Quantitative comparison of macular segmentation performance using identical retinal regions across multiple spectral-domain optical coherence tomography instruments

Sebastian M Waldstein, Bianca S Gerendas, Alessio Montuoro, Christian Simader, Ursula Schmidt-Erfurth

ABSTRACT

accuracy and clinical relevance.

before and after correction.

highest error impact.

INTRODUCTION

Purpose Comparison of optical coherence tomography

(OCT) segmentation performance regarding technical

Methods 29 eves were imaged prospectively with

Spectralis (Sp), Cirrus (Ci), 3D-OCT 2000 (3D) and RS-

3000 (RS) OCTs. Raw data were evaluated in validated

custom software. A 1 mm diameter subfield, centred on

the fovea, was investigated to compare identical regions

for each case. Segmentation errors were corrected on

each B-scan enclosed in this subfield. Proportions of

wrongly segmented A-scans were noted for inner and

outer retinal boundaries. Centre point thickness (CPT)

and central macular thickness (CMT) were compared

Results Segmentation errors occurred in 77% and

affected on average 29% of A-scans, resulting in mean

segmentation errors was 48% (Sp), 79% (Ci), 86% (3D)

and 93% (RS), p<0.001. Mean proportions of A-scans with wrong outer retinal boundary were 30% (Sp), 9%

(Ci), 23% (3D) and 10% (RS), p=0.006; proportions for the inner retinal boundary were 11% (Sp). 12% (Ci).

6% (3D) and 21% (RS), p=0.034. Mean deviations in

30/13 µm (3D) and 18/8 µm (RS), p=0.409/0.477.

OCT devices regarding technical accuracy and clinical

impact. Spectralis showed lowest error incidence but

Optical coherence tomography (OCT) is an innova-

tive in vivo imaging technique with worldwide

popularity, especially since the introduction of spec-

tral domain (SD) OCT.¹⁻³ A qualitative assessment

of the morphological changes that OCT visualises is

indispensable for the early diagnosis and accurate

follow-up of macular diseases.⁴ However, also the quantitative analysis of OCT data plays a pivotal

role in research studies and clinical practice.⁵ ⁶

Particularly central macular thickness (CMT) has

been heavily used as an outcome variable in major clinical trials evaluating pharmacological or laser

treatment of neovascular age-related macular degen-

practice as they permit a rapid evaluation of poten-

tial disease progression and/or treatment response.

Conclusions By comparison of identical regions, substantial differences were detected between the tested

CPT/CMT were 41/28 µm (Sp), 17/11 µm (Ci),

differences of 24/13 µm (CPT/CMT). The incidence of

Department of Ophthalmology, Christian-Doppler-Laboratory for Ophthalmic Image Analysis, Vienna Reading Center, Medical University of Vienna, Vienna, Austria

Correspondence to

Dr Christian Simader, Department of Ophthalmology, Christian-Doppler-Laboratory for Ophthalmic Image Analysis, Waehringer Guertel 18-20, Vienna 1090, Austria; optima@meduniwien.ac.at

SMW and BSG contributed equally.

Received 23 May 2014 Revised 10 October 2014 Accepted 8 December 2014 Published Online First 6 January 2015



To cite: Waldstein SM, Gerendas BS, Montuoro A, *et al. Br J Ophthalmol* 2015;**99**:794–800.

Waldstein SM, et al. Br J Ophthalmol 2015;99:794-800. doi:10.1136/bjophthalmol-2014-305573

eration (AMD) or retinal vascular disease.^{7–10} CMT macula clinic metrics are also frequently employed in clinical Medical U

Retinal thickness reports are provided by most commercially available OCT instruments. For thickness measurements, the devices use built-in automated segmentation algorithms that detect the retinal boundaries on OCT images. However, such automated segmentation algorithms are reported to fail frequently, particularly in the presence of pathology, with an output of incorrect thickness values as consequence.^{11–24} Unless segmentation errors are corrected manually, they may lead to wrong conclusions in clinical practice and may confound the results of clinical trials.²⁵

Importantly, the presence of a segmentation error may not automatically entail a clinically relevant retinal thickness deviation. It is, therefore, necessary to objectively quantify the presence and extent of segmentation errors (ie, the technical accuracy of the segmentation algorithm), as well as the impact of segmentation errors on thickness outputs (ie, the clinical relevance). Previous studies used subjective scoring scales,¹¹ ¹⁴ ¹⁶ ¹⁷ ²¹ ²² assessed thickness deviations,¹³ ¹⁸ or both,¹⁵ ²⁴ to investigate the severity of segmentation artefacts. As yet, an objective, quantitative method to evaluate the technical accuracy of segmentation algorithms is lacking. Moreover, previous studies including different instruments failed to ensure a comparison of identical regions.¹³ ¹⁴ ¹⁹ ²² A fair comparison of SD-OCT machines would, however, only be possible if exactly corresponding retinal loci were compared. A compelling quantitative inter-instrument comparison would be of value for clinicians and researchers when choosing between the several available SD-OCT devices for clinical practice or a research trial.

The aim of this study was to quantitatively compare the segmentation performance across four prevalent SD-OCT devices, focusing on both a measure of the technical accuracy of segmentation algorithms and the resulting clinical impact of segmentation errors.

PATIENTS AND METHODS

This prospective, comparative non-interventional case series was conducted in compliance with the tenets of the Declaration of Helsinki and was prospectively approved by the ethics committee at the Medical University of Vienna. Patients seen at the macula clinic at the Department of Ophthalmology, Medical University of Vienna, Austria, were included. Inclusion criteria were as follows: age 18 years or older, patient able and willing to give



written informed consent, presence of retinal disease involving thickening or thinning of the central retina in at least one eye, and clear optical media allowing OCT imaging in high quality. If a patient's compliance permitted, both eyes were eligible for inclusion. All patients underwent a complete ophthalmic examination and pupils were dilated before the imaging procedure.

OCT imaging procedure

Image acquisition followed an established standard procedure provided by the Vienna Reading Center. OCT imaging was performed by experienced, reading-centre certified operators, using each a commercially available 3D-OCT 2000 Mark II (Topcon, Tokyo, Japan), Cirrus HD-OCT (Carl Zeiss Meditec, Dublin, California, USA), RS-3000 (Nidek, Tokyo, Japan) and Spectralis (Heidelberg Engineering, Heidelberg, Germany) instrument under standardised conditions. Identical or similar raster scan patterns were chosen across the devices to enable an objective comparison. The specific scan protocols were as follows-3D-OCT 2000: '3D Macula' pattern with 128 sections (512 A-scans each) in a 6×6 mm area, Cirrus HD-OCT: 'Macular Cube' pattern with 128 sections (512 A-scans each) in a 6×6 mm area, RS-3000: 'Macula Map' pattern with 128 sections (512 A-scans each) in a 6×6 mm area and Spectralis: custom raster scan pattern with 49 or 25 sections (512 A-scans each; lower number of sections chosen in case of patient fatigue) in a 20°×20° field of view and automatic real-time averaging activated at 29 frames. In all instruments, an internal fixation target was provided. Each subject was scanned at all devices within 1 h, by the same operator and in random order, to counteract a potential systematic bias. Scans were immediately discarded and repeated if the operator noticed low scan quality upon initial review. In detail, the operator was instructed to inspect the entire OCT volume for excellent signal strength, even illumination and absence of motion or blink artefacts.

Grading software

OCT raw data were exported from each instrument and evaluated in custom, validated Vienna Reading Center software ('OCTAVO'). The software shows B-scan images as well as the device-provided segmentation lines in standardised display conditions. The segmentation lines can be manually adjusted at any location throughout the scan and can be hidden as needed in order to facilitate B-scan assessment. OCTAVO, furthermore, allows the grader to plot an Early Treatment Diabetic Retinopathy Study (ETDRS)-like grid and to display its centre point and the actual borders of the subfields on the corresponding B-scans (figure 1). The software provides, among other variables, centre point thickness (CPT, retinal thickness at the centre of the ETDRS grid) and CMT (average retinal thickness in the central 1 mm subfield of the grid). Thickness values can be calculated using either the device-provided segmentation only or using device-provided segmentation with manual corrections after reader input.

Grading procedure

Two masked certified reading supervisors of the Vienna Reading Center (SMW, BSG) evaluated all SD-OCT scans, strictly adhering to a standardised grading protocol. All scans of a particular eye were graded consistently by the same reader to counteract a potential annotation bias. In order to compare identical retinal loci across each tested SD-OCT device, the ETDRS grid was first manually replotted to the foveal centre. Afterwards, all errors in the device-provided segmentation lines were corrected on all B-scans enclosed within the central subfield (eg, 22 B-scans in Cirrus). Any visible deviations of the segmentation lines, however small these may be, were corrected manually. Only the contents of the central subfield were corrected and compared as CMT, one of the most important variables in clinical and research use, and precise manual segmentation



Figure 1 Example case with varying segmentation performance across the four spectral domain-optical coherence tomography (SD-OCT) devices. The segmentation boundaries are displayed in green in custom reading centre software. The central vertical yellow line marks the Early Treatment Diabetic Retinopathy Study-like grid centre point; turquoise vertical lines mark the borders of the central subfield. In this study, patients with neovascular age-related macular degeneration, misidentification of the retinal boundaries occurred in a range from 7% to 40% of possible A-scans within the central subfield (red lines represent segmentation errors).

Clinical science

correction in the entire volume stack would be impracticable. Correction of line errors was performed separately for the inner retinal border (ie, retina-vitreous interface) and the outer retinal border (ie, either above (3D-OCT 2000) or at the retinal pigment epithelium (Cirrus, RS-3000) or at Bruch's membrane (Spectralis)).

Assessment of technical accuracy of segmentation algorithms

A quantitative measure of the technical accuracy of the segmentation algorithms was devised as follows: each A-scan (equivalent to one pixel in y-direction on the OCT image) containing a segmentation error correction as per reader input was counted by the software. The sum of affected A-scans within the central subfield was divided by the total number of A-scans within the central subfield and expressed as percentage of affected A-scans. The percentage values were noted separately for segmentation errors at the inner and outer retinal borders.

Assessment of clinical impact of segmentation errors

Differences between uncorrected thickness values and manually corrected thickness values were used as a measure of the clinical impact of segmentation errors. A thickness deviation of 7 μ m or more (2× average nominal z-resolution of the evaluated devices) was set as cut-off point for clinical significance of segmentation errors, assuming that deviations below this value would not warrant manual correction in clinical practice.

Statistical evaluation

PASW Statistics (V.18.0.0, SPSS, Chicago, Illinois, USA) was used for statistical analysis. χ^2 Tests were used to compare categorical variables and analysis of variance was used to compare scale variables with assumed normal distribution between the SD-OCT devices. Linear regression was used to assess the relation between technical accuracy and clinical impact of segmentation errors. p Values ≤ 0.05 were considered statistically significant.

RESULTS

Twenty-nine eyes of 19 patients were included. The retinal pathological features were as follows: seven retinal vein occlusion, three diabetic macular oedema, three epiretinal membrane (grouped as 'oedematous disease'); six neovascular AMD, three lamellar hole, two Stargardt disease, one non-neovascular AMD (grouped as 'degenerative disease'); and four healthy eyes. The mean age of subjects was 50 ± 14 years; 10 were women and nine were men. Mean retinal thickness values in Cirrus HD-OCT were 296±150 µm (CPT) and 340±118 µm (CMT).

In the assessed replotted central subfield area, segmentation errors occurred in 76.7% of cases. In these cases, errors affected, on average, 22.5% of A-scans within the central subfield, resulting in a mean thickness difference of 18.8 μ m for CPT and 10.2 μ m for CMT. Clinically significant CMT deviations occurred in 25.9% of cases.

Incidence of segmentation errors across the devices

Proportions of cases with segmentation errors ranged from 48% (Spectralis) to 79% (Cirrus), 86% (3D-OCT 2000) and 93% (RS-3000) (p<0.001). Misidentification of the outer retinal border ranged from 27% (Spectralis) to 62% (3D-OCT 2000) (p=0.04), while misidentification of the inner retinal border ranged from 27% (Spectralis) to 93% (RS-3000) (p<0.001). Table 1 shows proportions of cases with segmentation errors across the tested devices with breakdowns for retinal borders and disease types.

Technical accuracy of segmentation algorithms

Among the cases with segmentation errors, the mean proportions of A-scans with erroneous segmentation were 21.0% (Cirrus), 28.9% (3D-OCT 2000), 30.6% (RS-3000) and 40.8% (Spectralis), p=0.237. Significant differences in the technical accuracy were found between the segmentation of the outer retinal boundary (mean 9.0% error (Cirrus) up to mean 30.4% error (Spectralis), p=0.034) and inner retinal boundary (mean 5.5% error (3D-OCT 2000) up to mean 20.8% error (RS-3000), p=0.006). Table 2 provides data on the technical accuracy with regard to retinal borders and disease types.

Table 1 Incidence of segmentation errors in central subfield across tested SD-OCT devices

	3D-OCT 2000 (%)	Cirrus HD-OCT (%)	RS-3000 (%)	Spectralis (%)	p Value*	Total (%)
				(,-,	P	
All eyes						
Any boundary	86.2	79.3	93.1	48.3	<0.001	76.7
Outer boundary	62.1	34.5	34.5	27.6	0.037	39.7
Inner boundary	41.4	65.5	93.1	27.6	<0.001	56.9
Healthy eyes						
Any boundary	50.0	75.0	100.0	50.0	0.362	68.8
Outer boundary	25.0	25.0	0.0	25.0	0.746	18.8
Inner boundary	25.0	75.0	100.0	25.0	0.077	56.3
Oedematous disease (R)	VO, DME, ERM)					
Any boundary	84.6	76.9	84.6	30.8	0.006	69.2
Outer boundary	61.5	23.1	30.8	15.4	0.064	32.7
Inner boundary	30.8	61.5	84.6	15.4	0.002	48.1
Degenerative disease (A	MD, macular hole, Stargardt)					
Any boundary	100.0	76.9	100.0	66.7	0.039	87.5
Outer boundary	75.0	50.0	50.0	41.7	0.388	54.2
Inner boundary	58.3	66.7	100.0	41.7	0.021	66.7

Bold p-values are <0.05 which is the formal significance level.

AMD, age-related macular degeneration; DME, diabetic macular oedema; ERM, epiretinal membrane; OCT, optical coherence tomography; RVO, retinal vein occlusion; SD, spectral domain.

Table 2 Mean percentages of affected A-scans within central subfield					
	3D-OCT 2000 (%)	Cirrus HD-OCT (%)	RS-3000 (%)	Spectralis (%)	p Value*
All eyes					
Any boundary	28.9	21.0	30.6	40.9	0.237
Outer boundary	23.5	9.0	9.85	30.4	0.006
Inner boundary	5.5	12.0	20.79	10.6	0.034
Healthy eyes					
Any boundary	18.9	12.3	16.7	9.8	0.793
Outer boundary	6.8	0.3	0.0	2.0	0.273
Inner boundary	12.2	11.9	16.7	7.7	0.864
Oedematous disease (RVC	D, DME, ERM)				
Any boundary	15.2	16.7	24.8	26.9	0.571
Outer boundary	14.7	6.7	4.8	23.3	0.370
Inner boundary	1.0	10.0	19.9	3.6	0.001
Degenerative disease (AM	ID, macular hole, Stargardt)				
Any boundary	43.3	28.0	40.7	55.7	0.394
Outer boundary	34.8	13.9	17.8	41.0	0.204
Inner boundary	8.5	14.4	22.9	14.7	0.370

Bold p-values are <0.05 which is the formal significance level. *Analysis of variance.

AMD, age-related macular degeneration; DME, diabetic macular oedema; ERM, epiretinal membrane; OCT, optical coherence tomography; RVO, retinal vein occlusion.

Clinical impact of segmentation errors

Mean deviations in CPT and CMT were 17.6 µm/8.4 µm (RS-3000), 16.7 µm/10.6 µm (Cirrus), 29.7 µm/13.0 µm (3D-OCT 2000) and 41.1 µm/28.1 µm (Spectralis). One eye with a large pigment epithelial detachment (figure 2) showed an extraordinarily severe segmentation error in Spectralis, resulting in a CPT/CMT deviation of 362 µm/336 µm. After exclusion of this outlier, mean deviations in CPT/CMT were 16.5 µm/4.4 µm for the Spectralis. Clinically relevant CMT deviation (over 7 µm) occurred in 17.2% (Cirrus HD-OCT), 20.7% (Spectralis), 27.6% (RS-3000) and 37.9% (3D-OCT 2000) (p=0.278). Table 3 shows an overview over the clinical impact of segmentation errors including breakdown for disease types.



Figure 2 Example case showing the highest retinal thickness deviation in the study (Spectralis optical coherence tomography, OCT). This patient with neovascular age-related macular degeneration presented with a large pigment-epithelial detachment (PED). The segmentation algorithm included in Spectralis OCT aims to detect Bruch's membrane, and should therefore include the PED into the measurement. However, the segmentation follows the retinal pigment epithelium in this case (lower green line). The segmentation was manually corrected to Bruch's membrane (yellow horizontal line). This misidentification resulted in a thickness error of 362 µm.

Impact of disease type on segmentation error incidence

Table 4 provides a comparison of the incidence of segmentation errors, technical accuracy and clinical impact among the evaluated disease types (healthy eyes, degenerative disease and oedematous disease). Generally, segmentation errors were found to be more frequent, severe and impactful in degenerative retinal disease.

Correlation between technical accuracy and clinical impact

Linear regression analysis was performed to assess the correlation between technical accuracy (area affected by segmentation errors in central subfield) and clinical impact (CMT deviation). Significant correlations were detected for all devices (regression lines are shown in figure 3). The percentage of affected A-scans had minimum impact on CMT in Spectralis (after exclusion of outlier), with a coefficient of 0.089 ± 0.016 (R²=0.74, p<0.001). 3D-OCT 2000 (R²=0.55, p<0.001) and RS-3000 (R²=0.46, p < 0.001) showed moderate correlations with coefficients of 0.523±0.098 and 0.357±0.077, respectively. Cirrus HD-OCT showed a strong correlation ($R^2=0.75$, p<0.001) with substantial impact of the area of segmentation error on CMT deviations (coefficient= 0.865 ± 0.108). Correlation between disease severity and segmentation error incidence.

Linear regression analysis was performed to assess the correlation between disease severity (ie, CMT) and the rate of segmentation errors as well as technical accuracy and clinical impact. No significant correlations were detected (all p values >0.05).

DISCUSSION

Our study was performed to quantitatively compare macular segmentation performance across four prevalent SD-OCT devices. We demonstrated the feasibility of a comparison of identical retinal loci and presented a novel measure for the technical accuracy of segmentation algorithms, that is, the proportion of error-affected A-scans. Our data indicate that significant differences exist between the devices regarding the incidence of errors and the technical accuracy, but not regarding the clinical impact of segmentation errors.

The incidence of any detectable alignment errors within the central subfield showed major differences between the tested **Clinical science**

Table 3 Mean centre point and central subfield thickness dev	viations
--	----------

	3D-OCT 2000	Cirrus HD-OCT	RS-3000	Spectralis	p Value*
All eyes					
Mean CPT deviation	29.7 μm	16.7 μm	17.6 μm	16.5 μm (41.1 μm†)	0.409
Mean CMT deviation	13.0 μm	10.6 μm	8.4 μm	4.4 μm (28.1 μm†)	0.477
Significant CMT deviation‡	37.9%	17.2%	27.6%	17.9% (20.7%†)	0.287
Healthy eyes					
Mean CPT deviation	17.0 μm	3.0 μm	10.0 μm	0.0 μm	0.008
Mean CMT deviation	1.5 μm	1.7 μm	1.8 μm	1.5 μm	0.994
Significant CMT deviation‡	0%	0%	0%	0%	n/a
Oedematous disease (RVO, DME, ERI	M)				
Mean CPT deviation	11.9 µm	17.3 μm	5.2 μm	7.5 μm	0.650
Mean CMT deviation	3.6 µm	6.4 μm	4.2 μm	2.8 μm	0.893
Significant CMT deviation‡	15.4%	7.7%	15.4%	7.7%	0.861
Degenerative disease (AMD, macular	hole, Stargardt)				
Mean CPT deviation	48.2 μm	20.1 μm	31.6 μm	26.3 μm	0.357
Mean CMT deviation	23.6 µm	17.5 μm	14.4 μm	6.1 μm	0.369
Significant CMT deviation‡	75.0%	33.3%	75.0%	41.7%	0.198

Bold p-values are <0.05 which is the formal significance level.

*Analysis of variance.

†Before exclusion of 1 outlier.

‡7 μm or more.

AMD, age-related macular degeneration; CMT, central macular thickness; CPT, centre point thickness; DME, diabetic macular oedema; ERM, epiretinal membrane; OCT, optical coherence tomography; RVO, retinal vein occlusion.

devices. It was lowest in Spectralis (about one-half of the evaluated eyes affected) and highest in the Nidek RS-3000, with segmentation errors present in almost all cases (93%). Among the cases with segmentation errors, Cirrus HD-OCT showed the least technical inaccuracy, with a mean of 21% affected A-scans within the central subfield. Using this particular measure, Spectralis OCT had worst outcomes with, on average, 41% A-scans affected. With a view to the potential clinical impact in CMT metrics, the Spectralis instrument showed the highest mean error (28.1 μ m) and the single highest CMT deviation (336 μ m). However, after exclusion of this extreme case, Spectralis delivered very reliable segmentation results (average 4 μ m deviation).

To the best of our knowledge, this is the first study to evaluate segmentation performance in the Nidek RS-3000 instrument. For this particular machine, our results show considerable technical inaccuracy, especially for detection of the inner retinal boundary. Of the affected cases, 27% showed clinically relevant segmentation errors. However, in terms of absolute CMT deviation, performance was generally acceptable with 8 μm error on average.

The results of our study should best be interpreted in synopsis of all its outcome variables. For the Spectralis instrument, although the chance of errors was rather low, errors consistently affected large proportions of A-scans and sometimes caused considerable thickness deviations. It can be assumed that this finding is due to a lower sampling rate in Spectralis. For the clinician, our results imply that it may be necessary to inspect Spectralis thickness outputs for errors, since these can be severe, and manual correction only requires little effort as the sampling density is low.

With regard to the other instruments, uniformly using the 512×128 pattern without B-scan averaging, a correction of segmentation errors in these machines does not seem feasible or worth the effort since thickness deviations are often within an acceptable range and manual correction of several B-scans might

Table 4	Incidence and	impact of	f segmentation	errors by	/ disease t	type (all	devices
---------	---------------	-----------	----------------	-----------	-------------	-----------	---------

	Healthy eyes	Degenerative disease	Oedematous disease	p Value
Incidence of segmentation errors				
Any boundary	9.5%	35.3%	31.0%	0.130*
Outer boundary	2.6%	22.4%	14.7%	0.017*
Inner boundary	7.8%	27.6%	21.6%	0.172*
Technical accuracy (Mean percentage of affected	A-scans)			
Any boundary	10.07%	36.10%	13.74%	<0.001†
Outer boundary	1.16%	22.87%	7.09%	0.001†
Inner boundary	8.91%	13.24%	6.64%	0.093†
Clinical impact				
Mean CPT deviation	5.19 μm	35.50 μm	7.53 μm	0.002†
Mean CMT deviation	1.13 μm	21.04 μm	3.08 µm	0.015†
Clinically significant CMT deviation (%)	0.0%	20.7%	5.2%	<0.001*

Bold p-values are <0.05 which is the formal significance level.

 $^{*}\chi^{2}$ test.

†Änalysis of variance. CMT, central macular thickness, CPT, centre point thickness.



Figure 3 Correlation of technical accuracy and clinical impact of segmentation errors. For each case, the technical accuracy measured in percentage of A-scans affected (x-axis) and the deviation in central macular thickness (CMT) (y-axis) are plotted and correlated. For Spectralis optical coherence tomography (OCT, top), even larger-area segmentation errors have little impact on CMT output; this device could, therefore, be considered more reliable. The other three instruments (bottom) show a stronger correlation between CMT deviation and segmentation error area.

be too laborious. In a research setting, however, manual inspection and correction of alignment errors remains mandatory, with up to 40% clinically relevant thickness deviations detected.

Not surprisingly, eyes in the 'degenerative disease' subgroup including AMD and lamellar macular holes, showed significantly poorer segmentation outcomes. It is therefore an unmet medical need to develop segmentation techniques providing robust outcomes even in the presence of severe disruptive pathology.

A limitation of our study is the use of a different scanning pattern in the Spectralis instrument. As the 512×128 on 6 mm×6 mm pattern is not available on the Spectralis machine, we employed the Vienna Reading Center standard pattern (512×49) for this instrument. Proportions were used instead of absolute numbers in all statistics in order to account for the difference in sampling density. A further limitation is the use of scan averaging (in accordance with our reading centre standard operating procedures) in the Spectralis instrument as compared with no averaging in the other SD-OCT machines. However, averaging in volume scans is currently only available in the Spectralis and it seems prudent to compare devices with regard to their maximum technical capabilities. Moreover, OCT signal quality as a potential influencing factor in segmentation performance was not evaluated in this study. Since no uniform signal quality measures are available in this study, only optimal quality scans were selected at each device during acquisition to remove such influence as much as possible.

In conclusion, our study demonstrated the feasibility of an objective comparison of segmentation performance using identical retinal loci in four SD-OCT machines. The Spectralis device showed the lowest error rates, but the highest impact of errors if present. This may be attributable to a lower sampling density in Spectralis. The remaining SD-OCT instruments demonstrated higher error rates, with generally low clinical impact of segmentation errors. Manual inspection and correction of segmentation failures may be clinically useful and feasible for Spectralis scans.

Contributors SMW: Design of the study, acquisition of data, analysis of data, literature research, writing, revision and approval of article. BSG: design of the study, acquisition of data, analysis of data, revision and approval of article. AM: analysis of data, revision and approval of article. CS: design of the study, provision of funding, revision and approval of article. US-E: revision and approval of article, provision of funding.

Funding The financial support by the Austrian Federal Ministry of Economy, Family and Youth and the National Foundation for Research, Technology and Development is gratefully acknowledged.

Competing interests None.

Ethics approval Ethics Committee, Medical University of Vienna.

Provenance and peer review Not commissioned; externally peer reviewed.

REFERENCES

- Huang D, Swanson EA, Lin CP, et al. Optical coherence tomography. Science 1991;254:1178–81.
- 2 Fercher AF, Hitzenberger CK, Drexler W, et al. In vivo optical coherence tomography. Am J Ophthalmol 1993;116:113–14.
- 3 Wojtkowski M, Leitgeb R, Kowalczyk A, et al. In vivo human retinal imaging by Fourier domain optical coherence tomography. J Biomed Opt 2002;7:457–63.
- 4 Geitzenauer W, Hitzenberger CK, Schmidt-Erfurth UM. Retinal optical coherence tomography: past, present and future perspectives. Br J Ophthalmol 2011;95:171–7.
- 5 Hee MR, Puliafito CA, Wong C, et al. Quantitative assessment of macular edema with optical coherence tomography. Arch Ophthalmol 1995;113:1019–29.
- 6 Hee MR, Baumal CR, Puliafito CA, et al. Optical coherence tomography of age-related macular degeneration and choroidal neovascularization. Ophthalmology 1996;103:1260–70.

Clinical science

- 7 Fung AE, Lalwani GA, Rosenfeld PJ, et al. An optical coherence tomography-guided, variable dosing regimen with intravitreal ranibizumab (Lucentis) for neovascular age-related macular degeneration. Am J Ophthalmol 2007;143:566–83.
- 8 Brown DM, Campochiaro PA, Singh RP, *et al*. Ranibizumab for macular edema following central retinal vein occlusion: six-month primary end point results of a phase III study. *Ophthalmology* 2010;117:1124–33.e1.
- 9 Campochiaro PA, Heier JS, Feiner L, et al. Ranibizumab for macular edema following branch retinal vein occlusion: six-month primary end point results of a phase III study. Ophthalmology 2010;117:1102–12.e1.
- 10 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;118:615–25.
- 11 Sadda SR, Wu Z, Walsh AC, et al. Errors in retinal thickness measurements obtained by optical coherence tomography. Ophthalmology 2006;113:285–93.
- 12 Krebs I, Haas P, Zeiler F, et al. Optical coherence tomography: limits of the retinal-mapping program in age-related macular degeneration. Br J Ophthalmol 2008;92:933–5.
- 13 Ho J, Sull AC, Vuong LN, et al. Assessment of artifacts and reproducibility across spectral- and time-domain optical coherence tomography devices. *Ophthalmology* 2009;116:1960–70.
- 14 Mylonas G, Ahlers C, Malamos P, et al. Comparison of retinal thickness measurements and segmentation performance of four different spectral and time domain OCT devices in neovascular age-related macular degeneration. Br J Ophthalmol 2009;93:1453–60.
- 15 Patel PJ, Chen FK, da Cruz L, et al. Segmentation error in stratus optical coherence tomography for neovascular age-related macular degeneration. Invest Ophthalmol Vis Sci 2009;50:399–404.

- 16 Keane PA, Mand PS, Liakopoulos S, *et al*. Accuracy of retinal thickness measurements obtained with Cirrus optical coherence tomography. *Br J Ophthalmol* 2009;93:1461–7.
- 17 Krebs I, Falkner-Radler C, Hagen S, et al. Quality of the threshold algorithm in age-related macular degeneration: stratus versus cirrus OCT. Invest Ophthalmol Vis Sci 2009;50:995–1000.
- 18 Han IC, Jaffe GJ. Evaluation of artifacts associated with macular spectral-domain optical coherence tomography. *Ophthalmology* 2010;117:1177–89.e4.
- 19 Giani A, Cigada M, Esmaili DD, et al. Artifacts in automatic retinal segmentation using different optical coherence tomography instruments. *Retina* 2010;30:607–16.
- 20 Modjtahedi S, Chiou C, Modjtahedi B, et al. Comparison of Macular Thickness Measurement and Segmentation Error Rate Between Stratus and Fourier-Domain Optical Coherence Tomography. Ophthalmic Surg Lasers Imaging 2010;41:301.
- 21 Malamos P, Ahlers C, Mylonas G, et al. Evaluation of segmentation procedures using spectral domain optical coherence tomography in exudative age-related macular degeneration. *Retina* 2011;31:453–63.
- 22 Krebs I, Smretschnig E, Moussa S, et al. Quality and reproducibility of retinal thickness measurements in two spectral domain optical coherence tomography machines. *Invest Ophthalmol Vis Sci* 2011;52:6925–33.
- 23 Matt G, Sacu S, Buehl W, et al. Comparison of retinal thickness values and segmentation performance of different OCT devices in acute branch retinal vein occlusion. Eye 2011;25:511–18.
- 24 Song Y, Lee BR, Shin YW, et al. Overcoming Segmentation Errors in Measurements of Macular Thickness Made By Spectral-Domain Optical Coherence Tomography. *Retina* 2012;32:569–80.
- 25 Keane PA, Liakopoulos S, Jivrajka RV, et al. Evaluation of optical coherence tomography retinal thickness parameters for use in clinical trials for neovascular age-related macular degeneration. *Invest Ophthalmol Vis Sci* 2009;50:3378–85.



Quantitative comparison of macular segmentation performance using identical retinal regions across multiple spectral-domain optical coherence tomography instruments

Sebastian M Waldstein, Bianca S Gerendas, Alessio Montuoro, Christian Simader and Ursula Schmidt-Erfurth

Br J Ophthalmol 2015 99: 794-800 originally published online January 6, 2015 doi: 10.1136/bjophthalmol-2014-305573

Updated information and services can be found at: http://bjo.bmj.com/content/99/6/794

These include:

References	This article cites 25 articles, 9 of which you can access for free at: http://bjo.bmj.com/content/99/6/794#BIBL
Email alerting service	Receive free email alerts when new articles cite this article. Sign up in the box at the top right corner of the online article.

Notes

To request permissions go to: http://group.bmj.com/group/rights-licensing/permissions

To order reprints go to: http://journals.bmj.com/cgi/reprintform

To subscribe to BMJ go to: http://group.bmj.com/subscribe/