Research Article | Vol. 11, No. 12 / 1 December 2020 / Biomedical Optics Express 7079

Biomedical Optics EXPRESS

Retinal layer thicknesses retrieved with different segmentation algorithms from optical coherence tomography scans acquired under different signal-to-noise ratio conditions

Tuomas Heikka,[1] Barry Cense,[2,3] and Nomdo M. Jansonius[1,*]

[1] *Department of Ophthalmology, University of Groningen, University Medical Center Groningen, Groningen, The Netherlands*
[2] *Department of Mechanical Engineering, Yonsei University, Seoul 03722, Republic of Korea*
[3] *Optical+Biomedical Engineering Laboratory, Department of Electrical, Electronic and Computer Engineering, University of Western Australia, Crawley, WA, Australia*
[*] *n.m.jansonius@umcg.nl*

Abstract: Glaucomatous damage can be quantified by measuring the thickness of different retinal layers. However, poor image quality may hamper the accuracy of the layer thickness measurement. We determined the effect of poor image quality (low signal-to-noise ratio) on the different layer thicknesses and compared different segmentation algorithms regarding their robustness against this degrading effect. For this purpose, we performed OCT measurements in the macular area of healthy subjects and degraded the image quality by employing neutral density filters. We also analysed OCT scans from glaucoma patients with different disease severity. The algorithms used were: The Canon HS-100's built-in algorithm, DOCTRAP, IOWA, and FWHM, an approach we developed. We showed that the four algorithms used were all susceptible to noise at a varying degree, depending on the retinal layer assessed, and the results between different algorithms were not interchangeable. The algorithms also differed in their ability to differentiate between young healthy eyes and older glaucoma eyes and failed to accurately separate different glaucoma stages from each other.

© 2020 Optical Society of America under the terms of the OSA Open Access Publishing Agreement

1. Introduction

Glaucoma is an age-related optic neuropathy. The primary site involved in this disease consists of the retinal ganglion cells (RGCs). Loss of RGC axons causes excavation of the optic nerve head (ONH) and thinning of the retinal nerve fiber layer (RNFL). The ganglion cell complex (GCC) consists of the RNFL, the ganglion cell layer (GCL) and the inner plexiform layer (IPL), while the GCL and the IPL together are referred to as GCIPL. In order to measure the thickness of these retinal layers *in vivo*, the currently most commonly used technique is optical coherence tomography (OCT) [1]. It has been shown that OCT-reported layer thicknesses depend largely on image quality; a lower image quality reduces the observed thickness [2]. This underestimation of layer thickness hampers the usage of OCT both for the detection of glaucoma (screening) and for progression detection, once glaucoma has been diagnosed (monitoring).

OCT uses infrared radiation, which passes through several different media when assessing the retina, and is therefore subject to noise from media opacities [3]. Media opacities, in elderly patients commonly caused by cataract, thus decrease image quality and indeed make the RNFL seem thinner [4]. A similar thinning with a decrease in image quality was found in healthy subjects, using neutral density filters [5]. On the other hand, RNFL thickness may be overestimated due to image intensity transformation of the spectrometer output, done as a pre-processing step in order to make low reflective layers more visible [6]. Finally, the changes

#399949 https://doi.org/10.1364/BOE.399949
Journal © 2020 Received 11 Jun 2020; revised 23 Aug 2020; accepted 24 Aug 2020; published 10 Nov 2020

Biomedical Optics EXPRESS

in the retinal tissue that occurs in glaucoma may compromise image quality [7,8]. In order to accurately measure RNFL thickness, it is of great interest to fully understand the influence of image quality and image processing on the reported RNFL thickness [9].

In this study, we aimed (1) to reproduce the previously reported degrading effect of poor image quality (low signal-to-noise ratio (SNR)) on the different layer thicknesses and (2) to compare different segmentation algorithms regarding their robustness against this degrading effect. For this purpose, we performed OCT measurements in healthy subjects and degraded the image quality by employing neutral density filters. We also analyzed OCT scans from glaucoma patients, in which the image quality is generally lower, amongst others due to a thinner RNFL with lower reflectivity. We analyzed the obtained images using (i) device-specific proprietary software, (ii) freely available segmentation software, being the Iowa reference algorithm [10], which was also used by [5], and (iii) the DOCTRAP software [11]. As a reference (iv), we followed the theoretical framework that we developed recently [6]. In that study, we introduced the full width at half maximum (FWHM) of the square of the first peak in the Fourier-transformed spectrometer output as an unbiased estimate of the RNFL thickness. We focused on the macular area, which is at least as relevant to glaucoma diagnostics as the peripapillary area [12–15] and allowed us to address different retinal layers. In a subsequent study, we will address the peripapillary area.

2. Methods

2.1. Study population

Twenty healthy subjects who fulfilled the inclusion criteria participated in this cross-sectional, observational study. The inclusion criteria were (1) age between 18 to 50, (2) best-corrected visual acuity of 1.0 or better in the study eye, (3) a spherical equivalent > -4 diopters (that is, high myopia was not allowed), (4) a normal visual field as tested with frequency doubling technology (FDT C20-1 screening mode; Carl Zeiss, Jena, Germany), (5) an intraocular pressure of 21 mmHg or below (non-contact tonometry), and (6) no ophthalmic history or first-degree relatives with a history of glaucoma. Mean (standard deviation [SD]) age was 28.5 (5.5) years; 45% were male. If both eyes were eligible, the right eye was chosen; this resulted in 19 right eyes and 1 left eye. The ethics board of the University Medical Center Groningen (UMCG) approved the study protocol. All participants provided written informed consent. The study followed the tenets of the Declaration of Helsinki.

We selected 102 subjects with glaucoma from the database of the Groningen Longitudinal Glaucoma Study [16]. Glaucoma patients were divided into three groups according to the mean deviation (MD) of their latest standard automated perimetry (HFA 30-2 SITA fast; Carl Zeiss Meditec AG, Jena, Germany) recording of their right eye. The three groups were early glaucoma (MD better than -6 dB), moderate glaucoma (MD between -12 and -6 dB), and severe glaucoma (over -12 dB). Within the database, the study patients have a randomly allotted study number. Randomization was done by selecting, for each group, the first 34 patients whose right eye fulfilled the inclusion criteria of spherical equivalent > -4 diopters and who had OCT scans taken with the Canon HS-100 device (see below). From the 102 glaucoma patients, 5 early, 5 moderate, and 9 severe patients were removed due to the presence of other eye diseases that might interfere with the segmentation of inner retinal layers (e.g., macular pucker, cystoid macular oedema), leaving 29 early, 29 moderate and 25 severe glaucoma patients. Mean (SD) age of the glaucoma patients was 72.8 (8.0), 72.7 (7.0), and 73.6 (9.7) years for early, moderate, and severe glaucoma patients, respectively; 53%, 56%, and 55% were male.

2.2. Data collection

Data collection in healthy subjects was done by a single operator (TH) using the Canon OCT-HS100 device (software version 4.2.1). Twelve OCT scans per subject were taken, three without

Biomedical Optics EXPRESS

a filter for reference, and three scans per different neutral density filter. Filters were used to attenuate the light, which lowers the SNR. Filter strengths used were 0.3, 0.6, and 0.9 optical density (Neutral density NIR filter; Edmund Optics, Barrington, USA), corresponding to 50%, 25%, and 12.5% of the light getting through in a single pass, respectively. Pupils were dilated with 0.5% tropicamide. Volume scans were taken from the macular area using the glaucoma scan mode (128 B-scans, each with 1024 A-scans in a 10×10 mm area centered at the fovea, with a depth resolution of 3 µm). The quality parameter (range: 0-10; the higher the better) provided by the device was also recorded. The OCT scans from the healthy subjects can be downloaded from dataverse.nl for comparison with future segmentation algorithms.

2.3. Regions of interest and segmentation algorithms

We analyzed the collected data for two different regions of interest (ROIs) with four different segmentation algorithms. We primarily focused on the RNFL; we also analyzed the GCIPL and total retinal thickness (TRT). The ROIs were (1) a circular area around the fovea with a diameter of 5 mm, excluding the fovea itself (eccentricity between 0.5 and 2.5 mm), named '5 mm ROI', and (2) a circular area with a diameter of 10 mm, connected to the 5 mm ROI (eccentricity between 2.5 and 5 mm), named '10 mm ROI'. These areas were chosen due to them being used by the Canon proprietary software.

The algorithms were (i) the device-specific proprietary software (Canon OCT-HS100, software version 4.2.1), (ii) the Iowa reference algorithm (Retinal Image Analysis Lab, Iowa Institute for Biomedical Imaging, Iowa City, IA, version 3.8.0; [10,17,18]). Scans were analyzed using the standalone desktop application called OCTExplorer, in which we chose the automated 10-layer macular OCT segmentation for the DICOM files exported from the Canon device. (iii) the Duke Optical Coherence Tomography Retinal Analysis Program (DOCTRAP software version 61.8; [11,19,20]. DOCTAP is an automated segmentation algorithm that runs on MATLAB. For the adjustable parameters, we manually chose the eye, scan orientation as vertical, 9.77 µm/pixel horizontal and 1.7 µm/pixel vertical pixel resolution. (iv) the FWHM of the square of the first peak in the linearly scaled intensity image. (details given below; [6]). For IOWA and DOCTRAP, we used the original, logarithmic images.

For algorithms with predefined regions (CANON and IOWA), we selected those regions that covered the above-mentioned ROIs as accurately as possible. CANON provides segmentation results for 8 areas; a ninth area corresponding to the inner 1 mm circle centred at the fovea is not reported by CANON and therefore discarded from all analyses (Fig. 1). For the 5 mm ROI and the 10 mm ROI, we took, for CANON, the average value from the inner 4 areas and outer 4 areas, respectively (Fig. 1(A)). For the IOWA algorithm, we selected the 10-2 grid, excluding the 4 squares centered at the fovea, as 5 mm ROI (Fig. 1(C)). The IOWA algorithm does not provide an equivalent of the 10 mm ROI. For algorithms without predefined regions (DOCTRAP and FWHM), we defined the two ROIs as depicted in Fig. 1(B). For each rectangular area of 1×0.5 mm in Fig. 1(B), we calculated the mean thickness from 6 B-scans of 100 A-scans each, and these mean thicknesses were subsequently averaged over the concerning ROIs.

For FWHM, we developed segmentation software using MATLAB (Mathworks, version R2018a, Natick, Massachusetts: The MathWorks Inc.). First, B-scan intensity values were transformed to (squared) spectrometer values assuming that the 0-255 grayscale intensity values correspond to a 50 dB dynamic range [21]. Dynamic range is device-specific and estimated to be between 40 to 60 dB [22,23]. The raw spectrometer values depicted as the linear-scale intensity I_{linear}values are, therefore:

$$I_{linear}(i,j) = 10^{(I_{\log}(i,j)/255)*5} \tag{1}$$

where $I_{linear}(i,j)$ is the transformed intensity value at a location (i,j) and $I_{\log}(i,j)$ the corresponding original log-scale intensity value.

Biomedical Optics EXPRESS

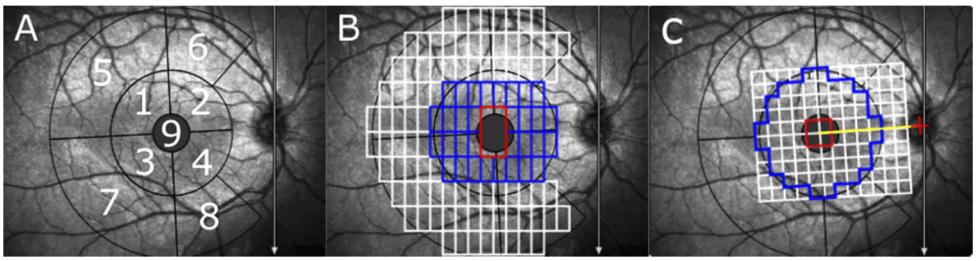

Fig. 1. A. Canon output image, 10 × 10 mm image size, where areas 1 to 4 correspond to the 5 mm ROI; numbers 5 to 8 correspond to the 10 mm ROI. The area corresponding to number 9 is excluded in all analyses. B. FWHM and DOCTRAP segmentation areas. Blue corresponds to the 5 mm ROI, white to the 10 mm ROI and the area inside the red rectangle is excluded. Each rectangle corresponds to 100 A-scans and 6 B-scans (1×0.5 mm). Arrows in A, B, and C indicate the B-scan scanning direction. C. 10-2 grid used in the IOWA algorithm (the area inside the blue border) covers an area corresponding to the 5 mm ROI used in A and B. [44]

For the linear-scale images, a median filter of 5×10 pixels was applied (5 along each A-scan; 10 successive A-scans) twice (first on the original image and then on the initial resulting image)

Table 1. Characteristics of the study population (mean with SD between brackets unless stated otherwise)[a]

	Healthy subjects n = 20		Glaucoma patients		
			Early (n=29)	Moderate (n=29)	Severe (n=25)
Age (years)	28.5 (5.5)		72.8 (8.0)	72.7 (7.0)	73.6 (9.7)
Gender (% male)	45		53	56	55
Visual acuity	1.2 (0.3)		0.67 (0.34)	0.91 (0.24)	0.85 (0.3)
IOP (mmHg)*	13.6 (2.8)		28.9 (11.9)	29.0 (7.2)	27.4 (10.5)
Visual field MD (dB)	-		-2.5 (1.7)	-9.3 (1.8)	-20.3 (5.0)
Image quality factor	0.0 OD	8.2 (0.7)	6.8 (1.1)	6.4 (1.0)	6.8 (1.1)
	0.3 OD	6.0 (0.7)			
	0.6 OD	3.9 (0.7)			
	0.9 OD	2.0 (0.5)			
Unsegmented RNFL in the FWHM algorithm (%)	0.0 OD	0.0	0.1	0.3	0.1
	0.3 OD	0.0			
	0.6 OD	0.0			
	0.9 OD	0.6			
Nonexistent RNFL thickness in the FWHM algorithm (%)	0.0 OD	2.4	22.2	30.9	33.8
	0.3 OD	5.4			
	0.6 OD	8.5			
	0.9 OD	9.5			
Unsegmented TRT in the FWHM algorithm (%)	0.0 OD	0.0	0.2	0.6	0.1
	0.3 OD	0.0			
	0.6 OD	0.2			
	0.9 OD	2.3			

[a]IOP = intraocular pressure; * = measured before the onset of treatment (all IOPs were normalized at the time point the scan was made); MD = mean deviation.

Biomedical Optics EXPRESS

to remove speckle noise from individual B-scans to align individual A-scans accurately. A-scans were aligned in the z-direction based on intensity threshold. The intensity threshold for detecting retinal structures was set to 20 times the mean intensity of the noise level (mean intensity in the first 50 pixels in each individual A-scan). Peaks or high reflective areas in an A-scan are usually in the RNFL and in the RPE. When two highly reflective areas, separated by a low reflective area defined as a minimum of 20 pixels (34 μm) in length below the set intensity threshold, were identified, the set intensity threshold was used to find the edge of the inner limiting membrane. If fewer than two peaks were found, the intensity threshold was lowered by fourfold (to 5 times the mean intensity of the noise level) in order to detect peaks in case of a weak signal due to ND filters or cataract. If only one peak, assumed to be the RPE, or no peaks were found with either of the intensity threshold levels, the A-scan was assigned as NaN (typically in case of, for example, a floater or a blink during image acquisition). A total of 100 consecutive linear-scale A-scans were averaged to construct an RNFL aligned B-scan. If all 100 consecutive A-scans had NaN value, the B-scan was considered missing. Table 1 presents the average percentage of missing locations per B-scan in the corresponding dataset.

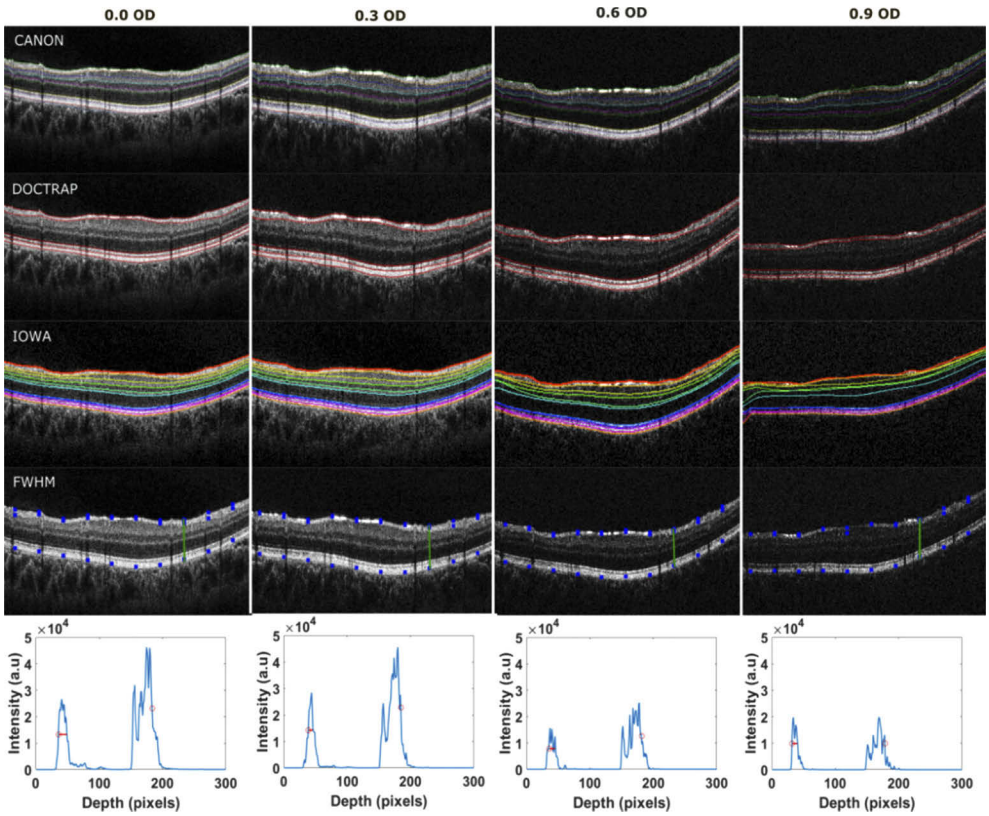

Fig. 2. Example segmentation results from different filters and different segmentation algorithms. From left to right: 0.0, 0.3, 0.6, and 0.9 OD filters. From top to bottom: CANON, DOCTRAP, IOWA, and FWHM. The green line in the FWHM B-scans represents the middle location of the averaged A-scan. Last row shows an averaged A-scan with the FWHM thickness (red line) and the TRT (red circles) fitted to the linear-scaled averaged A-scan (in blue). All images are collected from the same subject and were obtained from the same location in the retina.

Biomedical Optics EXPRESS

RNFL thickness was calculated, per B-scan, from the averaged 100 linear-scale A-scan as the FWHM above the noise floor (see Fig. 2). In order to avoid a too narrow thickness estimate from a noisy averaged A-scan, gaps of no more than 5 pixels (8.5 μm) were allowed between two consecutive peaks. When the distance was over 5 pixels, only the first peak was considered. In order to account for nonexistent RNFL in the different glaucoma conditions, RNFL thickness was defined to be zero if the maximum intensity of the RNFL peak was less than one-tenth of the maximum intensity of the RPE (see Table 1). TRT was estimated as the distance between the ILM border, used in FWHM estimation, and the half maximum of the RPE peak at the RPE/choroid border (TRT is marked with red circles in Figs. 2 & 3). If no RNFL peak was found, TRT was measured from the detected layer that fulfilled the intensity threshold, most likely the GCIPL. RNFL thicknesses from six vertically oriented B-scans were averaged in order to calculate the average RNFL thickness in a 1×0.5 mm rectangular area. These rectangular areas were attributed to the two ROIs described above (Fig. 1(B)).

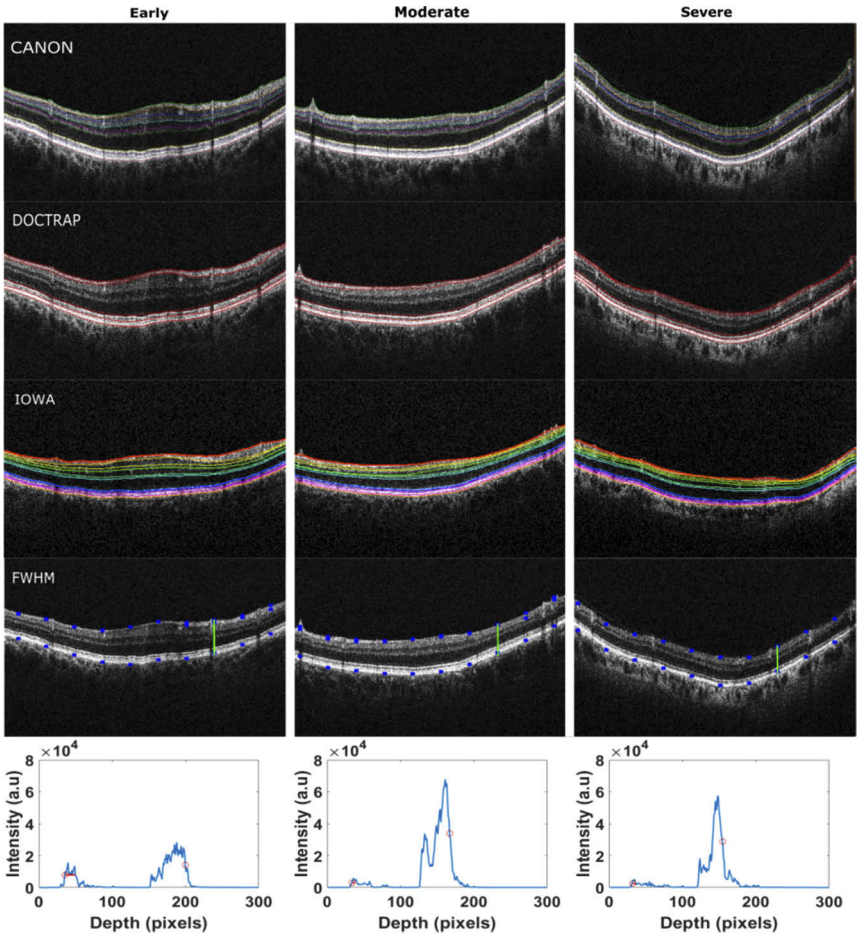

Fig. 3. Example segmentation results from different glaucoma conditions using different segmentation algorithms. From left to right the conditions are early, moderate, and severe. From top to bottom CANON, DOCTRAP, IOWA, and FWHM segmentation results. The green line in the FWHM B-scans represents the middle location of the averaged A-scan. Last row shows an averaged A-scan with the FWHM thickness (red line) and the TRT (red circles) fitted to the linear-scaled averaged A-scan (in blue).

Biomedical Optics EXPRESS

2.4. Data analysis

Mean RNFL thickness, GCIPL thickness, and TRT were calculated for both ROIs; the 10 mm ROI could not be analysed with the IOWA algorithm; the GCIPL thickness was only available for the Canon proprietary software and IOWA. Three images were taken per neutral density filter condition; all images were analysed with all algorithms. The mean thickness (calculated per filter condition and per algorithm) of the three images was used for the between-algorithm and between-filter condition analysis.

Data were presented as box plots showing for each algorithm and ROI, the RNFL thickness, GCIPL thickness, and TRTs as a function of optical density for healthy subjects and as a function of disease stage for glaucoma cases. We studied, for each algorithm separately, the influence of SNR with two-way ANOVA and the influence of disease presence and stage with one-way ANOVA. For post-hoc analysis, we used Bonferroni corrected t-tests. Repeatability was studied within, by using the three separate measurements, and between algorithms using the intraclass correlation coefficient (ICC) agreement. Because we expect a systematic difference between FWHM and the other algorithms (see Introduction section), we also used Pearson's correlation coefficient to compare the thickness measurements between FWHM and the other algorithms and we compared the thickness differences between FWHM and the other algorithms in the 0.0 OD condition using two-way ANOVA. The analysis was performed using R (version 3.6.0; R Foundation for Statistical Computing, Vienna, Austria). A P value of 0.05 or less was considered statistically significant.

3. Results

Figure 2 shows how the different algorithms (CANON, DOCTRAP, IOWA, and FWHM) segment the RNFL and the TRT, as a function of SNR (filter strength). All images are from the same subject and the same retinal location. Figure 3 illustrates the corresponding results as a function of glaucoma stage (early, moderate, and severe; from three different subjects).

Figures 4 and 5 show boxplots with RNFL thickness for the different algorithms as a function of filter strength (upper row) and disease stage (lower row), for the 5 mm (Fig. 4) and 10 mm (Fig. 5; not available for IOWA) ROI, respectively. Clearly, the RNFL is thinner in the 5 mm ROI than in the 10 mm ROI. FWHM yielded a thinner RNFL than the other algorithms in the no filter

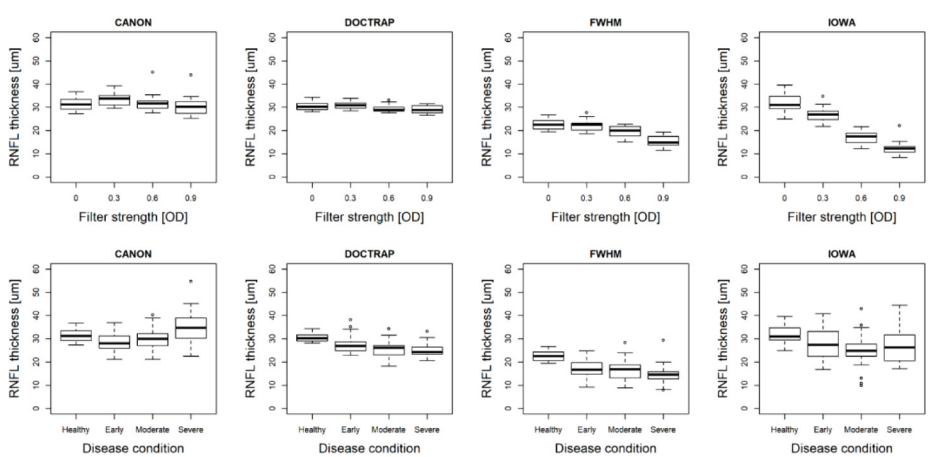

Fig. 4. RNFL thickness for the 5 mm ROI (between 0.5- and 2.5-mm eccentricity) using four different segmentation algorithms, as a function of optical density (upper row) and disease condition (lower row).

Biomedical Optics EXPRESS

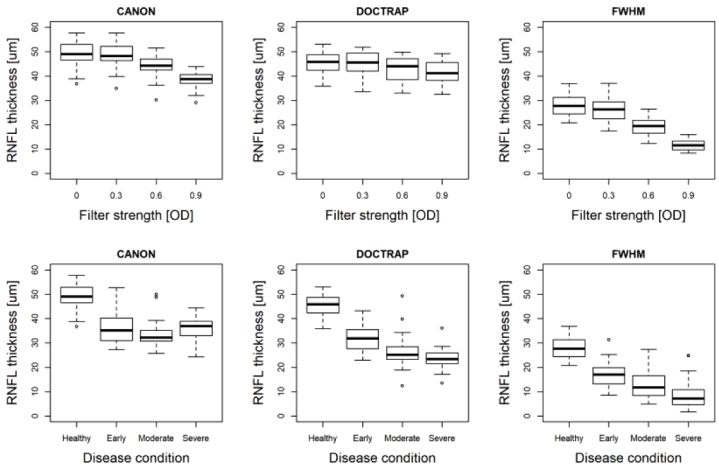

Fig. 5. RNFL thickness for the 10 mm ROI (between 2.5- and 5-mm eccentricity) using three different segmentation algorithms, as a function of optical density (upper row) and disease condition (lower row).

condition, for both the 5 mm ROI (median thickness 18.9 µm for FWHM versus 30.2, 31.3, and 31.0 µm for DOCTRAP, CANON and IOWA respectively; all P<0.001) and 10 mm ROI (27.7 µm for FWHM versus 45.8 and 49.0 µm for DOCTRAP and CANON, respectively; all P<0.001). The effect of glaucoma on the RNFL thickness was more pronounced for the 10 mm ROI than for the 5 mm ROI. Table 2 presents the corresponding Bonferroni corrected P values from the ANOVA analyses. All algorithms were more or less subject to layer thinning due to noise caused by the ND filters. For the 5 mm ROI, CANON was most robust regarding the applied filters but also failed to recognize glaucoma.

Figure 6 shows boxplots with GCIPL thickness for the two algorithms that provided this metric (CANON and IOWA) as a function of filter strength (upper row) and disease stage (lower row), for the 5 mm (both algorithms) and 10 mm (not available for IOWA) ROI. Obviously, the GCIPL is thicker in the 5 mm ROI than in the 10 mm ROI. Table 2 (middle part) presents the corresponding Bonferroni corrected P values from the ANOVA analyses. CANON but not IOWA was robust regarding the applied filters; both algorithms easily detected the presence of glaucoma.

Figure 7 shows boxplots with TRT for the different algorithms as a function of filter strength (upper row) and disease stage (lower row), for the 5 mm (DOCTRAP, IOWA, and FWHM) and 10 mm ROI (DOCTRAP and FWHM). TRT is thicker in the 5 mm ROI than in the 10 mm ROI. Table 2 (lower part) presents the corresponding Bonferroni corrected P values from the ANOVA analyses. TRT measurements were influenced by the applied filters for all algorithms. IOWA and FWHM appeared to be robust against noise up to the 0.3 OD filter; DOCTRAP up to the 0.6 OD filter. However, there was a slightly thicker TRT for the 0.3 OD filter for DOCTRAP, for both ROIs.

The last three columns of Table 2 show the suitability of the algorithms, layers, and ROIs for discriminating between the successive glaucoma stages. Regarding the layers, GCIPL and TRT were more suitable for discriminating between early and moderate glaucoma than RNFL; for RNFL, the 10 mm ROI was better than the 5 mm ROI. All current algorithms failed to discriminate between moderate and severe glaucoma, for all employed layers and ROIs, except for CANON assessing the GCIPL at 10 mm ROI. For CANON, the RNFL thickness seemed also significantly different between moderate and severe glaucoma at 5 mm ROI, but the RNFL thickness appeared to increase rather than decrease with disease severity, suggesting a chance

Biomedical Optics EXPRESS

Table 2. Post-hoc analyses of ANOVA results for the data presented in Figs. 4, 5, 6 and 7

Algorithm	Healthy[a]			Glaucoma[b]					
	0.3 OD	0.6 OD	0.9 OD	healthy vs early	healthy vs moderate	healthy vs severe	early vs moderate	early vs severe	moderate vs severe
RNFL at 5 mm ROI									
CANON	<0.001	1.0	1.0	0.27	1.0	0.63	1.0	<0.001	0.03
DOCTRAP	0.30	<0.001	<0.001	0.014	<0.001	<0.001	0.21	0.021	1.0
IOWA	<0.001	<0.001	<0.001	0.17	0.008	0.25	1.0	1.0	1.0
FWHM	1.0	<0.001	<0.001	<0.001	<0.001	<0.001	0.48	0.0039	0.43
RNFL at 10 mm ROI									
CANON	0.27	<0.001	<0.001	<0.001	<0.001	<0.001	0.53	1.0	0.94
DOCTRAP	0.29	<0.001	<0.001	<0.001	<0.001	<0.001	0.002	<0.001	0.35
FWHM	0.2	<0.001	<0.001	<0.001	<0.001	<0.001	0.13	<0.001	0.056
GCIPL at 5 mm ROI									
CANON	+	+	+	<0.001	<0.001	<0.001	<0.001	<0.001	0.07
IOWA	0.04	0.002	<0.001	<0.001	<0.001	<0.001	0.005	<0.001	1.0
GCIPL at 10 mm ROI									
CANON	0.93	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	0.03
TRT at 5 mm ROI									
DOCTRAP	<0.001	1.0	<0.001	<0.001	<0.001	<0.001	0.006	<0.001	1.0
IOWA	0.54	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	1.0
FWHM	0.85	<0.001	<0.001	0.0025	<0.001	<0.001	0.02	0.003	1.0
TRT at 10 mm ROI									
DOCTRAP	0.019	1.0	0.09	<0.001	<0.001	<0.001	0.0011	<0.001	1.0
FWHM	1.0	0.001	0.002	<0.001	<0.001	<0.001	<0.001	0.0011	1.0

[a]Bonferroni corrected paired t-test P values for comparing results without filter against different filter strengths,
[b]Bonferroni corrected t-test P values for comparing healthy subjects against different glaucoma conditions and different glaucoma conditions against each other, + ANOVA not significant

finding. Moreover, CANON failed for this combination of layer and ROI to discriminate between healthy and glaucoma.

Table 3 depicts the intraclass correlation coefficient (ICC) agreement for the RNFL, GCIPL and, TRT assessments, both between different segmentation algorithms and within each algorithm, the latter based on the three measurements performed for each filter strength. According to the interpretation guidelines by Cicchetti [24], TRT showed excellent agreement for all tested algorithms (ICC > 0.75). All four algorithms showed good (ICC > 0.60) repeatability for the RNFL thickness in the 5 mm ROI and excellent repeatability for the 10 mm ROI for essentially all filter strengths. For the GCIPL thickness, CANON showed excellent repeatability for all filter strengths whereas IOWA showed excellent repeatability only for 0.0 and 0.3 OD. Regarding the between algorithm comparisons, the agreement was excellent for TRT. For RNFL and GCIPL the algorithm comparisons were poor (ICC < 0.4), with the exception of a good agreement between CANON and DOCTRAP for the 10 mm ROI RNFL and an excellent agreement between CANON and IOWA for the 5 mm ROI GCIPL (up to 0.3 OD). Because there is a systematic difference in RNFL thickness between FWHM and the other algorithms (see above), compromising the ICC agreement, we also calculated Pearson's correlation coefficients for RNFL thickness between FWHM and CANON, IOWA, and DOCTRAP in the 5 mm RNFL ROI, yielding 0.58, 0.44, and

Biomedical Optics EXPRESS

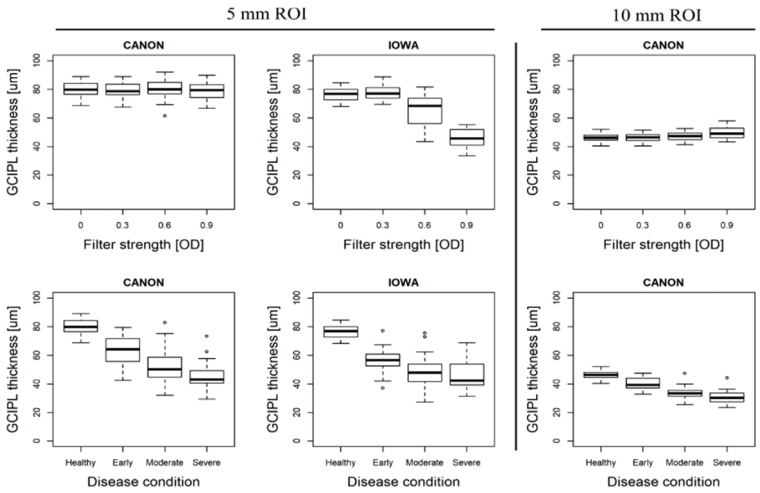

Fig. 6. GCIPL (combined GCL and IPL) layer thickness for the 5 mm ROI (between 0.5 and 2.5 mm eccentricity) using the CANON (left column) and IOWA (middle column) software, and for the 10 mm ROI (between 2.5 and 5 mm eccentricity) using the CANON software (right column), as a function of optical density (upper row) and disease condition (lower row).

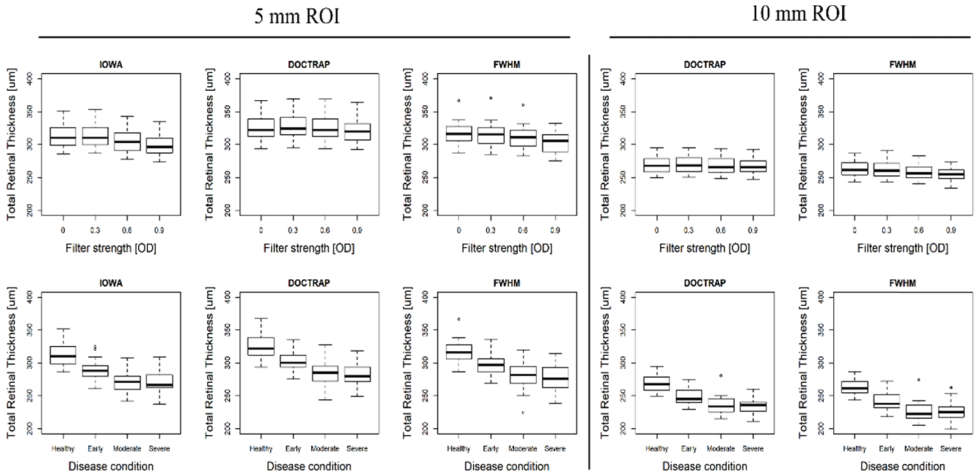

Fig. 7. Total retinal thickness (TRT) for the 5 mm ROI (eccentricity between 0.5 and 2.5 mm) using three different segmentation algorithms and for the 10 mm ROI (eccentricity between 2.5 mm and 5 mm) using two different segmentation algorithms, as a function of optical density (upper row) and disease condition (lower row).

0.48 respectively. For the 10 mm ROI, the correlation coefficient between FWHM and CANON was 0.67 and between FWHM and DOCTRAP 0.64.

Biomedical Optics EXPRESS

Table 3. Two-way intraclass correlation (ICC) depicting test-retest within and agreement between different segmentation algorithms[a]

				Healthy				Glaucoma		
Layer and ROI	Explanation	Algorithm 1	Algorithm 2	0.0 OD	0.3 OD	0.6 OD	0.9 OD	Early	Moderate	Severe
RNFL 5 mm	Test-retest	CANON	CANON	0.88	0.81	0.68	0.88	-	-	-
		DOCTRAP	DOCTRAP	0.92	0.83	0.79	0.70	-	-	-
		FWHM	FWHM	0.73	0.67	0.72	0.70	-	-	-
		IOWA	IOWA	0.73	0.63	0.66	0.19	-	-	-
	Comparison	CANON	DOCTRAP	0.69	0.42	0.25	0.34	0.79	0.19	0.06
		CANON	FWHM	0.05	0.03	0.02	0.02	0.05	0.02	0.01
		CANON	IOWA	0.56	0.10	0.02	0.01	0.44	0.26	0.11
		DOCTRAP	FWHM	0.03	0.04	0.02	0.00	0.05	0.09	0.07
		DOCTRAP	IOWA	0.46	0.17	0.02	0.01	0.56	0.12	0.14
		FWHM	IOWA	0.04	0.09	0.30	0.04	0.06	0.09	0.06
RNFL 10 mm	Test-retest	CANON	CANON	1	0.99	0.94	0.82	-	-	-
		DOCTRAP	DOCTRAP	0.97	0.97	0.97	0.91	-	-	-
		FWHM	FWHM	0.88	0.79	0.76	0.53	-	-	-
	Comparison	CANON	DOCTRAP	0.78	0.78	0.87	0.70	0.61	0.47	0.08
		CANON	FWHM	0.07	0.05	0.03	0.01	0.10	0.09	0.02
		DOCTRAP	FWHM	0.08	0.07	0.03	0.01	0.15	0.21	0.02
GCIPL 5 mm	Test-retest	CANON	CANON	0.97	0.96	0.91	0.91	-	-	-
		IOWA	IOWA	0.89	0.90	0.68	0.19	-	-	-
	Comparison	CANON	IOWA	0.77	0.80	0.11	0.00	0.71	0.85	0.48
GCIPL 10 mm	Test-retest	CANON	CANON	1	0.99	0.97	0.91	-	-	-
		DOCTRAP	DOCTRAP	1	1	1	0.98			
TRT - 5 mm	Test-retest	FWHM	FWHM	1	0.98	0.97	0.94			
		IOWA	IOWA	1	0.99	0.98	0.90			
	Comparison	DOCTRAP	FWHM	0.91	0.85	0.78	0.64	0.91	0.96	0.93
		DOCTRAP	I	0.81	0.78	0.65	0.51	0.7	0.74	0.8
		FWHM	I	0.92	0.93	0.87	0.83	0.81	0.83	0.87
TRT 10 mm	Test-retest	DOCTRAP	DOCTRAP	1	1	1	0.96			
		FWHM	FWHM	1	0.99	0.97	0.63			
	Comparison	FWHM	DOCTRAP	0.87	0.82	0.64	0.5	0.79	0.79	0.48

[a]TRT = Total retinal thickness, RNFL = Retinal nerve fiber layer, GCL = Ganglion cell layer, ROI = Region of interest, OD = Optical density

4. Discussion

Optical coherence tomography segmentation results of retinal layers are influenced by artificial noise - most layer thicknesses decrease with increasing neutral density filter strength. For RNFL, CANON is most robust against this degrading effect, followed by DOCTRAP and FWHM, and IOWA. FWHM gives a thinner RNFL than the other algorithms. For GCIPL, CANON but not IOWA is robust against the effect of the applied filters. For TRT, DOCTRAP, FWHM, and IOWA show a similar robustness. GCIPL and TRT are more suitable for discriminating between healthy and glaucoma, and between early and moderate glaucoma, than RNFL; for RNFL, the 10 mm

Biomedical Optics EXPRESS

ROI is better than the 5 mm ROI. The current algorithms mostly fail to discriminate between moderate and severe glaucoma, for all employed layers and ROIs.

4.1. Comparison with literature

Darma et al. showed, using the IOWA algorithm, that an optical density of 0.26 caused by an artificial filter produces a clinically meaningful decrease of 1 SD in the GCC thickness [5]. We found a similar layer thinning at 0.3 OD for the IOWA algorithm for both RNFL and GCIPL at the 5 mm ROI used by IOWA. Other algorithms were robust against the influence of noise at 0.3 OD, and showed a smaller decrease in layer thickness for higher filter strengths. Darma et al. also reported significant differences in TRT measurements, although these differences were within one SD from the baseline measurements even at 0.67 OD, and thus deemed not to be clinically meaningful [5]. In agreement with this, in our data different filter strengths had only a small influence on the TRT in all algorithms and ROIs assessed. Differences were significant, but the IQRs amply overlapped across the different filter strengths (Fig. 7).

Wolf-Schnurrbusch et al. compared six different OCT devices in their ability to measure central retinal thickness, and conclude that different OCT devices cannot be used interchangeably for the measurement of macular thickness, and that the differences between different devices are due to differences in the segmentation software [25]. A more recent study by Cho and Hwang showed, in four different OCT devices, that the macular retinal thickness in 32 healthy eyes differed by 0.7 to 27.3 µm, and only results between Spectralis OCT and RS-3000 Advance OCT showed good agreement [26]. Testing between devices is usually performed using the device-specific software, and therefore differences between devices can be attributed either to differences in the hardware or to differences in the built-in segmentation software. In our previous study, we showed that eleven different OCT devices produce comparable results, when the thickness is analyzed in an identical way, suggesting that differences due to hardware can possibly be eliminated [21]. Here we compared measurements from a single device using four different algorithms, and showed variable agreement between different algorithms in RNFL and GCIPL measurements, in both ROIs, which worsened even more in noisy conditions. Different algorithms, however, showed good agreement when assessing TRT, even in noisy conditions, possibly due to similarly defined ILM and choroid borders.

Grossly speaking, our data suggest that discriminating between moderate and severe glaucoma with OCT is difficult, indirectly indicating its limited value for progression detection in these stages. This is often attributed to a floor effect. Earlier, peripapillary RNFL (pRNFL) measurements have been shown to reach a measurement floor at a standard automated perimetry mean deviation somewhere between -10 and -15 dB [12,27,28]. It has been suggested, however, that macular measurements may be superior to peripapillary measurements for detecting change between moderate and severe glaucoma [29]. Our results suggest that detecting progression from RNFL, GCIPL, or TRT in either ROIs with any of the algorithms is challenging. Our findings showing that the mean GCIPL thickness does not produce significant differences in the late-stages of glaucoma is in agreement with findings by Sung et al. [13]. However some previous studies have shown that the GCIPL between the fovea and the optic disc, being an area that is spared relatively long in glaucoma, could provide valuable information on late disease progression [13–15]. We explored this in our data by looking at the GCIPL thickness in area 2 and 4 in Fig. 1(A). We found no clear differences between moderate and severe glaucoma for either IOWA (P=1.0) or CANON (P=0.05).

4.2. Limitations and strengths

A limitation of our study is that we were unable to measure exactly the same areas with all of the algorithms. Similarly, some algorithms were limited in the number of layers they could segment. While the Canon has an imaging area of 10×10 mm, other OCT devices imaging areas may be

Biomedical Optics EXPRESS

limited to an area of 6×6 mm. The measurement area in the Canon device is divided into eight regions (Fig. 1(A)), of which four are in the inner 5 mm ROI and the rest in the outer 10 mm ROI. IOWA analyzes preset areas within a 6 mm circle. We chose their 10-2 grid as it matches the 5 mm ROI as accurately as possible (Fig. 1(C)). DOCTRAP does not have segmentation for the GCL or IPL layers. For DOCTRAP we used the same grid as for the FWHM algorithm (Fig. 1(B)). Large ROIs were chosen as they have higher test-retest accuracy [30]. Although our primary outcome was the macular RNFL thickness, under- and overestimation biases of the RNFL thickness undoubtedly lead to an observed change in the underlying layers, mainly the GCL [6]. A strength of our study is that we compared four different algorithms under a wide range of SNR conditions in the same group of healthy subjects, and that we applied the same algorithms and methodology to subjects with different glaucoma stages. This contrasts with most earlier studies on the effect of SNR on OCT layer thickness assessment, which used only the device's proprietary algorithm or a single third-party algorithm. We used ND filters over computational SNR degradation. ND filters offer a physical proof of signal strength degradation and, importantly, they can be used in case of proprietary software inside the device where access to the raw images - needed for computational SNR degradation - is not possible. As all scans were acquired using the Canon device, the algorithm comparisons could be biased towards the CANON algorithm. In this regard it is worthwhile to mention that software differences seem to matter more than hardware differences in current commercial devices [21].

Importantly, the healthy subjects and glaucoma patients in our study differed regarding age. As such, any differences found between these two groups could be due to either glaucoma or ageing. The reason for the age difference is that we wanted to have clear optics, and limited variability therein, in our primary experiment, the effect of filter strength on OCT layer thickness assessment. This implies young healthy subjects. Despite the large database of the Groningen Longitudinal Glaucoma Study, it was not possible to find a reasonable number of age-similar glaucoma patients. Within the glaucoma patients, however, the stage subgroups were age-similar and differences in layer thicknesses between these groups could thus be attributed to glaucoma.

4.3. Implications

In this study we investigated whether low SNR images could be reliably segmented with the algorithms in question. As the answer - in general - is no, criteria for a good image become pivotal. The canon OCT provides an image quality index. The recommended value of this index is 7 or higher (scale: 0-10). However, the technician should not just observe the index, but should also pay attention to blink and motion. Interestingly, the value of 7 was only reached (on average) for the no filter condition in the healthy subjects (Table 1). As such, the cut-off point seems well-chosen. In glaucoma, however, the mean index value was below 7 (Table 1), indicating that clinicians often have to deal with scans that are formally suboptimal.

What are the clinical implications of our findings? First, in the past, apparent thinning of retinal layers has been observed in patients with cataract, where the RNFL thickness increased after cataract surgery [4,31,32]. Such a thickness increase could be due to improvement of the optical properties of the eye or to a real thickness change. Our study confirms the strong association between image quality and layer thicknesses as observed with OCT. Therefore, it is crucial to emphasize in a clinical setting that the best possible image quality has to be achieved during image acquisition. The acquisition time with OCT is only a few seconds, and averaging more B-scans or taking an entirely new scan doesn't take too much of the technician's or patient's time, but could greatly improve the accuracy of OCT interpretation. However, averaging is not possible in all OCT devices, and in devices where averaging is possible, excessive averaging increases the acquisition time, which in turn will result in more blink artifacts, motion artifacts, and discomfort for the patient. Recently, post-processing techniques that use deep learning for denoising the images have been proposed to achieve a better image quality for segmentation [33].

Biomedical Optics EXPRESS

Albeit promising, such post-processing methods are not a feasible option in a clinical setting; for this, the entire process has to be implemented in the device itself. For the time being, the most important message is that technicians should be instructed to check the acquired image and to repeat the test in case of a low SNR image, to see if the image quality can be improved by better scan alignment, and in case of blink or patient movement. Second, retinal layers have been reported to become thinner with age [34,35]. Apart from a real thinning, a degradation of the optical quality of the eye with age may have its impact as well. A thorough understanding of the effect of image quality on layer thickness assessment is pivotal for the interpretation of longitudinal data, both in healthy subjects - for understanding normal ageing - and in glaucoma patients - for an unbiased progression detection.

With current knowledge, macular RNFL thickness assessment is clearly sensitive to noise, and only at the 10 mm ROI, RNFL thickness seems to be a meaningful biomarker of disease. Hence, for studying changes in the inner retina with a macular OCT scan, the GCIPL seems to be preferable over mRNFL, being robust against noise and offering good disease detection and separation. This is in agreement with findings by Khawaja et al. from a large cohort of UK Biobank participants, where they concluded that for detecting glaucomatous and macular pathophysiologic processes, the GCIPL, over mRNFL and GCC, may be the superior inner retinal biomarker [35]. The GCC is usually used to study the usability of the macular OCT scan in glaucoma detection, as it has been shown to have the same diagnostic performance in detecting early, moderate and severe glaucoma as the pRNFL [36,37]. Kim et al. found that the GCC and RGCL thickness in the 3 to 6 mm diameter macular ROI were the best macular parameters for discriminating between early glaucoma and healthy. The only macular parameter that was not significantly different between glaucoma and healthy was the RNFL in their smallest ROI (1 to 2 mm diameter) [38]. In a population-based study by Springelkamp et al., the inferior part of the GCL offered the highest sensitivity at 97.5% specificity and area under the ROC-curve, compared to other macular or peripapillary parameters [39].

Total macular thickness has been suggested as a potential alternative to assess structural changes in patients with glaucoma [40]. However, earlier studies suggested that macular thickness does not outperform pRNFL in glaucoma detection [9,41,42]. Our findings suggest that TRT is both sensitive to disease induced changes and robust against noise, apparently outperforming mRNFL and GCIPL, and could therefore be a viable measure of glaucomatous damage. Hence, the role of TRT in glaucomatous damage and change detection needs further research.

What are the implications of our findings regarding our knowledge of the true RNFL thickness? In addition to layer thinning due to noise, retinal layers may be overestimated. Layer thickness overestimation has been shown in phantom eyes where a 49 µm layer thickness can be overestimated by 13-23 µm depending on the OCT device and their built-in algorithm used [43]. We confirmed this overestimation by performing phantom eye measurements similar to that of de Kinkelder et al. and developed a theoretical framework for this overestimation [6]. The FWHM approach was the spin-off of this work, and should be able to provide an unbiased RNFL thickness estimate, insensitive to noise. Clearly, FWHM yielded in the current study a thinner RNFL, indicating that indeed other algorithms overestimate this layer. FWHM thickness measurements provided on average a 11.9 µm thinner RNFL thickness in the 5 mm ROI and a 19.7 µm thinner thickness in the 10 mm ROI compared to the other algorithms. Whether this thickness difference can be tied to an actual overestimation, remains an open question since validating measurements in histology is extremely challenging. The phantom eye measurements, however, clearly support the hypothesis that the real RNFL is thinner than commonly reported [21,43]. As a practical method for clinical studies, FWHM was more robust against noise than IOWA, but did not surpass the other two algorithms.

Biomedical Optics EXPRESS

4.4. Conclusion

In conclusion, we showed that the four algorithms used in this study are all susceptible to noise at a varying degree, and the results between different algorithms are not interchangeable. The algorithms differ in their ability to differentiate between young healthy eyes and older glaucoma eyes, and they all fail to accurately separate different glaucoma stages from each other based solely on the RNFL thickness in the macular area. From the RNFL, GCIPL, and TRT measurements, the TRT was the most robust against noise, had the highest agreement between different algorithms, and was the best metric in separating early and moderate and early and severe glaucoma from each other. A cross-sectional screening study with age-similar healthy subjects and glaucoma patients and a longitudinal study in a large glaucoma cohort exploring the value of this parameter would be the logical next steps. None of the algorithms compared were able to separate moderate and severe glaucoma from each other, leaving a clear challenge in ocular imaging.

Funding

H2020 Marie Skłodowska-Curie Actions (661883); National Research Foundation of Korea (2019H1D3A2A02101784).

Disclosures

The authors declare no conflicts of interest

References

1. J. S. Schuman, "Quantification of Nerve Fiber Layer Thickness in Normal and Glaucomatous Eyes Using Optical Coherence Tomography," Arch. Ophthalmol. **113**(5), 586 (1995).
2. C. Y. L. Cheung, C. K. S. Leung, D. Lin, C.-P. Pang, and D. S. C. Lam, "Relationship between retinal nerve fiber layer measurement and signal strength in optical coherence tomography," Ophthalmology **115**(8), 1347–1351.e2 (2008).
3. J. M. Schmitt, S. H. Xiang, and K. M. Yung, "Speckle in optical coherence tomography," J. Biomed. Opt. **4**(1), 95–105 (1999).
4. D. W. Lee, J. M. Kim, K. H. Park, C. Y. Choi, and J. G. Cho, "Effect of media opacity on retinal nerve fiber layer thickness measurements by optical coherence tomography," J. Ophthalmic Vis. Res. **5**(3), 151–157 (2010).
5. S. Darma, P. H. B. Kok, T. J. T. P. van den Berg, M. D. Abràmoff, D. J. Faber, C. A. Hulsman, F. Zantvoord, M. P. Mourits, R. O. Schlingemann, and F. D. Verbraak, "Optical density filters modeling media opacities cause decreased SD-OCT retinal layer thickness measurements with inter- and intra-individual variation," Acta Ophthalmol. **93**(4), 355–361 (2015).
6. N. M. Jansonius, J. Cervantes, M. Reddikumar, and B. Cense, "Influence of coherence length, signal-to-noise ratio, log transform, and low-pass filtering on layer thickness assessment with OCT in the retina," Biomed. Opt. Express **7**(11), 4490–4500 (2016).
7. K. A. Vermeer, J. van der Schoot, H. G. Lemij, and J. F. de Boer, "RPE-normalized RNFL attenuation coefficient maps derived from volumetric OCT imaging for glaucoma assessment," Invest. Ophthalmol. Visual Sci. **53**(10), 6102–6108 (2012).
8. K. A. Vermeer, J. Mo, J. J. A. Weda, H. G. Lemij, and J. F. de Boer, "Depth-resolved model-based reconstruction of attenuation coefficients in optical coherence tomography," Biomed. Opt. Express **5**(1), 322–337 (2014).
9. Y. Nakatani, T. Higashide, S. Ohkubo, H. Takeda, and K. Sugiyama, "Evaluation of macular thickness and peripapillary retinal nerve fiber layer thickness for detection of early glaucoma using spectral domain optical coherence tomography," J. Glaucoma **20**(4), 252–259 (2011).
10. M. K. Garvin, M. D. Abràmoff, X. Wu, S. R. Russell, T. L. Burns, and M. Sonka, "Automated 3-D intraretinal layer segmentation of macular spectral-domain optical coherence tomography images," IEEE Trans. Med. Imaging **28**(9), 1436–1447 (2009).
11. S. J. Chiu, X. T. Li, P. Nicholas, C. A. Toth, J. A. Izatt, and S. Farsiu, "Automatic segmentation of seven retinal layers in SDOCT images congruent with expert manual segmentation," Opt. Express **18**(18), 19413–19428 (2010).
12. J.-C. Mwanza, D. L. Budenz, J. L. Warren, A. D. Webel, C. E. Reynolds, D. T. Barbosa, and S. Lin, "Retinal nerve fibre layer thickness floor and corresponding functional loss in glaucoma," Br. J. Ophthalmol. **99**(6), 732–737 (2015).
13. M. S. Sung, H. Heo, and S. W. Park, "Structure-function Relationship in Advanced Glaucoma After Reaching the RNFL Floor," J. Glaucoma **28**(11), 1006–1011 (2019).
14. K. E. Kim, K. H. Park, J. W. Jeoung, S. H. Kim, and D. M. Kim, "Severity-dependent association between ganglion cell inner plexiform layer thickness and macular mean sensitivity in open-angle glaucoma," Acta Ophthalmol. **92**(8), e650–e656 (2014).

Biomedical Optics EXPRESS

15. D. C. Hood, A. S. Raza, C. G. V. de Moraes, J. M. Liebmann, and R. Ritch, "Glaucomatous damage of the macula," Prog. Retinal Eye Res. **32**, 1–21 (2013).
16. G. P. Heeg, L. J. Blanksma, P. L. L. J. Hardus, and N. M. Jansonius, "The Groningen Longitudinal Glaucoma Study. I. Baseline sensitivity and specificity of the frequency doubling perimeter and the GDx nerve fibre analyser," Acta Ophthalmol. Scand. **83**(1), 46–52 (2005).
17. M. D. Abramoff, M. K. Garvin, and M. Sonka, "Retinal Imaging and Image Analysis," IEEE Rev. Biomed. Eng. **3**, 169–208 (2010).
18. K. Li, X. Wu, D. Z. Chen, and M. Sonka, "Optimal surface segmentation in volumetric images–a graph-theoretic approach," IEEE Trans. Pattern Anal. Mach. Intell. **28**(1), 119–134 (2006).
19. J. Y. Lee, S. J. Chiu, P. P. Srinivasan, J. A. Izatt, C. A. Toth, S. Farsiu, and G. J. Jaffe, "Fully Automatic Software for Retinal Thickness in Eyes With Diabetic Macular Edema From Images Acquired by Cirrus and Spectralis Systems," Invest. Ophthalmol. Visual Sci. **54**(12), 7595 (2013).
20. S. J. Chiu, J. A. Izatt, R. V. O'Connell, K. P. Winter, C. A. Toth, and S. Farsiu, "Validated automatic segmentation of AMD pathology including drusen and geographic atrophy in SD-OCT images," Invest. Ophthalmol. Visual Sci. **53**(1), 53–61 (2012).
21. T. Heikka, G. Ometto, T. Rowe, N. Jansonius, and D. Crabb, "Testing a phantom eye under various signal to noise ratio conditions using eleven different OCT devices," Biomed. Opt. Express **11**(3), 1306–1315 (2020).
22. N. Nassif, B. Cense, B. H. Park, S. H. Yun, T. C. Chen, B. E. Bouma, G. J. Tearney, and J. F. de Boer, "In vivo human retinal imaging by ultrahigh-speed spectral domain optical coherence tomography," Opt. Lett. **29**(5), 480–482 (2004).
23. W. Drexler and J. G. Fujimoto, *Optical Coherence Tomography: Technology and Applications* (Springer Science & Business Media, 2015), Chap 5.
24. D. V. Cicchetti, "Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology," Psychol. Assess. **6**(4), 284–290 (1994).
25. U. E. K. Wolf-Schnurrbusch, L. Ceklic, C. K. Brinkmann, M. E. Iliev, M. Frey, S. P. Rothenbuehler, V. Enzmann, and S. Wolf, "Macular Thickness Measurements in Healthy Eyes Using Six Different Optical Coherence Tomography Instruments," Invest. Ophthalmol. Visual Sci. **50**(7), 3432 (2009).
26. S. Y. Cho and Y. H. Hwang, "Comparison of Macular Retinal Thickness among Four Optical Coherence Tomography Devices in Healthy Young Subjects," J. Korean Ophthalmol. Soc. **60**(5), 434 (2019).
27. D. C. Hood, S. C. Anderson, M. Wall, and R. H. Kardon, "Structure versus function in glaucoma: an application of a linear model," Invest. Ophthalmol. Visual Sci. **48**(8), 3662–3668 (2007).
28. D. C. Hood and R. H. Kardon, "A framework for comparing structural and functional measures of glaucomatous damage," Prog. Retinal Eye Res. **26**(6), 688–710 (2007).
29. S. Moghimi, C. Bowd, L. M. Zangwill, R. C. Penteado, K. Hasenstab, H. Hou, E. Ghahari, P. I. C. Manalastas, J. Proudfoot, and R. N. Weinreb, "Measurement Floors and Dynamic Ranges of OCT and OCT Angiography in Glaucoma," Ophthalmology **126**(7), 980–988 (2019).
30. H. Springelkamp, K. Lee, W. D. Ramdas, J. R. Vingerling, A. Hofman, C. C. W. Klaver, M. Sonka, M. D. Abràmoff, and N. M. Jansonius, "Optimizing the information yield of 3-D OCT in glaucoma," Invest. Ophthalmol. Visual Sci. **53**(13), 8162–8171 (2012).
31. M. El-Ashry, S. Appaswamy, S. Deokule, and S. Pagliarini, "The effect of phacoemulsification cataract surgery on the measurement of retinal nerve fiber layer thickness using optical coherence tomography," Curr. Eye Res. **31**(5), 409–413 (2006).
32. P. H. B. Kok, T. J. T. P. van den Berg, H. W. van Dijk, M. Stehouwer, I. J. E. van der Meulen, M. P. Mourits, and F. D. Verbraak, "The relationship between the optical density of cataract and its influence on retinal nerve fibre layer thickness measured with spectral domain optical coherence tomography," Acta Ophthalmol. **91**(5), 418–424 (2013).
33. S. K. Devalla, G. Subramanian, T. H. Pham, X. Wang, S. Perera, T. A. Tun, T. Aung, L. Schmetterer, A. H. Thiéry, and M. J. A. Girard, "A Deep Learning Approach to Denoise Optical Coherence Tomography Images of the Optic Nerve Head," Sci. Rep. **9**(1), 14454 (2019).
34. W. K. Song, S. C. Lee, E. S. Lee, C. Y. Kim, and S. S. Kim, "Macular Thickness Variations with Sex, Age, and Axial Length in Healthy Subjects: A Spectral Domain–Optical Coherence Tomography Study," Invest. Ophthalmol. Visual Sci. **51**(8), 3913 (2010).
35. A. P. Khawaja, S. Chua, P. G. Hysi, S. Georgoulas, H. Currant, T. W. Fitzgerald, E. Birney, F. Ko, Q. Yang, C. Reisman, D. F. Garway-Heath, C. J. Hammond, P. T. Khaw, P. J. Foster, P. J. Patel, N. Strouthidis, U. K. Biobank Eye, and V. Consortium, "Comparison of Associations with Different Macular Inner Retinal Thickness Parameters in a Large Cohort: The UK Biobank," Ophthalmology **127**(1), 62–71 (2020).
36. N. R. Kim, E. S. Lee, G. J. Seong, J. H. Kim, H. G. An, and C. Y. Kim, "Structure–Function Relationship and Diagnostic Value of Macular Ganglion Cell Complex Measurement Using Fourier-Domain OCT in Glaucoma," Invest. Ophthalmol. Visual Sci. **51**(9), 4646 (2010).
37. N. A. Loewen, X. Zhang, O. Tan, B. A. Francis, D. S. Greenfield, J. S. Schuman, R. Varma, and D. Huang, "Combining measurements from three anatomical areas for glaucoma diagnosis using Fourier-domain optical coherence tomography," Br. J. Ophthalmol. **99**(9), 1224–1229 (2015).
38. H. J. Kim, K. H. Park, Y. K. Kim, and J. W. Jeoung, "Evaluation of Layer-by-Layer Segmented Ganglion Cell Complex Thickness for Detecting Early Glaucoma According to Different Macular Grids," J. Glaucoma **26**(8), 712–717 (2017).

Biomedical Optics EXPRESS

39. H. Springelkamp, K. Lee, R. C. W. Wolfs, G. H. S. Buitendijk, W. D. Ramdas, A. Hofman, J. R. Vingerling, C. C. W. Klaver, M. D. Abramoff, and N. M. Jansonius, "Population-Based Evaluation of Retinal Nerve Fiber Layer, Retinal Ganglion Cell Layer, and Inner Plexiform Layer as a Diagnostic Tool For Glaucoma," Invest. Ophthalmol. Visual Sci. **55**(12), 8428–8438 (2014).
40. K. R. Sung, J. H. Sun, J. H. Na, J. Y. Lee, and Y. Lee, "Progression Detection Capability of Macular Thickness in Advanced Glaucomatous Eyes," Ophthalmology **119**(2), 308–313 (2012).
41. G. Wollstein, H. Ishikawa, J. Wang, S. A. Beaton, and J. S. Schuman, "Comparison of three optical coherence tomography scanning areas for detection of glaucomatous damage," Am. J. Ophthalmol. **139**(1), 39–43 (2005).
42. F. A. Medeiros, L. M. Zangwill, L. M. Alencar, C. Bowd, P. A. Sample, R. Susanna Jr, and R. N. Weinreb, "Detection of glaucoma progression with stratus OCT retinal nerve fiber layer, optic nerve head, and macular thickness measurements," Invest. Ophthalmol. Visual Sci. **50**(12), 5741–5748 (2009).
43. R. de Kinkelder, D. M. de Bruin, F. D. Verbraak, T. G. van Leeuwen, and D. J. Faber, "Comparison of retinal nerve fiber layer thickness measurements by spectral-domain optical coherence tomography systems using a phantom eye model," J. Biophotonics **6**(4), 314–320 (2013).
44. K. Lee, Y. H. Kwon, M. K. Garvin, M. Niemeijer, M. Sonka, and M. D. Abràmoff, "Distribution of damage to the entire retinal ganglion cell pathway: quantified using spectral-domain optical coherence tomography analysis in patients with glaucoma," Arch. Ophthalmol. **130**(9), 1118–1126 (2012).