Repeatability of Manual Subfoveal Choroidal Thickness Measurements in Healthy Subjects Using the Technique of Enhanced Depth Imaging Optical Coherence Tomography

Wabeeda Rahman, Fred Kuanfu Chen, Jonathan Yeoh, Praveen Patel, Adnan Tufail, and Lyndon Da Cruz

PURPOSE. The aim of this study was to investigate the repeatability of manual measurements of choroidal thickness in healthy subjects imaged on spectral domain optical coherence tomography (OCT) using the enhanced depth imaging (EDI) technique.

METHODS. Fifty consecutive, healthy, young, adult volunteers with no known eye disease were enrolled prospectively. Two good-quality horizontal and vertical line scans through the fovea were obtained for each eye. Using the manual calipers provided by the software of the proprietary device, two experienced OCT readers measured the subfoveal choroidal thickness (SFCT) of the horizontal and vertical line scans for all eyes. The readers were masked to each other’s readings. Intraobserver, interobserver, and intrasession coefficients of repeatability (CRs) were calculated.

RESULTS. Mean (standard deviation [SD]) age of the study subjects was 38 (5) years (range, 30–49 years). Mean (SD) subfoveal choroidal thickness was 332 (90) μm (right eyes) and 332 (91) μm (left eyes). Intraobserver CR was approximately 23 (95% confidence interval [CI], 19–26) μm, whereas interobserver and intrasession CRs were greater at 32 (95% CI, 30–34) and 34 (95% CI, 32–36) μm, respectively. There was no significant difference in SFCT between all pairs of SFCT measurements except for the two inrasession vertical line scans.

CONCLUSION. A change of >32 μm was likely to exceed interobserver variability in SFCT. Future studies are required to estimate the repeatability of SFCT measurements in patients with chorioretinal pathology. (Invest Ophthalmol Vis Sci. 2011;52:2267–2271) DOI:10.1167/iovs.10-6024

Optical coherence tomography (OCT) is a noninvasive, noncontact imaging modality used to acquire high-resolution, cross-sectional scans of the retina. It has become a valuable tool for the diagnosis and management of chorioretinal disease. In the past, it has been difficult to image deeper structures such as the choroid with this method because of light scatter and decreased resolution and sensitivity with increasing depth. However, new OCT-based techniques for in vivo choroidal imaging have recently been described.2–5 Spaide3 has described the technique of enhanced depth imaging (EDI) OCT. The method involves placing the objective lens of the spectral-domain OCT device (Spectralis; Heidelberg Engineering, Heidelberg, Germany) closer to the eye such that an inverted image is obtained. By performing this maneuver, the deeper structures are placed closer to zero delay, allowing better visualization of the choroid. Its eye-tracking and image-averaging technology, high-speed scanning, reduced noise, and greater coverage of the macular area produces high-resolution, optical, cross-sectional images of the choroid and allows enhanced visualization of pathology. In addition, the thickness of the choroid can be measured at various locations within the macular region.

The ability to reliably image the choroid makes this an emerging area of study, particularly in ocular diseases that are known to involve the choroid or where the choroid is implicated in the pathophysiology of disease. As with any new method of clinical measurement, it is important to validate the repeatability of measurements with the device to understand how precise the tool is and, therefore, be able to distinguish true clinical change from measurement variability or error.

In contrast to retinal thickness measurements on OCT imaging, repeatability of choroidal thickness measurement using EDI OCT in healthy subjects has not been previously studied in detail. Only one study has examined the repeatability of choroidal thickness measurements (in 17 subjects) by reporting the correlation coefficient of measurements from two independent observers. However, Bland and Altman suggested that this method can be misleading and recommended the use of within-subject SD as an index of measurement error.

Therefore, the aim of this study was to investigate the repeatability of EDI OCT choroidal thickness measurements within and between observers and OCT scans in healthy subjects using the coefficient of repeatability as recommended by Bland and Altman. The normative data and variations in choroidal thickness across the macular region have been reported in a separate paper.

METHODS

The scans were obtained during a prospective study of EDI OCT choroidal imaging in healthy young adult volunteers with no known eye disease, the results of which are reported in a separate paper (Yeoh J, et al., unpublished data, 2010). The study was approved by the Institutional Review Board (Research Governance Committee, Moorfields Eye Hospital) and was conducted in adherence with the tenets of the Declaration of Helsinki. Fifty healthy subjects gave informed consent to participate in the study and had OCT scans between January and March 2010.

The OCT scans were performed by two retinal specialists (JY [OB1] and WR [OB2]), experienced at performing scans using the spectral-domain OCT device (Spectralis; Heidelberg Engineering). Two high-

From the Moorfields Eye Hospital, London, United Kingdom. Submitted for publication June 9, 2010; revised August 16 and October 3, 2010; accepted October 3, 2010.

Disclosure: W. Rahman, None; F.K. Chen, None; J. Yeoh, None; P. Patel, None; A. Tufail, None; L. Da Cruz, None

Corresponding author: Wabeeda Rahman, Moorfields Eye Hospital, London, United Kingdom; wabeeda.rahman@moorfields.nhs.uk.
Intrasession (Scan 2/ H11002 OB2V2). For estimation of the intrasession repeatability, SFCT measurements for the right and left eyes between the two observers were made. Interobserver repeatability was calculated using SFCT Observer 1 was masked to the previous choroidal thickness measurement, one OCT reader (OB1) also repeated the two horizontal SFCT measurements for the right eye (OB1H1 vs. OB1H2; OB1V1 vs. OB1V2). Intraobserver (OB1) RE OB2H1 and two vertical line scans obtained per eye. The SFCT was measured from the outer part of the hyperreflective line corresponding to the base of the RPE to the hyporeflective line or margin corresponding to the sclerchoroidal interface (Fig. 1). The measurements were performed with the readers masked to each other’s readings. In addition, one OCT reader (OB1) also repeated the two horizontal SFCT measurements for the right eye (OB1H1 and OB1H2), masked to the previous measurements made.

Three types of comparisons were made to calculate the intraobserver, interobserver, and intrasession repeatability, respectively. For estimation of the intrasession repeatability, two SFCT measurements of the 50 × 2 horizontal line scans of the right eye made by one observer (OB1) were used (OB1H1 vs. OB1H2; OB1H2 vs. OB1H1). Observer 1 was masked to the previous choroidal thickness measurements made. Interobserver repeatability was calculated using SFCT measurements for the right and left eyes between the two observers (OB1H1 vs. OB2H1; OB1H2 vs. OB2H2; OB1V1 vs. OB2V1; OB1V2 vs. OB2V2). For estimation of the intrasession repeatability, SFCT measurements of the two horizontal and vertical line scans (H1 vs. H2; V1 vs. V2) made by OB1 and OB2 were used.

Mean (SD) of the differences was calculated. Measurements were compared using the paired-samples t-test, and their relationships were assessed using the Pearson’s correlation coefficients.

In line with recommendation from Bland and Altman, the difference in SFCT was calculated and plotted against the mean value for all pairs of readings from the two groups being compared. The within-subject SD (Sw) derived from the mean square of the differences was used to calculate the coefficient of repeatability (CR) defined by Bland and Altman as \(1.96 \times \sqrt{(2S_w^2)}\) or 2.77 \(S_w\). In practical terms, the difference between two measurements for the same patient is expected to be less than the CR for 95% of pairs of observations. In addition, the 95% confidence intervals (CIs) for the estimated CR were calculated. For sample sizes of 100, 200, or 400 scans, with each scan subjected to two measurements, the 95% CIs are approximately \(\text{CR} \pm \text{CR} \times 0.139, \text{CR} \times 0.098, \text{or CR} \times 0.069\), respectively.

A two-tailed paired sample t-test was used to determine whether the measurements compared were significantly different. \(P < 0.05\) was considered statistically significant. Statistical analysis was performed using data analysis software (SPSS 12.0.1; SPSS Inc., Chicago, IL).

**RESULTS**

Data from the scans of both eyes of 50 healthy subjects were analyzed. The mean (SD) age of the study subjects was 38 (5) years (range, 30–49 years). There was an even number of men and women in the study. There were 22 Caucasian, 16 Asian, 8 Oriental, and 4 Afro-Caribbean subjects in the study. The mean (SD) subfoveal choroidal thickness was 332 (90) \(\mu m\) (range, 172–550 \(\mu m\)) and 352 (91) \(\mu m\) (range, 142–563 \(\mu m\)) for the right and left eyes, respectively. The mean (SD) axial length of the subjects’ eyes was 24.46 (1.12) mm (range, 22.24–26.96 mm) in the right eye and 24.45 (1.50) mm (range, 22.09–26.89 mm) in the left eye.

Intraobserver, interobserver, and intrasession differences and correlations among the pairs of measurements are summarized in Table 1. All these Pearson’s correlations were highly statistically significant \((P < 0.001)\). The coefficients of repeatability and the limits of agreement are shown in Table 2.

The intraobserver CR derived from H1 and H2 lines scans \((50 \times 2 = 100\) pairs of measurements) was approximately 23

![Figure 1](http://tvst.arvojournals.org/ on 10/14/2017)

**TABLE 1. Summary of Comparisons among Measurements, Observers, and Line Scans**

<table>
<thead>
<tr>
<th>Type of Comparison</th>
<th>Pairing of Scans</th>
<th>Mean (SD) Difference ((\mu m))</th>
<th>(P) (t-test)</th>
<th>Pearson’s Correlation Coefficient*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intraobserver (OB1’ − OB1)</td>
<td>OB1’H2 − OB1H2</td>
<td>+1.8 (11.9)</td>
<td>0.28</td>
<td>0.922</td>
</tr>
<tr>
<td></td>
<td>OB1’V2 − OB1V2</td>
<td>+0.8 (15.0)</td>
<td>0.71</td>
<td>0.988</td>
</tr>
<tr>
<td>Interobserver (OB2 − OB1)</td>
<td>OB2H1 − OB1H1</td>
<td>+2.6 (16.1)</td>
<td>0.26</td>
<td>0.984</td>
</tr>
<tr>
<td></td>
<td>OB2H2 − OB1H2</td>
<td>+4.1 (17.1)</td>
<td>0.10</td>
<td>0.983</td>
</tr>
<tr>
<td></td>
<td>OB2V1 − OB1V1</td>
<td>+0.8 (15.0)</td>
<td>0.71</td>
<td>0.988</td>
</tr>
<tr>
<td></td>
<td>OB2V2 − OB1V2</td>
<td>+2.4 (18.0)</td>
<td>0.35</td>
<td>0.984</td>
</tr>
<tr>
<td></td>
<td>LE OB2H1 − OB1H1</td>
<td>+1.1 (14.5)</td>
<td>0.59</td>
<td>0.988</td>
</tr>
<tr>
<td></td>
<td>LE OB2H2 − OB1H2</td>
<td>+2.7 (15.0)</td>
<td>0.20</td>
<td>0.987</td>
</tr>
<tr>
<td></td>
<td>LE OB2V1 − OB1V1</td>
<td>+1.5 (17.0)</td>
<td>0.54</td>
<td>0.982</td>
</tr>
<tr>
<td></td>
<td>LE OB2V2 − OB1V2</td>
<td>+3.8 (17.7)</td>
<td>0.13</td>
<td>0.981</td>
</tr>
<tr>
<td>Intrasession (Scan 2 − Scan 1)</td>
<td>OB1H2 − OB1H1</td>
<td>+2.7 (18.7)</td>
<td>0.31</td>
<td>0.979</td>
</tr>
<tr>
<td></td>
<td>OB1V2 − OB1V1</td>
<td>+7.6 (16.4)</td>
<td>0.00</td>
<td>0.987</td>
</tr>
<tr>
<td></td>
<td>LE OB1H2 − OB1H1</td>
<td>+1.4 (20.7)</td>
<td>0.64</td>
<td>0.975</td>
</tr>
<tr>
<td></td>
<td>LE OB1V2 − OB1V1</td>
<td>+2.8 (15.2)</td>
<td>0.20</td>
<td>0.986</td>
</tr>
<tr>
<td></td>
<td>RE OB2H2 − OB2H1</td>
<td>+4.2 (18.2)</td>
<td>0.11</td>
<td>0.981</td>
</tr>
<tr>
<td></td>
<td>RE OB2V2 − OB2V1</td>
<td>+4.4 (14.5)</td>
<td>0.04</td>
<td>0.990</td>
</tr>
<tr>
<td></td>
<td>LE OB2H2 − OB2H1</td>
<td>+0.2 (17.8)</td>
<td>0.92</td>
<td>0.982</td>
</tr>
<tr>
<td></td>
<td>LE OB2V2 − OB2V1</td>
<td>+0.5 (13.4)</td>
<td>0.80</td>
<td>0.989</td>
</tr>
</tbody>
</table>

LE, left eye; H, horizontal; OB, observer; RE, right eye; V, vertical.

* All \(P < 0.001\).


Intraobserver (OB1’ – OB1)

<table>
<thead>
<tr>
<th>Grouping of Scans (total n)</th>
<th>Mean Difference (μm)</th>
<th>95% Limit of Agreement (μm)</th>
<th>Coefficient of Repeatability (μm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>H1 (50)</td>
<td>+0.4*</td>
<td>−21.4 to +22.2</td>
<td>21.6</td>
</tr>
<tr>
<td>H2 (50)</td>
<td>−1.8*</td>
<td>−25.1 to +21.4</td>
<td>25.3</td>
</tr>
<tr>
<td>Interobserver (OB2 – OB1)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Horizontal (200)</td>
<td>+0.7*</td>
<td>−30.6 to +32.0</td>
<td>30.9</td>
</tr>
<tr>
<td>Vertical (200)</td>
<td>−1.7*</td>
<td>−34.8 to +31.4</td>
<td>32.6</td>
</tr>
<tr>
<td>Right eyes (200)</td>
<td>+1.3†</td>
<td>−31.3 to +33.9</td>
<td>32.6</td>
</tr>
<tr>
<td>Left eyes (200)</td>
<td>−2.5*</td>
<td>−33.6 to +29.0</td>
<td>31.5</td>
</tr>
<tr>
<td>Intrasession (Scan 2 – Scan 1)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Horizontal (200)</td>
<td>+2.0*</td>
<td>−34.9 to +38.9</td>
<td>37.0</td>
</tr>
<tr>
<td>Vertical (200)</td>
<td>+3.8t</td>
<td>−25.7 to +33.3</td>
<td>30.5</td>
</tr>
<tr>
<td>OB1 (200)</td>
<td>+3.6†</td>
<td>−31.5 to +38.7</td>
<td>35.7</td>
</tr>
<tr>
<td>OB2 (200)</td>
<td>+2.2†</td>
<td>−29.4 to +33.8</td>
<td>31.8</td>
</tr>
</tbody>
</table>

* P > 0.05; † P ≤ 0.05; ‡ P < 0.001; § P < 0.005.

(95% CI, 19–26 μm). A Bland-Altman plot of difference against mean SFCT showed no significant change in variability for the range of SFCT seen in healthy subjects (200–500 μm; Fig. 2).

The interobserver CR was approximately 32 μm (95% CI, 30–34 μm) for the entire set of H1, H2, V1, and V2 line scans for both eyes (50 × 4 × 2 = 400 pairs of measurements). The mean (SD) difference in SFCT measurements between the two observers was −0.5 (16.4) μm, paired t-test (P = 0.53). The interobserver CR for horizontal line scans, vertical line scans, right eye scans, and left eye scans are shown, separately, in Table 2. A Bland-Altman plot of difference against mean SFCT from the two observers’ measurements of all horizontal line scans is shown in Figure 3. A similar scatter distribution is seen for vertical scans.

The intraobserver CR was approximately 34 μm (95% CI, 32–36 μm) for the entire set of horizontal and vertical line scans for both observers and both eyes (50 × 2 × 2 × 2 = 400 pairs of measurements). The SFCT from the second scan was significantly greater than the SFCT from the first scan; the mean (SD) difference was 3 (18) μm, paired t-test (P = 0.001). However, this was not the case when only horizontal scans (H2-H1) were compared. Then the mean (SD) difference was 2 (19), paired t-test (P = 0.27; Table 2). A Bland-Altman plot of difference against mean SFCT from the two scans for all vertical line scans is shown in Figure 4. There was a bias toward greater SFCT from the second vertical scan compared with the first.

**DISCUSSION**

Our study showed that the mean SFCT in healthy subjects was 332 μm (range, 142–563 μm). In one histologic study, choroidal thickness in healthy maculae ranged from 84 to 193.5 μm, but, because histologic analysis is an invasive technique with shrinkage of tissue on fixation, it is not directly comparable to our OCT study. Margolis and Spaide used the same EDI OCT technique we used in our study and reported a mean (SD) SFCT of 287 (76) μm in healthy subjects. However, the mean age of their participants was 50.4 years compared with a mean age of 38 years in our study. Given that age is thought to be negatively correlated with SFCT, this may account for the difference in mean SFCT measurements between the two studies.

This study also provided an estimate of the repeatability of manual SFCT measurements using the spectral-domain OCT device (Spectralis; Heidelberg Engineering) in healthy subjects. These results showed that the intraobserver CR was approximately 22 to 23 μm, whereas interobserver and intrasession CRs were greater at approximately 32 to 37 μm. There was no significant difference in SFCT between all pairs of measurements except for the two intrasession vertical line scans.

Menke et al. studied the reproducibility of retinal thickness measurements in healthy persons using the spectral-domain OCT device (Spectralis; Heidelberg Engineering). They showed excellent intrasession reproducibility with almost no difference in retinal thickness measurements for all nine Early Treatment Diabetic Retinopathy Study areas. The mean difference between measurements was approximately 1 μm. The low variability was thought to be due to the eye-tracking features during scanning, use of the follow-up scanning protocol, improved resolution, imaging speed, scan coverage, and retinal segmentation algorithms. Direct comparison between the Menke et al. study and our study cannot be made for the

![Figure 2. Bland-Altman plot of difference against mean SFCT showed no significant change in intraobserver variability for the range of SFCT measurements seen in healthy subjects. Dashed lines: mean difference and 95% limits of agreement.](http://tvst.arvojournals.org/)

![Figure 3. Bland-Altman plot of difference against mean SFCT showed no significant change in interobserver variability for the range of SFCT measurements for all horizontal line scans. Dashed lines: mean difference and 95% limits of agreement.](http://tvst.arvojournals.org/)
that variability in SFCT measurements can arise from factors aware of any other studies that have reported on the repeat-
cause of the different statistical methodology used. We are not measured the same scan. Indeed, we found a CR of 23
/H11005
measurements performed by the two observers (subjects, there was a highly significant correlation between the
ment variability may be minimized if there is reliable segmen-
tion does not yet exist so our measurements were taken manually. Repeatability studies for retinal thickness have usually
adaptation software to a variable extent to improve image
the calipers can also vary depending on the magnification of
the suprachoroidal space, but often this line can be indistinct,
be related to scan quality or anatomic variation. Some choroi-
dal scans have a distinct hyporeflective line corresponding to
the sclerochoroidal interface on acquired images, which may
be used to place the calipers precisely at the fovea and
measurement to the factors described previously, the increased vari-
ability in SFCT measurements of the two horizontal or the
two vertical scans may be related to misplacement of the
line of the OCT slice during the second scan. Measurement
error because of scan misplacement might have been mini-
mized if the first scan was set as a reference image and the
follow-up protocol was used to acquire the second scan. The
intrasession CR might have been lower if this method had been used.

These CR estimates are useful for planning future clinical trials. Such reliable choroidal imaging techniques may allow better understanding and facilitate more in-depth studies of the pathology with underlying origins in the choroid or with asso-
ated choroidal changes. For example, they may be of benefit in longitudinal clinical studies investigating the pathogenesis of conditions such as neovascular age-related macular degenera-
tion, central serous chorioretinopathy, and idiopathic polypoi-
dal choroidal vasculopathy.19–21 Future OCT technologies may incorporate choroidal segmentation software to provide more reliable, automated data on choroidal thickness measurements at the macula. However, as seen in neurosensory retina bound-
ary segmentation, OCT readers need to be aware of the poten-
tial differences between devices in the software algorithm used to detect choriodoscleral and retinchoroidal boundaries, and the ability of the software to detect these boundaries in OCT images degraded by optical or motion artifacts and pathol-
ogy.22,23 Nevertheless, the results of this study help to improve our understanding of the repeatability of manual measure-
ments of SFCT and could help to validate the accuracy and precision of future automated segmentation algorithms.

The strengths of the study include its prospective design, which randomly recruited a large number of subjects with a wide range of choroidal thicknesses. Scans for the subjects were acquired consecutively (intrasession), which minimized the possibility of true change in SFCT caused by fluctuation in choroidal thickness. Such intersession variations in choroidal thickness and blood flow have been reported and may relate to diurnal variation, fluctuations in intraocular pressure, perfusion pressure, and circulating catecholamines.44–46 A limitation of this study was that scans were acquired from healthy subjects with no ocular pathology and did not reflect those seen in patients in a routine outpatient setting. For instance, the clarity of images in patients could be hindered by ocular pathology such as significant media opacity or masking of choroidal reflectance by epiretinal, intraretinal, subretinal, and choroidal hyperreflective lesions. In addition, some patients have poor fixation because of disease, leading to problems acquiring good-quality scans, making the accurate measurement of SFCT sometimes difficult. However, given the lack of intraobserver or intrasession repeatability studies for EDI OCT and the limited data on interobserver reproducibility, we chose to study healthy subjects. Additional studies are warranted to test the repeatability of choroidal thickness measurements with the spectral-domain OCT device (Spectralis; Heidelberg Engineer-
ing) in patients with chorioretinal pathology and in separate scanning sessions using the follow-up protocol.

In conclusion, our results indicate that a change of 23 μm is likely to represent true change when the same observer meas-
sures SFCT in healthy subjects. However, manual measure-
ments of SFCT can vary as much as 37 μm when EDI OCT scans are repeated within the same session without use of the
follow-up protocol. Further studies are required to determine whether these thresholds of SFCT change are also clinically significant.
References