tvst

Article

# Individualized Glaucoma Change Detection Using Deep Learning Auto Encoder-Based Regions of Interest

**Christopher Bowd[1], Akram Belghith[1], Mark Christopher[1], Michael H. Goldbaum[1], Massimo A. Fazio[2], Christopher A. Girkin[2], Jeffrey M. Liebmann[3], Carlos Gustavo de Moraes[3], Robert N. Weinreb[1], and Linda M. Zangwill[1]**

[1] Hamilton Glaucoma Center, Shiley Eye Institute, The Viterbi Family Department of Ophthalmology, UC San Diego, La Jolla, CA, USA
[2] School of Medicine, University of Alabama-Birmingham, Birmingham, AL, USA
[3] Bernard and Shirlee Brown Glaucoma Research Laboratory, Edward S. Harkness Eye Institute, Department of Ophthalmology, Columbia University Medical Center, New York, NY, USA

**Correspondence:** Christopher Bowd, Hamilton Glaucoma Center, Shiley Eye Institute, The Viterbi Family Department of Ophthalmology, UC San Diego, 9500 Gilman Dr, La Jolla, CA 92093-0946, USA. e-mail: cbowd@health.ucsd.edu

**Received:** September 18, 2020
**Accepted:** March 1, 2021
**Published:** July 22, 2021

**Keywords:** optical coherence tomography; imaging; progression; deep learning; artificial intelligence

**Citation:** Bowd C, Belghith A, Christopher M, Goldbaum MH, Fazio MA, Girkin CA, Liebmann JM, de Moraes CG, Weinreb RN, Zangwill LM. Individualized glaucoma change detection using deep learning auto encoder-based regions of interest. Transl Vis Sci Technol. 2021;10(8):19, https://doi.org/10.1167/tvst.10.8.19

**Purpose:** To compare change over time in eye-specific optical coherence tomography (OCT) retinal nerve fiber layer (RNFL)-based region-of-interest (ROI) maps developed using unsupervised deep-learning auto-encoders (DL-AE) to circumpapillary RNFL (cpRNFL) thickness for the detection of glaucomatous progression.

**Methods:** Forty-four progressing glaucoma eyes (by stereophotograph assessment), 189 nonprogressing glaucoma eyes (by stereophotograph assessment), and 109 healthy eyes were followed for ≥3 years with ≥4 visits using OCT. The San Diego Automated Layer Segmentation Algorithm was used to automatically segment the RNFL layer from raw three-dimensional OCT images. For each longitudinal series, DL-AEs were used to generate individualized eye-based ROI maps by identifying RNFL regions of likely progression and no change. Sensitivities and specificities for detecting change over time and rates of change over time were compared for the DL-AE ROI and global cpRNFL thickness measurements derived from a 2.22-mm to 3.45-mm annulus centered on the optic disc.

**Results:** The sensitivity for detecting change in progressing eyes was greater for DL-AE ROIs than for global cpRNFL annulus thicknesses (0.90 and 0.63, respectively). The specificity for detecting not likely progression in nonprogressing eyes was similar (0.92 and 0.93, respectively). The mean rates of change in DL-AE ROI were significantly faster than for cpRNFL annulus thickness in progressing eyes (−1.28 µm/y vs. −0.83 µm/y) and nonprogressing eyes (−1.03 µm/y vs. −0.78 µm/y).

**Conclusions:** Eye-specific ROIs identified using DL-AE analysis of OCT images show promise for improving assessment of glaucomatous progression.

**Translational Relevance:** The detection and monitoring of structural glaucomatous progression can be improved by considering eye-specific regions of likely progression identified using deep learning.

## Introduction

The detection of glaucomatous progression is one of the most challenging and most important aspects of glaucoma management. Most instruments that provide progression detection algorithms report global change or localized change within instrument defined regions.[1] In addition to loss of topographical information from spatial averaging in these regions, this one-size-fits-all strategy is not ideal because it is well-known that individual differences exist in anatomical structures (e.g., optic disc size, Bruch's membrane opening center to fovea angle, distribution of arcuate retinal nerve fiber layer [RNFL] bundles). The current project uses deep learning auto-encoders (DL-AE) to identify eye-specific regions-of-interest (ROI) in individual eyes followed over time that represent likely progression,

not likely progression, and no change relative to baseline measurements. These eye-specific results take into account both individual anatomical differences and individual variations in patterns and locations of glaucomatous progression that could easily be missed when relying on instrument-defined global or regional measurements.

In recent years, neural network algorithms (e.g., deep learning) have gained popularity owing in part to excellent classification performance in several complex domains, in particular for detecting change.[2] Two widely used classification approaches—supervised and unsupervised—can be used to detect change. The supervised approach uses data labeled with the desired outcome as the ground truth in a dataset for training the classifier. Among supervised models, the convoluted neural network[3] and deep belief network[4] are the most widely used deep learning classifiers but both need a large amount of training data to reach optimal performance. Because large datasets are readily available for detecting glaucoma from fundus photographs and optical coherence tomography (OCT) images and predicting the severity of visual field (VF) damage, convoluted neural networks have been used successfully for these tasks.[5–10]

The unsupervised approach uses data that are not labeled with a desired outcome. This point is a particular advantage for detecting disease-related change in glaucoma because training datasets with large numbers of eyes identified as progressing from glaucoma are more difficult to compile. Among unsupervised models, the DL-AE[11,12] is a deep learning strategy that has been used for change detection. Specifically, the DL-AE is a hierarchical deep neural network structure composed of multilayer auto-encoders that decrease the input to a select few representations and then reconstruct the features as output (see the Methods section for more details). An AE in its simplest form is an unsupervised hidden one-layer neural network in which the output layer is set to be equal to the input layer. The DL-AE aims to reconstruct the original input as accurately as possible in the output layer. Failure to accurately reconstruct the original input can be used as an indicator of change.

The purpose of this report was to use a DL-AE to detect glaucoma-related change from OCT optic nerve head high-resolution cube scans and to compare the results to OCT circumpapillary RNFL (cpRNFL) thickness measurements derived from the same cube scans for successfully classifying disease-related change (progression) as defined by standardized review of serial optic disc stereophotographs for progression of glaucomatous optic neuropathy (PGON; i.e., the study-specific ground truth). We hypothesized that DL-AE models used to identify eye-specific regions of interest for progression detection would detect PGON more accurately than the average change across cpRNFL thickness measurements.

## Methods

For this prospective longitudinal cohort study, participants were selected from two ongoing studies designed to evaluate optic nerve structure and visual function in glaucoma: the UC San Diego-based Diagnostic Innovations in Glaucoma Study (DIGS) and the African Descent and Glaucoma Evaluation Study (ADAGES). The three-site ADAGES collaboration includes the Hamilton Glaucoma Center at the Shiley Eye Institute, University of California, San Diego (La Jolla, California; data coordinating center); the Edward S. Harkness Eye Institute, Columbia University Medical Center (New York, New York); and the Department of Ophthalmology, University of Alabama, Birmingham (Birmingham, Alabama). The protocols for DIGS and ADAGES are identical and have been described elsewhere.[13] Informed consent was obtained from each participant and each institution's Human Subjects Committee approved all methodology. All methods adhered to the tenets of the Declaration of Helsinki for research involving human subjects and to the Health Insurance Portability and Accountability Act. DIGS and ADAGES are registered as cohort clinical trials [http://www.clinicaltrials.gov (identifiers NCT00221897 and NCT00221923; September 14, 2005)].

### Participants

At study entry, eligible participants had best-corrected visual acuity of 20/40 or better, spherical refractive error of less than 5.0 diopters, cylinder of less than 3.0 diopters, and open angles on gonioscopy. All participants were at least 18 years old. Participants were excluded if they had a history of intraocular surgery (except for uncomplicated cataract surgery or uncomplicated glaucoma surgery). Eyes with coexisting retinal disease, uveitis, or nonglaucomatous optic neuropathy also were excluded. Diabetic participants with no evidence of retinal involvement were included. Both healthy and glaucoma eyes were included in the longitudinal analyses.

All subjects underwent an annual comprehensive ophthalmologic examination, including a review of medical history, best-corrected visual acuity, slit-lamp biomicroscopy, dilated funduscopic examination, and stereoscopic optic disc photography. Semiannual

examination included intraocular pressure (IOP), OCT optic nerve head imaging, and VF testing.

In addition to the eligibility criteria, healthy subjects were required to have an IOP of less than 22 mm Hg with no history of elevated IOP or IOP-lowering medication, and at least two reliable normal VFs (pattern standard deviation within 95% confidence limits and a glaucoma hemifield test result within normal limits). Glaucoma eyes had at least three consecutive and reliable (defined elsewhere in this article) VF examinations with either pattern standard deviation of 5% or greater or a glaucoma hemifield test result outside of the 99% normal limits.

In this report, we included four groups of participants. The first three of four groups required OCT follow-up of 3 or more years with four or more images available from independent imaging sessions.

1. Progressing glaucoma eyes: The first group was composed of 44 progressing eyes of 33 glaucoma patients followed for an average (95% confidence interval [CI]) of 4.4 years (4.0–4.6) years (all available eyes graded as progressing with the required follow-up duration and number of images available as described elsewhere in this article). Progression was determined by an assessment of stereophotographs. Specifically, the baseline and each follow-up photograph were assessed for PGON by two observers (masked to participant identification and diagnosis) using a stereoscopic viewing device to evaluate digitized paired images on a 21-inch or larger computer monitor. PGON was defined as a decrease in the neuroretinal rim width, the appearance of a new RNFL defect, or the enlargement of a preexisting RNFL defect. A third observer adjudicated any disagreement in assessment between the first two observers.[14] Because these graded eyes represented ground truth for glaucomatous progression, they were used to assess the sensitivity of the two progression detection methods evaluated.
2. Nonprogressing glaucoma eyes: A second group of 303 nonprogressing glaucoma eyes from 189 patients followed for an average (95% CI) of 3.9 (3.8-4.0) years was used to train and test the deep learning model. Two observers reviewed the stereophotographs as described above and did not find evidence of PGON. These eyes were used to assess the specificity of the two progression detection methods evaluated.
3. Healthy eyes: A third group of 109 eyes from 59 healthy subjects followed for an average (95% CI) of 3.2 (2.9-4.4) years was used to estimate aging effects. Healthy participants were recruited from the general population through advertisement, from referring practices, and from the staff and employees from the Shiley Eye Institute, University of California, San Diego. These eyes were used to assess the rate of change attributable to natural aging detected by the two progression detection methods evaluated.
4. Stable glaucoma eyes: A fourth group of 50 glaucoma stable eyes from 27 early, moderate, and advanced glaucoma patients with five serial OCT examinations imaged every week for approximately 5 weeks was used to train the DL-AE models. We assumed that in this short follow-up period, glaucoma-related changes were not likely to occur (i.e., eyes had stable, nonprogressing disease). These eyes were used to decrease the likelihood that changes owing to measurement variability or segmentation errors would be classified as glaucomatous progression.

## VF Testing

All patients underwent VF testing using the Swedish Interactive Thresholding Algorithm Standard 24-2 strategy on the Humphrey Field Analyzer (Carl Zeiss Meditec Inc., Dublin, CA) at baseline and during follow-up for classification purposes. All VFs were evaluated by UC San Diego Visual Field Assessment Center personnel based on a standardized protocol. VFs with more than 33% fixation losses or false-negative errors or more than 20% false-positive errors were automatically excluded. VFs exhibiting a learning effect (i.e., follow-up VFs showing consistent improvement compared with baseline VFs) also were excluded. VFs were further reviewed qualitatively for lid and rim artifacts, fatigue effects, evidence that the VF results were attributable to a disease other than glaucoma (e.g., homonymous hemianopia), and inattention.

## Spectral-Domain OCT

Study eyes were imaged with SD-OCT (Spectralis; Heidelberg Engineering GmbH, Heidelberg, Germany; software version 5.2.0.3). Tissue thickness measurements from the high-resolution optic nerve head cube scan protocol (73 B-scans with 768 A-scans each) were used to train and test DL-AE analyses to detect individualized ROI-related change over the whole image and to detect change in cpRNFL thickness within a 2.22 mm to 3.45 mm annulus centered on optic disc. Quality assessment of OCT scans was completed by experienced examiners in the UC San Diego Imaging Data Evaluation and Assessment Center who were masked to the subjects' results of the other tests

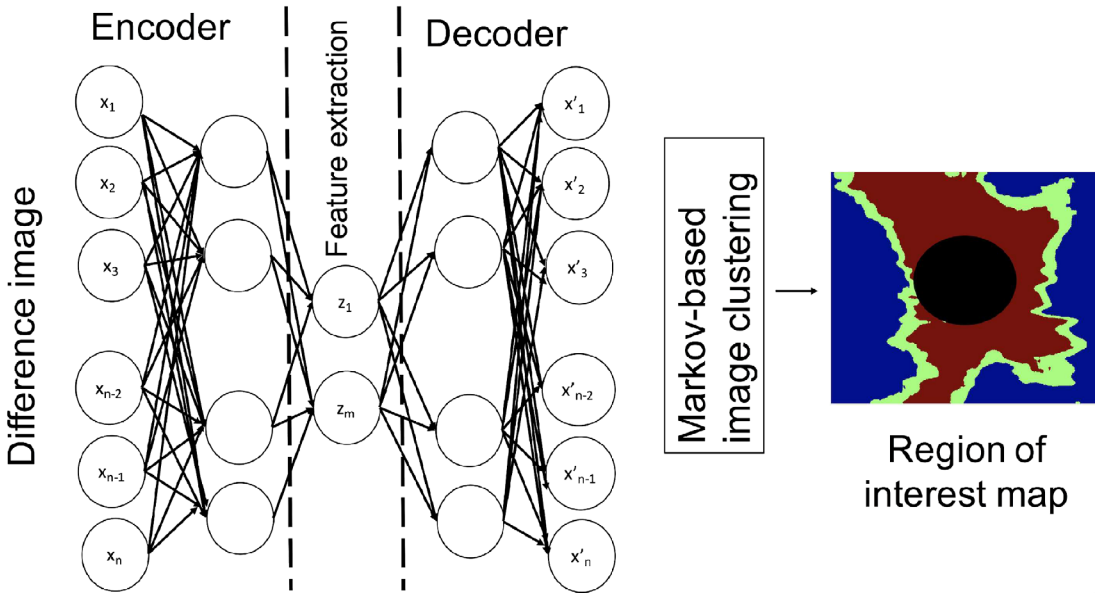


**Figure 1.** Framework of DL-AE region of interest (ROI) map generation.

according to standard protocols. To be included, all OCT images required a signal strength of more than 15 dB. There was no statistical difference in mean signal strength between training (29.3 dB; 95% CI, 27.1–31.5 dB) and testing datasets (29.7 dB; 95% CI, 27.4–31.8 dB) ($P = 0.88$).

## Retinal Layer Segmentation

Raw three-dimensional SD-OCT images were exported to a numerical computing language (MATLAB; MathWorks, Natick, MA). The San Diego Automated Layer Segmentation Algorithm (SALSA) was used to automatically segment the RNFL layer.[15] In brief, we assumed that each B-scan consists of several retinal layers (e.g., the Bruch's membrane layer, RNFL). Because the intraretinal layers have different thicknesses, each layer can be defined by a curve modeling its skeleton and a filter or set of filters modeling its thickness. To segment the different layers, it is sufficient to estimate their skeletons and the hyperparameters of the filters.

## Region of Interest Map Generation

### DL-AE for Change Detection

Autoencoders are a specific type of feed-forward neural networks in which the input is the same as the output. They compress the input into a lower dimensional code and then reconstruct the output from this representation. A detailed introduction to autoencoders can be found in.[16] In brief, the goal of the DL-AE is to reconstruct the output (x′) to match the input (x) as closely as possible. Hence, the reconstruction image is the output of the DL-AE model. For the change detection task, the input is the difference image between the first image and each follow-up image. Specifically, the model input is the pairwise image difference between any pair of SALSA-defined RNFL thickness map and the SALSA RNFL maps from nonprogressing or stable glaucoma eyes. Therefore, the model will learn the RNFL changes attributable to aging (from healthy eyes) and machine and segmentation variabilities (from stable glaucoma eyes). Because the model was never trained on progressing eyes, we expect the model to have more difficulty reconstructing the output to match the input. In other words, the difference between the reconstructed image and the input image (i.e., the reconstruction error; discussed elsewhere in this article) is higher in progressing eyes compared with nonprogressing eyes.

Specifically, consider an eye with $N$ number of follow-up visits. For each visit, we calculate the RNFL thickness map using SALSA. We use the difference images between any pair of RNFL thickness maps of a given eye to estimate the ROI map. If an eye has $N$ visits, we can generate $C(N,2)$ pairs of combinations where $C(.,.)$ is the combinatorial function [e.g., $C(4,2) = 6$]. $M$ elements from the difference images (the difference between each pair of the input image and the reconstructed image through the DL-AE) for each corresponding pixel are stacked together to form the input vector x, which is then fed to a five-layer DL-AE. The goal of the DL-AE is to reconstruct output x′ to match x as closely as possible. Because the number of nodes gets progressively smaller (encoder) and then larger again (decoder) (Fig. 1), the model is forced to use the most important features of the image to be

**Table 1.** Layer Types, Layer Sizes, Activation Functions, and Patch Dimensions for Each Autoencoder Layer

| Layer | Layer Type | Layer Size | Activation Function | Patch Dimensions |
|---|---|---|---|---|
| 0 | Input | | | 50 × 512 × 51 2 × 1 |
| 1 | Convolution | 5 × 5 × 32 | tanh | 50 × 256 × 256 × 32 |
| 2 | Convolution | 5 × 5 × 64 | relu | 50 × 128 × 128 × 6 4 |
| 3 | Convolution | 5 × 5 × 128 | relu | 50 × 64 × 64 × 128 |
| 4 | Deconvolution | 5 × 5 × 64 | relu | 50 × 128 × 128 × 64 |
| 5 | Deconvolution | 5 × 5 × 1 | tanh | 50 × 512 × 512 × 1 |

able to reconstruct the input using compressed data. The final layer has the same number of nodes as the input. The difference between the original input vector x and the reconstruction x' is called the reconstruction error. The DL-AE learns to minimize this reconstruction error. The DL-AE uses the reconstruction error as the anomaly score. Data points with high reconstruction error are considered to be anomalies or, in this application, examples of glaucomatous progression. Because only data representing normal instances are used to train the DL-AE, the DL-AE will reconstruct normal data very well. However, it will fail to do so when presented with anomalous data, which the DL-AE has not encountered during training.

In this article, we used the Python deep learning library Keras.[17] We used Glorot uniform to initialize the weight matrix and tanh activation function for the hidden layers. The model was compiled using the Adam optimization algorithm, which usually gives better results than the simpler stochastic gradient descent algorithm.[12] To train the network, the Adam method was exploited to minimize the loss function (mean squared error) by using a learning rate $\alpha = 10-3$ and a batch size of 50. We trained the DL auto-encoder end to end for 150 epochs. We selected the best model with lowest reconstruction error. Table 1 shows layer types, sizes, activation functions, and patch dimensions for each autoencoder layer.

To train the model, we used 100 nonprogressing glaucoma eyes and 50 stable glaucoma eyes.

Because the model is only trained using nonprogressing glaucoma eyes, the change detection problem can be formulated as an abnormality detection problem where the model will have a higher reconstruction error for progressing eyes ($\|x\text{-}x'\|^2$). In other words, because (1) the magnitude of change in progressing eyes is higher than in nonprogressing eyes and stable glaucoma eyes and (2) the model was not trained on progressing eye combinations, we expected the model to have more difficulty reconstructing the output to match the input of progressing eyes compared with the nonprogressing eyes. This difficulty was quantified as a reconstruction error image using Markov-based image segmentation strategies described below. The difference images were then used to estimate the ROI maps using Markov-based segmentation.

### Markov-Based Image Segmentation

The reconstruction error image ($\|x\text{-}x'\|^2$) was used to estimate the ROI map. We used a Markov random field–based segmentation algorithm to exploit the statistical correlation of intensities among the neighboring pixels.[18] The Markov random field is a stochastic process that specifies the local characteristics of an image and is combined with the given data to reconstruct the true image. The Markov random field of prior contextual information is a powerful method for modeling spatial continuity (i.e., we assumed that glaucomatous change appears in continuous areas rather than in scattered pixels).

Each reconstruction error image was then segmented into three regions based on the magnitude of the error. Examples of these regions are shown in Figure 1. Region 1 is the no change region (blue color), region 2 is the likely progression region (green color), and region 3 is the not likely progression region (red color). The likely progression region in the ROI map is the region of interest where we expect the glaucomatous changes to occur.

The DL-AE regions of interest were defined in the parapapillary region and average cpRNFL thickness was measured in the parapapillary region within a 2.22-mm to 3.45-mm annulus surrounding the optic disc. The parapapillary region was chosen to avoid the higher variability of RNFL thickness measurements closer to the disc margin which occurs owing in part to the higher prevalence of vessels close to the disc margin and the resultant shadowing that makes RNFL segmentation challenging. Comparing the DL-AE and cpRNFL thickness results in the parapapillary region is clinically relevant to the cpRNFL thickness measurements provided by most OCT instruments.

Finally, we used the most frequent class for each pixel of each ROI map to generate a map used to estimate RNFL thickness loss. Instead of following

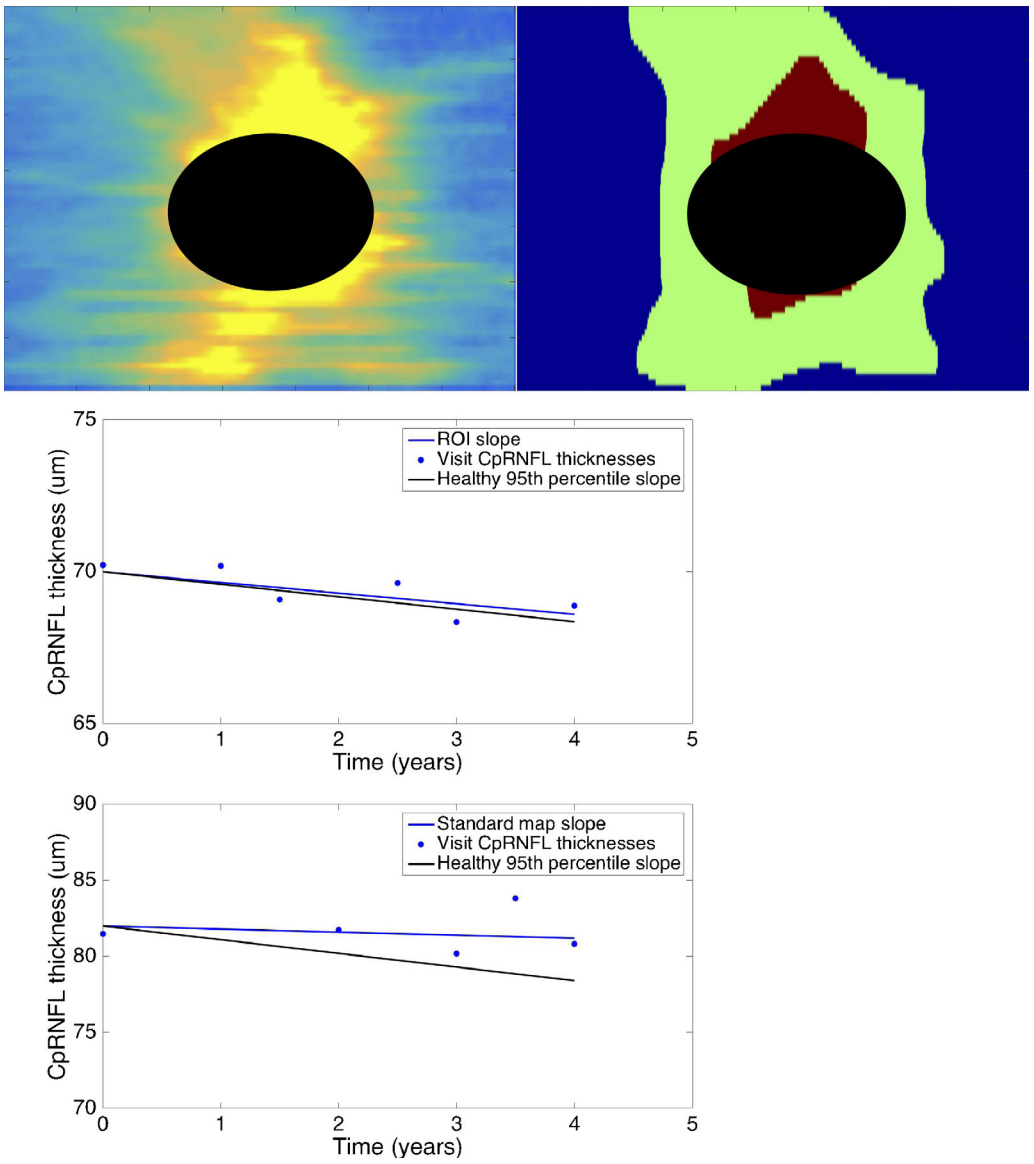


**Figure 2.** Example of a PGON eye incorrectly classified as not-likely progression by the DL-AE ROI map. The RNFL thickness map (top left), the region of interest map (top right): the blue color is the nonchanged region, the red color is the not likely progression region and the green color is the likely progression region. The rate of change in RNFL thickness in the green likely progression region of interest was not significantly faster (blue line) than that observed in healthy eyes over time (black line) (middle). The cpRNFL annulus thickness rate of change also was not significantly faster than that observed in healthy eyes over time (bottom).

the RNFL thickness in the whole RNFL thickness image or using a standard grid, RNFL loss in the likely progression region (i.e., the ROI) was used to classify overall images as progressing and not progressing.

## Definition of Progression

ROI progression for DL-AE models was determined from the likely progressing region. Using maximum likelihood estimation linear regression, OCT cpRNFL progression was defined as the coefficient associated with the variable "Time," describing the slope (or rate of change) per year of the dependent variables and its p-value reflects whether this change was significant compared with a flat slope (i.e., no progression). Statistical significance was defined at a $P$ value of less than 0.05.

Rates of progression (decrease in thickness by microns per year) for the DL-AE models were defined as the average change in all pixels within the DL-AE ROI map defined as likely progressed (green colored region in Fig. 1) over time. An eye was identified as having a likely progression region if the rate of RNFL loss was significantly ($P < 0.05$) different from zero and

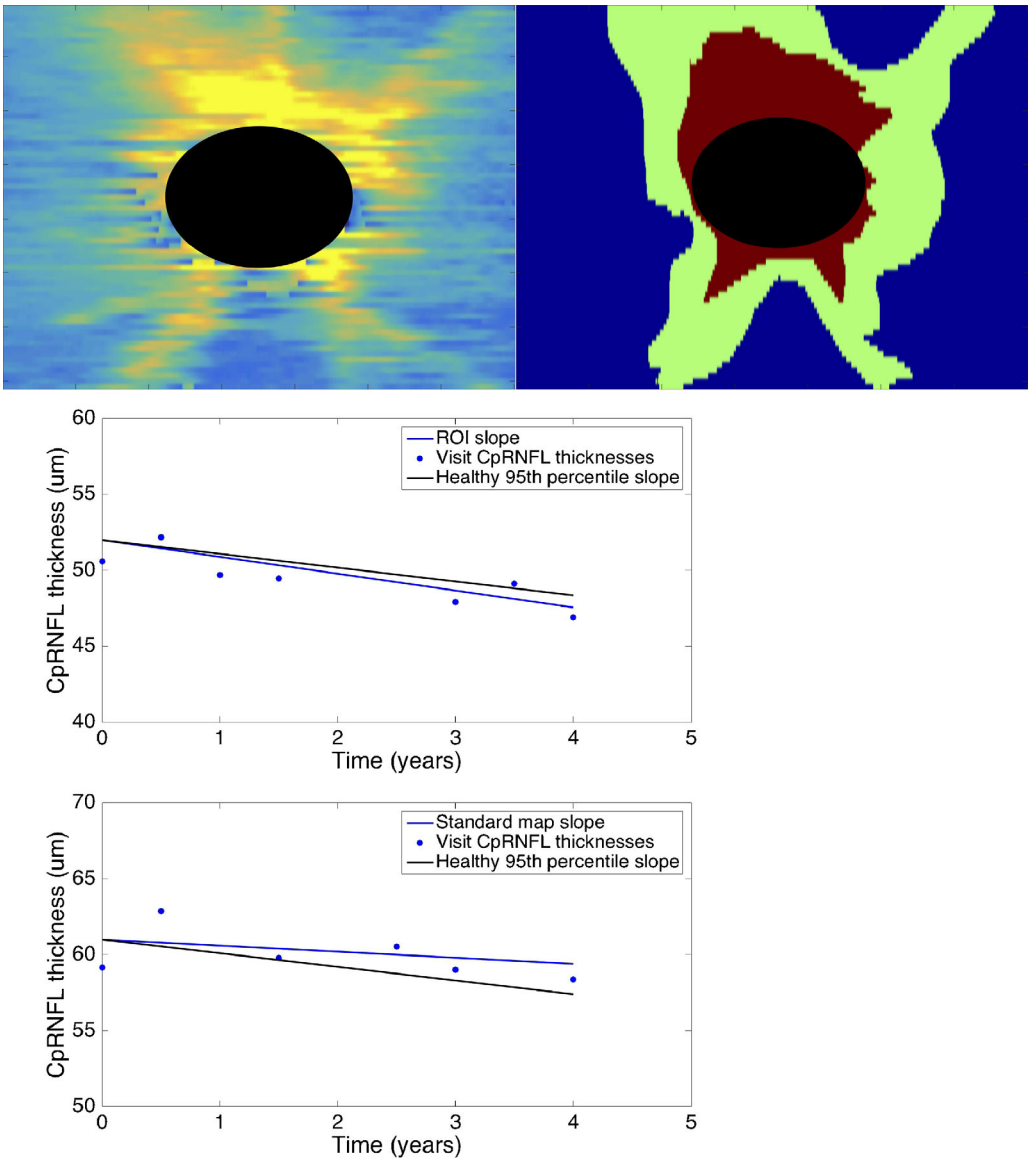


**Figure 3.** Example of a PGON eye correctly classified as likely progression by the DL-AE ROI map. The RNFL thickness map (top left), the region of interest map (top right): the blue color is the nonchanged region, the red color is the not-likely progression region and the green color is the likely progression region. The rate of change in RNFL thickness in the green likely progression region of interest was significantly greater (blue line) than that observed in healthy eyes over time (black line) (middle). The cpRNFL annulus thickness rate of change was incorrectly classified as not-likely progression (bottom).

faster or greater than the fifth percentile (single tail) of the healthy group. It is important to note that the fifth percentile cut-off is unique for each eye because the DL-AE ROI map output is eye specific. Therefore, the fifth percentile is calculated for each eye using results from the healthy group. Figures 2 and 3 demonstrate ROI maps that incorrectly and correctly identify regional RNFL progression in PGON eyes, respectively. For the eye in Figure 2, the slope of change in DL-AE ROI RNFL thickness was not significantly greater (i.e., was not more negative) than that observed in healthy eyes over time (Fig. 2c). Therefore, this PGON eye was incorrectly classified as not progressing. This incorrect classification also was the case for the cpRNFL annulus (Fig. 2d). For the eye in Figure 3, the slope of change in DL-AE ROI RNFL thickness was significantly greater than that observed in healthy eyes over time (Fig. 3c); the eye was correctly classified as progressing. However, change within the cpRNFL annulus thickness was not identified correctly (Fig. 3d).

The rates of progression for cpRNFL thickness within the 2.22-mm to 3.45-mm annulus (derived from the optic disc cube scan, outer circle shown in Fig. 4) were defined as the average change across all pixels.

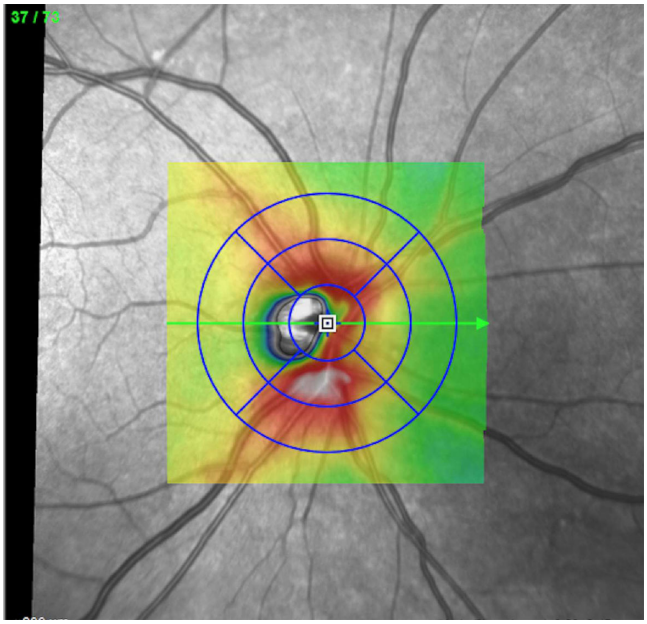

**Figure 4.** Circumpapillary RNFL thickness scan region obtained within a 2.22 mm to 3.45 mm annulus centered on optic nerve (outer circle) derived from high resolution optic disc cube.

Similar to the DL-AE ROI progression, an eye was identified as a progressing glaucoma eye if the rate of RNFL loss was significantly ($P < 0.05$) different from zero and faster or greater than the fifth percentile of the healthy group.

## Statistical Analysis

Descriptive statistics to compare demographic characteristics by group (healthy and glaucoma participants) were completed using $\chi^2$ tests to compare categorical variables and $t$ tests to compare continuous variables. Mixed effects models were used to calculate the mean rates of change (slopes) for DL-AE ROI and cpRNFL annulus thickness loss from baseline. Models included group (healthy versus glaucoma), time, and the interaction term Group × Time. $P$ values of less than 0.05 were considered statistically significant. The models were adjusted for age, and the correlation between eyes within subjects. Statistical analysis was performed using SAS, Version 9.2 (SAS Institute, Cary, NC).

# Results

A summary of the demographic and clinical characteristics at baseline of each study group are shown in Table 2. Progressing glaucoma patients were significantly older than healthy participants (mean age = 61.4 and 54.3 years; respectively; $P < 0.001$) and significantly younger than nonprogressing glaucoma patients in the training group (mean age, 71.7 years; $P < 0.001$). Progressing, nonprogressing and healthy individuals were similar with respect to gender ($P = 0.17$). Progressing glaucoma eyes had worse baseline VF mean deviation (MD) compared with healthy eyes (mean MD, −7.09 and −0.43 dB respectively; $P < 0.001$) but similar baseline VF MD to nonprogressing glaucoma eyes in the training group (mean MD, −6.87 dB; $P = 0.52$) and had longer mean follow-up than healthy subjects (4.4 and 3.2 years, respectively; $P < 0.001$), but similar follow-up to nonprogressing glaucoma eyes in the training group (3.8 years; $P = 0.12$). The glaucoma progressing eyes, nonprogressing eyes in the training group, and the healthy eyes were similar with respect to axial length ($P = 0.30$), but the progressing glaucoma eyes had thinner central corneas compared with the other eyes (all comparisons $P \leq 0.02$).

## Between-Group Differences in Baseline OCT Measurements

The baseline global DL-AE RNFL thickness measurements in the ROI and baseline global cpRNFL annulus thickness measurements are presented in Table 3. Compared with nonprogressing and progressing glaucoma eyes, healthy eyes had significantly thicker baseline RNFL for both the DL-AE ROI (72.7 µm, 60.5 µm, and 95.7 µm, respectively) and the cpRNFL annulus (74.6 µm, 63.1 µm, and 99.9 µm, respectively; both comparisons $P < 0.001$). Progressing glaucoma eyes had significantly thinner baseline RNFL thickness for both the DL-AE ROI and the cpRNFL annulus compared with nonprogressing glaucoma eyes (74.6 µm vs 63.1 µm and 72.7 µm vs 60.5 µm, respectively; both comparisons $P < 0.001$).

## Glaucoma Progression Detection

Using the RNFL rate of change criterion, the sensitivity, or number of correctly identified PGON eyes, was higher for the DL-AE ROI (40 eyes [sensitivity, 0.90; 95% CI, 0.85–0.96] compared with the cpRNFL annulus (28 eyes [sensitivity, 0.63; 95% CI, 0.57–0.69]). The specificity, or number of correctly identified nonprogressing glaucoma eyes, for the DL-AE ROI (77 eyes [specificity, 0.92; 95% CI, 0.88–0.97]) was similar to the cpRNFL annulus (78 eyes [specificity, 0.93; 95% CI, 0.90–0.98]).

**Table 2.** Clinical and Demographic Characteristics of the Healthy, Progressing, and Nonprogressing Glaucoma Eyes

| | (A) Healthy Group, Mean (95% CI) | (B) Nonprogressing Group, Mean (95% CI) | (C) Progressing Group, Mean (95% CI) | Analysis of Variance *P* Value | Post Hoc *P* Values* |
|---|---|---|---|---|---|
| No. of patients (eyes) | 59 (109) | 52 (84) | 42 (44) | | |
| Age (years) | 54.3 (51.0 to 57.6) | 71.7 (69.9 to 73.5) | 61.4 (56.3 to 66.4) | <0.001 | A < C < B |
| Female (%) | 70% | 66% | 68% | 0.17 | |
| Axial length ($mm^2$) | 23.2 (23.4 to 23.7) | 23.9 (23.7 to 24.2) | 24.1 (23.7 to 24.4) | 0.3 | |
| IOP (mm Hg) | 15.4 (14.8 to 15.9) | 15.3 (14.3 to 16.0) | 14.2 (12.4 to 16.0) | 0.45 | |
| CCT (um) | 553.4 (546.0 to 560.9) | 536.3 (527.4 to 545.3) | 525.4 (504.8 to 543.9) | 0.02 | C = B < A |
| Mean Follow-up (years) | 3.2 (2.9 to 3.4) | 3.8 (3.6 to 4.0) | 4.4 (4.0 to 4.6) | <0.001 | A<B<C |
| VF MD (dB) | −0.43 (−0.64 to 1.35) | −6.87 (−8.34 to −5,39) | −7.09 (−9.96 to −4.23) | <0.001 | B = C < A |
| Median number of visits (min, max) | 5 (4, 7) | 6 (4, 9) | 6 (4, 10) | <0.001 | A < B = C |

*Tukey test: Alpha = 0.05.

**Table 3.** Differences in Baseline RNFL Thickness in Healthy, Progressing Glaucoma, and Nonprogressing Glaucoma Eyes

| | Baseline RNFL Thickness (μm), Mean (95% CI) | | | | |
|---|---|---|---|---|---|
| Model | (A) Healthy Eyes | (B) Nonprogressing Glaucoma Eyes | (C) Progressing Glaucoma Eyes | Analysis of Variance *P* Value | Post Hoc *P* Values* |
| Deep learning auto-encoder region of interest | 95.7 (92.5-96.3) | 72.7 (68.4-76.9) | 60.5 (55.0-65.9) | <0.001 | C < B < A |
| cpRNFL annulus | 99.8 (97.5-102.0) | 74.6 (70.4-78.9) | 63.1 (57.6-68.6) | <0.001 | C < B < A |

*Tukey test: Alpha = 0.05.

## RNFL Rate of Change

The mean rates of RNFL loss in progressing glaucoma eyes, nonprogressing glaucoma eyes, and healthy eyes for the DL-AE ROI and the cpRNFL annulus, stratified by follow-up time, are presented in Table 4. Comparative distributions of rates of change between healthy and progressing glaucoma eyes and healthy and nonprogressing glaucoma eyes with a follow-up of 3 or more years are shown in Figures 5 and 6, respectively. Rates of change for both the DL-AE ROI and the cpRNFL annulus were significantly different from zero for all study groups (i.e., progressive change, on average, was observed in all groups).

Specifically, for follow-up of 3 or more years, the mean rates of RNFL thinning within the DL-AE ROI defined as likely progressed (green colored region in Fig. 1) were −1.28 μm/y (95% CI, −1.38 to −1.15 μm/y; $P < 0.001$) in progressing glaucoma eyes, −1.03 μm/y (95% CI, −1.12 to −0.93 μm/y; $P < 0.001$) in nonprogressing glaucoma eyes and −0.80 μm/y (95% CI,−0.88 to −0.71 μm/y; $P < 0.001$) in healthy eyes. The mean rate of change in progressing eyes was significantly faster than in nonprogressing and healthy eyes ($P < 0.001$). The mean rate of change in progressing eyes also was faster than in nonprogressing eyes ($P = 0.03$). In the cpRNFL annulus, the mean rates of change were −0.83 μm/y (95%

**Table 4.** Differences in the Rate of RNFL Thickness Loss (µm/year) Between Healthy, Progressing Glaucoma Eyes and Nonprogressing Glaucoma Eyes Stratified by Follow-Up Time

| Model | RNFL Thinning (µm/y) Mean (95% CI) | | | Analysis of Variance *P* Value | Post Hoc *P* Values* |
|---|---|---|---|---|---|
| | (A) Healthy Eyes | (B) Nonprogressing Glaucoma Eyes | (C) Progressing Glaucoma Eyes | | |
| DL-AE region of interest | | | | | |
| Likely progressing (1+year follow-up) | −0.12 (−0.42 to 0.1) | −0.36 (−0.62 to 0.01) | −0.51 (−0.77 to −0.12) | | C < A = B |
| Likely progressing (2-year follow-up) | −0.31 (−0.62 to −0.12) | −0. 78 (−0.94 to −0.64) | −1.08 (−1.17 to −0.99) | <0.001 | C < B = A |
| Likely progressing (≥3-year follow-up) | −0.80 (−0.88 to −0.71) | −1.03 (−1.12 to −0.93) | −1.28 (−1.38 to −1.15) | <0.001 | C < B < A |
| cpRNFL annulus | | | | | |
| Global (1-year follow-up) | 0.08 (−0.14 to 0.36) | −0.11 (−0.32 to 0.41) | −0.13 (−0.21 to 0.08) | | C = B = A |
| Global (2-year follow-up) | −0.05 (−0.14 to 0.1) | −0.23 (−0.42 to −0.12) | −0.31 (−0.38 to −0.21) | | C = B < A |
| Global (≥3-year follow-up) | −0.61 (−0.69 to −0.50) | −0.78 (−0.88 to −0.67) | −0.83 (−0.93 to −0.72) | <0.001 | C = B < A |

*Tukey test: Alpha = 0.05.

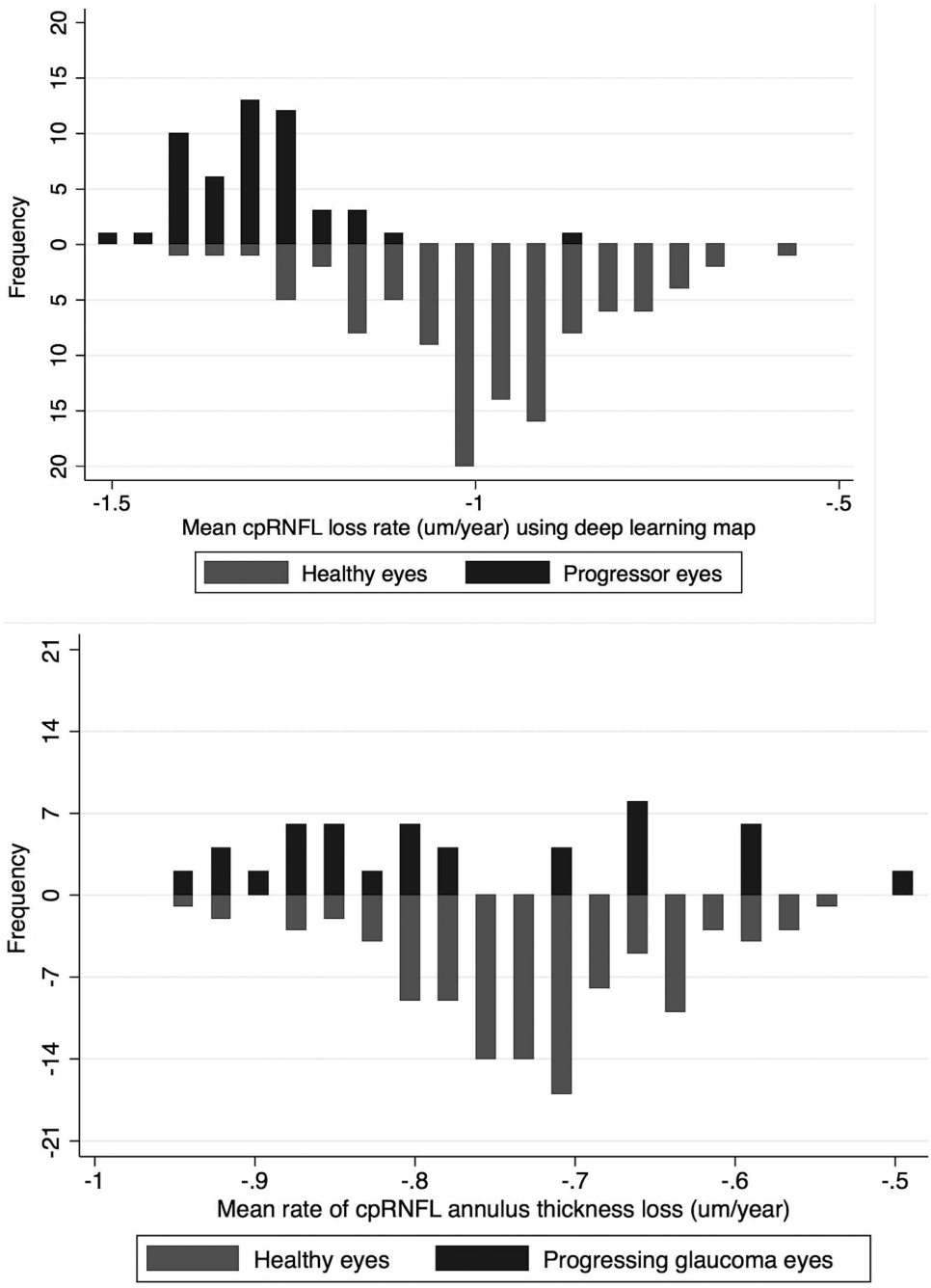


**Figure 5.** Mean RNFL loss rate between healthy and progressing glaucoma eyes using the deep learning auto encoder region of interest map (top) and the cpRNFL annulus thickness map (bottom).

CI, −0.93 to −0.72 µm/y; $P < 0.001$) in progressing glaucoma eyes, −0.78 µm/y (95% CI, −0.88 to −0.67 µm/y; $P < 0.001$) in nonprogressing glaucoma eyes, and −0.61 µm/y (95% CI, −0.69 to −0.50 µm/y; $P < 0.001$) in healthy eyes. Although the progressing and nonprogressing glaucoma eyes showed a faster rate of change than healthy eyes, there was no statistical difference between the rate of change in nonprogressing glaucoma and healthy eyes using this method ($P = 0.55$).

To further investigate and compare rates of change between progression detection methods we stratified results by follow-up times and glaucoma severity. After 1-year of follow-up, using the RNFL rate of change criterion, the sensitivity, or number of correctly identified PGON eyes, was higher

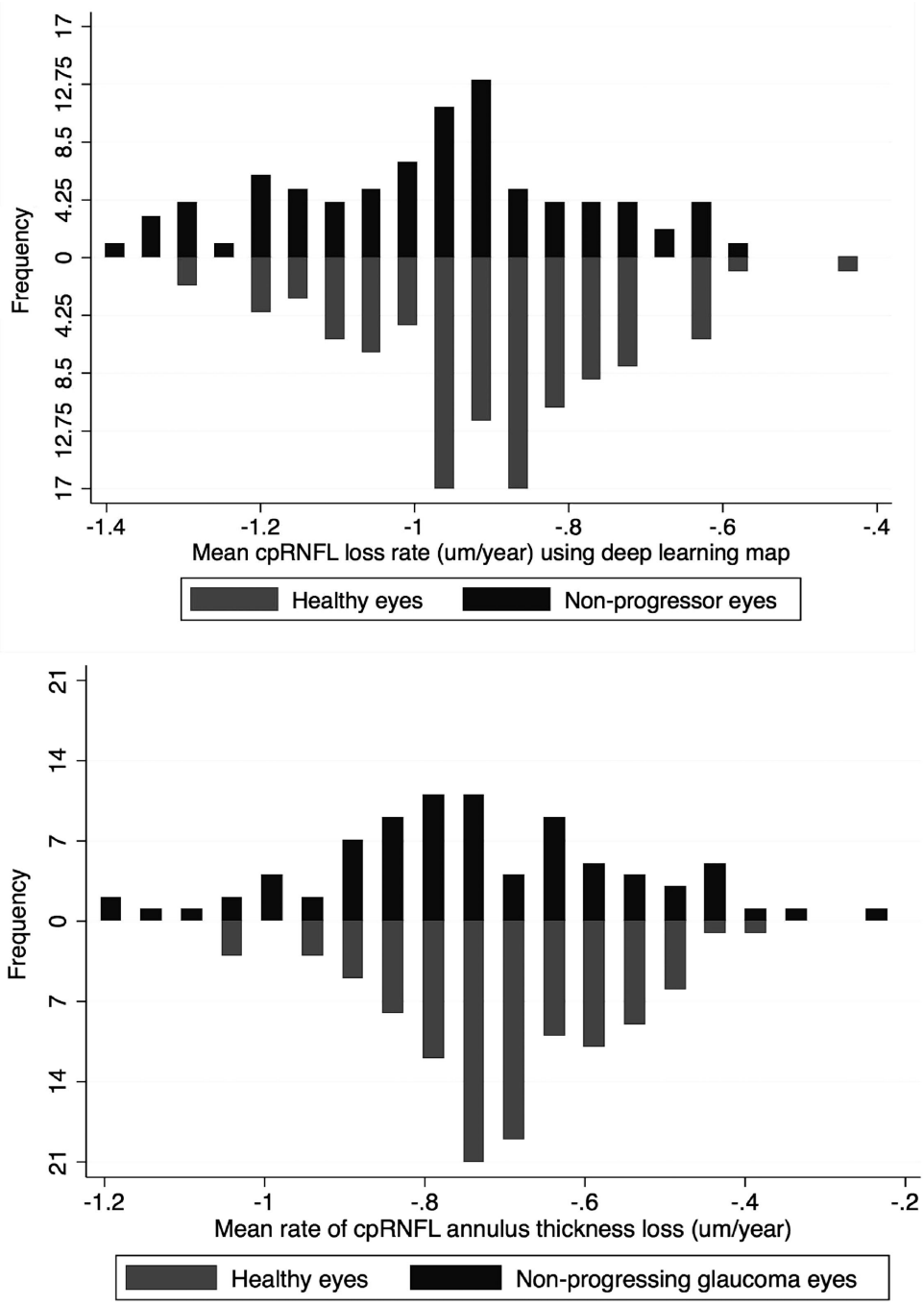


**Figure 6.** Mean RNFL loss rate between healthy and nonprogressing glaucoma eyes using the deep learning auto encoder region of interest map (top) and the cpRNFL annulus thickness map (bottom).

for the DL-AE ROI compared with the cpRNFL annulus (22 eyes [sensitivity, 0.50; 95% CI, 0.43–0.59] vs 10 eyes [sensitivity, 0.22; 95% CI, 0.10–0.34], respectively). The specificity, or number of correctly identified nonprogressing glaucoma eyes, for the DL-AE ROI was similar to the cpRNFL annulus (105 eyes [specificity, 0.70; 95% CI, 0.52–0.78] vs 102 eyes [specificity, 0.68; 95% CI, 0.54–0.81], respectively).

After 2 years of follow-up, using the RNFL rate of change criterion, sensitivity was higher for the DL-AE ROI compared with the cpRNFL annulus (33 eyes [sensitivity, 0.76; 95% CI, 0.64–0.82] vs 22 eyes [sensitivity, 0.51; 95% CI, 0.37–0.59], respectively). The specificity for the DL-AE ROI was similar to the cpRNFL annulus (124 eyes [specificity, 0.83; 95% CI, 0.76–0.94] vs 122 eyes [specificity, 0.81; 95% CI, 0.74–0.92], respectively).

The rates of RNFL loss stratified by follow-up time in healthy, nonprogressing glaucoma eyes and progressing glaucoma eyes for DL-AE ROI and the cpRNFL annulus are shown in Table 4.

Regarding model performance and disease severity, when eyes were divided into early glaucoma (VF MD of −6 dB or greater, 25 progressing eyes and 91 nonprogressing eyes) and moderate to advanced glaucoma with (VF MD of less than −6 dB, 19 progressing eyes and 59 nonprogressing eyes) patterns of results were similar across groups using both analysis strategies.

Using the RNFL rate of change criterion the sensitivity for identifying PGON eyes was higher for the DL-AE ROI compared with the cpRNFL annulus in both early glaucoma and moderate to advanced glaucoma. In early glaucoma eyes, sensitivity was 0.89 (95% CI, 0.78–0.94) for the DL-AE ROI compared with 0.58 (95% CI, 0.40–0.66) for the cpRNFL annulus. In moderate to advanced glaucoma eyes sensitivity was 0.92 (95% CI, 0.81–0.97) for the DL-AE ROI compared with 0.67 (95% CI, 0.52–0.74) for the cpRNFL annulus. The specificity for identifying nonprogressing glaucoma eyes was similar using both methods. In early glaucoma eyes specificity was 0.93 (95% CI, 0.84–0.99) for the DL-AE ROI compared with 0.94 (95% CI, 0.86–0.99) for the cpRNFL annulus. In moderate to advanced glaucoma eyes sensitivity was 0.88 (95% CI, 0.79–0.94) for DL-AE ROI compared with 0.89 (95% CI, 0.79–0.95) for the cpRNFL annulus.

## Discussion

This study demonstrated that using a DL-AE to generate an individualized RNFL thickness change map for each eye, more known progressing eyes could be identified (increased sensitivity) compared with using average cpRNFL thickness measurements in all eyes and when eyes were dichotomized into early and moderate to advanced disease, while maintaining similar specificity. In addition, observed rates of change in RNFL thickness per year were faster (i.e., greater) in progressing glaucoma eyes and nonprogressing glaucoma eyes using DL-AE ROI compared with cpRNFL annulus thickness. Finally, significant change over time was detected in not only progressing and nonprogressing (defined by stereophotograph assessment) glaucoma eyes, but also in healthy eyes.

Specifically, we formalized the change detection task as an abnormality detection task where we trained the model on nonglaucomatous changes (e.g., aging, variability) leading to a high reconstruction error with the DL-AE when the observed change was greater than what the model observed during training. Then, we used a Markov-based segmentation algorithm to classify each pixel of the RNFL thickness map into three regions: a no change region, a likely progression region, and a not likely progression region. The global mean RNFL rate of loss in the likely change region was calculated and compared with the aging RNFL loss from healthy eyes. To be considered as progressing, these eyes were required to have statistically significant slopes of change in estimated cpRNFL loss that were faster than average age-related losses obtained from longitudinal follow-up of a group of healthy eyes. Moreover, the results of this study not only help our understanding of the different glaucoma progression patterns, they also highlight the variability in patterns of glaucoma change and the limitations of using a standard global parameter to calculate the rate of loss. Overall, we identified 40% more eyes with glaucomatous progression using the DL-AE ROI model compared with the average cpRNFL annulus thickness derived from the same optic disc cube scan patterns.

Several other studies have demonstrated that using an individualized region of interest to monitor glaucomatous progression outperforms instrument defined global cpRNFL thickness measurements.[19–21] Hood et al.[19] and Thenappan et al.[20] demonstrated that the width of an ROI, defined using the 1% lower normative limits of cpRNFL thickness, increased significantly over time in eyes where the global cpRNFL did not exhibit a statistically significant decrease. Regarding comparison to results presented herein, Hood et al.[19] required glaucomatous eyes to have observable disc hemorrhages, which makes the reported results less generalizable to the current results. In contrast with Thenappan et al.,[20] where only an automatically outlined ROI based on objective criteria of RNFL loss compared with normative databases was used to detect glaucomatous changes, Wu et al.[21] used both automatically and manually outlined ROIs to define progressive changes, because an automatic ROI approach may not always capture an appropriate ROI. The manual ROI approach involved subjectively defining areas of glaucomatous defect in the second of two longitudinal scans observed on the OCT fundus projection images, OCT en face slab images, OCT RNFL thickness plots, OCT RNFL thickness deviation plots, OCT RNFL circle b-scan images, and OCT cpRNFL thickness profile plots (modified TSNIT plot) and comparing these areas with baseline images post hoc (graders were masked to information available in the baseline images). The manual ROI approach was more efficient for detecting progressive cpRNFL loss compared with

the automatic ROI approach and global cpRNFL thickness measurements. In the current study, we used an automated approach that used DL-AEs to identify the ROI with the highest probability of change with similar results.

The limitations of the current study include the relatively small sample size and the limited age range of the healthy participants. For these reasons, this method may not be generalizable to other populations. Moreover, we only used healthy and nonprogressing eyes to define the ROI because of the small number of progressing eyes available. If more progressing eyes were available, we could use the DL-AE for feature extraction to classify an eye into progressing versus nonprogressing groups. In addition, the current study did not include a large number of advanced glaucoma eyes in which it is challenging to detect cpRNFL thinning owing to the presence of a measurement floor (e.g.,[22–26]). Including more advanced glaucoma eyes in a follow-up study would increase confidence in the current method to detect change across the full glaucoma severity continuum. However, it is possible that the current method is best applied to detecting progression in early to moderate disease (given the more diffuse nature of advanced structural loss), which would not greatly decrease its clinical usefulness.

It is also possible that progression close to the disc margin was missed using both progression detection methods. Because variability influences the ability to detect progression, we measured both the DL-AE regions of interest and average cpRNFL thickness in the parapapillary region where measurement variability is decreased. To confirm a decrease in variability in the parapapillary region we calculated the coefficient of variation (CV) of RNFL thickness in the stable glaucoma group within the 2.22-mm to 3.45-mm annulus currently used and within a 1.0-mm to 2.22-mm annulus (i.e., closer to the optic disc margin). The results showed that the global, superior, temporal, inferior, and nasal CV were higher within the 1.0-mm to 2.22-mm annulus (CV, 2.45, 2.91, 2.72, 3.13, and 3.41, respectively) compared with the 2.22-mm to 3.45-mm annulus (CV, 1.31, 1.76, 1.69, 1.92, and 2.01, respectively). The higher RNFL thickness variability closer to the optic disc would limit the ability of the autoencoder, as well as other RNFL thickness-based methods to detect small changes in this region

Finally, it may be the case that false-positive progression classifications decreased the specificity of progression detection in our study. In the current study, progression was defined if the rate of RNFL loss was significantly different from zero and faster or greater than the fifth percentile of the healthy eyes. A recent study using computer simulation suggested that using this criterion for progression may result in approximately a 10% to 20% false-positive progression events (defined using trend based analysis of average OCT RNFL thickness measurements) in healthy eyes over 4 years depending on frequency of testing.[27] However, we report a relatively high level of specificity (i.e., likely not decreased significantly owing to a significant number of false positives) for both progression methods investigated (92% and 93%). Even if false positives owing to the criteria chosen to define progression contributed to our results, the comparison between the DL-AE ROI and cpRNFL annulus measurement-based progression rates is valid because results using both methods theoretically would be affected similarly. In addition, it is possible that the less than perfect specificity could be the result of an improved change detection method detecting actual disease-related change in nonprogressing glaucoma eyes that was not observed by subjective stereophotograph assessment.

In conclusion, by tailoring the analysis to the individual patient, individualized ROIs identified using unsupervised DL-AE analysis of OCT images show promise for improving assessment of glaucomatous progression. These results clearly support the usefulness of automated, patient-specific monitoring of glaucomatous progression.

## Acknowledgments

Supported by the National Institutes of Health/National Eye Institute, Bethesda, Maryland: R21 EY027945, T32 EY026590, R01 EY026574, R01 EY019869, R01 EY027510, P30 EY022589, R01 EY029058 and an unrestricted grant from Research to Prevent Blindness, New York, New York.

Disclosure: **C. Bowd**, None; **A. Belghith**, None; **M. Christopher**, None; **M.H. Goldbaum**, None; **M.A. Fazio**, Heidelberg Engineering (F); **C.A. Girkin**, Heidelberg Engineering (F); **J.M. Liebmann**, Aerie Pharmaceuticals (C), Alcon (C), Allergan (C), Bausch & Lomb (C, F), Carl Zeiss Meditec (C, F), Eyenovia (C), Galimedix Therapeutics (C), Heidelberg Engineering (C, F), Optovue (F), Novartis Pharmaceuticals (C), Reichert (C), Topcon (F); **C.G. de Moraes**, Belite Bio (C), Carl Zeiss Meditec (C), Galimedix Therapeutics (C), Heidelberg Engineering (F), Novartis (C), Perfuse Therapeutics (C), Topcon Medical Systems (F); **R.N. Weinreb**, Aerie Pharmaceuticals (C), Allergan (C), Bausch & Lomb (C, F), Carl Zeiss Meditec (F),

Centervue (F), Eyenovia (C), Heidelberg Engineering (F), Implantdata (C), Konan Medical (F), Novartis (C), Optovue (F), Research to Prevent Blindness (F); **L.M. Zangwill**, Carl Zeiss Meditec (F), Heidelberg Engineering (F), Optovue (F), Topcon Medical Systems (F)

## References


1. Vianna JR, Chauhan BC. How to detect progression in glaucoma. *Prog Brain Res.* 2015;221:135–158.
2. Gong M, Zhao J, Liu J, Miao Q, Jiao L. Change detection in synthetic aperture radar images based on deep neural networks. *IEEE Trans Neural Netw Learn Syst.* 2016;27(1):125–138.
3. Khan SH, He X, Porikli F, Bennamoun M. Forest change detection in incomplete satellite images with deep neural networks. *IEEE Trans Geosci Remote Sens.* 2017;55(9):5407–5423.
4. Zhang H, Gong M, Zhang P, Su L, Shi J. Feature-level change detection using deep representation and feature change analysis for multispectral imagery. *IEEE Geosci Remote Sens.* 2016;13(11):1666–1670.
5. Phan S, Satoh S, Yoda Y, et al. Evaluation of deep convolutional neural networks for glaucoma detection. *Jpn J Ophthalmol.* 2019;63(3):276–283.
6. Liu H, Li L, Wormstone IM, et al. Development and Validation of a Deep Learning System to Detect Glaucomatous Optic Neuropathy Using Fundus Photographs. *JAMA Ophthalmology.* 2019;137(12):1353–1360.
7. Diaz-Pinto A, Morales S, Naranjo V, Kohler T, Mossi JM, Navea A. CNNs for automatic glaucoma assessment using fundus images: an extensive validation. *Biomed Eng Online.* 2019;18(1):29.
8. Christopher M, Bowd C, Belghith A, et al. Deep learning approaches predict glaucomatous visual field damage from OCT optic nerve head en face images and retinal nerve fiber layer thickness maps. *Ophthalmology.* 2019;127(3):346–356.
9. Christopher M, Belghith A, Bowd C, et al. Performance of deep learning architectures and transfer learning for detecting glaucomatous optic neuropathy in fundus photographs. *Sci Rep.* 2018;8(1):16685.
10. An GZ, Omodaka K, Hashimoto K, et al. Glaucoma diagnosis with machine learning based on optical coherence tomography and color fundus images. *J Healthc Eng.* 2019;2019:4061313.
11. Gong M, Yang H, Zhang P. Feature learning and change feature classification based on deep learning for ternary change detection in SAR images. *ISPRS J Photogramm Remote Sens.* 2017;129:212–225.
12. Yuan X, Huang B, Wang Y, Yang C, Gui W. Deep learning-based feature representation and its application for soft sensor modeling with variable-wise weighted SAE. *IEEE Trans Industr Inform.* 2018;14(7):3235–3243.
13. Sample PA, Girkin CA, Zangwill LM, et al. The African Descent and Glaucoma Evaluation Study (ADAGES): design and baseline data. *Arch Ophthalmol.* 2009;127(9):1136–1145.
14. Medeiros FA, Zangwill LM, Bowd C, Sample PA, Weinreb RN. Use of progressive glaucomatous optic disk change as the reference standard for evaluation of diagnostic tests in glaucoma. *Am J Ophthalmol.* 2005;139(6):1010–1018.
15. Belghith A, Bowd C, Medeiros FA, et al. Does the location of Bruch's membrane opening change over time? Longitudinal analysis using San Diego Automated Layer Segmentation Algorithm (SALSA) stability of Bruch's membrane opening location over time. *Invest Ophthalmol Vis Sci.* 2016;57(2):675–682.
16. Chollet F. Building autoencoders in keras. Available at: https://blogkerasio/building-autoencoders-in-kerashtml. 2019. Accessed.
17. Chollet F. Keras. Available at: https://githubcom/fchollet/keras. 2019. Accessed October, 2019.
18. Belghith A, Bowd C, Medeiros FA, Balasubramanian M, Weinreb RN, Zangwill LM. Glaucoma progression detection using nonlocal Markov random field prior. *J Med Imaging (Bellingham).* 2014;3(1):1–9.
19. 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.
20. Thenappan A, De Moraes CG, Wang DL, et al. Optical coherence tomography and glaucoma progression: a comparison of a region of interest approach to average retinal nerve fiber layer thickness. *J Glaucoma.* 2017;26(5):473–477.
21. Wu Z, Thenappan A, Weng DSD, Ritch R, Hood DC. Detecting glaucomatous progression with a region-of-interest approach on optical coherence tomography: a signal-to-noise evaluation. *Transl Vis Sci Technol.* 2018;7(1):19.
22. Bowd C, Zangwill LM, Weinreb RN, Medeiros FA, Belghith A. Estimating optical coherence tomography structural measurement floors to improve detection of progression in advanced glaucoma. *Am J Ophthalmol.* 2017;175:37–44.

23. Hood DC, Kardon RH. A framework for comparing structural and functional measures of glaucomatous damage. *Prog Retin Eye Res.* 2007;26(6):688–710.
24. Moghimi S, Bowd C, Zangwill LM, et al. Measurement floors and dynamic ranges of OCT and OCT angiography in glaucoma. *Ophthalmology.* 2019;126(7):980–988.
25. Mwanza JC, Budenz DL, Warren JL, et al. Retinal nerve fibre layer thickness floor and corresponding functional loss in glaucoma. *Br J Ophthalmol.* 2015;99(6):732–737.
26. Mwanza JC, Kim HY, Budenz DL, et al. Residual and dynamic range of retinal nerve fiber layer thickness in glaucoma: comparison of three OCT platforms. *Invest Ophthalmol Vis Sci.* 2015;56(11):6344–6351.
27. Wu ZC, Saunders LJ, Zangwill LM, Daga FB, Crowston JG, Medeiros FA. Impact of normal aging and progression definitions on the specificity of detecting retinal nerve fiber layer thinning. *Am J Ophthalmol.* 2017;181:106–113.