bonin J, Yadav SK, Beckers I, et al. Modular deep neural networks for automatic quality control of retinal optical coherence tomography scans. *Comput Biol Med.* 2021;141:104822.
 26. Kauer J, Gawlik K, Zimmermann HG, et al. Automatic quality evaluation as assessment standard for optical coherence tomography. Paper presented at: Advanced Biomedical and Clinical Diagnostic and Surgical Guidance Systems XVII, February 26, 2019; San Francisco, California, United States.
 27. Petzold A, Albrecht P, Balcer L, et al. Artificial intelligence extension of the OSCAR-IB criteria. *Ann Clin Transl Neurol.* 2021;8(7):1528–1542.
 28. Yadav SK, Reitebuch U, Polthier K. Mesh denoising based on normal voting tensor and binary optimization. *IEEE Trans Vis Comput Graph.* 2017;24(8):2366–2379.
 29. Yadav,SK . Surface denoising based on the variation of normals and retinal shape analysis (Doctoral dissertation), 2018. Available at: <https://refubium.fu-berlin.de/handle/fub188/22984>.
 30. 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(6):1151–1158.
 31. Na JH, Sung KR, Baek S, Sun JH, Lee Y. Macular and retinal nerve fiber layer thickness: which is more helpful in the diagnosis of glaucoma? *Invest Ophthalmol Vis Sci.* 2011;52(11):8094–8101.
 32. Nouri-Mahdavi K, Nowroozizadeh S, Nassiri N, et al. Macular ganglion cell/inner plexiform layer measurements by spectral domain optical coherence tomography for detection of early glaucoma and comparison to retinal nerve fiber layer measurements. *Am J Ophthalmol.* 2013;156(6):1297–1307.e1292.
 33. 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(11):6788–6795.
 34. Blumberg DM, Dale E, Pensec N, et al. Discrimination of Glaucoma Patients From Healthy Individuals Using Combined Parameters From Spectral-domain Optical Coherence Tomography in an African American Population. *J Glaucoma.* 2016;25(3):e196–e203.
 35. Chen MJ, Chang YF, Kuo YS, Hsu CC, Ko YC, Liu CJ. Macular ganglion cell-inner plexiform vs retinal nerve fiber layer measurement to detect early glaucoma with superior or inferior hemifield defects. *J Chin Med Assoc.* 2019;82(4):335–339.
 36. Kim KE, Park KH. Macular imaging by optical coherence tomography in the diagnosis and management of glaucoma. *Br J Ophthalmol.* 2018;102(6):718–724.
 37. Lee WJ, Kim YK, Park KH, Jeung JW. Trend-based Analysis of Ganglion Cell-Inner Plexiform Layer Thickness Changes on Optical Coherence Tomography in Glaucoma Progression. *Ophthalmology.* 2017;124(9):1383–1391.
 38. Thenappan A, Tsamis E, Zemborain ZZ, et al. Detecting Progression in Advanced Glaucoma: Are Optical Coherence Tomography Global Metrics Viable Measures? *Optom Vis Sci.* 2021;98(5):518–530.
 39. Chan A, Duker JS, Ko TH, Fujimoto JG, Schuman JS. Normal macular thickness measurements in healthy eyes using Stratus optical coherence tomography. *Arch Ophthalmol.* 2006;124(2):193–198.
 40. Liu X, Shen M, Yuan Y, et al. Macular Thickness Profiles of Intraretinal Layers in Myopia Evaluated by Ultrahigh-Resolution Optical Coherence Tomography. *Am J Ophthalmol.* 2015;160(1):53–61.e52.
 41. Asaoka R, Murata H, Hirasawa K, et al. Using Deep Learning and Transfer Learning to Accurately Diagnose Early-Onset Glaucoma From Macular Optical Coherence Tomography Images. *Am J Ophthalmol.* 2019;198:136–145.
 42. Hood DC, Xin D, Wang D, et al. A Region-of-Interest Approach for Detecting Progression of Glaucomatous Damage With Optical Coherence Tomography. *JAMA Ophthalmol.* 2015;133(12):1438–1444.
 43. Hood DC, Raza AS. On improving the use of OCT imaging for detecting glaucomatous damage. *Br J Ophthalmol.* 2014;98(Suppl 2):ii1–ii9.
 44. De Moraes CG, Muhammad H, Kaur K, Wang D, Ritch R, Hood DC. Interindividual Variations in Foveal Anatomy and Artifacts Seen on Inner Retinal Probability Maps from Spectral Domain OCT Scans of the Macula. *Transl Vis Sci Technol.* 2018;7(2):4.