Relationships of Retinal Structure and Humphrey 24-2 Visual Field Thresholds in Patients With Glaucoma

Hrvoje Bogunović,¹ Young H. Kwon,^{2,3} Adnan Rashid,¹ Kyungmoo Lee,¹ Douglas B. Critser,³ Mona K. Garvin,^{1,4} Milan Sonka,^{1,3} and Michael D. Abramoff¹⁻⁵

¹Department of Electrical and Computer Engineering, University of Iowa, Iowa City, Iowa, United States

²Stephen A. Wynn Institute for Vision Research, University of Iowa, Iowa City, Iowa, United States

³Department of Ophthalmology and Visual Sciences, University of Iowa, Iowa City, Iowa, United States

⁴Department of Biomedical Engineering, University of Iowa, Iowa City, Iowa, United States

⁵Iowa City VA Health Care System, Iowa City, Iowa, United States

Correspondence: Michael D. Abramoff, University of Iowa, 11205 Pomerantz Family Pavilion, Iowa City, IA 52242, USA; michael-abramoff@uiowa.edu.

Submitted: October 15, 2014

Accepted: November 21, 2014

Citation: Bogunović H, Kwon YH, Rashid A, et al. Relationships of retinal structure and Humphrey 24-2 visual field thresholds in patients with glaucoma. *Invest Ophthalmol Vis Sci*. 2015;56:259–271. DOI:10.1167/iovs.14-15885

PURPOSE. To determine relationships between spectral-domain optical coherence tomography (SD-OCT) derived regional damage to the retinal ganglion cell-axonal complex (RGC-AC) and visual thresholds for each location of the Humphrey 24-2 visual field, in all stages of open-angle glaucoma.

METHODS. Patients with early, moderate, and advanced glaucoma were recruited from a tertiary glaucoma clinic. Humphrey 24-2 and 9-field Spectralis SD-OCT were acquired for each subject. Individual OCT volumes were aligned, nerve fiber layer (NFL), ganglion cell and inner plexiform layers (GCL+IPL) cosegmented. These layers were then partitioned into 54 sectors corresponding to the 24-2 grid. A Support Vector Machine was trained independently for each sector to predict the sector threshold, using these structural properties.

RESULTS. One hundred twenty-two consecutive subjects, 43 early, 39 moderate, and 40 advanced, glaucoma were included (122 eyes). Average correlation coefficient (*R*) was 0.68 (0.47–0.82), and average root mean square error (RMSE) was 6.92 dB (3.93–8.68 dB). Prediction performance averaged over the entire field, superior hemifield, and inferior hemifield had *R* (RMSE) values of 0.77 (3.76), 0.80 (5.05), and 0.84 (3.80) dB, respectively.

CONCLUSIONS. Predicting individual 24-2 visual field thresholds from structural information derived from nine-field SD-OCT local NFL and GCL+IPL thicknesses using the RGC-AC concept is feasible, showing the potential for the predictive ability of SD-OCT structural information for visual function. Ultimately, it may be feasible to complement and reduce the burden of *subjective* visual field testing in glaucoma patients with predicted function derived *objectively* from OCT.

Keywords: perimetry, visual field, image analysis, OCT, ganglion cell

Glaucoma is a progressive disease of the optic nerve and, if left untreated, can lead to irreversible loss of vision.¹ Glaucoma results in apoptosis of the retinal ganglion cells (RGCs) with corresponding nerve fiber loss.^{2,3} We have previously proposed the term retinal ganglion cell-axonal complex (RGC-AC) to stress the distributed nature of the loss to multiple neighboring ganglion cells and corresponding axons,⁴ which leads to characteristic glaucomatous visual field (VF) loss, while damage to the part of the RGC-AC within the optic nerve head (ONH) leads to characteristic cupping.

It is important to detect glaucoma early, as well as to monitor changes in glaucomatous damage. The clinical standard for detection of disease and its progression has been automated perimetry (e.g., Humphrey 24-2 VF), and clinical assessment of the optic nerve cup.⁵ However, once moderate VF loss occurs (in the range of 12–15 dB mean deviation [MD] loss or more), perimetric test-retest variability rises substantially^{6–9} and limits a reliable determination of VF change. The limitations of reliability and reproducibility for VF measurement as the main parameter in the assessment of glaucoma damage

inhibit optimal patient care and research into improved treatments.

Spectral-domain optical coherence tomography (SD-OCT) allows unprecedented, patient-friendly, three-dimensional spatial resolution of the nerve fiber layer (NFL), ganglion cell layer (GCL), and inner plexiform (IPL) retinal layers.^{10–12} Automated, three-dimensional image analysis allows objective quantification of the peripapillary and macular NFLs and GCLs, with resolutions of 1 micron, below the typical 4 micron axial resolution of the SD-OCT scanner.^{13–18} To date, these localized peripapillary NFL measurements have not correlated well with visual function.¹⁹ Attempts to improve this structure-function correlation so far have not included the RGC-AC concept, utilizing regional local measurements of GCL and NFL along the RGC-AC, as it traverses up from a set of ganglion cells through the NFL to form the nerve fiber bundle (NFB) and proceeding to the ONH (Fig. 1), all of which are affected in glaucoma subjects.⁴ We hypothesize that novel quantitative metrics of RGC-AC morphology allow improved prediction of visual function and that thickness of SD-OCT derived layers allows quantification of the amount of glaucomatous damage along the

FIGURE 1. The retinal ganglion cell-axonal complex consists of a set of ganglion cells and their axons, shown in gray, located in a single 24-2 based retinal region. Local RGC-AC structural indices can be calculated for its GCL, which forms the RGC-AC origin (block outlined in *bright green*), where RGC-AC function is measured, for its patient-specific NFL trajectory (adjoining regions in *black*) and for its terminal ONH wedge shaped region (in *dark green*)—the last is not in the present study. The RGC-AC has multiple segments: A ganglion cell body segment localized in the RGC layer; multiple NFL segments localized in the retinal NFL between the ganglion cell body and ONH in a patient-specific trajectory; and an ONH segment located in the neural rim of the ONH.

RGC-AC, as we assume that measured VF thresholds are related to RGC-AC layer thicknesses, forming a proxy for the number of axons or ganglion cells.

Leveraging our previously developed and validated automated algorithms for segmentation and registration of SD-OCT (in the public domain as the Iowa Reference Algorithms [<http://www.biomed-imaging.uiowa.edu/downloads/>]), we have developed automated registration and segmentation of nine-field OCT of the posterior pole,²⁰ covering all of the Humphrey 24-2 VF test locations, and allowing precise quantification of the local GCL and NFL thickness at each of the 54 test locations. Combined with the RGC-AC concept and its axonal trajectory according to Garway-Heath et al.,²¹ these analysis techniques allow aggregate glaucoma damage assessment for each RGC-AC.

The purpose of the present study is to determine the correlations between SD-OCT derived measures of regional damage to the proximal RGC-AC—retinal GCL thickness and NFL thickness along the entire RGC-AC trajectory—and the visual threshold assessments for each of the 54 test locations on the 24-2 grid derived from a Humphrey 24-2 VF exam, in patients with all stages of open-angle glaucoma.

METHODS

Subjects/Participants

In this prospective study, inclusion criteria were age 18 to 85, diagnosed with glaucoma suspect or open-angle glaucoma according to the following definitions:

- Glaucoma suspect: suspicious optic nerve appearance (enlarged cupping on clinical examination) with normal VF and intraocular pressure (IOP; ≤ 21 mm Hg) or normal optic disc appearance on biomicroscopy and normal VF, but with elevated IOP (> 21 mm Hg); and
- Open-angle glaucoma: primary or secondary open-angle glaucoma (e.g., exfoliative or pigmentary) with an open iridocorneal angle, glaucomatous optic disc and/or NFL defects on biomicroscopy, and VF changes (regardless of IOP level). Glaucomatous optic discs were identified as those with either diffuse or focal thinning of the neuroretinal rim. Visual field abnormalities were considered to be glaucomatous if they were consistent with the

optic nerve examination and had either (i) a typical NFL distribution, or (ii) a glaucoma hemifield test outside the normal limits. This diagnosis was made by fellowship trained glaucoma specialists according to the above definitions.

Subject's fundus visualization sufficient on indirect ophthalmoscopy to allow OCT; able to undergo perimetry Humphrey 24-2 VF Swedish Interactive Thresholding Algorithm (SITA) with sufficient reliability (false positive and false negative error $< 25\%$, fixation loss $< 33\%$); perimetry obtained within a 3-month period of SD-OCT imaging; and perimetry free of artifacts, such as rim effects. Exclusion criteria were a history of angle closure or combined mechanism glaucoma, or any nonglaucomatous optic neuropathy, corneal or retinal diseases that can affect VF, cataracts or any other disease with visual acuity $< 20/40$; and OCT of unsuitable quality determined by visual observation. Subjects were recruited matching age and disease severity in one of three approximately equally sized severity groups, based on the MD of the 24-2 VF threshold testing:

- Early glaucoma (including glaucoma suspects) < 6 dB loss;
- Moderate glaucoma 6 to 12 dB loss; and
- Advanced glaucoma > 12 dB.

One eye of each subject chosen to reflect adequate representation of each of the severity groups was further analyzed.

Data Collection

Standard automated perimetry (SAP) based on SITA Standard 24-2 VF protocol was performed with the Humphrey Field Analyzer (HFA II; Carl Zeiss Meditec, Inc., Dublin, CA, USA), which evaluates the VF as threshold assessments at 54 different retinal locations. For OCT image acquisition, a nine-field per eye protocol was used, where a subject sequentially fixates on a spot 12.5° apart in a 3×3 grid pattern. This protocol takes approximately 6 minutes per eye and covers 60° on the retina, sufficiently large enough to include the 60° area probed with 24-2 VF test. Each field is imaged with SD-OCT (Spectralis; Heidelberg Engineering, Inc., Heidelberg, Germany, $768 \times 61 \times 496$ voxels, $9.53 \times 8.07 \times 1.92$ mm³, with voxel size of $12.41 \times 132.22 \times 3.87$ μ m³) using eye tracking mode. The device

FIGURE 2. Multifield alignment. (a) Nine SLO fields imaged consecutively, corresponding OCT volumes not shown. (b) Mosaic of SLO image covering the nine retinal subfields mutually registered from the nine single field images. (c) Corresponding wide-field composite OCT image (en face projection shown).

additionally acquires 2D scanning laser ophthalmoscopy (SLO) fundus image (768×768 pixels, 9.5×9.5 mm² with pixel size of 12.41×12.41 μ m²), automatically coregistered with the OCT image by the device. Both the raw VF data, exported from the HFA II as integer threshold data, as well as the raw OCT volumes, exported as .vol format, were deidentified, and stored in our XNAT ophthalmology research database.²²

The study protocol was approved by the Institutional Review Board of the University of Iowa and adhered to the tenets of the Declaration of Helsinki; written informed consent was obtained from all participants. Subjects with missing OCT fields were excluded from further analysis.

Multifield Registration and Intraretinal Layer Segmentation

From the set of nine spatially overlapping OCT images (Fig. 2a) a widefield mosaic was created (Fig. 2b). We created the mosaic by performing 2D en face, affine alignment of the imaged fields, based on matched Speeded Up Robust Features²³ key points on the corresponding SLO images. The SLO images were used because they have higher spatial resolution and more pronounced texture than en face projections of the OCT volume data. The 2D affine transformations resulting from registration of SLO images were then applied to the corresponding OCT images, so that their relative positions and rotations were known.

The OCT images aligned in space were then passed to the segmentation step where the retinal layers of all the fields are cosegmented, taking into account the possible mutual displacements along the depth-direction (*z*-axis). The cosegmentation is an extension of the automated graph-search based multisurface segmentation method.^{24,25} In the cosegmentation, all nine OCT images are segmented simultaneously, prior to any merging, to fully incorporate the duplicate information in the overlapping areas of the volumes. This is achieved by imposing a soft intrasurface-interfield constraint in the overlapped area of each pair of overlapping images. The constraints penalize deviations from the expected surface height differences, taken as the depth-axis shifts that produce the maximum cross-correlation of pairwise-overlapped areas, assuring consistent image segmentation across the overlapped areas. Finally, after the cosegmentation, the images and the segmented surfaces are stitched together using the Voronoi

diagram based parcelization to obtain a wide-field composite OCT (Fig. 2c) and the corresponding layer thicknesses. An illustration of cosegmentation and an example of the obtained wide-field segmentation and layer thicknesses are shown in Figures 3 and 4, respectively.

Predictive Model of Visual Function From Retinal Layer Thickness

In order to locally measure structural properties and predict VF threshold assessments, we partition the wide-field retinal area into elements of a subject-specific structure-derived grid (called S-Grid), consisting of 54 sectors, that resembles the 24-2 SAP matrix (Fig. 5a). The positioning of the S-Grid on the retina is defined by two points, the fovea and the ONH center. In the nine-field segmentation, the fovea and the ONH center are identified automatically^{4,13,14} (Fig. 5b). Then we center the S-Grid on the fovea while the ONH center is positioned on the center of sector 26. An example of such a constructed S-Grid, overlaid on a composite OCT image, is shown in Figure 5c.

Once each A-scan in the nine-field OCT has been assigned to a corresponding S-Grid sector, the average GCL+IPL and NFL thickness values for that sector are computed as the mean layer thickness from all A-scans in that sector (from a total of 2030 A-scans per sector). Any missing thickness information (when segmentation was considered unreliable), was bilinearly interpolated from the four neighboring sectors.

Independent predictive models are built for each sector of the S-Grid, except for sectors 26 and 35 that cover the ONH area. Each sector-specific predictive model is based on the structural properties of the associated RGC-AC that originates at the sector of interest. Thus the properties are composed of the average sector thickness of the GCL+IPL layer and the average NFL thickness over the sector, as well as the NFL thicknesses of all the S-Grid sectors along the estimated RGC-AC trajectory. In this pilot study, we used an approximation of the RGC-AC trajectory locations based on the NFB model of Garway-Heath et al.²¹ Obviously, individual differences in the trajectories of the RGC-ACs over the retina exist, but were ignored in this pilot study. For a sector-specific model, the NFL of sectors that belong in the same Garway-Heath region are included in its RGC-AC trajectory (Fig. 6). Any NFL sectors in the Garway-Heath region that are temporal of the originating

FIGURE 3. Multifield, multilayer cosegmentation. *Top three rows* show a single B-scan slice for three volumes out of the nine total, which horizontally overlap. Overlapping OCT volumes (aligned using the transformation using the SLO images) are segmented simultaneously (cosegmentation). *Bottom row:* Cosegmentation result shown on single composite B-scan (wide-field, spanning 60 degrees horizontally), after flattening and stitching of composite OCT volumes.

sector are ignored, because the RGC-AC trajectory should not run temporally.

Thus, for each subject, the measured structural properties (i.e., GCL+IPL and NFL sector thicknesses) of the subset of sectors along the RGC-AC trajectory, as described above for each sector-specific predictive model, form a feature vector. Each sector-specific predictive model is trained using these feature vectors to predict the continuous output represented by the VF threshold (see below) for that S-Grid sector. The predictive model is nonlinear and implemented as a support vector regression machine (SVM)²⁶ with a radial basis function kernel. SVM is a powerful and popular supervised learning model where the input feature vector is first mapped onto a higher-dimensional space using a nonlinear mapping defined by the kernel, and then a linear regression model is constructed in such a hyperdimensional space.²⁷

Training and Evaluating the Predictive Model

To train and evaluate the predictive model, it is important to know the visual function data at each of the S-Grid sectors for each subject individually. This requires registering the 24-2 VF grid containing the measured threshold values at each location in the grid to the composite wide nine-field OCT image. For the registration we use a similarity transformation (translation, rotation, and isotropic scaling), uniquely defined by two pairs of corresponding points, specifically, the pair formed by the fovea and fixation center and the pair formed by the center of the ONH and the center of the blind spot. The blind spot center was manually located on each VF printout by a consensus between two ophthalmologists (MDA and YHK). Such registration appropriately matches each location in the (VF) functional data to the corresponding area in the structural data, assuming any in- or ex-cyclotorsion was constant at both acquisition of the OCT and perimetry for the study eye.

FIGURE 4. Wide-field layer thickness maps of NFL (*left*) and GCL+HPL (*right*) of a single subject showing glaucomatous loss in the inferotemporal region. The maps represent fully 3D layer thickness derived from the merged nine-field OCT analysis.

FIGURE 5. Subject-specific structure-derived grid (S-Grid) construction. (a) S-Grid containing 54 sectors. (b) Composite OCT showing the automatically determined fovea and center of the ONH. (c) S-Grid aligned on the underlying structural OCT data using foveal and ONH landmarks. Though the S-Grid resembles the 24-2 Humphrey grid, it is a subdivision of the structural OCT data based on structural landmarks.

FIGURE 6. Construction of a structural feature vector for two example sectors of interest: 20 and 47. (a) Each sector in the subject-specific S-Grid is color grouped according to its Garway-Heath region (vertically flipped to correspond to structural OCT space). (b) Ganglion cell and inner plexiform layer thickness map showing sectors 20 and 47. (c) Nerve fiber layer thickness map showing the sectors 20 and 47 (in red) as well as their corresponding sectors along the RGC-AC trajectory originating at sectors of interests (mean NFL thickness is shown in each sector, instead of A-scan thickness, for clarity).

However, the test locations of 24-2 VF grid may not match sector centers of the S-Grid. This happens when the blind spot on the VF deviates from the expected location ($15^\circ, -3^\circ$). In such situations the VF threshold values at the S-Grid sector centers were obtained by bilinear interpolation. A correction has been applied to position the predicted test locations closer to the fovea center with respect to measured structural data, as suggested by Hood et al.²⁸ The central four test locations (corresponding to sectors 23, 24, 32, 33) and the surrounding 12 test locations have been moved by 2.4° and 1.8° toward the fovea center, respectively.

Once trained, the model produces a set of predicted thresholds at the 54 S-Grid sectors. As the S-Grid is made to resemble the 24-2 VF grid, we present the predicted VF thresholds as a grayscale map by simulating the printout produced by the HFA II perimetry device. These steps are shown for a representative example in Figure 7.

Statistical Analysis

For comparison purposes all left eye scans were mirrored to conform to the scans of the right eye. For descriptive analysis of the population sample used, we computed the mean and the standard deviation of thickness values of NFL and GCL+IPL, and the measured VF thresholds across S-Grid sectors, stratified according to glaucoma severity subgroup (early, moderate, advanced).

The performance of the predictive model was evaluated using leave-one-out cross-validation. For qualitative evaluation, grayscale maps of the measured and predicted VF were created. To quantitatively evaluate the prediction at each sector of the S-Grid, we computed the Pearson correlation coefficient (R), and the root mean square error (RMSE). Finally, we evaluated the aggregate performance for predicting average VF areas (superior hemifield, inferior hemifield, and entire field) by displaying the measured-predicted scatter plots and computing R and RMSE.

RESULTS

We recruited 142 consecutive patients from the Glaucoma Service at the University of Iowa. Of the 142 subjects, 20 subjects were excluded because they had incomplete imaging or functional studies. A random eye from the remaining subjects, 122 eyes (67 OD, 55 OS) of 122 subjects, was studied.

Mean (SD) age was 65.4 (10.4) years and 48 (39%) were male. The cohort included 111 (self-identified) Caucasian subjects, 4 African-American, 3 Asian-American, 1 Native American subjects (race was unknown or undisclosed for remaining 3). Among the 111 Caucasian subjects, four subjects identified as Hispanic. 43 had early (including five glaucoma suspects), 39 had moderate, and 40 had advanced glaucoma.

Average image acquisition time for all nine OCT fields was 6 minutes (range, 4–11 minutes), and was well tolerated by all subjects. Anecdotally, subjects in our study preferred the nine-field OCT to the VF testing, and some indicated it as an important reason for them to participate.

The sector average GCL+IPL thickness had the mean (SD), for early glaucoma 43.2 (8.9) μm (range, 37.2–74.1 μm), for moderate glaucoma 41.3 (7.4) μm (range, 36.7–72.2 μm), and for advanced glaucoma 41.9 (5.1) μm (range, 38.1–63.5 μm). The sector average NFL thickness had the mean (SD), for early glaucoma subjects 30.3 (14.8) μm (range, 12.9–71.6 μm), for moderate glaucoma subjects 25.2 (12.3) μm (range, 10.3–57.9 μm), and for advanced glaucoma subjects 22.5 (9.9) μm (range, 10.8–51.8 μm). The sector average VF threshold had the mean (SD), for early glaucoma subjects 26.4 (2.4) dB (range, 19.7–30.3 dB), for moderate glaucoma subjects 21.3 (4.3) dB (range, 9.3–29.8 dB), and for advanced glaucoma subjects 13.6 (6.0) dB (range, 2.9–25.0 dB). Correspondingly, the regional differences in these values for specific sectors are shown in Figure 8. This figure illustrates that in general, GCL+IPL and NFL thicknesses decrease as glaucoma becomes more advanced, although qualitatively the differences between groups are not as pronounced as the VF thresholds. Striking is the increasing asymmetry between inferior and superior hemifields (average visual thresholds) as glaucoma is more severe, see the third (actual values) and fourth columns (simulated Humphrey 24-2 printout), while superior and inferior NFL and GCL+IPL thicknesses, first and second column, lack such asymmetry.

Next, we compared the *actual* VF thresholds from Humphrey VF to *predicted* threshold derived from the RGC-AC based sector-specific predictive models. Quantitatively (Fig. 9), the correlation R across the VF grid sectors ranged from 0.47 to 0.82, with a mean of 0.68. The RMSE ranged from 3.93 to 8.68 dB, with a mean of 6.92 dB. Finally, performance of predicting mean VF over three regions (Fig. 10) produced correlations of 0.77, 0.80, and 0.84, corresponding to the

FIGURE 7. Example of presentation of measured and predicted VF threshold data for a single subject as used in this study. (a) Scan of the subject's actual HFA II printout of the left eye (mirrored to right eye format). (b) Grayscale map directly derived from the raw threshold data exported from Humphrey Field Analyzer shows how the HFA II printout is simulated. (c) Grayscale map of measured raw threshold data, after interpolation of the VF grid to the S-grid. (d) Grayscale map of predicted VF thresholds for all 54 sectors in the S-grid, from the sector-specific predictive models. (e) Numerical values of measured thresholds as displayed in (c). (f) Numerical values of predicted threshold values as displayed in (d). Grayscale maps are consistently shown in the standard manner (i.e., superior field of grayscale represents superior VF and inferior retina).

entire VF, superior hemifield (inferior retina), and the inferior hemifield (superior retina), respectively. The dB scale for the thresholds was chosen deliberately.

For qualitative comparison, analogous to Figure 8, in Figure 11 we show the sector-specific average predicted visual thresholds, which also show the asymmetry between superior and inferior hemifield. This appears to be so, even though the predictions are entirely based on the GCL+IPL and NFL thickness inputs, which do not show this asymmetry, as shown in Figure 8. Figure 12 shows a qualitative evaluation of prediction performance for all 122 study eyes, with many examples of a close match between the actual and predicted thresholds.

DISCUSSION

The results of this study show that there is substantial correlation between SD-OCT derived measures of the RGC-AC—namely, retinal GCL thickness and NFL thickness—and the VF threshold assessments for each of the 54 test locations on the 24-2 grid derived from a single Humphrey 24-2 VF exam, in patients at all stages of glaucoma.

Specifically, the correlation between the measured Humphrey 24-2 perimetric thresholds using structural measurements of the RGC-AC trajectory from nine-field Spectralis OCT was good, ranging from 0.47 to 0.82, depending on the sector location (Fig. 9). By aggregating over multiple sectors—for the superior and inferior hemifield and the

FIGURE 8. Mean (SD) values for all sectors across the studied population stratified according to glaucoma severity (rows). *First column:* Population average thickness in each sector of both GCL and IPL layers. *Second column:* Population average thickness of NFL in each sector. *Third column:* Population average 24-2 VF threshold (dB) of each sector shown in retinotopic orientation (i.e., inferior hemifield thresholds are superior in this figure). *Fourth column:* A grayscale map of the population average of each sector's VF threshold oriented according to the clinical standard (i.e., inferior hemifield thresholds are inferior in the figure). The labels S, I, T, and N denote superior, inferior, temporal, and nasal retina, respectively.

entire 60° field—correlations improved ($R = 0.77$ – 0.84 ; Fig. 10). Prediction error reported as RMSE ranged across locations from 3.93 dB to 8.68 dB, comparable to the perimetric test-retest variability as described in the literature.²⁹ The predictions were completely automated and obtained by machine learning employing an SVM as a nonlinear regression model. A modified version of our previously published and publicly available Iowa Reference Algorithms was used to simultaneously segment the NFL, GCL, and IPL thickness at each of the 54 S-grid sectors that span the 60° retinal area tested with the 24-2 VF test

protocol. Based on these results, we make several observations:

1. The traditional “plateau” or leveling off between actual and predicted VF correlation is not pronounced, even though the latter was derived from structure (Fig. 10). When we linearized the scale, the correlation became worse; we believe training the classifier on function in dB scale is responsible for this. Had we trained the algorithm in the linear scale, the results might have been different;

FIGURE 9. Performance of the predictive models for each S-Grid sector, expressed as: (a) R , and (b) RMSE. Prediction is best in fovea and the superotemporal quadrant of the retina.

FIGURE 10. Performance of the predictive models averaged over: (a) all S-grid sectors in the entire field, (b) all S-grid sectors in the superior hemifield, and (c) all S-grid sectors in the inferior hemifield.

2. The correlations between structure and function were largest in the fovea and the superotemporal retina, corresponding to the inferonasal field. The larger structure-function correlation temporally is likely because the RGC-AC trajectories are longer, leading to

more NFL sectors being pooled for function prediction. Conversely, closer to the optic nerve, the RGC-AC trajectories are shorter, with fewer sectors leading to lower correlation. However, this does not explain why the superotemporal retina has better structure-function

FIGURE 11. Mean (SD) values for all predicted thresholds across the studied population stratified according to glaucoma severity (rows). *First column:* Predicted population average 24-2 VF threshold (dB) of each sector shown in retinotopic orientation (i.e., inferior hemifield thresholds are superior in this figure). *Second and third columns:* Grayscale maps of the predicted and measured population average of each sector's VF threshold oriented according to the clinical standard (i.e., inferior hemifield thresholds are inferior in the figure). The labels S, I, T, and N denote superior, inferior, temporal, and nasal retina, respectively.

FIGURE 12. Visual fields presented as simulated HFA II printouts for each of the 122 subjects used in the study. Each pair displays the actual perimetry measured thresholds (*left*) and the predicted (*right*) VF thresholds for that eye. The pairs are sorted column-wise (*top-to-bottom, left-to-right*) by the decreasing MD and stratified (*top-bottom*) by early, moderate, and advanced glaucoma. The fields of the subject illustrated in Figure 7 are outlined.

correlation than the corresponding inferotemporal retina (Fig. 10). It is possible that foveal-disc cyclo-torsion affects the superior retina differently than inferior retina as we try to correlate with function, or that there is more redundancy in the superior than the inferior retina because the inferior hemifield is more important for survival in primates. Another possible reason could be the tendency to obtain higher correlations when there is a wider range of measurements;

3. The range of retinal layer thickness of the three glaucoma groups seemed less pronounced than the differences in VF thresholds (Fig. 8). Possibly, this is explained by the differences in their distribution (linear versus logarithmic); and
4. The asymmetry between superior and inferior hemifield threshold as glaucomatous damage becomes more severe, visible in Figure 8, is also present in the predicted VFs (Fig. 11). It has been reported before that inferior optic nerve rim (corresponding to the superior VF) tends to show earlier and more progressive damage compared to the superior rim.³⁰⁻³² Remember that this asymmetry is absent in the structural information (Fig. 8) though the prediction is entirely derived from this structural information. Likely, the sector-specific predictive models interpret the same NFL and GCL+IPL inputs differently depending on their location.

Previous studies have shown limited associations between structure derived from OCT and function derived from perimetry in patients with glaucoma.^{33,34} By focusing on the prediction of a glaucoma diagnosis—for example, in a screening setting—rather than predicting individual field thresholds as in our study, and combining multiple structural parameters such as ganglion cell-inner plexiform, retinal nerve fiber, and ONH metrics derived from OCT, reasonable performance can be obtained (Springelkamp, et al. *IOVS* 2014;55:ARVO E-Abstract 4740).³⁵ However, for the management of glaucoma, rather than the prediction of a glaucoma diagnosis, estimates of local thresholds are important.

In a recent study, Zhang, et al.^{36,37} have used principal component analysis and multiple regression of average peripapillary NFL thickness and macular NFL and GCL+IPL thickness to successfully predict the 24-2 thresholds and determine the performance of this approach to predict a diagnosis of glaucoma, with a specificity of 98% and sensitivity of 78% at the level of VF hemifields. This study was performed on both controls and subjects with glaucoma (severity not specified). However, they only looked at predictive performance at the level of the hemifield, while we studied performance for each VF location independently. Another difference is that we used an explicit model of the RGC-AC, with regional NFL and GCL thicknesses derived from wide-field OCT as input, while they used a model-free approach with average NFL and GCL thicknesses. We assume that their regression derived model would exhibit similar bundle patterns.

The correlation values were not as high as the Food and Drug Administration (FDA) would like in order for structural measures to replace perimetric testing. Nevertheless, the results represent a first step toward “smarter” structure-function correlation, as we are integrating structure information along the entire RGC-AC NFB trajectory pattern in the retina, rather than a small cross-section of it near the optic disc (e.g., using peripapillary NFL thickness). This latter, more common, approach has yielded relatively poor structure function correlations so far. This lower structure

function correlation has been attributed to the differences in dynamic range and scale (linear versus logarithmic), stage of glaucoma, or the presence of a glial component to the NFL thickness.^{36,37} While these factors all undoubtedly play a role, another reason is the challenge in precisely matching a wedge of peripapillary NFL bundle to a corresponding patch of retina where the RGC body is located. The present study takes the next step of defining an RGC-AC trajectory that encompasses a much greater area of the retina to be able to correlate with the individual test locations of Humphrey 24-2. Our hypothesis is that measured VF threshold values are directly related to layer thicknesses, as a proxy of the number of axons or ganglion cells, as measured with OCT, and hence we did not rely on age corrected threshold values. We utilized the RGC+IPL layer thickness and adjacent NFL thickness along the RGC-AC trajectory pathway as described in Garway-Heath et al.²¹ However, the localization of the RGC-AC trajectory has striking differences between subjects^{38,39} and potentially, patient-specific localization of the RGC-AC trajectory may improve correlation and predictive performance.

Though we have previously published methods for automated determination of the ONH rim tissue and Bruch's membrane opening for Cirrus OCT, the ONH was not used in this pilot study. We are currently testing⁴ the hypothesis that prediction performance will improve if ONH derived metrics are included in the prediction vector.

A major limitation of this pilot study is that the reference standard by which the prediction algorithm is both trained and tested is derived from a single 24-2 VF. Most of our subjects were experienced VF testers from a tertiary, academic glaucoma clinic. Nevertheless, the intravisit variability of the 24-2 SITA VF is substantial.^{7,8,40} A more noisy reference standard necessarily negatively affects the measured performance.⁴¹ Future studies will require a more robust reference standard created from multiple VFs for a single eye. This will also allow the reproducibility of repeat Humphrey 24-2 versus predicted VF from repeat OCTs to be estimated. Another limitation of this pilot study is that our prediction algorithm was evaluated in a leave-one-out fashion, instead of separate training and test sets. Finally, we did not have a normative structural dataset for comparison. Thus, we could not compute “deviation from normative values” for structural information to compare with age corrected MD and other normative data based perimetric values. We plan to address some of these shortcomings with future studies.

In summary, our results show that predicting individual 24-2 VF thresholds from structural information, as derived from wide-field SD-OCT local NFL and GCL+IPL thicknesses measurements and using the RGC-AC concept, is feasible. The results show the potential for the predictive ability of SD-OCT structural information for visual function. Ultimately, it may be feasible to complement and reduce the burden of *subjective* VF testing in glaucoma patients with the predicted function derived *objectively* from OCT structure.

Acknowledgments

The research patients were generously provided by Wallace L.M. Alward, MD, John H. Fingert, MD, PhD, Anna S. Kitzmann, MD, and Khadija S. Shahid, OD, with the help of Teresa Kopel.

Supported in part by National Institutes of Health Grants R01 EY019112, R01 EY018853, and R01 EB004640; the Department of Veterans Affairs; and the Marlene S. and Leonard A. Hadley Glaucoma Research Fund. YHK was supported by Clifford M. & Ruth M. Altermatt Professorship.

Disclosure: **H. Bogunović**, None; **Y.H. Kwon**, None; **A. Rashid**, None; **K. Lee**, None; **D.B. Critser**, None; **M.K. Garvin**, None; **M. Sonka**, P; **M.D. Abramoff**, P

References

- Weinreb RN, Anton-Lopez A. *Glaucoma Definition and Classification*. Oxford Medical Publications. Oxford; New York: Oxford University Press; 1999.
- Danias J, Shen F, Kavalarakis M, et al. Characterization of retinal damage in the episcleral vein cauterization rat glaucoma model. *Exp Eye Res*. 2006;82:219–228.
- Kwon YH, Fingert JH, Kuehn MH, Alward WL. Primary open-angle glaucoma. *N Engl J Med*. 2009;360:1113–1124.
- Lee K, Kwon YH, Garvin MK, Niemeijer M, Sonka M, Abramoff MD. Distribution of damage to the entire retinal ganglion cell pathway: quantified using spectral-domain optical coherence tomography analysis in patients with glaucoma. *Arch Ophthalmol*. 2012;130:1118–1126.
- AAO PPP Glaucoma Panel, Hoskins Center for Quality Eye Care. Primary Open-Angle Glaucoma PPP. 2010. Available at: <http://one.aao.org/preferred-practice-pattern/primary-openangle-glaucoma-ppp-october-2010>. Accessed December 8, 2014.
- Heijl A, Lindgren A, Lindgren G. Test-retest variability in glaucomatous visual fields. *Am J Ophthalmol*. 1989;108:130–135.
- Henson DB, Chaudry S, Artes PH, Faragher EB, Ansons A. Response variability in the visual field: comparison of optic neuritis, glaucoma, ocular hypertension, and normal eyes. *Invest Ophthalmol Vis Sci*. 2000;41:417–421.
- Wall M, Johnson CA, Kutzko KE, Nguyen R, Brito C, Keltner JL. Long- and short-term variability of automated perimetry results in patients with optic neuritis and healthy subjects. *Arch Ophthalmol*. 1998;116:53–61.
- Werner EB, Petrig B, Krupin T, Bishop KI. Variability of automated visual fields in clinically stable glaucoma patients. *Invest Ophthalmol Vis Sci*. 1989;30:1083–1089.
- Wojtkowski M, Leitgeb R, Kowalczyk A, Bajraszewski T, Fercher AE. In vivo human retinal imaging by Fourier domain optical coherence tomography. *J Biomed Optics*. 2002;7:457–463.
- de Boer JF, Cense B, Park BH, Pierce MC, Tearney GJ, Bouma BE. Improved signal-to-noise ratio in spectral-domain compared with time-domain optical coherence tomography. *Optics Letters*. 2003;28:2067–2069.
- Wojtkowski M, Srinivasan V, Fujimoto JG, et al. Three-dimensional retinal imaging with high-speed ultrahigh-resolution optical coherence tomography. *Ophthalmology*. 2005;112:1734–1746.
- Garvin MK, Abramoff MD, Lee K, Niemeijer M, Sonka M, Kwon YH. 2-D pattern of nerve fiber bundles in glaucoma emerging from spectral-domain optical coherence tomography. *Invest Ophthalmol Vis Sci*. 2012;53:483–489.
- Lee K, Niemeijer M, Garvin MK, Kwon YH, Sonka M, Abramoff MD. Segmentation of the optic disc in 3-D OCT scans of the optic nerve head. *IEEE Trans Med Imaging*. 2010;29:159–168.
- Abramoff MD, Garvin M, Sonka M. Retinal imaging and image analysis. *IEEE Rev Biomed Engin*. 2010;3:169–208.
- Mwanza JC, Gendy MG, Feuer WJ, Shi W, Budenz DL. Effects of changing operators and instruments on time-domain and spectral-domain OCT measurements of retinal nerve fiber layer thickness. *Ophthalmic Surg Lasers Imaging*. 2011;42:328–337.
- Mwanza JC, Oakley JD, Budenz DL, Chang RT, Knight OJ, Feuer WJ. Macular ganglion cell-inner plexiform layer: automated detection and thickness reproducibility with spectral domain-optical coherence tomography in glaucoma. *Invest Ophthalmol Vis Sci*. 2011;52:8323–8329.
- Abramoff MD, Mullins RF, Lee K, et al. Human photoreceptor outer segments shorten during light adaptation. *Invest Ophthalmol Vis Sci*. 2013;54:3721–3728.
- Hood DC, Anderson SC, Wall M, Raza AS, Kardon RH. A test of a linear model of glaucomatous structure-function loss reveals sources of variability in retinal nerve fiber and visual field measurements. *Invest Ophthalmol Vis Sci*. 2009;50:4254–4266.
- Bogunovic H, Sonka M, Kwon Y, Kemp P, Abramoff MD, Xiaodong W. Multi-surface and multi-field co-segmentation of 3-D retinal optical coherence tomography. *IEEE Trans Med Imag*. 2014;32:2242–2253.
- Garway-Heath DF, Poinoosawmy D, Fitzke FW, Hitchings RA. Mapping the visual field to the optic disc in normal tension glaucoma eyes. *Ophthalmology*. 2000;107:1809–1815.
- Wahle A, Lee K, Harding AT, et al. Extending the XNAT archive tool for image and analysis management in ophthalmology research. *Proc. SPIE 8674, Medical Imaging 2013: Advanced PACS-based Imaging Informatics and Therapeutic Applications*, 86740M; 2013.
- Bay H, Ess A, Tuytelaars T, Van Gool L. Speeded-Up Robust Features (SURF). *Comput Vis Image Underst*. 2008;110:346–359.
- Garvin MK, Abramoff MD, Wu X, Russell SR, Burns TL, Sonka M. Automated 3-D intraretinal layer segmentation of macular spectral-domain optical coherence tomography images. *IEEE Trans Med Imaging*. 2009;28:1436–1447.
- Quellec G, Lee K, Dolejsi M, Garvin MK, Abramoff MD, Sonka M. Three-dimensional analysis of retinal layer texture: identification of fluid-filled regions in SD-OCT of the macula. *IEEE Trans Med Imaging*. 2010;29:1321–1330.
- Pochet NL, Suykens JA. Support vector machines versus logistic regression: improving prospective performance in clinical decision-making. *Ultrasound Obstet Gynecol*. 2006;27:607–608.
- Sonka M, Hlavac V, Boyle R. *Image Processing, Analysis, and Machine Vision*. 4th ed. Stamford, CT: Cengage Learning; 2014.
- Hood DC, Raza AS, de Moraes CGV, Liebmann JM, Ritch R. Glaucomatous damage of the macula. *Prog Retin Eye Res*. 2013;32:1–21.
- Artes PH, Iwase A, Ohno Y, Kitazawa Y, Chauhan BC. Properties of perimetric threshold estimates from full threshold, SITA standard, and SITA fast strategies. *Invest Ophthalmol Vis Sci*. 2002;43:2654–2659.
- Jonas JB, Fernandez MC, Sturmer J. Pattern of glaucomatous neuroretinal rim loss. *Ophthalmology*. 1993;100:63–68.
- Strouthidis NG, Gardiner SK, Sinapis C, Burgoyne CF, Garway-Heath DE. The spatial pattern of neuroretinal rim loss in ocular hypertension. *Invest Ophthalmol Vis Sci*. 2009;50:3737–3742.
- See JLS, Nicolela MT, Chauhan BC. Rates of neuroretinal rim and peripapillary atrophy area change: a comparative study of glaucoma patients and normal controls. *Ophthalmology*. 2009;116:840–847.
- Artes PH, Chauhan BC. Longitudinal changes in the visual field and optic disc in glaucoma. *Prog Retin Eye Res*. 2005;24:333–354.
- Zhu H, Crabb DP, Schlottmann PG, et al. Predicting visual function from the measurements of retinal nerve fiber layer structure. *Invest Ophthalmol Vis Sci*. 2010;51:5657–5666.

35. 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.
36. Zhang X, Bregman CJ, Raza AS, De Moraes G, Hood DC. Deriving visual field loss based upon OCT of inner retinal thicknesses of the macula. *Biomed Opt Express*. 2011;2:1734–1742.
37. Zhang X, Raza AS, Hood DC. Detecting glaucoma with visual fields derived from frequency-domain optical coherence tomography. *Invest Ophthalmol Vis Sci*. 2013;54:3289–3296.
38. Jansonius NM, Nevalainen J, Selig B, et al. A mathematical description of nerve fiber bundle trajectories and their variability in the human retina. *Vision Res*. 2009;49:2157–2163.
39. Jansonius NM, Schiefer J, Nevalainen J, Paetzold J, Schiefer U. A mathematical model for describing the retinal nerve fiber bundle trajectories in the human eye: average course, variability, and influence of refraction, optic disc size and optic disc position. *Exp Eye Res*. 2012;105:70–78.
40. Wall M, Woodward KR, Doyle CK, Artes PH. Repeatability of automated perimetry: a comparison between standard automated perimetry with stimulus size III and V, matrix, and motion perimetry. *Invest Ophthalmol Vis Sci*. 2009;50:974–979.
41. Quellec G, Abramoff MD. Estimating maximal measurable performance for automated decision systems from the characteristics of the reference standard. Application to diabetic retinopathy screening. Paper presented at: *Engineering in Medicine and Biology Society (EMBC), 2014 36th Annual International Conference of the IEEE*; August 26–30, 2014; Chicago, IL. <http://ieeexplore.ieee.org/xpl/articleDetails.jsp?reload=true&arnumber=6943552>. Accessed December 24, 2014.