



Agreement Between Trend-Based and Qualitative Analysis of the Retinal Nerve Fiber Layer Thickness for Glaucoma Progression on Spectral-Domain Optical Coherence Tomography

Atalie C. Thompson · Ang Li · Sanjay Asrani

Received: April 13, 2021 / Accepted: May 27, 2021 / Published online: July 1, 2021
© The Author(s) 2021

ABSTRACT

Introduction: To evaluate the agreement between trend-based analysis and qualitative assessment of the retinal nerve fiber layer (RNFL) thickness for glaucomatous progression on spectral-domain optical coherence tomography (SDOCT).

Methods: Retrospective review of 190 eyes from 103 patients with glaucoma or suspected glaucoma that underwent SDOCT imaging during four consecutive clinic visits. Trend-based progression was characterized by a significantly negative slope. Progression by qualitative analysis was determined by review of raw SDOCT B-scans.

Results: The slope was significantly greater in those with progression than without progression for both trend-based and qualitative analysis ($p < 0.001$). However, the qualitative grading classified a significantly greater proportion of eyes as progressing compared to trend-based analysis in both the superotemporal (ST) (23.2% vs. 10.5%, $p = 0.001$) and inferotemporal (IT) RNFL (27.4% vs 8.4%, $p < 0.001$). The trend-based and qualitative

classifications of progression showed poor agreement in both the ST ($\kappa = 0.0135$) and IT RNFL ($\kappa = 0.1222$). The agreement between trend-based and qualitative analysis was lower for eyes with artifacts (ST = 58.11%; IT = 68.7%) than those without artifacts (ST = 80.2%; IT = 74.8%). Moreover, among eyes with artifacts, there was no significant difference in slope between those qualitatively categorized as progressing versus not progressing ($p > 0.05$).

Conclusions: Poor agreement was found between a trend-based and qualitative analysis of change in RNFL on SDOCT. Careful qualitative review of SDOCT imaging may identify specific areas of glaucoma progression not captured by trend-based methods, especially in the presence of artifacts. Such an approach may also prove useful for detecting glaucoma progression in a clinical setting when there are few data points available.

Keywords: Glaucoma progression; Qualitative; Retinal nerve fiber layer; Spectral-domain optical coherence tomography; Trend-based

A. C. Thompson · S. Asrani (✉)
Department of Ophthalmology, Duke University,
Box 3802, Durham, NC 27710, USA
e-mail: sanjay.asrani@duke.edu

A. Li
Cleveland Clinic, Cleveland, OH, USA

Key Summary Points

Why carry out this study?

Artifacts in automated segmentation may be present on 20–50% of spectral-domain optical coherence tomography (SDOCT) scans of the retinal nerve fiber layer (RNFL), and have the potential to adversely impact the accuracy of trend-based analysis for glaucoma progression detection.

Qualitative review can assess for the presence of such artifacts and enable clinicians to correct errors as well as evaluate whether progression has truly occurred.

The objective of this study was to determine whether trend-based analysis agrees with qualitative assessment of the RNFL profile and raw B-scan from SDOCT, and to identify whether artifacts may potentially explain any disagreement.

What was learned from the study?

Glaucoma progression by qualitative versus trend-based SDOCT analysis showed poor agreement.

More eyes were detected as progressing by qualitative than by trend-based analysis, particularly in the presence of artifacts.

Careful qualitative review of SDOCT imaging may identify specific areas of glaucoma progression not captured by trend-based methods, especially in the presence of artifacts.

DIGITAL FEATURES

This article is published with digital features, including a summary slide, to facilitate understanding of the article. To view digital features

for this article go to <https://doi.org/10.6084/m9.figshare.14680812>.

INTRODUCTION

Glaucoma is a progressive optic neuropathy characterized by loss of retinal ganglion cells that results in specific patterns of thinning of the retinal nerve fiber layer (RNFL) and neuroretinal rim over time [1]. In early disease stages, such changes in the structure of the optic nerve precede detectable visual field loss [2, 3]. Early intervention by reduction of intraocular pressure may help to prevent vision loss from glaucomatous progression [4, 5].

Spectral-domain optical coherence tomography (SDOCT) has become the de facto technique for analyzing glaucomatous structural changes in the RNFL, due to its reliability and reproducibility [2, 4, 6]. In particular, if the superotemporal (ST) or inferotemporal (IT) circumpapillary RNFL thickness is low relative to normative values, suspicion for early glaucoma is raised [7–10]. Although defining early glaucoma based on characteristic changes in the RNFL on SDOCT has become routine in clinical practice, there are still no agreed-upon guidelines for detection of glaucomatous progression over time.

Numerous clinical studies have used trend-based analysis to define glaucoma progression, in which a statistically significant and negative slope drawn across consecutive follow-up visits is considered evidence of change [11–15]. While some commercially available software programs provide a linear trend line for the change in the average measures of RNFL over time, some clinicians will review the output from the automated segmentation to determine whether progression has occurred over consecutive visits. Automated segmentation software provides a two-dimensional linear profile of the circum-papillary RNFL with a change analysis that compares the current thickness to the prior or baseline visit.

However, artifacts can impact the accuracy of segmentation and may be present on 20–50% of OCT images [16–18]. Thus, review of the raw OCT B-scan is important when assessing for the

presence of progression so that artifacts can be identified and accounted for. Qualitative assessments that incorporate review of the automated segmentation profile as well as the raw B-scan may possibly be better able to detect glaucomatous progression than trend-based methods that assume that the quantification of RNFL thickness is correct. Moreover, such qualitative interpretations may be particularly useful in a clinical setting where the number of time points available for making a clinical decision are much shorter than what is optimal for trend-based analysis.

The objective of this study was to determine whether trend-based analysis agrees with qualitative assessment of the RNFL profile and raw B-scan from SDOCT, and to identify whether artifacts may potentially explain any disagreement.

METHODS

This was a retrospective chart review of consecutive patients aged 18 years or older with glaucoma or suspected glaucoma diagnosis who presented to a tertiary referral center over a 1-month period between June 1 and June 30, 2019. The study was approved by the Duke University Institutional Review Board with a waiver of informed consent due to the retrospective nature of the study. The protocol adhered to the tenets of the Declaration of Helsinki and was compliant with the Health Insurance Portability and Accountability Act of 1996 (HIPAA).

All included eyes underwent Spectralis SDOCT (version 5.4.7.0.; Heidelberg Engineering) imaging of the circumpapillary RNFL during the presenting visit as well as during at least three prior visits that were spaced from 6 to 12 months apart per the clinical practice patterns. If any SDOCT scans were obtained within 3 months of intraocular surgery or if the patient had a history of uveitis, the eye was excluded. Quality of the scans was not considered in the exclusion criteria; however, all scans had been previously reviewed by the attending (S.A.) at the time of image acquisition and repeated if they were of very poor quality per the clinic's

protocol. The circumpapillary RNFL thickness scans were obtained using circular scans with a diameter of 12° around the optic nerve. The standard summary printout includes the straightened profile of the raw scan with delineated boundaries of the RNFL as well as the global mean and average sectoral thickness values of the six regions: temporal, nasal, superotemporal, superonasal, inferotemporal, and inferonasal. The sectoral average values for the superotemporal and inferotemporal thickness were recorded for each of the visits.

All images were reviewed by a single experienced grader (S.A.) for any evidence of change between visits. The reviewer was masked to the results of the trend-based analysis.

The raw SDOCT B-scans (without segmentation lines) of the peripapillary retina were examined for evidence of artifacts and segmentation errors. Since subtle changes in OCT RNFL thickness may be missed by automated segmentation software, careful qualitative inspection of the peripapillary scans was performed to determine whether true progression had occurred at any time point across the four scans. Qualitative (true) progression was defined as a decrease in the RNFL thickness based on inspection of the raw B-scans by examining all four scans using a flicker image method and a change (i.e. red) from the automated segmentation profile. The machine's software allowed for the flicker method to be applied to aligned raw images. A single assessment was provided for the entire series of four consecutive scans. If progression was not found during evaluation of the automated segmentation profile or the raw B-scan images, then this was considered negative for progression.

Baseline demographic and clinical data were also collected, including sex, race, age, baseline ST and IT RNFL thickness, and glaucoma type and stage based on International Classification of Diseases (ICD)-9 and ICD-10 codes.

Statistical Analysis

Progression was assessed and compared between the ST and IT RNFL sectors using quantitative trend-based analysis and qualitative subjective

analysis of the SDOCT B-scans. Trend-based progression was characterized by a linear slope that was statistically significantly negative ($p < 0.05$) by ordinary least squares regression (OLS) which was performed for each eye. Generalized estimating equations were used to compare the difference in the slope for those that progressed and did not progress by either definition of progression, i.e. trend-based analysis or qualitative analysis. In addition, the linear slopes from OLS were compared among those with and without progression by either the trend-based or the qualitative definition, stratified by the presence or absence of artifacts. The agreement between quantitative trend-based and qualitative definitions of progression was assessed using the kappa statistic. Finally, the difference in the proportion categorized as

progressing by trend-based versus qualitative analysis was determined with estimation of a bootstrapped 95% confidence interval to account for more than one eye in study subjects. All statistical analyses were conducted in Stata 15.1 (StataCorp). A p value < 0.05 was considered statistically significant.

RESULTS

A total of 190 eyes from 103 patients with glaucoma or suspected glaucoma diagnoses were included in the study. Table 1 shows the baseline demographic and clinical characteristics of the study cohort. The mean age at baseline was 72.3 years (range 36–100). The median follow-up times at the second, third, and fourth clinic visits were 11, 20, and 30 months, respectively. Of 190 eyes, 5 (2.6%) were suspected of having glaucoma, and 64 (33.7%) had mild, 80 (42.1%) moderate, and 41 (21.6%) severe glaucoma. The most common diagnosis was open-angle glaucoma (50.5%), followed by normal-tension (18.95%) and chronic angle-closure glaucoma (17.9%).

Table 2 displays the slope of change in RNFL in the ST or IT sectors by each method. The slope in those with and without progression based on trend-based analysis was -3.33 ± 2.36 vs. -0.41 ± 3.32 $\mu\text{m}/\text{year}$ in the ST RNFL ($p < 0.001$) and -4.16 ± 2.25 vs. -0.94 ± 3.51 $\mu\text{m}/\text{year}$ in the IT RNFL ($p < 0.001$). The slope in those with and without progression based on qualitative assessment of the raw B-scan was -2.35 ± -0.22 vs. -0.22 ± 2.62 $\mu\text{m}/\text{year}$ in the ST RNFL ($p < 0.001$) and -2.68 ± 4.29 vs. -0.66 ± 3.03 $\mu\text{m}/\text{year}$ in the IT RNFL ($p < 0.001$).

Artifacts were present in 38.9% (74/190) of the ST RNFL and 35.3% (67/190) of the IT RNFL eyes with glaucoma or suspected glaucoma diagnoses. Among those eyes with artifacts, the slope in those with and without progression based on trend-based analysis was -3.31 ± 1.52 vs. -0.40 ± 2.36 $\mu\text{m}/\text{year}$ (GEE $P < 0.001$) in the ST RNFL and -5.13 ± 2.83 vs. -0.19 ± 2.85 $\mu\text{m}/\text{year}$ (GEE $P < 0.001$) in the IT RNFL (Table 3A). Among those eyes without artifacts, the slope in those with and without progression based on

Table 1 Baseline demographic and clinical characteristics

Age, years, mean (SD)	72.2 (13.4)
Female sex, n/N (%)	57/103 (55.3%)
SDOCT RNFL, mean μm (SD)	
ST RNFL sector	89.8 (28.7)
IT RNFL sector	87.3 (31.0)
Glaucoma stage, n/N (%)	
Suspected	5/190 (2.63%)
Mild	64/190 (33.68%)
Moderate	80/190 (42.11%)
Severe	41/190 (21.58%)
Glaucoma type, n/N (%)	
Open-angle glaucoma suspected	2/190 (1.05%)
Narrow-angle glaucoma suspected	3/190 (1.58%)
Open-angle glaucoma	96/190 (50.5%)
Normal-tension glaucoma	36/190 (18.95%)
Chronic angle-closure glaucoma	34/190 (17.9%)
Secondary glaucoma	19/190 (10.0%)

SDOCT spectral-domain optical coherence tomography, RNFL retinal nerve fiber layer, ST superotemporal, IT inferotemporal, SD standard deviation

Table 2 Comparison of RNFL slope ($\mu\text{m}/\text{year}$) in eyes with and without progression by either trend-based analysis or qualitative analysis

	Trend-based analysis		GEE <i>P</i> value	Qualitative analysis		GEE <i>P</i> value
	With progression	Without progression		With progression	Without progression	
ST sector	-3.33 ± 2.36	-0.41 ± 3.32	< 0.001	-2.35 ± -0.22	-0.22 ± 2.62	< 0.001
IT sector	-4.16 ± 2.25	-0.94 ± 3.51	< 0.001	-2.68 ± 4.29	-0.66 ± 3.03	< 0.001

RNFL retinal nerve fiber layer, *GEE* generalized estimating equations, *ST* superotemporal, *IT* inferotemporal

Table 3 A Comparison of RNFL slope ($\mu\text{m}/\text{year}$) in eyes with and without progression by trend-based analysis and by qualitative analysis in SDOCT scans with artifacts. **B** Comparison of RNFL slope ($\mu\text{m}/\text{year}$) in eyes with and without progression by trend-based analysis and by qualitative analysis in SDOCT scans without artifacts

A						
	Trend-based analysis		GEE <i>P</i> value	Qualitative analysis		GEE <i>P</i> value
	With progression	Without progression		With progression	Without progression	
ST sector	-3.31 ± 1.52	-0.40 ± 2.36	< 0.001	-0.49 ± 3.05	-0.95 ± 2.13	0.473
IT sector	-5.13 ± 2.83	-0.19 ± 2.85	< 0.001	-0.14 ± 3.40	-0.62 ± 2.95	0.629

B						
	Trend-based analysis		GEE <i>P</i> value	Qualitative analysis		GEE <i>P</i> value
	With progression	Without progression		With progression	Without progression	
ST sector	-3.35 ± 3.08	-0.41 ± 3.80	0.005	-4.80 ± 5.5	0.15 ± 2.8	< 0.001
IT sector	-3.84 ± 2.06	-1.37 ± 3.8	< 0.001	-4.15 ± 4.1	-0.68 ± 3.1	< 0.001

RNFL retinal nerve fiber layer, *GEE* generalized estimating equations, *ST* superotemporal, *IT* inferotemporal

trend-based analysis was -3.35 ± 3.08 vs. $-0.41 \pm 3.80 \mu\text{m}/\text{year}$ (*GEE P* = 0.005) in the ST sector and -3.84 ± 2.06 vs. $-1.37 \pm 3.78 \mu\text{m}/\text{year}$ (*GEE P* < 0.001) in the IT sector (Table 3B).

Among those eyes with artifacts, the slope in those with and without progression based on qualitative analysis was -0.49 ± 3.05 vs. $-0.95 \pm 2.13 \mu\text{m}/\text{year}$ (*GEE P* = 0.473) in the ST quadrant and -0.14 ± 3.40 vs. $-0.62 \pm 2.95 \mu\text{m}/\text{year}$ (*GEE P* = 0.629) in the IT quadrant (Table 3A). Among those eyes without artifacts, the slope in those with and without progression based on qualitative analysis was -4.80 ± 5.54 vs. $0.15 \pm 2.78 \mu\text{m}/\text{year}$ (*GEE P* < 0.001) in the ST quadrant and -4.15 ± 4.10 vs. $-0.68 \pm 3.09 \mu\text{m}/\text{year}$ (*GEE P* < 0.001) in the IT quadrant (Table 3B).

Table 4 shows the difference in proportion as well as the agreement between trend-based and qualitative analysis for identifying glaucoma progression. The trend-based criteria classified 10.5% (20/190) of all eyes with progression in the ST RNFL, whereas the qualitative grading classified 23.2% (44/190) of all eyes as truly progressing (bootstrap *p* = 0.001). The agreement between the trend-based and qualitative classifications was 71.58% for the ST RNFL, with a low kappa of 0.0135 (*p* = 0.42). Similarly, in the IT RNFL, the trend-based criteria classified 8.4% (16/190) of all eyes with progression, whereas the qualitative grading classified 27.4% (52/190) of all eyes as truly progressing (bootstrap *p* < 0.001). The agreement between the trend-based and qualitative classifications for

Table 4 **A** Comparison of proportion of all eyes categorized as progressing by trend-based versus qualitative analysis. **B** Comparison of proportion of eyes with artifacts categorized as progressing by trend-based versus qualitative analysis. **C** Comparison of proportion of eyes without artifacts categorized as progressing by trend-based versus qualitative analysis

A						
	Trend-based analysis	Qualitative analysis	Difference in proportions	Agreement	Kappa	Kappa <i>P</i> value
ST sector	20/190 (10.5%)	44/190 (23.2%)	0.001	71.58%	0.0135	0.42
IT sector	16/190 (8.4%)	52/190 (27.4%)	< 0.001	72.36%	0.1222	0.02
B						
	Trend-based analysis	Qualitative analysis	Difference in proportions	Agreement	Kappa	Kappa <i>P</i> value
ST sector	10/74 (13.5%)	25/74 (33.8%)	0.003	58.11%	-0.0976	0.84
IT sector	4/67 (6.0%)	19/67 (28.4%)	< 0.001	68.7%	-0.0130	0.56
C						
	Trend-based analysis	Qualitative analysis	Difference in proportions	Agreement	Kappa	Kappa <i>P</i> value
ST sector	10/116 (8.6%)	19/116 (16.4%)	0.069	80.2%	0.11	0.11
IT sector	12/123 (9.8%)	33/123 (26.8%)	< 0.001	74.8%	0.20	0.005

ST superotemporal, IT inferotemporal

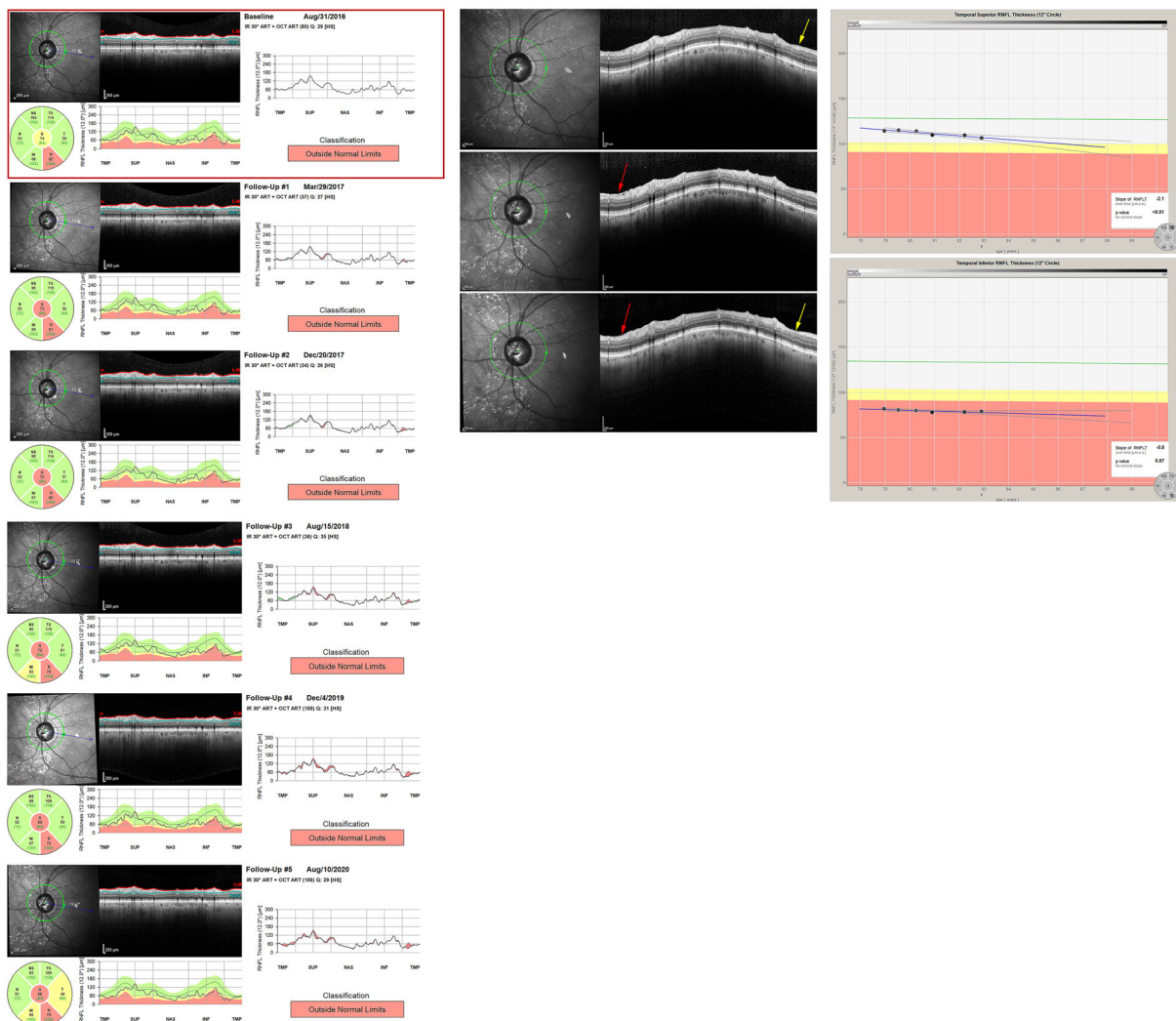


Fig. 1 The left panels show consecutive retinal nerve fiber layer change profiles over time which were generated by automated segmentation. The middle panels show the raw optical coherence tomography B-scan images. The right panels show the trend line and p value. In the left panels, the superotemporal area of red suggests possible progression by the change profile. However, close inspection of the raw B-scan in the middle panels reveals an epiretinal membrane that subsequently resolves (red arrows), suggesting that there is not true progression in the superotemporal quadrant according to qualitative analysis. The right superior panel demonstrates the trend line which is

significantly negative for the superotemporal quadrant ($p = 0.01$). Thus, the superotemporal quadrant is miscategorized with progression if relying on trend-based analysis from automated segmentation, for it is shown to have false progression by qualitative analysis. In the middle panels, the yellow arrows highlight an area of true progression in the inferotemporal quadrant that is also seen in the change profile in the left panels. However, the right bottom panel shows that the trend-based analysis was not significantly negative in the inferotemporal quadrant ($p = 0.07$). This was an example of a false negative for progression if relying on trend-based analysis

the IT RNFL was 72.36%, with a low kappa of 0.1222 ($p = 0.02$). Qualitative analysis detected a greater proportion of eyes as progressing than trend-based analysis among those with artifacts

than among those without artifacts (Table 4B, 4C). The agreement between trend-based and qualitative analysis was lower among those eyes

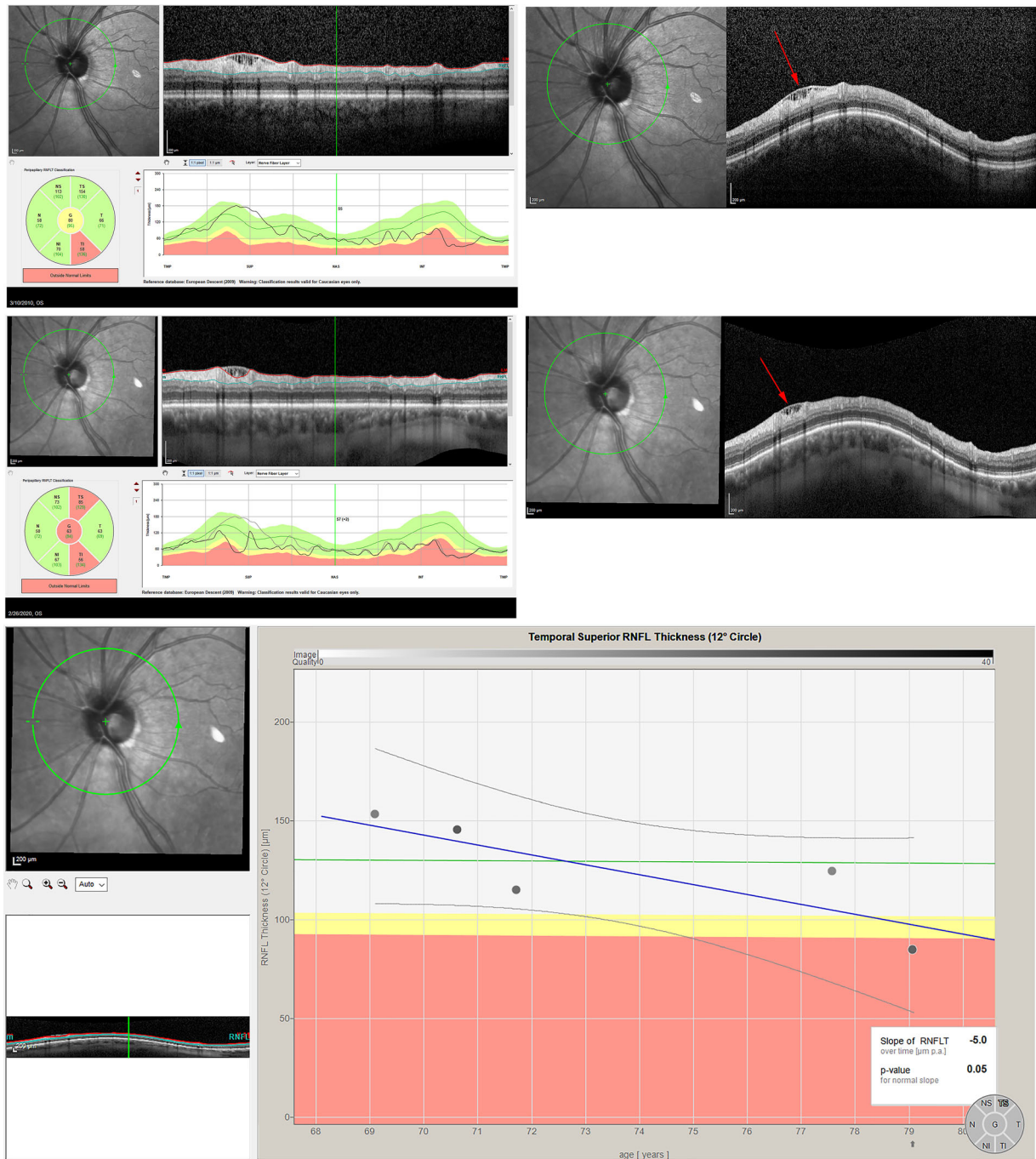


Fig. 2 The top left panels show a change in the profile of the superotemporal nerve fiber layer between the first and the most recent visit, concerning for possible progression. The bottom panel shows a significantly negative slope by trend-based

analysis. However, the raw optical coherence tomography B-scans in the upper right panels demonstrate that this change is due to a decrease in traction from the epiretinal membrane (red arrows), and is thus an example of false progression

with artifacts (ST 58.11%; IT 68.7%) than those without artifacts (ST 80.2%; 74.8% IT).

Figure 1 shows an example where the ST RNFL profile from automated segmentation

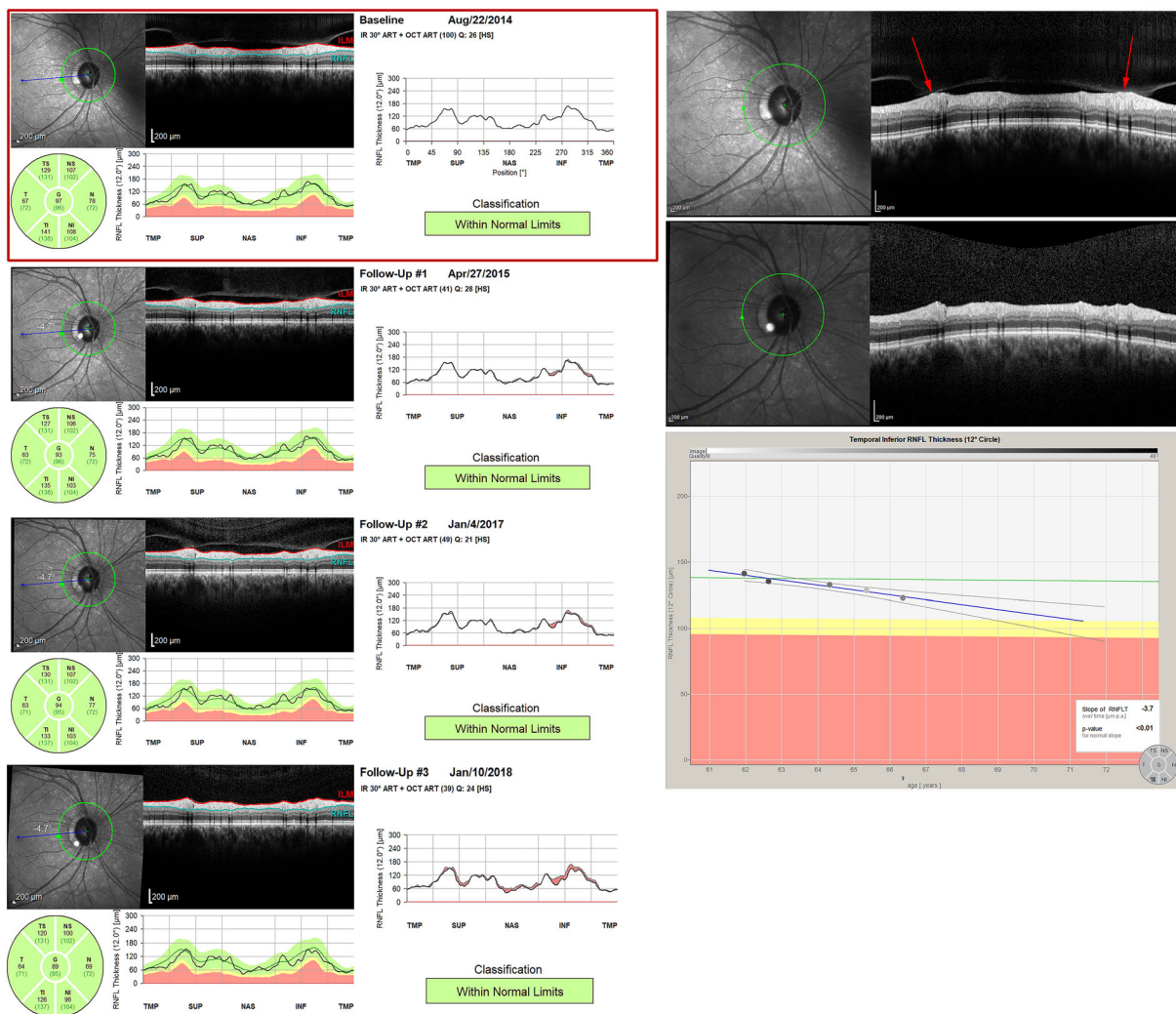


Fig. 3 The left panels show a red area of change in the retinal nerve fiber layer profile of the superotemporal and inferotemporal quadrants. The bottom right panel also shows a significantly negative slope by trend-based analysis.

suggests progression (red change) and has a significantly negative slope by trend-based analysis. However, close inspection of the raw OCT B-scan reveals that this is an example of false progression due to an epiretinal membrane that subsequently resolves. In the IT RNFL of the same eye, the slope is not significantly negative even though there is a real change in the RNFL on both the change profile and the raw OCT B-scan. Figure 2 also demonstrates false progression by trend-based analysis because of fluctuation in the ST RNFL thickness

However, qualitative assessment of the raw optical coherence tomography B-scans in the upper right panels demonstrate that the change is due to a release of vitreous traction (red arrows) and is thus an example of false progression

due to an epiretinal membrane. Figure 3 shows false progression that occurs in both the ST and IT RNFL due to release of vitreous traction.

DISCUSSION

Accurate detection of progression is critical to the timely management of glaucoma progression so that appropriate therapy can be initiated to slow or even halt the disease process. Over the past decade, OCT has risen to the forefront

of available imaging modalities for glaucoma not only because it is highly reproducible but also because it can identify subtle structural changes in the peripapillary RNFL that may precede corresponding detectable visual field loss [11]. Although SDOCT is the most widely adopted modality for monitoring structural changes in glaucoma, there is no consensus as to the best way to determine whether glaucoma progression has in fact occurred. Trend-based methods have been widely utilized to define progression in clinical studies. However, a trend line can be adversely impacted if the RNFL thickness measurements quantified by automated segmentation software are not correct. Artifacts are often present on SDOCT, and can adversely impact the precision of automated segmentation and thus the accuracy of quantified RNFL thickness measurements [16–18]. Application of qualitative methods can help to ensure the accuracy of SDOCT segmentation by inspecting the raw B-scan for artifacts. Moreover, qualitative assessment of the scan can afford detection of glaucoma progression even when there are only a small number of follow-up scans available, as is often the case following an intervention performed in clinical practice, whereas trend-based methods may be better suited to detection of progression when there is longer follow-up available. In our study, we found that there were significantly more eyes detected as progressing by qualitative than trend-based analysis. Some of this difference may be explained by the presence of artifacts, since qualitative methods can detect progression even in the presence of artifacts. Thus, application of qualitative review of the raw B-scan to SDOCT interpretation may improve the ability to accurately diagnose glaucoma progression.

Discrepancies in qualitative and quantitative methods of RNFL analysis may occur if qualitative methods are better suited to detecting patterns of change that are typical of glaucoma. Wu et al. [19] recently reviewed 409 eyes for glaucoma progression by a qualitative method for widefield OCT versus conventional quantitative global circumferential RNFL thickness. In their study, quantitative methods missed cases with characteristic patterns of glaucomatous

damage, whereas qualitative methods missed cases if there was a general reduction in RNFL thickness but not characteristic patterns. In other words, qualitative methods were more likely to detect patterns of glaucomatous progression that were missed when relying on global RNFL thickness parameters. Moreover, quantitative methods may detect false positives if a change in global RNFL occurred in the absence of a typical pattern of change.

Loss of global RNFL is not specific to glaucoma and can occur in other non-glaucomatous optic neuropathies as well as normal aging [15, 20]. Additionally, since we noticed that the most common cause of loss of global RNFL is posterior vitreous separation, especially in the nasal sectors, we did not evaluate the global RNFL in our study. Rather we focused on the superotemporal and inferotemporal sectors, as changes in these locations are most characteristic of glaucomatous progression, especially in early disease [7–10]. Nevertheless, we still found substantial disagreement between the quantitative trend-based and qualitative methods, with a poor kappa score in both the ST and IT RNFL sectors. Moreover, qualitative methods detected a significantly greater proportion of cases than trend-based methods ($p < 0.001$). In our study, one possible explanation for this disagreement is the presence of artifacts which can impact the quantitative values used for trend-based analysis. Traditional trend-based analysis relies on the values provided by automated segmentation of the RNFL under the assumption that they are accurate. However, artifacts have been widely reported in the literature across different SDOCT platforms. Poor-quality scans due to errors during image acquisition and subtle retinal pathologies such as epiretinal membranes and vitreomacular traction can contribute to segmentation errors [16–18, 21–25], even in healthy subjects [25, 26].

In our study, careful inspection of the raw B-scan images during qualitative review enabled us to identify artifacts that could affect the interpretation of progression if relying solely on the output from automated segmentation. Artifacts were present on over one third of the ST (38.9%; 74/190) and IT (35.3%; 67/190)

RNFL in eyes with glaucoma or suspected glaucoma diagnoses. Our study suggests that the poor agreement between trend-based and qualitative methods is at least partially explained by the presence of artifacts, since the agreement between the two methods was lower among eyes with artifacts (ST 58.1%; IT 68.7%) than eyes without artifacts (ST 80.2%; IT 74.8%), as is seen in Table 4B. This is because trend-based progression was a binary categorization defined by a significantly negative slope with a p value < 0.05 , regardless of the presence or absence of artifacts.

Application of a trend-based definition of progression could lead to cases of true progression being missed if the artifact led to artifactual thickening of the RNFL, thus masking the underlying progression. Even though the definition of progression by qualitative analysis was not based on the RNFL slope values, we also found that the slope was significantly more negative among eyes that progressed than those that did not progress by the qualitative criteria when there were no artifacts present. However, among eyes with artifacts, the slope did not differ significantly between those that progressed and those that did not progress by qualitative analysis. Such cases were miscategorized as non-progressors by the trend-based definition because the RNFL thickness slope from automated segmentation was not significantly negative due to the presence of artifact. This finding suggests that qualitative analysis was able to detect additional cases of progression in the presence of artifacts. This may help explain why significantly more cases were detected by qualitative methods than by trend-based methods. Similarly, in a few cases, the trend-based analysis overcalled progression because it only relied on a significantly negative slope, which could be due to artifactual changes rather than true progression.

Limitations

This study has several limitations including its retrospective design. In the absence of a true gold standard for determining progression on OCT, we cannot definitively know how many

true cases exist. Comparison to alternative structural criteria, such as optic disc photos, was not performed, as the grading of disc photos is also an inherently subjective process and may be less sensitive than OCT for glaucoma detection. Moreover, disc photos are rarely collected in clinical practice currently. Comparison to Humphrey visual field data would also not be useful for verifying true progression since changes in the OCT can precede changes in the visual field by months to years. Changes in OCT RNFL structure and standard automated perimetry (SAP) have been shown to have poor agreement for glaucomatous progression [11, 13]. SAP is also measured on a logarithmic scale, which makes it less sensitive to small changes and more difficult to use for verification of OCT changes [27].

Unlike the trend-based progression software in Cirrus, we did not average two RNFL scans to establish the baseline. Moreover, in this study the slopes and p values were estimated using RNFL data from the Heidelberg Spectralis, but the commercial software does not provide this analysis. The analysis was also based on only four time points, but it is possible that longer follow-up may show that trend analysis is able to detect more cases of progression since it may be better suited to data sets with a greater number of consecutive time points. The qualitative evaluation of the raw B-scan also has certain drawbacks since it was performed by a single expert grader (S.A.) rather than multiple graders, and can be prone to subjectivity. However, despite its subjective nature, qualitative assessment may offer unique advantages over trend-based analysis, since progression can be detected in the presence of artifacts or when there are only a small number of follow-up studies available, as is often the case in clinical practice. Moreover, in clinical practice we are often trying to determine whether progression has occurred over shorter time intervals rather than over multiple years of data. This is because with each diagnosis of progression and assessment of response to therapeutic intervention, the baseline RNFL should be reset to the time therapy was escalated. Thus, our study underscores that the qualitative approach may be beneficial when longer-term data are not

available, as is often the case in the clinical setting. As AI systems for progression are developed, training such algorithms on images labeled with qualitative assessments performed by expert graders may improve their accuracy. The fact that qualitative assessment showed higher detection rates than trend analysis may also have important implications for its application in randomized clinical trials, since it could decrease the required sample size and shorten the necessary duration of follow-up before detection of progression. Future analyses will include development and application of a rigorous subjective grading system by multiple graders, and assessment of the structure–function relationship by evaluation of visual field progression.

CONCLUSIONS

In conclusion, we found that there was poor agreement between qualitative and trend-based analysis for detection of glaucoma progression on SDOCT. Moreover, qualitative analysis detected a significantly greater proportion of eyes as progressing than did trend-based analysis, which may be partly explained by the ability to discriminate cases in the presence of artifacts through careful review of the SDOCT B-scan. Future studies may also find that trend-based approaches could be enhanced if combined with a qualitative review of the raw imaging to ensure accurate segmentation of the circum-papillary RNFL.

ACKNOWLEDGEMENTS

Funding. Dr. Thompson receives grant support from the American Glaucoma Society and an NEI 1-K23 EY030891-01. No funding or sponsorship was received for the publication of this article.

Authorship. All named authors meet the International Committee of Medical Journal Editors (ICMJE) criteria for authorship for this article, take responsibility for the integrity of

the work as a whole, and have given their approval for this version to be published.

Author's Contributions. All authors contributed to the study conception and design. Material preparation and data analysis was performed by Atalie Thompson. Data collection was performed by Ang Li and Sanjay Asrani. The first draft of the manuscript was written by Atalie Thompson and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

Prior Presentation. This study was presented at ARVO, May 1, 2021.

Disclosures. Dr. Asrani receives lecture honoraria from Heidelberg Engineering, Franklin, MA and from Sophia Laboratories, Mexico City. He is a consultant to Alievio. Dr. Thompson and Dr. Li have nothing to disclose. Dr. Thompson is now affiliated with Wake Forest Baptist Health, Department of Ophthalmology, Winston-Salem, NC, USA.

Compliance with Ethics Guidelines. The study was approved by the Duke University Institutional Review Board with a waiver of informed consent due to the retrospective nature of the study. The protocol adhered to the tenets of the Declaration of Helsinki and was compliant with HIPAA.

Data Availability. The data sets generated during and/or analyzed during the current study are available from the corresponding author on reasonable request.

Open Access. This article is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License, which permits any non-commercial use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons

licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc/4.0/>.

REFERENCES

- Weinreb RN, Aung T, Medeiros FA. The pathophysiology and treatment of glaucoma: a review. *JAMA*. 2014;311(18):1901–11.
- Hood DC, Kardon RH. A framework for comparing structural and functional measures of glaucomatous damage. *Prog Retin Eye Res*. 2007;26(6):688–710.
- Medeiros FA, Alencar LM, Zangwill LM, Bowd C, Sample PA, Weinreb RN. Prediction of functional loss in glaucoma from progressive optic disc damage. *Arch Ophthalmol (Chicago III: 1960)*. 2009;127(10):1250–6.
- Medeiros FA, Alencar LM, Zangwill LM, Sample PA, Weinreb RN. The Relationship between intraocular pressure and progressive retinal nerve fiber layer loss in glaucoma. *Ophthalmology*. 2009;116(6):1125–33.e1–3.
- Nouri-Mahdavi K, Caprioli J. Measuring rates of structural and functional change in glaucoma. *Br J Ophthalmol*. 2015;99(7):893–8.
- Medeiros FA, Zangwill LM, Alencar LM, Bowd C, Sample PA, Susanna R Jr, et al. Detection of glaucoma progression with stratus OCT retinal nerve fiber layer, optic nerve head, and macular thickness measurements. *Invest Ophthalmol Vis Sci*. 2009;50(12):5741–8.
- Jonas JB, Fernández MC, Stürmer J. Pattern of glaucomatous neuroretinal rim loss. *Ophthalmology*. 1993;100(1):63–8.
- Jonas JB, Mardin CY, Schlötzer-Schrehardt U, Naumann GO. Morphometry of the human lamina cribrosa surface. *Invest Ophthalmol Vis Sci*. 1991;32(2):401–5.
- Diniz-Filho A, Abe RY, Zangwill LM, Gracitelli CP, Weinreb RN, Girkin CA, et al. Association between intraocular pressure and rates of retinal nerve fiber layer loss measured by optical coherence tomography. *Ophthalmology*. 2016;123(10):2058–65.
- Radius RL. Regional specificity in anatomy at the lamina cribrosa. *Arch Ophthalmol (Chicago, III: 1960)*. 1981;99(3):478–80.
- Zhang X, Dastiridou A, Francis BA, Tan O, Varma R, Greenfield DS, et al. Comparison of glaucoma progression detection by optical coherence tomography and visual field. *Am J Ophthalmol*. 2017;184:63–74.
- Alencar LM, Zangwill LM, Weinreb RN, Bowd C, Sample PA, Girkin CA, et al. A comparison of rates of change in neuroretinal rim area and retinal nerve fiber layer thickness in progressive glaucoma. *Invest Ophthalmol Vis Sci*. 2010;51(7):3531–9.
- Leung CK, Liu S, Weinreb RN, Lai G, Ye C, Cheung CY, et al. Evaluation of retinal nerve fiber layer progression in glaucoma a prospective analysis with neuroretinal rim and visual field progression. *Ophthalmology*. 2011;118(8):1551–7.
- Leung CK, Chiu V, Weinreb RN, Liu S, Ye C, Yu M, et al. Evaluation of retinal nerve fiber layer progression in glaucoma: a comparison between spectral-domain and time-domain optical coherence tomography. *Ophthalmology*. 2011;118(8):1558–62.
- Wu Z, 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;181:106–13.
- Giani A, Cigada M, Esmaili DD, Salvetti P, Luccarelli S, Marziani E, et al. Artifacts in automatic retinal segmentation using different optical coherence tomography instruments. *Retina (Philadelphia, Pa)*. 2010;30(4):607–16.
- Ho J, Sull AC, Vuong LN, Chen Y, Liu J, Fujimoto JG, et al. Assessment of artifacts and reproducibility across spectral- and time-domain optical coherence tomography devices. *Ophthalmology*. 2009;116(10):1960–70.
- Moore DB, Jaffe GJ, Asrani S. Retinal nerve fiber layer thickness measurements: uveitis, a major confounding factor. *Ophthalmology*. 2015;122(3):511–7.
- Wu Z, Weng DSD, Rajshekhar R, Thenappan A, Ritch R, Hood DC. Evaluation of a qualitative approach for detecting glaucomatous progression using wide-field optical coherence tomography scans. *Transl Vis Sci Technol*. 2018;7(3):5.

20. Vianna JR, Danthurebandara VM, Sharpe GP, Hutchison DM, Belliveau AC, Shuba LM, et al. Importance of normal aging in estimating the rate of glaucomatous neuroretinal rim and retinal nerve fiber layer loss. *Ophthalmology*. 2015;122(12):2392–8.
21. Han IC, Jaffe GJ. Evaluation of artifacts associated with macular spectral-domain optical coherence tomography. *Ophthalmology*. 2010;117(6):1177–89.e4.
22. Chong GT, Lee RK. Glaucoma versus red disease: imaging and glaucoma diagnosis. *Curr Opin Ophthalmol*. 2012;23(2):79–88.
23. Liu Y, Simavli H, Que CJ, Rizzo JL, Tsikata E, Maurer R, et al. Patient characteristics associated with artifacts in Spectralis optical coherence tomography imaging of the retinal nerve fiber layer in glaucoma. *Am J Ophthalmol*. 2015;159(3):565–76.e2.
24. Asrani S, Essaid L, Alder BD, Santiago-Turla C. Artifacts in spectral-domain optical coherence tomography measurements in glaucoma. *JAMA Ophthalmol*. 2014;132(4):396–402.
25. Hwang YH, Lee JY, Kim YY. The effect of head tilt on the measurements of retinal nerve fibre layer and macular thickness by spectral-domain optical coherence tomography. *Br J Ophthalmol*. 2011;95(11):1547–51.
26. Kim NR, Lim H, Kim JH, Rho SS, Seong GJ, Kim CY. Factors associated with false positives in retinal nerve fiber layer color codes from spectral-domain optical coherence tomography. *Ophthalmology*. 2011;118(9):1774–81.
27. Medeiros FA, Zangwill LM, Bowd C, Mansouri K, Weinreb RN. The structure and function relationship in glaucoma: implications for detection of progression and measurement of rates of change. *Invest Ophthalmol Vis Sci*. 2012;53(11):6939–46.