tvst

Article

Matching Misaligned Spectralis OCTs to a Reference Scan in Pediatric Glaucoma with Poor Fixation and Nystagmus

John P. Kelly[1,2], Francine M. Baran[1,2], James O. Phillips[1,3], and Avery H. Weiss[2]

[1] Roger H. Johnson Vision Clinic, Seattle Children's Hospital, Division of Ophthalmology, Seattle, WA, USA
[2] University of Washington, Department of Ophthalmology, Seattle, WA, USA
[3] University of Washington School of Medicine, Department of Otolaryngology, Seattle, WA, USA

Correspondence: John P. Kelly, Roger H. Johnson Vision Clinic Seattle Children's Hospital, OA.5.342, Seattle, WA, 98105, USA. e-mail: john.kelly@seattlechildrens.org

Received: January 21, 2020
Accepted: August 25, 2020
Published: September 22, 2020

Keywords: optical coherence tomography; nystagmus; pediatric ophthalmology; optic nerve

Citation: Kelly JP, Baran FM, Phillips JO, Weiss AH. Matching misaligned spectralis OCTs to a reference scan in pediatric glaucoma with poor fixation and nystagmus. Trans Vis Sci Tech. 2020;9(10):21, https://doi.org/10.1167/tvst.9.10.21

Purpose: Poor fixation or nystagmus in children causes misalignment errors when measuring circumpapillary retinal nerve fiber layer (cpRNFL) thickness by simultaneous scanning laser ophthalmoscope imaging/optical coherence tomography (SLO/OCT). We investigated a method to assess cpRNFL from misaligned SLO/OCT scans.

Methods: Heidelberg Spectralis SLO/OCT scans from a single clinical examination were retrospectively analyzed when automated eye tracking was unreliable. Retinal layer thickness was measured at overlapping match locations between a reference and misaligned scans based on the position data from simultaneously acquired SLO images. Three layers were segmented: cpRNFL, internal limiting membrane to outer nuclear layer (ILM-ONL), and total retinal thickness (TR). Accuracy was defined as the difference in thickness between the reference and misaligned scans at their match locations after correction for scan angle.

Results: Thirty-five subjects, evaluated for glaucomatous nerve loss, met inclusion criteria. Group-averaged accuracy was −2.7, 1.4, and 0.3 µm for cpRNFL, ILM-ONL, and TR thickness, respectively. Across all layers, interobserver intraclass correlation coefficients ranged from 0.97 to 0.63 and the maximum Bland-Altman 95% limits of agreement were −21.6 to 20.7 µm. Variability was greatest for cpRNFL thickness and least for TR thickness. Increased variability was associated with lower signal-to-noise ratio but not with image-motion indices of shear, rotation, and scale.

Conclusions: Retinal layer thickness can be compared to a reference cpRNFL OCT scan when poor fixation and nystagmus causes misalignment errors. The analysis can be performed post hoc using multiple misaligned scans from standard SLO/OCT protocols.

Translational Relevance: Our method allows for assessment of cpRNFL in children who fail eye tracking.

Introduction

Assessment of retinal nerve fiber layer (RNFL) thickness by spectral domain optical coherence tomography (OCT) is important for managing glaucoma or secondary glaucoma in pediatric subjects. In children with poor fixation, the RNFL can be measured by a circumpapillary scan (cpRNFL) because of its rapid acquisition time. Precisely centering the cpRNFL scan onto the optic disc is critical for accurate comparison of a subject's cpRNFL to standardized cpRNFL measurements.[1,2] Decentration of the scan due to eye movements or poor fixation is one of the most common artifacts encountered with OCT imaging.[3] Significant changes in retinal layer thickness occurs when the cpRNFL scan is misaligned by about 42 µm in the vertical or horizontal direction.[2,4] Detection of disease progression also requires precise alignment of follow-up cpRNFL scans to a baseline (reference) to eliminate erroneous changes in RNFL thickness because of position artifacts. To maintain accurate alignment to the retina, the Spectralis OCT (Heidelberg Engineering GmbH, Heidelberg, Germany) incorporates simultaneous scanning laser ophthalmoscope imaging (SLO) and automated eye tracking (TruTrack). TruTrack

Copyright 2020 The Authors

This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License.

Active Eye Tracking is a patented technology that uses a second laser beam to actively track the eye during OCT scanning to avoid motion artifact. TruTrack allows averaging of b-scans after internal compensation for eye movement artifacts, fixation errors, and head tilt. The number of averaged b-scans after TruTrack compensation is referred to as "ART."

Fixation in patients with childhood glaucoma can be limited for multiple reasons such as inattention, nystagmus related to central vision loss, and manifest latent nystagmus due to subnormal binocularity. Nystagmus may emerge, or increase in intensity, when monocularly viewing a target in the OCT machine. The large amplitude eye movements in these subjects causes significant position errors when real-time eye tracking is unreliable. The difficulty of the task is shown by simulating circumpapillary b-scan motion artifacts based on eye movements in children with manifest latent nystagmus.[5] In Figure 1A, a healthy control child maintains alignment of multiple circumpapillary scans with a reference scan (blue circle) in the presence of small fixational eye movements. In contrast, Figures 1B and 1C show that circumpapillary scans rarely, if ever, align with the reference scan in subjects with nystagmus (also see Discussion). Therefore subjects with large eye movements or nystagmus are excluded from clinical studies because of artifacts.[6,7]

Sophisticated methods have proposed motion artifact correction using multiple orthogonal OCT volumes.[8–11] Unfortunately, these methods incorporate custom volumetric OCT scanning methods or use hardware that is unavailable on commercial devices. Moreover, custom methods may not be applicable to pediatric subjects given the large eye-movement artifacts shown in Figure 1. We sought to develop a method that measures retinal layer thickness in misaligned cpRNFL scans due to eye movement artifacts using a Spectralis device in a clinical setting. The analysis proposes that multiple misaligned scans acquired during poor fixation will have overlapping matching locations to a reference scan like that shown in Figures 1B and 1C. Assuming the collection of misaligned scans have good coverage of the reference scan position, it is possible to extract cpRNFL thickness from these misaligned scans at the matching locations. A critical component is to devise a method that can determine where the reference and misaligned scans would match in retinal position and then segment retinal layers only at these matching locations.

Another goal of this study is to assess the accuracy of the methodology from retrospective records in a clinical setting. To accomplish this goal, we compare retinal layer thickness (including cpRNFL thickness)

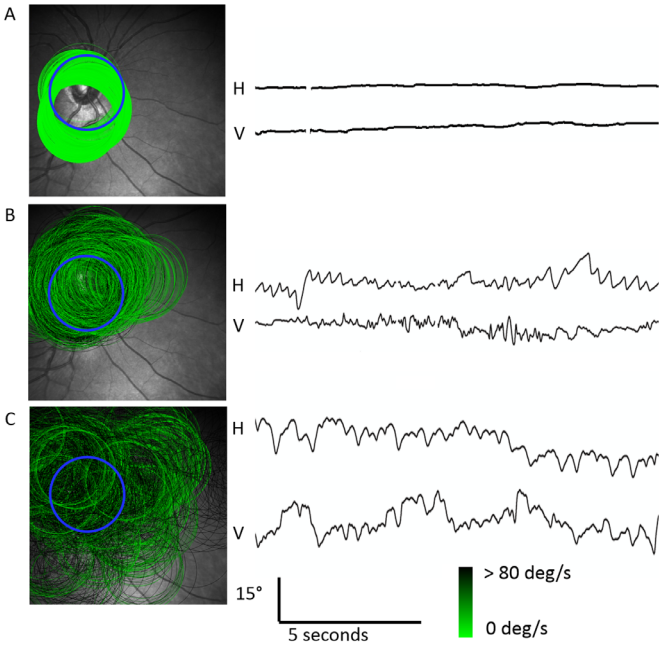

Figure 1. Simulation of circumpapillary scan positions over time in 3 subjects with different levels of fixation. The left side shows the same SLO image with a circumpapillary reference scan (*blue circle*) placed at identical positions. In (A), the child with stable fixation maintains circumpapillary scan positions near the reference scan at velocities < 20 degrees/second (*green circles*, color coded by velocity). (B) A subject with moderate amplitude nystagmus would have poor positioning with respect to the reference scan due to both horizontal and vertical eye movements. (C) Subject with large amplitude nystagmus and re-fixation saccades have large variation in scan locations and eye velocity such that accurate positioning is not possible for the entire recording. Traces to the right are corresponding video-oculography recordings of horizontal (*H*) and vertical (*V*) eye movement positions. For the horizontal trace, upward deflections are rightward movements and downward deflections are leftward movements. For the vertical trace, upward deflections are upward movements and downward deflections are downward movements. The fovea position is determined manually when it is visible (*blue square*).

at the matching locations between a chosen reference scan and all misaligned scans. We use data from a single imaging session to remove variability due to disease progression Furthermore, we assessed clinical factors and motion artifacts that might contribute to measurement errors in this analysis. Under ideal circumstances there will be no difference in layer thickness between the reference and the misaligned data.

Methods

Retrospective review of patient records was approved by the Institutional Review Board of Seattle Children's Hospital and conformed to the requirements of the United States Health Insurance Portability and

Privacy Act. Subjects were younger than 18 years of age (except when older subjects were referred to our service), who were selected from a cohort of records seen at Seattle Children's Hospital from December 2009 to July 2013. All subjects were being observed for risk of glaucomatous-related damage to the RNFL.

Imaging was obtained by the Spectralis OCT/SLO (Heidelberg Engineering, Germany). Our standard imaging protocol incorporates a cpRNFL scan (approximately 3.5 mm diameter, 768 × 497 pixels). Simultaneous SLO images were acquired in approximately 96 milliseconds (768 × 768 pixels; 30° × 30° field) at a nominal resolution of 8.7 × 8.7 mm or 11.3 µm/pixel. OCT images were acquired in approximately 20 milliseconds using high-speed mode (40,000 A-scans/second; 768 a-scans per b-scan) with nominal 3.9 µm/pixel in the axial dimension and approximately 10 to 14 µm/pixel in the lateral dimension. For all analyses in this study, each subject's data used the data, position coordinates, and conversion factors that were supplied by Heidelberg Eye Explorer (HEYEX) software version 1.10.4.0.

Our standard protocol in children with poor fixation or nystagmus has been to attempt "Tru-Track" automated eye tracking with a cpRNFL scan. If TruTrack failed after multiple attempts (i.e., before the Spectralis software automatically aborted eye tracking due to laser exposure duration), we then acquired numerous single cpRNFL scans until either there were at least two scans centered near the optic disc, the child became too inattentive to continue, or the Spectralis automatically aborted because of laser-exposure duration. In these subjects, it was not possible to predict how accurately the scans were centered on the optic disc until post hoc. Furthermore, volumetric scans could not be acquired if the TruTrack component failed. Note that the Spectralis device could have an ART value greater than 1 for single cpRNFL scans after pressing the "acquire" button, which was related to random factors out of our control such as nystagmus-related foveation periods or brief moments of adequate fixation. Head movement was reduced by gently holding the child's forehead to the head-rest bar if needed.

Subject selection was aided by searching through a Spectralis database with a combination of codes (nystagmus, latent nystagmus, increased intra-ocular pressure, large cup-to-disc ratio, congenital glaucoma, aphakic glaucoma, acquired glaucoma, anterior segment dysgenesis, aniridia, Stickler's syndrome, Sturge-Weber syndrome, and anterior uveitis). The inclusion criteria for this study were (1) subject visits in which the Spectralis aborted TruTrack (as described above) and (2) a minimum of five SLO/OCT images showing that automated eye tracking was aborted (e.g., SLO images have a dashed green line indicating strong eye movements). Intermittent ART averaging (determined by internal Spectralis criteria) was not used in the inclusion or exclusion criteria. Subjects with severe optic nerve loss were excluded if the global cpRNFL thickness was less than 30 µm (as scored by HEYEX automated segmentation) to avoid having the floor effect influence the analysis of the method. The eye with the greater number of OCT/SLO images was analyzed from each subject.

All imaging data in this study were acquired only during a single session to ensure there were no changes to the optic nerve. The image quality score, which is the estimated signal-to-noise ratio (SNR) in dB units, was not used in the exclusion criteria. A single reference scan was chosen for each subject judged to have a combination of minimum apparent motion-artifacts on the SLO and OCT, having the closest alignment with the center of the optic disc, and the highest SNR on OCT. Whether a reference scan had partial averaging by ART was not used in the inclusion criteria. The reference scan could be misaligned from optimal centering on the optic nerve but served as the comparison for all misaligned scans.

Retinal Layer Segmentation

Software was developed that matched sections of misaligned OCT scans to the reference cpRNFL scan based on SLO coordinates, which then segmented retinal layers at matching locations.[12] The software was a modification to a previous release (http://faculty.washington.edu/jokelly/manualoct). The program imported raw XML/TIFF files from HEYEX software and had a graphical-user-interface for image processing and real-time segmentation while simultaneously viewing corresponding SLO and OCT images (Fig. 2). Images were excluded from the analysis if there were no matching loci to a reference scan, or the user deemed the SLO/OCT images had excessive noise, shadows, or image clipping that prevented detection of retinal layers. Under all conditions, the user viewed the raw SLO and OCT images and subsequent segmentation at an increased magnification of 1.5 to better view retinal layer boundaries (simulating the Spectralis HEYEX interface). The user could then apply custom contrast-luminance windowing or median filtering. Initial segmentation of retinal layers on the OCT image were performed in the background with the aid of the OpenCV C++ library (source code and documentation at www.opencv.org). For the purpose of segmentation, the OCT image was downsampled ×4 and then underwent a processing stream of

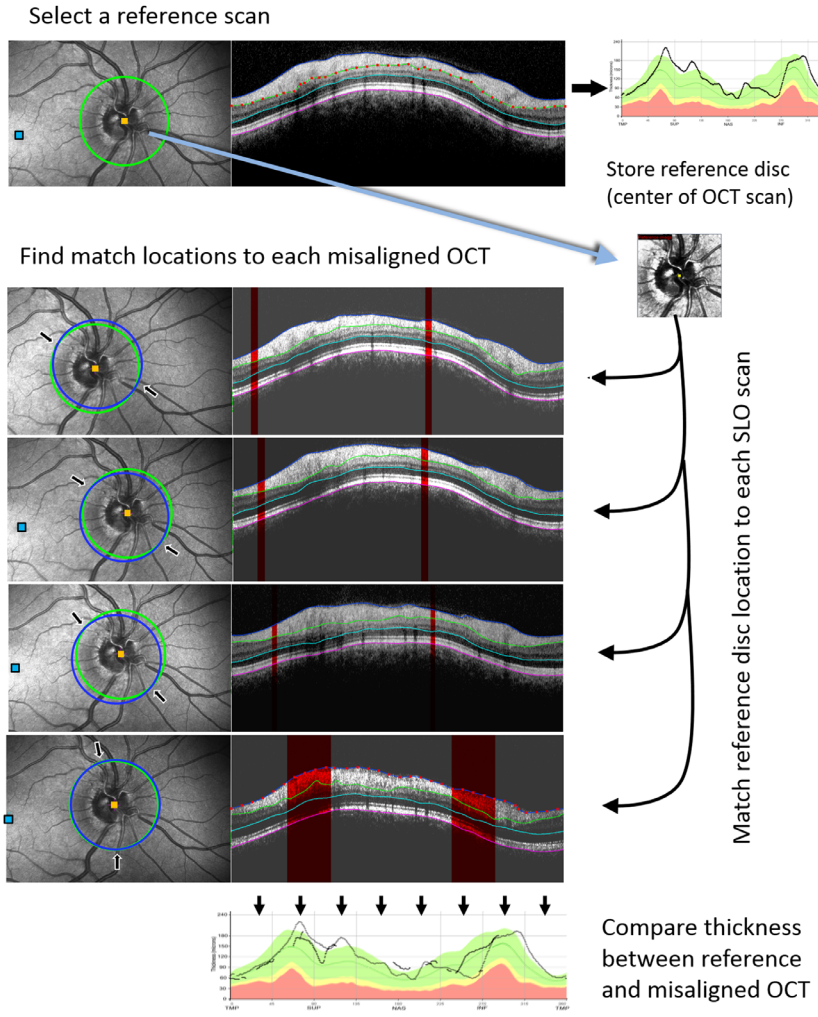

Figure 2. Schematic of the methodology. *Top*, a reference SLO/OCT scan is selected for each subject and is segmented into three layers by the software as shown by control points (red dots on OCT image) with spline fitting. At right the thickness of the retinal nerve fiber layer serves as a standard for comparisons and is plotted on the Spectralis normative template. The software stores a copy of the optic disc image, centered on the reference circumpapillary scan, and uses the image to find its location on all other misaligned scans (center of the match location is shown by the orange box). The match of the optic disc determines where the reference scan would lie on all other scans and is also used to quantify image distortion due to eye movements. The software then finds matching loci on the SLO image between the reference scan (green circle) with the misaligned OCT scan position (*blue circle*). The match regions on the SLO are indicated by arrows, while the red regions show the corresponding matching locations on the OCT (*right side*). The user ensures that layer segmentation is accurate only in the highlighted red regions as these locations will be analyzed. Layer thickness for matching regions are accumulated across subsequent misaligned OCT scans, which are then plotted along with the reference (*gray line*) in the bottom plot. The difference in thickness between the reference and all misaligned scans are used to assess accuracy. Note, not all data are shown for clarity.

two-dimensional (2D) median blur, k-means segmentation, erode function, then adaptive thresholding and then image dilation. Each OCT scan was segmented into three different retinal thickness: (1) the internal limiting membrane (ILM) to retinal nerve fiber layer (cpRNFL), (2) the ILM to the outer nuclear layer (ILM-ONL), and (3) total retinal from the ILM to Bruch's membrane (TR). The initial searches started at 32 equal positions across the OCT scan. A control point was placed on the full-scale image at each of the 32 locations, and then each layer boundary was fit by Catmull-Rom splines. Each layer thickness was then derived from the vertical position offsets of the spline fit after appropriate scaling. Layer segmentations could be automatically adjusted in real-time by the user using contrast windowing or the user could manually reposition, add, or delete spline points to ensure an accurate match to the layer boundary (simulating the Spectralis HEYEX software interface). There was no correction for axial motion.

Matches Between Reference and Misaligned Scans

After exporting all files from HEYEX, each subject's session started with the chosen reference scan. The user clicked the fovea location so that all measurements were corrected for rotation of the center of the cpRNFL scan with the fovea. After the retinal layers were segmented and corrected for angle, a 256- × 256-pixel SLO image of the reference optic disc (centered on the cpRNFL scan) was stored to find its location in all subsequent misaligned images for that subject. In this manner, all segmentations were limited only to b-scan and a-scan locations that matched the reference OCT scan location (e.g., a series of locations highlighted in red in Figure 2). The matching alignment of the reference optic disc image to all subsequent SLO images was determined by the OpenCV "Features Accelerated Segment Test, adapted pyramid feature-detection, with random sample consensus outlier removal" algorithm (FAST-RANSAC). The matching algorithm created a perspective (homography) transformation matrix from which translation, rotation, scale, and image shear coefficients of a rigid body could be derived documentation at https://docs.opencv.org/master/d9/dab/tutorial_homography.html). If the FAST-RANSAC matching failed, the search used a normalized cross-correlation assuming a rigid 2D affine transform (then image shear was assigned a value of 0.0, and horizontal and vertical scaling were assigned a value of 1.0). The user always had an option to override the automated location match by manually setting the position of center of the reference scan. Under all circumstances, an image of the reference optic disc and its center were available for visually inspecting the alignment. The user then clicked the fovea location so that all measurements could be corrected for rotation of the fovea with respect to the center of the cpRNFL scan. If the fovea was not visible, it was taken to be of the same angle as the reference scan.

The search algorithm stepped the angle α in increments of 2π/OCT pixel width. An array of α,x,y coordinates derived from the SLO image for the reference were generated for all values of α as follows:

$$\text{Ref}\ (\alpha, x, y) = (\text{eC} * \text{cpR} * \text{COS}(\alpha) + \text{Rx}, \text{cpR} * - \text{SIN}(\alpha) + \text{Ry}).$$

Then an array of α,x,y coordinates for the misaligned image were generated as follows:

$$\text{mOCT}(\alpha, \text{x}, \text{y}) = (\text{eC} * \text{cpR} * \text{COS}(\alpha) + \text{Sx}, \text{cpR} * - \text{SIN}(\alpha) + \text{Sy}),$$

where eC is -1 for the right eye and 1 for the left eye, cpR is the radius of the circular scan, Rx, Ry and Sx, Sy are the center coordinates of the reference and misaligned cpRNFL scans, respectively (all values exported by HEYEX software). A brute-force search then found matching locations that met a distance criterion allowing for slight variations in position was set by absolute differences between the reference and misaligned coordinates (ABS(mOCT (α',x) − Ref (α,x)) ≤ 1.5, and ABS(mOCT (α',y) − Ref (α,y)) ≤ 1.5). We then corrected for angle $\alpha'' = (\alpha', \text{Sx}, \text{Sy}) - (\text{fovea angle}, \text{Rx}, \text{Ry})$.

SLO/OCT scans were analyzed by two graders (author JPK with eight years of OCT experience and author FMB with eight years of experience treating glaucoma but inexperienced with manually segmenting OCTs). The graders were masked to details of the subject's clinical findings. Interobserver agreement was performed by having FMB repeat all measurements on all subjects. The final analysis only used the data scored by JPK rather than pooling data from both graders. To address small head tilts and software derived translation and ONH-fovea-angle rotations, all variability and accuracy measures were matched for cpRNFL scan angle to within ±5°.

Data were analyzed using Excel 360 (Microsoft Corporation, Redmond, WA, USA) or SPSS (version 12; IBM Corp. Armonk, NY, USA). Accuracy was defined as the difference in thickness between matching points on the misaligned scan and the reference scan after correction for scan angle. If the process was without any systematic error, the average for each subject would be 0.0 for all layers while the standard deviation assessed measurement variability. Variability and accuracy were evaluated by coefficient of variation (COV), intraclass correlation coefficient (ICC), and Bland-Altman plots. Clinical and imaging factors were assessed in a multiple regression model to predict the variability of thickness measurements. The dependent variable was standard deviation of accuracy. The independent variables were (1) logMAR visual acuity, (2) cup-to-disc ratio, (3) spherical equivalent refractive error, (4) ART value of the reference SLO, (5) ART value of the reference OCT, (6) number of match location, (7) SNR of the reference OCT, and (8) average cpRNFL thickness of the reference.

Results

From a total of 294 relevant database records, 40 subjects were identified who met inclusion criteria. Five of these subjects were excluded due to inadequate imaging and the remaining 35 subjects were analyzed (Table). Best corrected visual acuity, in log minimum

Table. Patient Data

Primary etiology (no.)	
Anirdia	1
Anterior segment dysgenesis	4
Aphakia	8
Congenital glaucoma	11
Glaucoma suspect	2
Juvenile glaucoma	4
Pseudophakia	3
Sturge-Weber	1
Uveitis, aphakia	1
Sex (no.)	
Male	19
Female	16
Tested eye (no.)	
Right	18
Left	17
Age (years)	
Mean	9.2
SD	5.1
Minimum	2
Maximum	26
Log MAR	
Mean	0.57
SD	0.45
Minimum	−0.12
Maximum	1.70
Spherical equivalent (diopters)	
Mean	0.67
SD	7.77
Minimum	−14.25
Maximum	18.75
Cup-to-disc ratio	
Mean	0.48
SD	0.25
Minimum	0.10
Maximum	0.95
Mean OCT quality score	
Mean	19.4
SD	6.1
Minimum	6.8
Maximum	33.8
No. matching loci to reference	
Mean	313
SD	356
Minimum	40
Maximum	1944

Log MAR, visual acuity in logarithm of minimum and of resolution.

angle of resolution (logMAR), ranged from −0.12 to 1.70. Therefore the study included subjects with good vision having poor fixation and/or latent nystagmus, and subjects with poor fixation due to significant visual loss. Spherical refractive error was uncorrelated with best corrected logMAR. The average cpRNFL thickness of the reference image was uncorrelated with spherical refractive error, best corrected logMAR, clinical cup-to-disc ratio, SNR, ART values, and number of matching locations (all correlations; $P > 0.05$). To avoid intersubject sample bias, the Table reports each subject's averaged value. The ART values and SNR, which were not used in inclusion/exclusion criteria, varied greatly from 1 to 100 for ART, and from 2 to 39 for SNR.

A total of 487 SLO/OCT images were available from the 35 subjects. Of the 487 images, 119 were excluded, 35 images were used for the reference, and 333 images were used for misaligned scans. From the 333 misaligned image pairs, there were 10959 matching a-scan locations to their corresponding reference locations. As this was a retrospective analysis, some subjects had short testing times that resulted in fewer scans (average number of scans per each subject = 24; minimum 6; maximum 60) and thus a smaller number of matching locations (Table). The automated optic disc alignment was accepted in 309 of 333 SLO images (92.8%), of which 266 images were successfully aligned by the FAST-RANSAC algorithm (79.9%). In 20 of 35 subjects, the reference scan had some level of ART averaging as set by internal Spectralis acquisition criteria even though TruTrack was aborted (median = 4, first quartile = 4, third quartile = 9).

The group averaged accuracy was -2.7, 1.4, and 0.3 µm for cpRNFL, ILM-ONL, and TR thickness, respectively (Fig. 3). These values were not significantly different from 0.0 (one-sample t-test; p > 0.172 for all). The absolute differences between matching points and the reference scan after correction for scan angle averaged 18.9, 15.8, and 13.4 µm for cpRNFL, ILM-ONL, and TR thickness, respectively. Standard deviations across subjects varied up to 38 µm, but averaged 22.4, 19.8, and 16.6 µm for cpRNFL, ILM-ONL, and TR thickness, respectively. Accuracy of the cpRNFL measurements were significantly different than ILM-ONL or TR measurements (*t*-test; $P = 0.008$ and $P < 0.001$, respectively). COV was lowest for TR thickness (5%, range 2%–9%) and highest for cpRNFL thickness (25%, range of 12%–44%). COV for ILM-ONL thickness was 11% (range 6%–21%). The number of scans from each subject and ART values were not significantly correlated with accuracy or COV for all layer segmentations (all r^2 values < 0.132; $P > 0.05$).

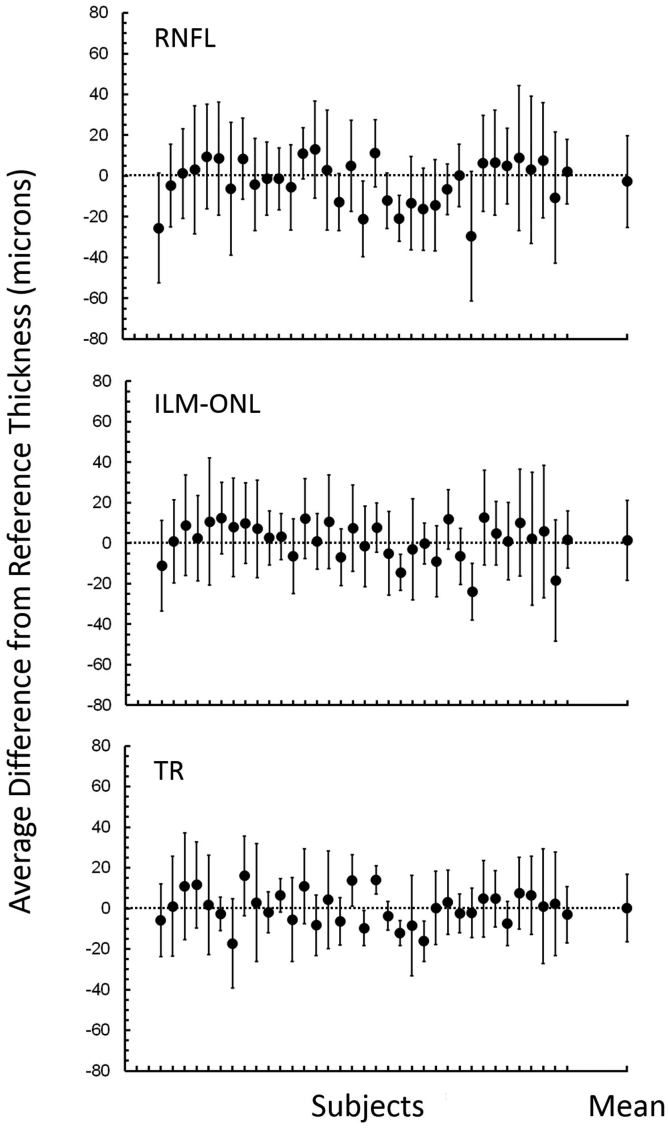

Figure 3. Accuracy of the analysis represented by average difference in retinal layer thickness of misaligned scans from the reference scan. The top plot shows differences for the RNFL. The *middle plot* shows differences for the thickness from the ILM-ONL. The *bottom plot* shows differences for total retinal thickness from ILM to Bruch's membrane (TR). Individual subjects are plotted along the abscissa and the group mean on the right. Error bars are 1 standard deviation of the difference. The *dotted line* is the ideal outcome of 0.0.

Figure 4 shows that variability was larger in superior and inferior quadrants when indexed by absolute value of accuracy. Compared to the temporal quadrant, there were larger absolute differences in the superior, and inferior quadrant (*t*-test; $P < 0.0001$ for all measured layers). Furthermore, the temporal quadrant had lower absolute differences compared to the nasal quadrant for cpRNFL and ILM-ONL (*t*-test; $P < 0.0001$) but not for TR ($P = 0.49$). Although these values were statistically significant, standard deviations overlapped between quadrants.

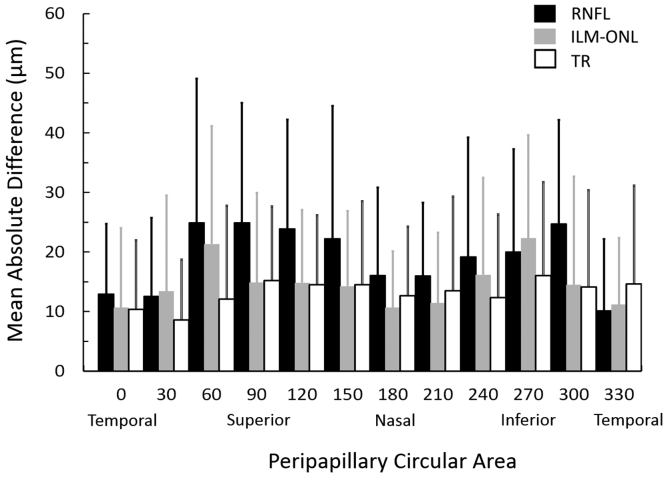

Figure 4. Mean of absolute difference in retinal layer thickness between match locations on misaligned scans and the corresponding reference scan with respect to relative circumpapillary scan angle. *Black columns* represent RNFL. *Gray columns* represent ILM-ONL. *White columns* represent total retinal thickness from ILM to Bruch's membrane (TR). The plot is based on data from all subjects, with a minimum of 600 data points for each column.

Interobserver Measurements

For comparisons of reference scans between JPK and FMB, the ICC was 0.88 for cpRNFL, 0.94 for ILM-ONL, and 0.97 for TR thickness (95% confidence intervals [CI] < 0.02). To avoid intersubject sample bias, each subject had their measurements averaged for each retinal layer and the resulting Bland-Altman plots (interobserver repeatability) are shown in Figures 5A and 5B. The results for the ILM-ONL segmentation are not shown but were similar. The average difference (bias) from the Bland-Altman analysis was −1.8, −4.3, and −1.5 µm for cpRNFL, ILM-ONL, and TR thickness, respectively. All but five subjects had values that fell within the 95% CI of the limits of agreement.

For interobserver comparisons of misaligned scans of all subjects, there was a total of 7733 locations that matched between observers FMB and JPK. The interobserver ICC was 0.63 for cpRNFL, 0.74 for ILM-ONL, and 0.77 for TR thickness (95% CI < 0.03). Again, to avoid intersubject sample bias, each subject had their measurements averaged for each retinal layer, and the resulting Bland-Altman plots are shown in Figures 5C and 5D. For repeatability of the misaligned scans, the bias from the Bland-Altman analysis was −2.6, −1.7, and 0.4 µm for cpRNFL, ILM-ONL, and TR thickness, respectively. All but six subjects had values that fell within the 95% CI of the limits of agreement. The results from ILM-ONL segmentation are not shown but were similar.

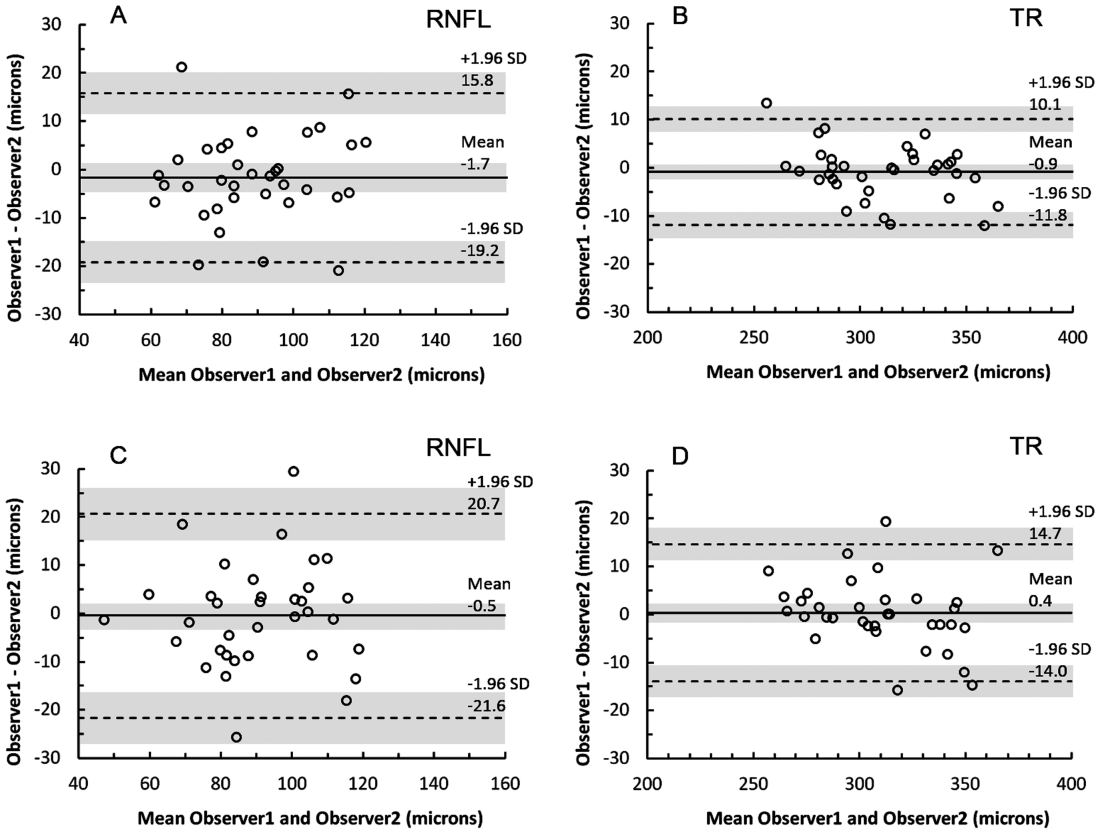

Figure 5. Bland-Altman plots showing comparisons between two observers for scoring mean RNFL and (B) mean total retinal (TR) thickness. The top row (A, B) shows comparisons between reference images. The bottom row (C, D) shows comparisons of matched locations on misaligned images. The *solid line* indicates the average difference, the *dashed lines* indicate the 95% limits of agreement. The *shaded boxes* represent the 95% CI on bias, upper, and lower limits of agreement. All data points were matched for circular scan angle with respect to the fovea.

Analysis of Measurement Error

A multiple regression model examined possible factors that contributed to cpRNFL thickness variability. The results showed an adjusted multiple $R^2 = 0.30$ ($P = 0.02$). Factors with standardized beta weights having a P value < 0.05 were (1) average cpRNFL thickness of the reference, (2) SNR, and (3) logMAR visual acuity (values of 0.510, −0.414, and 0.351, respectively). Individual correlations of these three factors with a significant beta weight were low ($r^2 < 0.16$ for all; $P > 0.05$) and therefore these factors accounted for only a small fraction of the total variability. The multiple regressions were also poor for both ILM-ONL and TR thickness measurement variability (adjusted multiple $R^2 = 0.13$ and 0.11, respectively; both $P > 0.16$) and none of the beta weights were significant. A multiple regression examined the relationship of COV with the independent variables age, logMAR, cup-to-disc ratio, spherical equivalent error, ART value of the reference OCT, mean ART, and SNR. For cpRNFL, the adjusted multiple R^2 was 0.47 ($P = 0.001$). Factors with standardized beta weights having a P value < 0.05 were age (0.50) and SNR (−0.45). For ILM-ONL, the adjusted multiple R^2 was 0.32 ($P = 0.01$); again, age and SNR were the only factors with a statistically significant standardized beta weight. For TR, the adjusted multiple R^2 was 0.26, which was not statistically significant ($P = 0.254$).

Another source of variability is the difference in acquisition speed between the SLO image and the corresponding OCT image (approximately 96 vs 20 milliseconds). In the presence of rapid eye movements, there could be random position offsets or image distortions between the OCT scan and the corresponding SLO scan due to their different acquisition speeds. The potential effect of this eye movement artifact on the matching cpRNFL thickness is shown in Figure 6. The matching-location and reference data are represented in the standard HEYEX cpRNFL plot format. Note there are subsets of matching-location data that showed an apparent lateral shift, or offset, in thickness near the arcades (arrows). We hypothesize the apparent shift between the reference and corresponding

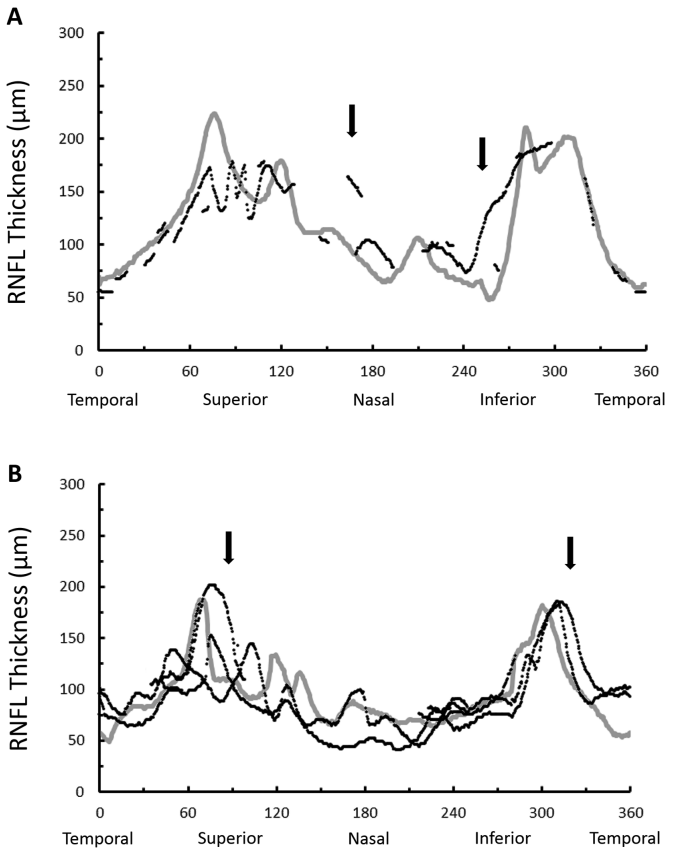

Figure 6. Examples of measurement errors of RNFL thickness from matching locations on misaligned circumpapillary scans (*black points*) compared to the reference scan (*gray points*). Data from subject 7 are plotted in (A), whereas data from subject 29 are plotted in (B). *Arrows* point to areas with large errors likely related to motion artefact occurring between acquisition times of the OCT with respect to the corresponding SLO image.

match-locations resulted from a rapid eye movement that occurred during the asynchronous periods between the OCT and SLO imaging. Specifically, if the SLO image takes 96 milliseconds while the OCT image takes 20 milliseconds, a rapid eye movement that occurred in the 76-millisecond lag time will distort the matching locations. The Spectralis instrument does not store eye movement data, therefore we estimated motion artifacts (image rotation, shear, horizontal and vertical scale) from the homography transformation matrices required to align the reference optic disc to the misaligned SLO image. There was no significant correlation between measurement variability of cpRNFL thickness with image rotation and image shear ($r^2 = 0.04$ and 0.06, respectively). There was no correlation between image scaling and measurement variability of any retinal thickness measurement ($r^2 < 0.001$ for all analyses). Similar correlations were seen for ILM-ONL thickness and TR thickness without statistical significance (both $P > 0.05$). The linear correlations remained poor if all raw data points were analyzed instead of subject averages ($r^2 \leq 0.04$ for all analyses). Overall, the amount of variance explained (<7%) was too low for these factors to be of clinical importance.

Discussion

This study demonstrates a technique to measure retinal layer thickness in children with poor fixation and nystagmus who fail eye tracking on the Spectralis OCT. This work is relevant for subjects in which nystagmus and poor fixation cause decentration of the cpRNFL scan or misalignment to a reference scan. We assessed measurement variability of the technique, and potential sources of error that can limit the ability of the method to detect accurate changes in retinal layer thickness. Factors that contributed to variability of cpRNFL measurements were average cpRNFL thickness, low SNR, and reduced visual acuity. However, these factors accounted for only a fraction of the total variance. Alternatively, displacement of the OCT scan relative to its implied location on the corresponding SLO image due to rapid eye movements can potentially generate large errors in matched-location thickness measurements. Such motion artifacts might be apparent when plotting the matched-location measurements on top of the reference data (e.g., Fig. 6). Therefore our analysis still requires careful review of the data for accuracy.

The methodology uses standard cpRNFL protocols. Scans can be acquired repeatedly until the area of the cpRNFL location is deemed adequately covered. Given the current status of the software, the number of acquired cpRNFL scans must be subjectively estimated at acquisition time or must be determined after a post-hoc review. We estimate a minimum of 20 to 30 scans would be necessary. The increased acquisition time would add more than 30 seconds of testing time (possibly several minutes depending on subject cooperation). Nonetheless, these subjects would typically have more imaging anyway due to their eye movements. A significant drawback of our analysis is that it requires much longer analysis time for manual segmentation, and the possibility that scans will not fully cover all 360° of the cpRNFL. Therefore a practical application of our method would improve by integration with acquisition hardware and robust automated segmentation algorithms; both would require further development.

Although the primary purpose of this article was to develop a novel methodology and assess its accuracy,

there are several relevant applications. Our methodology could be adapted for longitudinal assessment when cpRNFL scans are misaligned at follow-up. In this circumstance, changes in cpRNFL thickness from a baseline reference should still reflect localized or global RNFL loss along portions of matching circumpapillary scans. Relative changes like that shown in Figure 6 could be reviewed from the data overlaid onto a baseline. Age-appropriate normative cpRNFL data are currently available for sectors only[13,14] and are not directly comparable to the plots shown in this study. We did not compare our subjects to the pediatric normative data since many reference scans were not accurately aligned and thus could not be used as a gold standard. Additionally, our analysis may be useful for analyzing cpRNFL data from children excluded from clinical studies because of their eye movement artifacts.[6,7] Further studies will be required to address the application of the methodology and determine the sensitivity/specificity to detect cpRNFL loss when the reference scan is misaligned from the optic disc center.

The contrast between the RNFL and ganglion cell layers is typically reduced in the presence of eye movements and optic nerve atrophy, which also increases segmentation error. To address this issue, we compared segmentations between cpRNFL, IPL-ONL, and TR. Our measurements of misaligned scans found lower variability for TR compared to the cpRNFL thickness. This outcome was expected as the ILM and Bruch's membrane form boundaries with higher contrast. Therefore when image quality is poor, TR thickness is expected be more accurate than cpRNFL thickness. Although the average error of the match-location measurements was close to zero, measurement variability could be significant due to poor fixation and nystagmus. For comparison, variability of cpRNFL thickness in control children ranges about 2 to 12.5 µm,[15,16] which is larger than that in control adults, all with presumed stable fixation.[17,18] In contrast, our values of intrasession COV of cpRNFL sectors are larger than that reported in children and adults with presumed stable gaze (0.8% to 11.7%), although larger COV can be found in subjects with more severe glaucoma.[6,16,19] Our average interobserver test-retest differences of cpRNFL thickness were close to the range reported by Ghasia et al.[16] (3.06 to 9.80 µm) in pediatric subjects with glaucoma/glaucoma suspect, who presumably had stable gaze on a Spectralis OCT. Although Ghasia et al.[16] did not report a Bland-Altman analysis, our 95% limits of agreement ranged from −21.6 to 20.7 µm at maximum, which is slightly higher than that reported for comparisons between Cirrus and Spectralis OCT machines for adults having glaucoma.[17]

The faster a-scan rate of swept-source OCT might ameliorate motion artifacts in our subjects but at the cost of reduced SNR.[20] To address this, we calculated the theoretical limit of sampling in order for an OCT image to be completely free of eye movement artifacts. Assuming there are approximately 280 µm per visual degree and a single cpRNFL scan takes 20 milliseconds, then the eye should be moving less than 2°/second to maintain an accurate lateral resolution of 12 µm across the entire OCT image. Figure 1 plots simulations of cpRNFL scan movements overlaid on the fundus based on actual eye movement recordings in a control child and three children with latent-manifest nystagmus (technical details of eye movement recording are reported by Weiss et al.[5]). The top of Figure 1 shows the healthy control child maintains fixation with eye velocities below 2°/second within in a position window of ±2° for 88% of the 15 second recording (color intensity of the cpRNFL scan is scaled by eye velocity). There is ample time for optimal imaging or eye tracking in the control. In contrast, the three subjects with latent-manifest nystagmus have eye velocities ≤ 2 degrees/second for less than 1% of the time during the entire eye movement recording. Using en-face volumetric OCT imaging makes the situation more complicated. Assuming 200,000 a-scans/second with a volume of 768 × 768 a-scans, then acquisition time is 2.9 seconds (assuming minimum lag time for raster fly back). Based on the duration of eye movement data in subjects with nystagmus, the longest duration of stable gaze (≤ 2°/second) is 168 milliseconds, predicting only a small fraction of the volume area will have optimal imaging across multiple en face volume acquisitions. Furthermore, accurate *en-face* volume reconstruction would be extremely difficult given large amplitude fluctuations in horizontal and vertical eye position across a duration of 2.9 seconds. Thus our analysis indicates that eye movement artifacts will continue to be an important source of measurement error in these subjects even with faster a-scan speeds.

There are several important caveats in this study. The methods provided in this study are for research purposes and cannot replace standard clinical procedures. The intent of the study is exploratory in nature. The data are based on retrospective post-hoc data from a relatively small group of subjects seen at a tertiary hospital. The methodology is dependent on a reference cpRNFL scan that may not be accurately aligned to the optic nerve, in which case the analysis is only applicable for relative changes at follow-up. Although it could be argued that an ideal cpRNFL scanning should be acquired at an older age when fixation/attention improves, it is not possible to guarantee there will be improvement in fixation with age nor is it possible to

be certain the cpRNFL was stable over time. The aim of this research was to develop a strategy to objectively monitor a child for optic nerve damage when either visual field testing is unreliable, or it is impractical to wait for a child's fixation to improve.

Acknowledgments

The authors thank Kristina Tarczy-Hornoch and Michelle Cabrera for comments on a previous version.

Supported by an unrestricted grant from grant from the Peter LeHaye, Barbara Anderson, and William O. Rogers Endowment Funds.

Disclosure: **J.P. Kelly**, None; **F.M. Baran**, None; **J.O. Phillips**, None; **A.H. Weiss**, None

References

1. Cheung CY, Yiu CK, Weinreb RN, Lin D, Li H, Yung AY, et al. Effects of scan circle displacement in optical coherence tomography retinal nerve fibre layer thickness measurement: a RNFL modelling study. *Eye (Lond)*. 2009;23:1436–1441.
2. Shin JW, Shin YU, Uhm KB, et al. The Effect of Optic Disc Center Displacement on Retinal Nerve Fiber Layer Measurement Determined by Spectral Domain Optical Coherence Tomography. *PLoS One*. 2016;11:e0165538.
3. Liu Y, Simavli H, Que CJ, 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;159:565–576.
4. Campbell RJ, Coupland SG, Buhrmann RR, Kertes PJ. Effect of eccentric and inconsistent fixation on retinal optical coherence tomography measures. *Arch Ophthalmol*. 2007;125:624–627.
5. Weiss AH, Kelly JP, Hopper RA, Phillips JO. Crouzon Syndrome: relationship of eye movements to pattern strabismus. *Invest Ophthalmol Vis Sci*. 2015;56:4394–4402.
6. Rajjoub RD, Trimboli-Heidler C, Packer RJ, Avery RA. Reproducibility of retinal nerve fiber layer thickness measures using eye tracking in children with nonglaucomatous optic neuropathy. *Am J Ophthalmol*. 2015;159:71–77.
7. Perucho-González L, Martínez de la, Casa JM, Sáenz-Francés F, Morales-Fernandez L, Méndez-Hernández CD, Sánchez-Jean R, García-Feijoó J. Retinal nerve fiber layer thickness in children with primary congenital glaucoma measured by spectral domain optical coherence tomography. *J AAPOS*. 2019;23:94.
8. Ricco S, Chen M, Ishikawa H, Wollstein G, Schuman J. Correcting motion artifacts in retinal spectral domain optical coherence tomography via image registration. *Med Image Comput Comput Assist Interv*. 2009;12:100–107.
9. Kraus MF, Liu JJ, Schottenhamml J, Chen CL, Budai A, Branchini L, Ko T, Ishikawa H, Wollstein G, Schuman J, Duker JS, Fujimoto JG, Hornegger J. Quantitative 3D-OCT motion correction with tilt and illumination correction, robust similarity measure and regularization. *Biomed Opt Express*. 2014;5:2591–2613.
10. Salmon AE, Cooper RF, Langlo CS, Baghaie A, Dubra A, Carroll J. An automated reference frame selection (ARFS) algorithm for cone imaging with adaptive optics scanning light ophthalmoscopy. *Trans Vis Sci Tech*. 2017;6(2):9.
11. Baghaie A, Yu Z, D'Souza RM. Involuntary eye motion correction in retinal optical coherence tomography: Hardware or software solution? *Med Image Anal.* 2017;37:129–145.
12. Kelly JP, Baran F, Phillips JO, Weiss AH. Optical coherence tomography in optic nerve hypoplasia: correlation with optic disc diameter, nerve fiber layer thickness, and visual function. *J Neuroophthalmol*. 2018;38:312–319.
13. Turk A, Ceylan OM, Arici C, et al. Evaluation of the nerve fiber layer and macula in the eyes of healthy children using spectral domain optical coherence tomography. *Am J Ophthalmol*. 2012;153:552–559.
14. Yanni SE, Wang J, Cheng CS, et al. Normative reference ranges for the retinal nerve fiber layer, macula, and retinal layer thicknesses in children. *Am J Ophthalmol*. 2013;155:354–360.
15. Dave P, Jethani J, Shah J. Asymmetry of retinal nerve fiber layer and posterior pole asymmetry analysis parameters of spectral domain optical coherence tomography in children. *Semin Ophthalmol*. 2017;32:443–448.
16. Ghasia FF, El-Dairi M, Freedman SF, Rajani A, Asrani S. Reproducibility of spectral-domain optical coherence tomography measurements in adult and pediatric glaucoma. *J Glaucoma*. 2015;24:55–63.
17. Tan BB, Natividad M, Chua KC, Yip LW. Comparison of retinal nerve fiber layer measurement between 2 spectral domain OCT instruments. *J Glaucoma*. 2012;21:266–273.

18. Suh MH, Yoo BW, Park KH, Kim H, Kim HC. Reproducibility of spectral-domain optical coherence tomography RNFL map for glaucomatous and fellow normal eyes in unilateral glaucoma. *J Glaucoma*. 2015;24:238–244.
19. Garas A, Vargha P, Hollo G. Reproducibility of retinal nerve fiber layer and macular thickness measurement with the RTVue-100 optical coherence tomography. *Ophthalmology*. 2010;117:738–746.
20. Lee SY, Kwon HJ, Bae HW, et al. Frequency, type and cause of artifacts in Swept-Source and cirrus HD optical coherence tomography in cases of glaucoma and suspected glaucoma. *Curr Eye Res* 2016;41:957–964.