Repeatability of Spectralis OCT Measurements of Macular Thickness and Volume in Diabetic Macular Edema

Oliver Comyn, Ling Zhi Heng, Felicia Ikeji, Kanom Bibi, Philip G. Hykin, James W. Bainbridge, and Praveen J. Patel

Purpose. We estimated coefficients of repeatability for Spectralis optical coherence tomography (OCT)–derived automated retinal thickness and volume measurements in subjects with center-involving diabetic macular edema (DME).

Methods. A total of 50 eyes of 50 consecutive patients with center-involving DME underwent four consecutive "fast" volume scans at a single session using one OCT device operated by one of two experienced operators. Bland-Altman coefficients of repeatability (CR) were calculated for automated retinal thickness measurements in the nine Early Treatment of Diabetic Retinopathy Study (ETDRS) subfields, center point thickness, and total macular volume. Scans were evaluated for significant automated retinal boundary detection error and revised estimates for CR calculated with these scans excluded.

Results. CR in the central subfield was 8.03 μm (95% confidence interval [CI] 7.70–8.35 μm). In other subfields, CR ranged from 6.54 to 18.25 μm. Scan sets from 13 subjects had significant boundary detection error; reanalysis with these excluded yielded a CR for the central subfield of 7.44 μm with CR for all other subfields <8 μm.

Conclusions. Retinal thickness measurements in subjects with DME obtained using Spectralis OCT are considerably less variable than has been reported with other devices. Changes in central subfield thickness >8 μm can be considered more indicative of true clinical change rather than measurement variability. This finding informs clinical practice and clinical trial design. (*Invest Ophthalmol Vis Sci.* 2012;53:7754–7759) DOI:10.1167/iovs.12-10895

Spectral domain optical coherence tomography (SD-OCT) technology now is available widely and has led to an improved understanding of the structural changes that occur in macular diseases, while providing a means of assessing quantitatively the effects of treatments on retinal thickness and volume. Inclusion criteria for clinical trials and retreatment protocols in trials of intravitreal agents for macular disease frequently incorporate a quantitative OCT parameter. Studying repeatability allows true clinical change to be distinguished from naturally occurring measurement variability. Clinicians then can define a threshold to recognize when a condition has changed, which subsequently can be used in clinical trials and clinical practice to determine the need for further treatment or to identify a therapeutic response.

Diabetic macular edema (DME) is the leading cause of vision loss in a working age population.[1,2] Clinical trials of anti-VEGF agents have demonstrated the superiority of bevacizumab and ranibizumab over conventional laser therapy in this condition.[3–6] These trials have used quantitative OCT measures in their retreatment criteria. Repeatability of OCT retinal thickness measurements has been studied in diabetic macular edema using time domain devices, with estimates for the coefficient of repeatability (CR) for retinal thickness in the central subfield of 21[7] and 38 μm.[8] Comparative studies have failed to demonstrate a statistically significant improvement in CR for newer spectral domain devices evaluating retinal thickness in DME.[9,10] The Spectralis SD-OCT device has the capability to track eye movement and can place follow-up scans automatically in the same retinal location, potentially leading to highly repeatable measurements of retinal thickness. However, to the best of our knowledge, there have been no studies evaluating the repeatability of retinal thickness measurements in eyes with DME using the Spectralis OCT.

The aim of our study was to estimate CRs for retinal thickness in the nine Early Treatment of Diabetic Retinopathy Study (ETDRS) subfields and total macular volume for subjects with center-involving DME using the Spectralis SD-OCT device. The results of our study will help in defining optimal retinal thickness retreatment criteria for clinical trials and in distinguishing true change from measurement variability in clinical practice.

Methods

We included in the study 50 eyes of 50 consecutive subjects with center-involving DME undergoing screening for recruitment to clinical trials. If both eyes of a patient were eligible for inclusion, the eye with worse visual acuity was included. The study was performed according to the tenets of the Declaration of Helsinki and all subjects had signed an informed consent form before participating in the study. Subjects were not excluded from OCT scanning if they had media opacity or vitreoretinal interface abnormalities that could affect the automated retinal segmentation algorithm.

OCT Imaging

All measurements were taken by two experienced OCT technicians (FI, KB) certified for obtaining clinical trial images for reading centers using a single Spectralis OCT instrument running HRA2/Spectralis Family Acquisition Module 5.4.7.0 (Heidelberg Engineering GmbH, Heidel-

From the NIHR Moorfields Biomedical Research Centre, Moorfields Eye Hospital NHS Foundation Trust, and UCL Institute of Ophthalmology, London, United Kingdom.

Supported by the National Institute for Health Research (NIHR) Biomedical Research Centre based at Moorfields Eye Hospital NHS Foundation Trust and UCL Institute of Ophthalmology. The authors alone are responsible for the content and writing of this paper.

Submitted for publication September 4, 2012; revised October 22, 2012; accepted October 22, 2012.

Disclosure: **O. Comyn**, None; **L.Z. Heng**, None; **F. Ikeji**, None; **K. Bibi**, None; **P.G. Hykin**, None; **J.W. Bainbridge**, None; **P.J. Patel**, Heidelberg (C)

Corresponding author: Oliver Comyn, NIHR Moorfields Biomedical Research Centre, 2nd Floor RDCEC, Moorfields Eye Hospital NHS Foundation Trust, City Road, London EC1V 2PD, United Kingdom; olivercomyn@doctors.org.uk.

Investigative Ophthalmology & Visual Science, November 2012, Vol. 53, No. 12

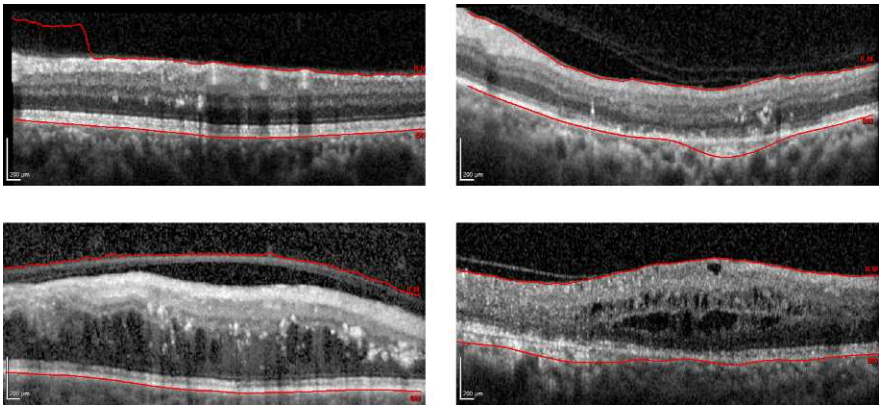

Figure 1. Examples of automated retinal boundary detection error: misplacement of ILM line is shown on the *left*. Misplacement of BM line is shown on the *right*.

berg, Germany). Subjects underwent four consecutive scans in a single session using the "fast macular volume" preset, consisting of a 25-line horizontal raster scan covering 20° × 20°, centered on the fovea. Scans were obtained in high speed mode with the automated real time feature enabled and set to 8 frames. After the first scan, subjects were asked to sit back from the machine, the scan was set as "reference" and three further scans were obtained in "follow-up" mode with the subject sitting back from the machine between each scan acquisition.

Statistical Analysis

Retinal thickness and volume measurements for the nine ETDRS subfields were obtained from Heidelberg Eye Explorer software (version 1.7.0.0; Heidelberg Engineering GmbH) and transferred to an Excel spreadsheet for calculation of coefficients of repeatability using methods described by Bland and Altman.[11] Specifically, in each subfield the variance of the four measurements for every patient was calculated initially to obtain S_w, the within subject standard deviation, from the square root of the average of the variances across all n subjects. CR then was calculated by $1.96 \times \sqrt{}(2\times$ average variance). The width of the 95% confidence interval (CI) for CR then was obtained by $1.96 \times \frac{Sw}{\sqrt{2n(m-1)}}$, where n = number of subjects and m = number of times the test was performed.[12]

Segmentation Error

Scans were evaluated manually for the presence of inner or outer retinal boundary detection error by one experienced observer (OC). The automated Spectralis segmentation algorithm defines the inner retinal boundary at the location of the inner limiting membrane (ILM) and the outer retinal boundary at Bruch's membrane (BM). For the purposes of our study, significant automated boundary detection error was defined by us as the misplacement of either of these boundaries continuously over a section of scanned retina of 1 mm or greater. This was deemed to have occurred when the automated line clearly and unambiguously deviated from the hyper-reflective interface representing Bruch's membrane, or deviated from the clearly visualized inner retinal boundary, for example by following hyper-reflective interfaces anterior to the retina. Examples of this are shown in Figure 1. Subjects were excluded from the subanalysis if boundary detection error was present in any of the four scan sets obtained.

Results

Subject Characteristics

A total of 50 subjects completed the study, of which 18 were female. The mean age was 61.7 years (range 30–82), and 31 subjects were Caucasian, 12 Asian-Indian, 3 Afro-Caribbean, and 4 from other ethnic groups. Mean ETDRS best-corrected visual acuity was 67 letters (Snellen equivalent 20/40^{-3}, ETDRS letter score range of 47–82).

Automated Retinal Thickness and Volume Measurements

The mean thickness and volume values for each macular subfield with repeated measurements are summarized in

Table 1. Mean and Standard Deviation of Retinal Thickness (μm) for Four Repeated Measurements in Nine ETDRS Subfields ($n = 50$)

Measurement	1		2		3		4	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Center	476.7	102.4	476.5	102.0	476.8	102.7	476.6	103.2
Inner								
Superior	456.9	94.4	455.2	90.8	455.3	90.7	455.1	90.9
Temporal	455.7	87.3	454.6	85.6	454.6	86.2	454.9	85.7
Inferior	435.2	70.4	434.5	70.3	434.0	71.6	434.5	71.8
Nasal	441.5	84.0	440.7	82.8	441.1	83.2	440.8	83.6
Outer								
Superior	371.2	83.7	372.2	84.2	371.8	83.9	372.6	85.1
Temporal	368.6	83.8	368.1	81.9	368.2	82.3	368.6	82.4
Inferior	348.8	66.9	349.5	66.4	348.8	67.3	349.4	66.9
Nasal	374.3	61.0	373.7	59.4	372.9	60.1	373.8	59.9
Center point thickness	473.6	125.2	475.0	126.0	475.2	125.8	473.9	126.1

Table 2. Mean and Standard Deviation of Retinal Volume (mm^3) for Four Repeated Measurements in Nine ETDRS Subfields ($n = 50$)

Measurement	1		2		3		4	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Center	0.37	0.08	0.37	0.08	0.37	0.08	0.37	0.08
Inner								
Superior	0.72	0.15	0.71	0.14	0.72	0.14	0.72	0.14
Temporal	0.72	0.14	0.71	0.13	0.71	0.13	0.71	0.13
Inferior	0.68	0.11	0.68	0.11	0.68	0.11	0.68	0.11
Nasal	0.69	0.13	0.69	0.13	0.69	0.13	0.69	0.13
Outer								
Superior	1.93	0.45	1.93	0.45	1.94	0.45	1.94	0.46
Temporal	1.88	0.43	1.88	0.42	1.88	0.42	1.88	0.42
Inferior	1.81	0.36	1.81	0.36	1.81	0.36	1.81	0.36
Nasal	1.92	0.34	1.92	0.33	1.92	0.33	1.92	0.34
Total macular volume	10.73	1.73	10.72	1.69	10.72	1.69	10.73	1.70

Tables 1 and 2. For retinal thickness and volume in each subfield, for center point thickness and for total macular volume, mean values for each patient were plotted against standard deviation to demonstrate no correlation between variability and the magnitude of the measurement. Results for central subfield thickness are shown in Figure 2; plots for other subfields similarly demonstrated no relationship.

The CR for retinal thickness in the central subfield was 8.03 μm (95% CI 7.70–8.35 μm). CR in the other subfields ranged from 6.54 to 18.25 μm (Table 3). CR for macular volume was 0.08 mm^3 or lower for all subfields and 0.21 mm^3 for total macular volume.

Scans with Retinal Boundary Detection Error

Significant automated retinal boundary detection error as described above was present in 13 (26%) subjects. For the remaining 37 subjects without significant segmentation error, the CR was now 7.44 μm (95% CI 7.09–7.79 μm) in the central subfield. Estimates for CR in the remaining eight subfields ranged from 3.97 μm (outer nasal) to 7.23 μm (inner inferior), demonstrating that the exclusion of scans with significant segmentation error improved the repeatability of retinal thickness measures (reducing the CR). Results for all subfields before and after the exclusion of scans with boundary detection error are shown in Figures 3A to 3B.

Prevalence of Media Opacity and Vitreomacular Interface Abnormality

Mild-to-moderate cataract was present in 7 of 13 (53.8%) subjects with and in 21 of 37 (56.8%) without significant boundary detection error ($P = 1.00$, Fisher's exact test). Similar numbers in both groups were pseudophakic (3/13 [23.1%] vs. 8/37 [21.6%], $P = 1.00$). More subjects who had boundary detection error were found to have vitreomacular interface abnormalities (epiretinal membrane, vitreomacular traction, or thickened posterior hyaloid membranes) than those without boundary detection error (8/13 [61.5%] vs. 14/37 [37.8%], $P = 0.197$).

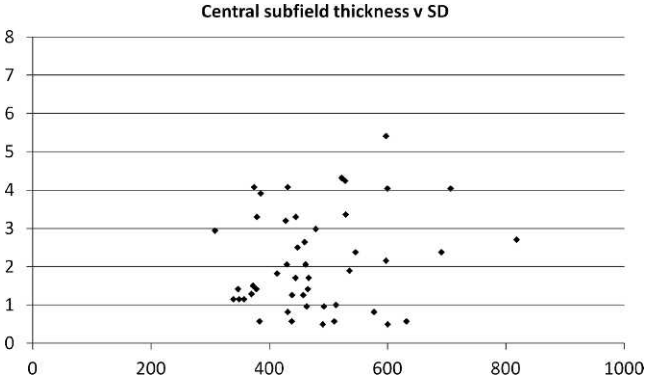

Figure 2. Mean central subfield thickness for each patient plotted against standard deviation.

Discussion

OCT imaging enables rapid noninvasive interrogation of macular morphology and thickness, and is well established as an essential imaging modality in the assessment of the patient with macular disease. SD-OCT technology has resulted in faster image acquisition time and higher resolution images compared to older time domain (TD) technology. The Spectralis OCT combines SD-OCT technology with eye tracking and line scan averaging improving the signal-to-noise ratio, potentially enhancing the ability of segmentation algorithms to detect the true inner and outer retinal boundaries. As change in OCT-based macular thickness measurement is one of the criteria for retreatment with pharmacotherapy in eyes with DME, in clinical practice and in clinical trials,[4,6,13] it is important to establish the repeatability of OCT-derived macular thickness measurements. A knowledge of repeatability would better

Table 3. CR for Nine ETDRS Subfields (μm), Center Point Thickness (μm), and Total Macular Volume (mm^3)

Subfield	Thickness/μm		Volume	
	CR	CI	CR	CI
Center	8.03	7.70–8.35	0.01	0.01–0.01
Inner				
Superior	18.25	17.51–19.00	0.03	0.03–0.03
Temporal	7.67	7.36–7.98	0.01	0.01–0.01
Inferior	14.77	14.16–15.37	0.02	0.02–0.02
Nasal	9.75	9.35–10.15	0.02	0.02–0.02
Outer				
Superior	14.14	13.56–14.71	0.08	0.08–0.09
Temporal	6.54	6.27–6.81	0.04	0.04–0.05
Inferior	10.16	9.75–10.58	0.05	0.05–0.05
Nasal	11.67	11.19–12.14	0.07	0.06–0.07
Center point	21.62	20.74–22.51		
Total macular volume			0.21	0.20–0.22

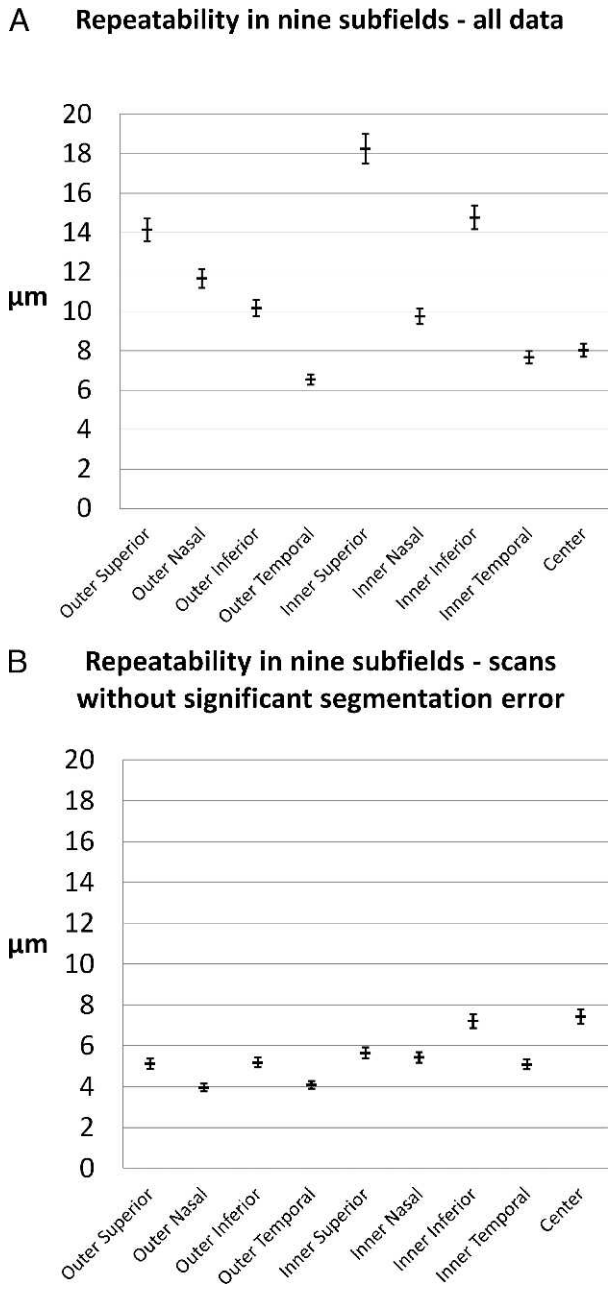

Figure 3. **(A)** CRs for nine ETDRS subfields with 95% CIs shown ($n = 50$). **(B)** CRs with CIs for nine ETDRS subfields once scans with significant segmentation error had been excluded ($n = 37$).

allow physicians to identify true clinical change from measurement variability. To our knowledge, this is the first report evaluating the repeatability of retinal thickness and volume measurements from the Spectralis OCT in eyes with DME. Previous studies of repeatability in macular disease have yielded poorer repeatability estimates (higher CR values) than the current study (Table 4). Evaluation of repeatability in neovascular age-related macular degeneration using Zeiss Stratus TD-OCT has estimated CR for the central macular subfield at 67 µm, representing 23% of the total macular thickness.[14] The Diabetic Retinopathy Clinical Research Network (DRCR.net) study, examining subjects with DME using the Zeiss OCT-3 machine, reported the half-width of the 95% CI for change (equivalent to the Bland Altman CR) to be 38 µm.[8]

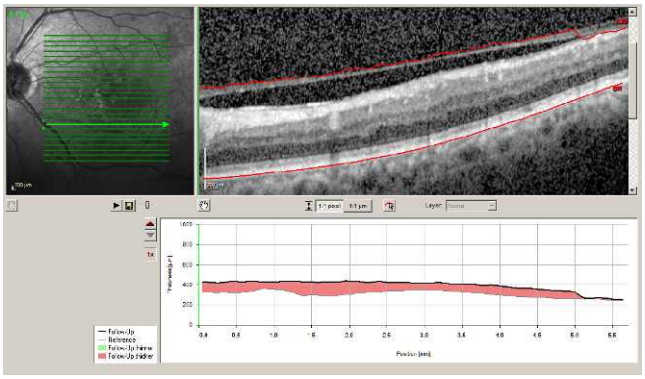

Figure 4. Example of scan affected by significant segmentation error. The automated boundary detection algorithm has identified incorrectly a hyper-reflective line anterior to the ILM as the inner retinal boundary.

Evaluation of SD-OCT devices has yielded similar results. The Cirrus SD-OCT was found to have a CR of 42.4 µm in the central subfield for subjects with neovascular AMD,[15] which improved to 26.1 µm once scans with segmentation error had been removed. In DME, no significant difference was found between TD and SD devices in two separate studies.[9,10] Although this study was not designed to compare the Spectralis OCT directly with a TD device, the estimates obtained for CR are an order of magnitude better than those reported previously in TD devices and in the SD devices that have been used in comparative studies.

Retinal boundary segmentation error has been shown to have an impact on the repeatability of retinal thickness measurements in neovascular AMD.[14] Excluding scan sets affected by segmentation error in one of six radial line scans obtained on Zeiss Stratus OCT improved the CR from 67 to 50 µm. Decreased scan acquisition time has led to the widespread use of raster line scanning techniques, which have increased the number of scans used to derive retinal thickness measurements. This increases the possibility of at least one scan being affected by segmentation error, and in the current study, 16 (32%) scan sets were affected by minor degrees of segmentation error that did not meet our criteria of misplacement of a retinal boundary over at least 1 mm length of scan. This definition was chosen to capture the scans that were most affected by boundary detection error, such as in Figure 4. In this scan, the inner retinal boundary has been placed on a hyperreflective interface anterior to the inner retina; the graphic representation of change below shows that in a previous scan in the series this segmentation error did not occur.

Patients with severe cataract (i.e., grade 3 or greater) were excluded from the study. Mild-to-moderate cataract was found in a similar number of subjects in the subgroup with as in the group without segmentation error. It was more common to find vitreomacular interface abnormalities in subjects whose scans were affected by segmentation error, but with the number of subjects included in the study this did not reach statistical significance. The contribution of vitreomacular interface abnormalities to segmentation error on OCT scans merits further study.

Although excluding scans with segmentation error did not have a large effect on CR in the central subfield, other subfields with higher CRs initially had improved repeatability when these scans were excluded. This suggests that when examining scans to identify clinical change, if one scan set contains segmentation error, numerical retinal thickness values should

Table 4. Summary of Previous Studies Reporting CR in Macular Disease Using TD- and SD-OCT Devices

Study	Disease	Sample Size (eyes)	Device	Central (A1) Subfield CR
Massin et al. (2001)[7]	DME	10	TD (Humphrey OCT)	<21 μm*
Krzystolik et al. for the DRCR.net[8]	DME	212†	TD (Zeiss OCT3)	38 μm
Patel et al.[14]	nvAMD	51	TD (Stratus OCT)	67 μm
Forooghian et al.[9]	DME	33	TD (Stratus OCT)	17.9 μm
			SD (Cirrus HD-OCT)	19.0 μm
Domalpally et al.[10]	DME	63	TD (Stratus OCT)	27.4 μm
			SD (Topcon 3D OCT 1000)	20.1 μm
Parravano et al.[19]	nvAMD	49	SD (Cirrus HD-OCT)	42.4 μm

All studies used the 6 radial line scan protocol (6 mm length) with TD devices (termed "fast macular thickness map" by Zeiss) and the 128 horizontal line protocol (512 A scans per line, 6 × 6 mm) with SD devices.

* Exact figure not reported.

† 107 patients.

be interpreted with a greater degree of caution than if all scans have undergone correct automated retinal boundary detection.

Strengths of this study include the sample size and number of repeated measurements, which ensure that the study is powered to estimate S_w to within 11% of the true population value.[12] Additionally, this study has examined the impact of segmentation error on repeatability, although the investigators acknowledge that there are no accepted protocols for defining this and using a different definition could have yielded different results. One potential weakness of this study is that the OCT macular thickness measurements were taken as part of clinical trial involvement and not specifically for an assessment of repeatability. However, this also could be viewed as a strength as the patients and OCT scanning methods would be more in line with "real world" practice with more generalizable repeatability assessments generated as a result of this pragmatic approach. Importantly, we did not seek to exclude DME eyes with co-existent cataract or vitreomacular interface abnormalities, which frequently co-exist with DME and may increase the chance of segmentation error. Previous studies have excluded patients with conditions that may affect results adversely, meaning that their results from a carefully selected cohort may not be applicable to the DME population in general.[9,16] A further study excluded subjects with high degrees of refractive error or astigmatism.[15] Again, we included these patients to evaluate how the OCT device copes with the entire range of subjects that may be present in clinical situations.

Studies using TD-OCT devices have demonstrated diurnal variation in macular thickness in subjects with DME, with retinal thickness decreasing throughout the day. Estimates for this change range from 13[17] to 49 μm[18] in the central subfield. This previously has been of limited clinical relevance as the magnitude of this change is similar to the inherent test–retest variability of TD-OCT devices reported in Table 4. Our finding with the Spectralis OCT of a smaller CR than the estimated diurnal variation suggests that this phenomenon may be important clinically given the greater precision of measurements taken with this device.

In summary, we reported excellent intrasession repeatability of retinal thickness values using the Spectralis OCT device in eyes with DME. Our results suggested that a retinal thickness change of greater than 8 μm in the central 1 mm subfield is more indicative of clinical change rather than measurement variability. The results of our study may be used to design retreatment criteria in clinical trials. In clinical practice, the results can be used to distinguish true clinical change from measurement variability.

References

1. Klein R, Knudtson MD, Lee KE, Gangnon R, Klein BE. The Wisconsin Epidemiologic Study of Diabetic Retinopathy XXIII: the twenty-five-year incidence of macular edema in persons with type 1 diabetes. *Ophthalmology*. 2009;116:497–503.
2. Bunce C, Wormald R. Leading causes of certification for blindness and partial sight in England & Wales. *BMC Public Health*. 2006;6:58.
3. Michaelides M, Kaines A, Hamilton RD, et al. A prospective randomized trial of intravitreal bevacizumab or laser therapy in the Management of Diabetic Macular Edema (Bolt study) 12-month data: report 2. *Ophthalmology*. 2010;117:1078–1086.
4. Elman MJ, Aiello LP, Beck RW, et al. Randomized trial evaluating ranibizumab plus prompt or deferred laser or triamcinolone plus prompt laser for diabetic macular edema. *Ophthalmology*. 2010;117:1064–1077.
5. Mitchell P, Bandello F, Schmidt-Erfurth U, et al. The RESTORE Study: ranibizumab monotherapy or combined with laser versus laser monotherapy for diabetic macular edema. *Ophthalmology*. 2011;118:615–625.
6. Massin P, Bandello F, Garweg JG, et al. Safety and efficacy of ranibizumab in diabetic macular edema (RESOLVE Study): a 12-month, randomized, controlled, double-masked, multicenter phase II study. *Diabetes Care*. 2010;33:2399–2405.
7. Massin P, Vicaut E, Haouchine B, Erginay A, Paques M, Gaudric A. Reproducibility of retinal mapping using optical coherence tomography. *Arch Ophthalmol*. 2001;119:1135–1142.
8. Krzystolik MG, Strauber SF, Aiello LP, et al. Reproducibility of macular thickness and volume using Zeiss optical coherence tomography in patients with diabetic macular edema. *Ophthalmology*. 2007;114:1520–1525.
9. Forooghian F, Cukras C, Meyerle CB, Chew EY, Wong WT. Evaluation of time domain and spectral domain optical coherence tomography in the measurement of diabetic macular edema. *Invest Ophthalmol Vis Sci*. 2008;49:4290–4296.
10. Domalpally A, Gangaputra S, Peng Q, Danis RP. Repeatability of retinal thickness measurements between spectral-domain and time-domain optical coherence tomography images in macular disease. *Ophthalmic Surg Lasers Imaging*. 2010; 41(suppl):S34–S41.
11. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet*. 1986;1:307–310.
12. Bland M. How can I decide the sample size for a repeatability study? University of York. [User Web Areas]. 2010. Available at:

http://www-users.york.ac.uk/~mb55/meas/sizerep.htm. Accessed May 17, 2010.

13. Nguyen QD, Brown DM, Marcus DM, et al. Ranibizumab for diabetic macular edema: results from 2 phase III randomized trials: RISE and RIDE. *Ophthalmology*. 2012;119:789–801.
14. Patel PJ, Chen FK, Ikeji F, et al. Repeatability of stratus optical coherence tomography measures in neovascular age-related macular degeneration. *Invest Ophthalmol Vis Sci*. 2008;49: 1084–1088.
15. Parravano M, Oddone F, Boccassini B, et al. Reproducibility of macular thickness measurements using Cirrus SD-OCT in neovascular age-related macular degeneration. *Invest Ophthalmol Vis Sci*. 2010;51:4788–4791.
16. Lammer J, Scholda C, Prunte C, Benesch T, Schmidt-Erfurth U, Bolz M. Retinal thickness and volume measurements in diabetic macular edema: a comparison of four optical coherence tomography systems. *Retina*. 2011;31:48–55.
17. Danis RP, Glassman AR, Aiello LP, et al. Diurnal variation in retinal thickening measurement by optical coherence tomography in center-involved diabetic macular edema. *Arch Ophthalmol*. 2006;124:1701–1707.
18. Kotsidis ST, Lake SS, Alexandridis AD, Ziakas NG, Ekonomidis PK. 24-Hour variation of optical coherence tomography-measured retinal thickness in diabetic macular edema. *Eur J Ophthalmol*. 2012;22:785–791.
19. Parravano M, Oddone F, Boccassini B, et al. Reproducibility of macular thickness measurements using Cirrus SD-OCT in neovascular age-related macular degeneration. *Invest Ophthalmol Vis Sci*. 2010;51:4788–4791.