Results and Repeatability of Retinal Thickness Measurements From Certification Submissions

Ronald P. Danis, MD; Marian R. Fisher, PhD; Ericka Lambert, BA; Anne Goulding, BA; Deborah Wu, MD; Li-Yin Lee, MS

Objectives: To present the results for subgroups defined by center point (CP) measurement and to assess the repeatability of the Fast Retinal Thickness Map analysis results from the Stratus OCT3 machine.

Methods: Two hundred eighty-one replicate OCT3 scans from 134 operators’ certification submissions to a reading center were analyzed, including scans from eyes that were reported to be normal and eyes with exudative age-related macular degeneration and with macular edema due to diabetic retinopathy or retinal vascular occlusion.

Results: The mean (SD) of the CP was 284 (150) µm and the center subfield (CC) was 301 (130) µm. The CP CR was 49 µm and the CC CR was 27 µm. The CR increased by increasing retinal thickness for the CP and the CC within arbitrarily defined subgroups. For the 87 eyes with a session 1 CP of 175 µm or less, the CP CR was 17 µm and the CC CR was 10 µm.

Conclusions: Among experienced operators, given the same operator, machine, and eye at the same sitting, OCT3 retinal thickness maps appear to have a CR that is likely to be less than the clinically important difference.

Arch Ophthalmol. 2008;126(1):45-50

Optical coherence tomography (OCT) was introduced commercially as a new retinal imaging modality in 1995. It has quickly become a standard tool for clinical diagnostic and research purposes for glaucoma and macular diseases. The 10-µm axial resolution of third-generation OCT allows objective measurement of retinal thickness, assessment of retinal morphology, and detection of subretinal fluid. Clinicians and researchers have come to rely on the numeric measurements generated by the manufacturer’s scanning software for detection and longitudinal assessment of macular thickening and for clinical research study eligibility and outcomes. Optical coherence tomography is also used for the evaluation of choroidal neovascularization due to age-related macular degeneration (AMD) and other causes. The study of retinal morphology and topography using OCT in exudative AMD is more commonly adjunctive to other assessments, such as fluorescein angiography. Laser light attenuation from OCT scanning deep to the retina renders choroidal features less demonstrable than retinal morphology.

The majority of the current literature reporting OCT measurements in normal and diseased maculas describes first- and second-generation machines. Measurement repeatability of the third-generation device is expected to be improved slightly over earlier versions because of greater rapidity of scan acquisition (which avoids movement artifact) and software enhancements.

We assessed measurement repeatability in the setting of a large series of operator certification submissions to a centralized reading center from multiple clinical sites. These certification submissions consisted of consecutive scans of the same patient at the same visit taken by the same technician using the same equipment. Operators were applying for certification to participate in clinical trials of exudative AMD or macular edema (ME) due to diabetic retinopathy or retinal vascular occlusion.

Methods

The clinical site equipment was the commercially available third-generation Stratus OCT3 (Carl Zeiss Meditec, Inc, Dublin, California). The scan protocol included the Fast Macular Thickness Map algorithm, which features 6 scan...
Table 1. Summary of Session 1 OCT Macular Thickness Measurements in Clinical Trial Certification Submissions

<table>
<thead>
<tr>
<th>CP at Session 1, µm</th>
<th>Designated as Diseased, No. (%)</th>
<th>Mean (SD)</th>
<th>Center Subfield, µm</th>
</tr>
</thead>
<tbody>
<tr>
<td>≤175</td>
<td>87 (5)</td>
<td>154 (13)</td>
<td>193 (18)</td>
</tr>
<tr>
<td>176-225</td>
<td>56 (14)</td>
<td>196 (13)</td>
<td>227 (16)</td>
</tr>
<tr>
<td>226-325</td>
<td>45 (82)</td>
<td>272 (34)</td>
<td>283 (31)</td>
</tr>
<tr>
<td>326-425</td>
<td>48 (100)</td>
<td>379 (31)</td>
<td>381 (38)</td>
</tr>
<tr>
<td>≥426</td>
<td>45 (100)</td>
<td>554 (107)</td>
<td>536 (101)</td>
</tr>
<tr>
<td>Total</td>
<td>281 (51)</td>
<td>284 (150)</td>
<td>301 (130)</td>
</tr>
</tbody>
</table>

Abbreviations: CP, center point; OCT, optical coherence tomography.

RESULTS

From May 2003 through November 2004, there were 314 OCT3 submissions for operator certification. Of these, 281 scans (139 normal eyes and 142 diseased eyes [66 AMD, 76 ME]) from 134 operators met the aforementioned accuracy and completeness criteria.

Subgroups defined by the center point (CP) at session 1 were created arbitrarily to facilitate data presentation and evaluation (Table 1). All of the scans with a CP of 325 µm or more were from diseased eyes; 93% of the scans with a CP of 175 µm or less were from normal eyes.

For the 139 normal scans, the mean (SD) CP was 173 (8) µm; for the 66 AMD scans, 383 (18) µm; and for the 76 ME scans, 399 (22) µm. Instructions for the Stratus
OCT3 retinal thickness tabular output note the fovea minimum parameter with a normal range of 135 to 215 µm; 11 of the normal eyes exceeded this range and 11 of the diseased eyes (6 AMD, 5 ME) were within this range.

For each of the 281 eyes, the session 1 CP vs session 2 CP was plotted (Figure 1A). As summarized in Table 2, the CP coefficient of repeatability was 49 µm and the mean difference between session 1 and session 2 CP was 2 µm with a 95% confidence interval of −1 to 5 µm. There was no difference in the CP coefficient of repeatability between eyes with AMD and ME (data not shown). The CP coefficient of repeatability increased as the mean CP of the subgroup defined by the CP at session 1 increased. The average of session 1 CP and session 2 CP was plotted against the difference of session 1 CP minus session 2 (Figure 1B), and the CP coefficient of reliability for all 281 eyes and by subgroup are shown. The CP coefficient of reliability (17 µm) for the subgroup with a CP of 175 µm or less was different from the subgroup with a CP of 176 to 225 µm (18 µm) and that was different from the subgroup with a CP of 226 to 325 (57 µm) (P < .001).

For each of the 281 eyes, the session 1 CP vs the session 2 CP was plotted with a line segment whose length is 2 SDs and the midpoint is the CP (Figure 2). As extreme examples, in Figure 2, there is an eye with a session 1 CP of 441 µm (SD = 10 µm) and a session 2 CP of 286 µm (SD = 115 µm) and there is another eye with a session 1 CP of 571 µm (SD = 20 µm) and a session 2 CP of 428 µm (SD = 167 µm). These 2 eyes and the 43 other eyes whose CP at session 1 was 426 µm or more had a CP coefficient of reliability of 78 µm and a mean difference of 2 µm with a 95% confidence interval of −1 to 5 µm. There was no difference in the CP coefficient of repeatability between eyes with AMD and ME (data not shown). The center subfield coefficient of reliability was 27 µm and the center subfield mean difference between session 1 and session 2 was 2 µm with a 95% confidence interval of −1 to 2 µm (Table 3). There was no difference in the center subfield coefficient of repeatability between eyes with AMD and ME (data not shown). The center subfield coefficient of repeatability for the subgroup with a CP of 175 µm or less (10 µm) was different from the subgroup with a CP of 176 to 225 µm (18 µm) (P < .001). For the subgroup with a CP of 175 µm or less, the subset of 62 eyes with a standard deviation of 5% or less of the CP did not differ in the center subfield repeatability coefficient from the subset of 25 eyes with a standard deviation of more than 5% of the CP (center subfield coefficients of reliability of 8 µm and 13 µm).

For each of the 281 eyes, the session 1 CP and center subfield were plotted (Figure 3). These measurements were not independent. The estimate of how much larger the center subfield was than the CP was 57 µm for all 281 eyes and 75 µm for the subset with a CP of 325 µm.

### Table 2. Session 1 CP Minus Session 2 CP: OCT Macular Thickness Measurements in Clinical Trial Certification Submissions

<table>
<thead>
<tr>
<th>CP at Session 1, µm</th>
<th>Total</th>
<th>SD ≤ 5% of CP</th>
<th>SD &gt; 5% of CP</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Sample Size</td>
<td>Descriptors</td>
<td>Sample Size</td>
</tr>
<tr>
<td>≤ 175</td>
<td>87</td>
<td>17 (-3) [-5 to 1]</td>
<td>62</td>
</tr>
<tr>
<td>176-225</td>
<td>56</td>
<td>28 (2) [-2 to 6]</td>
<td>33</td>
</tr>
<tr>
<td>226-325</td>
<td>45</td>
<td>57 (9) [-7 to 10]</td>
<td>18</td>
</tr>
<tr>
<td>326-425</td>
<td>48</td>
<td>61 (1) [-8 to 10]</td>
<td>28</td>
</tr>
<tr>
<td>≥ 426</td>
<td>45</td>
<td>78 (12) [0 to 24]</td>
<td>36</td>
</tr>
<tr>
<td>Total</td>
<td>281</td>
<td>49 (2) [-1 to 5]</td>
<td>177</td>
</tr>
</tbody>
</table>

Abbreviations: CP, center point; OCT, optical coherence tomography; SD, standard deviation.

*Coefficient of repeatability (mean difference of session 1 CP minus session 2 CP) [95% confidence interval for mean difference].
or less. The solid line is the CP plus a constant (arbitrarily chosen as 25 µm), included as a common frame of reference in Figure 3, and indicates that as the CP increases, the estimated line for the center subfield differs (largest visual discrepancy in the upper right portion of Figure 3).

For each of the retinal thickness subgroups, the macular map sectoral thickness diagram summarizes the session 1 mean (SD) CP, followed by the coefficient of repeatability and the 95% confidence interval on mean difference (session 1 CP minus session 2 CP) (Figures 4, 5, 6, 7, and 8).

**Table 3. Session 1 CP Minus Session 2 Center Subfield: OCT Macular Thickness Measurements in Clinical Trial Certification Submissions**

<table>
<thead>
<tr>
<th>CP at Session 1, µm</th>
<th>Total</th>
<th>SD &lt; 5% of CP</th>
<th>SD &gt; 5% of CP</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Sample Size</td>
<td>Descriptors</td>
<td>Sample Size</td>
</tr>
<tr>
<td>≤ 175</td>
<td>87</td>
<td>10 (−1) [−2 to 0]</td>
<td>62</td>
</tr>
<tr>
<td>176-225</td>
<td>56</td>
<td>18 (2) [−1 to 4]</td>
<td>33</td>
</tr>
<tr>
<td>226-325</td>
<td>45</td>
<td>28 (1) [−3 to 5]</td>
<td>18</td>
</tr>
<tr>
<td>326-425</td>
<td>48</td>
<td>40 (0) [−6 to 5]</td>
<td>28</td>
</tr>
<tr>
<td>≥ 426</td>
<td>45</td>
<td>38 (2) [−4 to 7]</td>
<td>36</td>
</tr>
<tr>
<td>Total</td>
<td>281</td>
<td>27 (2) [−1 to 2]</td>
<td>177</td>
</tr>
</tbody>
</table>

Abbreviations: See Table 2.

*a Coefficient of repeatability (mean difference of session 1 CP minus session 2 CP) [95% confidence interval for mean difference].

Figure 3. Center point (CP) data vs center subfield (CC) data for 281 Stratus OCT3 (Carl Zeiss Meditec, Inc, Dublin, California) scans. The solid line is CC=CP+25 and the red dashed line is CC=0.86CP+57.

1990s, in recent years OCT has been in a more dynamic environment of evolving equipment and software. Published reports of OCT3 reproducibility are limited by small samples and variable methods. In 1 study, 10 normal subjects had 1 eye imaged 6 times per day on 3 different occasions. High reproducibility with the Fast Macular Thickness Map was reported, with an intraclass correlation of 88% and intervisit/intravisit standard deviations of less than 4 µm at the center subfield. In another study using OCT3, 10 normal subjects had 1 eye imaged twice by different operators on 3 separate occasions. Analysis of variance demonstrated no difference in macular measurements. In 22 eyes with diabetic ME, repeated scanning with OCT3 yielded a reproducibility coefficient of 37 µm or less for the center subfield, which was considered good. Studies of reproducibility of macular scanning with earlier-generation OCT

**COMMENT**

The numeric measurements generated by OCT scanning software are used by clinicians and researchers for detection and longitudinal assessment of macular thickening and for clinical research study eligibility and outcomes. However, the peer-reviewed literature provides relatively few data on the repeatability of such measurements. In contrast to fundus photography in the 1980s and 1990s, in recent years OCT has been in a more dynamic environment of evolving equipment and software. Published reports of OCT3 reproducibility are limited by small samples and variable methods. In 1 study, 10 normal subjects had 1 eye imaged 6 times per day on 3 different occasions. High reproducibility with the Fast Macular Thickness Map was reported, with an intraclass correlation of 88% and intervisit/intravisit standard deviations of less than 4 µm at the center subfield. In another study using OCT3, 10 normal subjects had 1 eye imaged twice by different operators on 3 separate occasions. Analysis of variance demonstrated no difference in macular measurements. In 22 eyes with diabetic ME, repeated scanning with OCT3 yielded a reproducibility coefficient of 37 µm or less for the center subfield, which was considered good. Studies of reproducibility of macular scanning with earlier-generation OCT

Figure 4. Macular map of sectoral thickness results for the subgroup (n=87) with a center point (CP) of 175 µm or less. Subgroups based on session 1 CP measurement. Order from left to right is outer nasal, inner nasal, center subfield (CC), inner temporal, and outer temporal sectors and top to bottom is outer superior, inner superior, inner inferior, and outer inferior sectors. Top values are sector mean (SD). Bottom values are sector coefficient of repeatability and 95% confidence interval on mean difference of session 1–session 2.
machines also are limited by small samples and variable methods, but with generally similar results.7-9 Interoperator variability has been noted to be low in 2 studies of 109 and 25 patients,10 whereas 1 study of 20 eyes suggested systematic difference based on macular thickness.11 In a study using a single OCT3 machine and a single operator, Polito et al12 report 9 µm as the coefficient of repeatability for the CP in healthy retinas (n=10; mean CP, 180 µm) and 20 µm, in eyes with clinical diabetic ME (n=15; mean CP, 394 µm). Given that these results are from not 1 but 134 different operators and not 1 but multiple clinical sites and that the patient groups were not directly comparable, it is not surprising that these coefficients of repeatability are larger (ie, 17 µm in the 87 eyes with a CP of 175 µm or less and 49 µm in all 281 eyes).

The results of the current study might be considered an ideal scenario to document repeatability of the OCT3 in a multicenter context; we examined replicate imaging studies, including diverse disease representation, a large number of trained and experienced OCT operators, and a situation whereby operators would be expected to submit only their highest-quality efforts (since these submissions would be used to determine certification for performance of OCT in clinical trials).

Figure 5. Macular map of sectoral thickness results for the subgroup (n=56) with a center point (CP) of 176 to 225 µm. See Figure 4 for description of sectors and values.

Figure 6. Macular map of sectoral thickness results for the subgroup (n=45) with a center point (CP) of 226 to 325 µm. See Figure 4 for description of sectors and values.

Figure 7. Macular map of sectoral thickness results for the subgroup (n=48) with a center point (CP) of 326 to 425 µm. See Figure 4 for description of sectors and values.

Figure 8. Macular map of sectoral thickness results for the subgroup (n=45) with a center point (CP) of 426 µm or more. See Figure 4 for description of sectors and values.
The critical issue is specification of the clinically acceptable difference and whether it is expected to increase as the CP and/or center subfield increases. The coefficient of repeatability, based on this subset of submissions, provides an estimate of how different 2 measurements must be before one can conclude that, with 95% probability, a change has occurred by natural history or by some intervention.

Using arbitrarily defined subgroups based on session 1 CP, the CP coefficient of repeatability was at most 10% of the CP for both the subgroup with a CP of 175 µm or less (coefficient of repeatability, 17 µm) and the subgroup with a CP of 176 to 225 µm (28 µm) and then increased to approximately 20% of the CP for other subgroups (coefficient of repeatability, 57 µm, 61 µm, and 78 µm). For eyes in the subgroup with a CP of 175 µm or less that also had a standard deviation of 5% or less of the CP, the CP coefficient of repeatability was 12 µm.

Using the arbitrarily defined subgroups based on session 1 CP, the center subfield coefficient of repeatability was 10 µm for the subgroup with a CP of 175 µm or less and increased to 18 µm for the subgroup with a CP of 176 to 225 µm, 28 µm for the subgroup with a CP of 226 to 325 µm, 40 µm for the subgroup with a CP of 326 to 425 µm, and 38 µm for the subgroup with a CP of 426 µm or more. These coefficient of repeatability estimates seem plausible given the report of 22 eyes with diabetic ME that had a reproducibility coefficient of 37 µm or less for the center subfield, which was considered good.

The standard deviation of the CP is calculated from the 6 CP measurements from each of the underlying scans of the retina map. Therefore, it is expected that when the underlying scans differ from one another enough to produce a relatively large standard deviation, this reflects variability throughout the study. Causes of such variability include eye movement due to fixation or motor instability, or variations in boundary line detection, which are more likely to occur with poor signal strength or with very thick retinas with abnormal anatomy.1 Approximately two-thirds of the submissions had a standard deviation of 5% or less of the CP; this was 80% in the subgroup with a CP of 426 µm or more and testifies to the excellent technique of the operators and cooperation of the patients. As an internal metric in these data, a standard deviation of more than 5% of the CP was most useful in the subgroup of eyes with a CP of 175 µm or less; the CP coefficient of repeatability for those 25 eyes was 26 µm compared with 12 µm for the subset of 62 eyes with a standard deviation of 5% or less of the CP.

Based on these analyses in a multicenter context of the same eye at the same visit using the same equipment and the same operator, a single OCT3 scan without artifact obtained by an experienced technician seems to have a coefficient of repeatability that is likely to be less than the clinically important difference.

Submitted for Publication: December 13, 2005; final revision received March 5, 2007; accepted March 18, 2007.

Correspondence: Ronald P. Danis, MD, Department of Ophthalmology and Visual Sciences, University of Wisconsin–Madison, 406 Science Dr, Ste 400, Madison, WI 53711 (rdanis@opth.wisc.edu).

Financial Disclosure: None reported.

Additional Contributions: Tim Hess, MS, provided the R programming for Figures 4, 5, 6, 7, and 8. We appreciate the insightful comments from the reviewer.

REFERENCES