ORIGINAL COMMUNICATION

In vivo assessment of retinal neuronal layers in multiple sclerosis with manual and automated optical coherence tomography segmentation techniques

Michaela A. Seigo · Elias S. Sotirchos · Scott Newsome · Aleksandra Babiarz · Christopher Eckstein · E'Tona Ford · Jonathan D. Oakley · Stephanie B. Syc · Teresa C. Frohman · John N. Ratchford · Laura J. Balcer · Elliot M. Frohman · Peter A. Calabresi · Shiv Saidha

Abstract Macular optical coherence tomography (OCT) segmentation, enabling quantification of retinal axonal and neuronal subpopulations, may help elucidate the neuroretinal pathobiology of multiple sclerosis (MS). This study aimed to determine the agreement, reproducibility, and visual correlations of retinal layer thicknesses measured by different OCT segmentation techniques, on two spectral-domain OCT devices. Macular scans of 52 MS patients and 30 healthy controls from Spectralis OCT and Cirrus HD-OCT were segmented using fully manual (Spectralis), computer-aided manual (Spectralis and Cirrus), and fully automated (Cirrus) segmentation techniques. Letter acuity was recorded. Bland-Altman analyses revealed low mean differences across OCT segmentation techniques on both devices for ganglion cell + inner plexiform layers (GCIP; 0.76–2.43 μm), inner nuclear + outer plexiform layers (INL + OPL; 0.36–1.04 μm), and outer nuclear layers including photoreceptor segment (ONL + PR; 1.29–3.52 μm) thicknesses. Limits of agreement for GCIP and ONL + PR thicknesses were narrow. Results of fully manual and computer-aided manual segmentation were comparable to those of fully automated segmentation. MS patients demonstrated macular RNFL, GCIP, and ONL + PR thinning compared to healthy controls across OCT segmentation techniques, irrespective of device ($p < 0.03$ for all). Low-contrast letter acuity in MS correlated significantly and more strongly with GCIP than peripapillary RNFL thicknesses, regardless of the segmentation method or device. GCIP and ONL + PR thicknesses, measured by different OCT devices and segmentation techniques, are reproducible and agree at the individual and cohort levels. GCIP thinning in MS correlates with visual dysfunction. Significant ONL + PR thinning, detectable across OCT segmentation techniques and devices, strongly supports ONL pathology in MS. Fully automated, fully manual and computer-assisted manual OCT segmentation techniques compare closely, highlighting the utility of accurate and time-efficient automated segmentation outcomes in MS clinical trials.

M. A. Seigo and E.S. Sotirchos contributed equally to this work and are co-first authors.

M. A. Seigo · E. S. Sotirchos · S. Newsome · A. Babiarz · E. Ford · S. B. Syc · J. N. Ratchford · P. A. Calabresi · S. Saidha (✉)
Department of Neurology, Johns Hopkins University School of Medicine, 600 N Wolfe Street, Pathology 627, Baltimore, MD 21287, USA
e-mail: ssaidha2@jhmi.edu; shivsaidha@physicians.ie

P. A. Calabresi
e-mail: calabresi@jhmi.edu

C. Eckstein
Department of Neurology, University of South Alabama College of Medicine, Mobile, AL, USA

J. D. Oakley
Voxeleron LLC, Pleasanton, CA, USA

T. C. Frohman · E. M. Frohman
Departments of Neurology and Ophthalmology, University of Texas Southwestern Medical Center, Dallas, TX, USA

L. J. Balcer
Departments of Neurology and Ophthalmology, University of Pennsylvania School of Medicine, Philadelphia, PA, USA

 Springer

Abbreviations

OCT	Optical coherence tomography
TD-OCT	Time domain OCT
SD-OCT	Spectral domain OCT
MS	Multiple sclerosis
RRMS	Relapsing remitting MS
SPMS	Secondary progressive MS
PPMS	Primary progressive MS
CNS	Central nervous system
ON	Optic neuritis
AMT	Average macular thickness
RNFL	Retinal nerve fiber layer
pRNFL	Peripapillary RNFL
mRNFL	Macular RNFL
GCL	Ganglion cell layer
FAS	Fully automated segmentation
FMS	Fully manual segmentation
CAMS	Computer-aided manual segmentation
HC	Healthy control
ART	Automatic real time
ILM	Inner limiting membrane
IPL	Inner plexiform layer
INL	Inner nuclear layer
OPL	Outer plexiform layer
ELM	External limiting membrane
IS/OS	Inner and outer photoreceptor segment junction
IS	Inner photoreceptor segments
BM	Bruch's membrane
GCIP	Ganglion cell layer + inner plexiform layer
ONL	Outer nuclear layer
ONL + PR	Outer nuclear layer + photoreceptor segments
RPE+	Retinal pigment epithelium
ETDRS	Early Treatment Diabetic Retinopathy Study
ICC	Intraclass correlation coefficient
LOA	Limits of agreement

Introduction

Optical coherence tomography (OCT) is a non-invasive, quantitative, and reproducible imaging technique that utilizes low-coherence, near-infrared light to generate high-resolution, cross-sectional images of tissues such as the retina [1–4]. Compared to older, third-generation time-domain OCT (TD-OCT), fourth-generation spectral-domain OCT (SD-OCT) has better axial scan resolution (4–6 vs. 8–10 μm), faster acquisition speeds (up to 60-times faster than TD-OCT), and better reproducibility [3–6]. OCT enables in vivo study and quantitative characterization of the effects of a wide variety of disease processes on the retina, an anatomically isolated and easily

accessible structure of the anterior visual pathway. There is growing interest in the utility of OCT in multiple sclerosis (MS), an inflammatory demyelinating disorder of the central nervous system (CNS) with predilection to affect the optic nerves clinically and subclinically [7, 8]. Post-mortem analyses reveal that 94–99% of MS patients have demyelinating plaques in their optic nerves, irrespective of optic neuritis (ON) history [9, 10]. Optic nerve demyelin-ation is thought to result in retrograde degeneration of its constituent fibers, leading to atrophy of the retinal nerve fiber layer (RNFL), the innermost retinal layer, from where these axons originate. Degeneration of the retinal nerve fibers in turn leads to death of retinal ganglion cells, the neurons from which these axons are derived [11].

OCT studies in MS have primarily concentrated on conventional OCT measures, namely the average peripap-illary RNFL thickness (pRNFL) and average macular thickness (AMT). pRNFL thickness primarily reflects ret-inal axonal integrity [12], whereas AMT is a non-specific composite measure of all retinal layers. In addition to ganglion cell layer (GCL) and RNFL atrophy, recently described inner and outer nuclear layer pathology in MS may also contribute to AMT reductions [13, 14].

OCT segmentation has the potential to isolate discrete axonal and neuronal retinal layers based on variability in tissue reflectivity resulting from differing layer composi-tions, enabling their quantitative assessment. To date, fully manual segmentation-derived measures of the GCL on Spectralis SD-OCT images have been assessed in a small cohort of MS patients [15]. While additional retinal layers are visually discriminable on scans from this device, there has been a paucity of further exploration of this technique in MS. Moreover, reproducibility of retinal layer thickness measures derived from this technique remain largely undetermined in MS. Fully automated Cirrus HD-OCT segmentation-derived composite measures of retinal layers have also been assessed in MS, in a more thorough, but nonetheless limited fashion [13, 14]. While measures generated with this technique demonstrate excellent reproducibility [13], this technique has not been expanded to isolate all visually discriminable retinal layers. Fur-thermore, this segmentation technique identifies thinning of the deeper retinal neuronal layers in MS, which may be reflective of primary retinal neuronopathy [13, 14]. While these findings largely agree with pathological [16] and electrophysiological [17, 18] MS data, they remain to be recapitulated with other OCT segmentation techniques, and other OCT devices. Overall, it is unclear how manual and automated methods of OCT segmentation compare to one another, within and between OCT devices.

In this study our primary objectives were (1) to deter-mine the agreement of individual and composite retinal layer thickness measurements derived from fully manual

(FMS), computer-aided manual (CAMS), and fully automated (FAS) segmentation techniques of images acquired with two commonly utilized SD-OCT devices (Cirrus HD-OCT and Spectralis OCT) in MS and healthy controls, and (2) to determine the intra-rater and inter-rater reproducibility of these same thickness measures derived from fully manual segmentation of Spectralis images. Utilizing these different segmentation techniques on both SD-OCT devices our secondary objectives were (1) to explore whether thinning of individual retinal layers can be identified in MS, and (2) to explore correlations between retinal layer thicknesses and visual function measures.

Materials and methods

Study population

The study protocol was approved by the Institutional Review Board of Johns Hopkins University, and written informed consent was obtained from all study participants. Patients were recruited from the Johns Hopkins Multiple Sclerosis Center by convenience sampling. Diagnosis of MS was confirmed by the treating neurologist (PAC), based on the McDonald criteria [19]. The MS disease subtype was classified as relapsing-remitting (RRMS), secondary progressive (SPMS), or primary progressive (PPMS). Healthy controls (HCs) were recruited from among Johns Hopkins staff. Individuals with spherical refractive error of ±6.0 diopters, history of ocular surgery, glaucoma, hypertension, diabetes, or any other apparent ocular pathology were excluded from the study. Patients within 3 months of acute ON were also excluded to minimize the confounding effect of ON-associated RNFL edema on OCT measurements.

OCT

Retinal imaging was performed by experienced technicians on Cirrus HD-OCT model 4000, software version 5.0 (Carl Zeiss Meditec, Dublin, CA, USA) and Spectralis OCT, software version 5.2.4 (Heidelberg Engineering, Heidelberg, Germany), as described in detail elsewhere [4, 20]. Briefly, Cirrus macular and optic disc data were obtained using the macular cube 512 × 128 and optic disc cube 200 × 200 protocols, respectively. Only scans with signal strengths ≥7 and without artifact were included in the study. Spectralis macular scans were obtained using the fast macular protocol. Spectralis macular scans included in this study had an automatic real time (ART) of 16, signal strength ≥20 dB, and were devoid of artifact. Spectralis optic disc data were obtained using the RNFL-N protocol, with an ART of 100. Cirrus and Spectralis scans were obtained in random order on the same day.
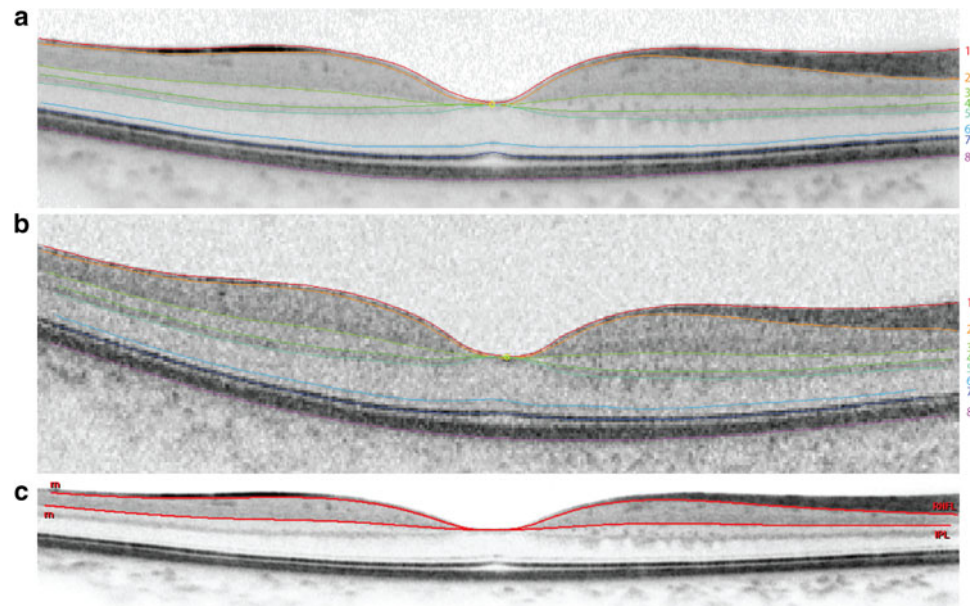
OCT segmentation

The central horizontal macular B-scan (traversing the fovea) of participants' right eyes was manually segmented. Computer-aided manual segmentation (CAMS) was performed on both Spectralis and Cirrus scans. Fully manual segmentation (FMS) was performed only on Spectralis scans due to software constraints. All manual segmentations were performed by the same segmenter, unless stated otherwise. Segmenters were blinded to disease status. Fully automated three-dimensional segmentation (FAS) was also performed on Cirrus macular cube scans (automated segmentation software for Spectralis scans is unavailable at our institution). FAS measures average thicknesses of retinal layers in an annulus of inner radius 0.54 mm and outer radius 2.4 mm, centered on the fovea, providing measures derived from relatively global macular sampling, compared to CAMS and FMS. To segment the target layers of interest on a single horizontal scan by CAMS and FMS takes approximately 20–30 min and 30–40 min, respectively. Given their time-intensiveness, CAMS and FMS segmentation of all constituent macular B scans was considered impractical.

As described in detail elsewhere [21, 22], CAMS, written in MATLABv7.4 (Mathworks, Natick, MA), employs a "point-and-click" method in which the segmenter places points along retinal layer borders of interest. A sophisticated spline algorithm connects the points in a smooth and continuous fashion, and the generated boundary may be toggled on and off while the segmenter adjusts points until the border is satisfactorily defined.

In FMS, the segmenter visually identifies the layer boundaries and adds points along an interpolated line that adjusts as each point is placed. Generated boundary lines may not be toggled on and off with FMS. The segmenter may magnify the image at the expense of image quality. All CAMS and FMS boundary identifications were reviewed by an expert physician (SS) to ensure anatomical accuracy. Thicknesses of the layers of interest were calculated across a macular width of 5.0 mm, centered on the fovea, for both CAMS and FMS.

On CAMS and FMS, eight borders were defined (Fig. 1): (1) the inner limiting membrane (ILM); (2) the outer boundary of the macular RNFL (mRNFL); (3) the outer boundary of the inner plexiform layer (IPL); (4) the outer boundary of the inner nuclear layer (INL); (5) the outer boundary of the outer plexiform layer (OPL); (6) the external limiting membrane (ELM); (7) the junction between the inner and outer photoreceptor segments (IS/OS); and (8) Bruch's membrane (BM). Using these borders, the thicknesses of the following individual and composite layers were defined: 1–2 mRNFL; 2–3 GCL + IPL (GCIP); 3–4 INL; 3–5 INL + OPL; 5–6 outer

Fig. 1 Example of OCT segmentation techniques. **a** and **b**: Spectralis CAMS (**a**) and Cirrus CAMS (**b**) borders: (*1*) inner limiting membrane; (*2*) outer boundary of retinal nerve fiber layer; (*3*) outer boundary of inner plexiform layer; (*4*) outer boundary of inner nuclear layer; (*5*) outer boundary of outer plexiform layer; (*6*) external limiting membrane; (*7*) inner/outer photoreceptor segment junction; (*8*) Bruch's membrane. *Layers* 1–2 mRNFL; 2–3 GCIP; 3–4 INL; 3–5 INL + OPL; 5–6 ONL; 6–7 IS; 5–8 PR; 7–8 RPE+. **c** Spectralis FMS borders example: (RNFL) outer boundary of retinal nerve fiber layer and (IPL) outer boundary of inner plexiform layer

nuclear layer (ONL); 6–7 inner photoreceptor segments (IS); 5–8 ONL + photoreceptor segments (ONL + PR); and 7–8 retinal pigment epithelium (RPE+).

Cirrus FAS has been described in detail by our group elsewhere [13, 14]. Briefly, segmentation software automatically identifies the ILM, the outer boundaries of the RNFL, IPL, and OPL, and the inner boundary of the RPE, yielding average thickness measurements for (1) mRNFL; (2) GCIP; (3) INL + OPL; and (4) ONL + PR (by this protocol defined as being between the outer boundary of the OPL and the inner boundary of the RPE).

Visual acuity

Monocular visual function testing was performed using retro-illuminated eye charts in a darkened room prior to OCT examination. High-contrast (100%) Early Treatment Diabetic Retinopathy Study (ETDRS) charts (at 4 m) and low-contrast (2.5 and 1.25%) Sloan Letter Charts (at 2 m) were used. The total number of correct letters identified on each chart was recorded (maximum score of 70 letters) to determine letter-acuity scores for each contrast level.

Statistical analysis

Statistical analyses were performed using Stata 11 (Stata-Corp, College Station, TX) and SPSS 19 (IBM, Armonk, NY). Only right eyes were included in analyses, since CAMS and FMS were only performed on right eyes (due to the time-intensiveness of these procedures). Bland-Altman analyses and interscanner agreement indices were used to evaluate agreement between segmentation methods [23].

Interscanner agreement indices have been previously utilized to quantify interscanner variation between different MRI machines and conventional OCT measures between OCT devices [20, 24, 25]. If x_a is the measurement on machine a, and x_b is the measurement on machine b, the interscanner agreement index is defined as:

$$\text{Interscanner agreement index} = 1 - \frac{|x_a - x_b|}{(x_a + x_b)/2}$$

Pearson's correlation was used to assess correlation of CAMS and FMS measures with FAS measures. Intra-rater and inter-rater reproducibilities were assessed with intra-class correlation coefficients (ICC) using a two-way model for absolute agreement. The ICC is used to assess the consistency of measurements made by multiple observers (or repeated measurements by the same observer). A higher ICC (maximum: 1.0) represents better agreement. In general, an ICC with a lower limit of the 95% confidence interval ≥0.75 is considered to be consistent with excellent reproducibility [26].

Exploratory comparisons of retinal layer thickness measures between MS patients and healthy controls were performed with multivariate linear regression, adjusting for age and sex. Comparisons among MS patients were also adjusted for disease duration. As visual function data were not normally distributed, monocular letter-acuity score comparisons were performed with the Wilcoxon rank sum test, and correlations of retinal layer thicknesses with letter-acuity were assessed utilizing Spearman's rank correlation. Statistical significance was defined as $p < 0.05$. Correction for multiple comparisons was not performed, as examined variables were related, and therefore likely to be correlated.

Results

Study population

Eighty-two participants [52 MS patients (47 RRMS, 2 SPMS, 3 PPMS), mean age = 40.6, range = 22–68] and 30 healthy controls (mean age = 39.5, range = 23–56) were included in the study. There were no significant differences in age or sex ratios between groups (Table 1).

Table 1 Summary of demographics and clinical data

	MS	HC
Subjects, *n* (%)	52	30
RRMS	47 (90.4)	
SPMS	2 (3.8)	
PPMS	3 (5.8)	
Age (SD)	40.6 (10.0)	39.5 (9.9)
Range	22–68	23–56
Sex, *n* (%)		
Male	9 (17)	5 (17)
Female	43 (83)	25 (83)
Disease duration (years) (SD)	9 (8.0)	
History of optic neuritis, *n* (%)	16 (30.8)	
Letters read correctly (SD)[a]		
100% contrast	57.1 (5.3)	59.4 (6.2)
2.5% contrast	25.8 (11.9)	31.5 (9.9)
1.25% contrast	14.2 (10.3)	19.5 (11.3)
Peripapillary RNFL thickness (μm)		
Cirrus HD-OCT	89.4 (10.3)	94.4 (10.9)
Spectralis OCT	88.6 (9.1)	96.7 (10.1)

RRMS relapsing remitting multiple sclerosis, *SPMS* secondary progressive multiple sclerosis, *PPMS* primary progressive multiple sclerosis, *SD* standard deviation, *RNFL* retinal nerve fiber layer

[a] Visual acuity data were obtained for *n* = 42 MS patients and *n* = 20 HCs

OCT segmentation comparisons

Bland-Altman analyses revealed that, in general, manual intra-retinal macular segmentation performed with the same segmentation technique (CAMS) on both OCT devices (i.e., comparison of Cirrus CAMS and Spectralis CAMS) yielded the lowest mean differences (an indicator of agreement across the cohort) between calculated thickness measures (mean differences in μm: mRNFL = 1.94; GCIP = 0.76; INL + OPL = 0.36; ONL = −0.03; ONL + PR = −1.29; RPE+ = 0.17) (Table 2). Considering the average thicknesses of these layers (Table 3), all except the mRNFL agreed excellently across the cohort. Similarly, with the exception of the mRNFL, agreement across the cohort was excellent for these layers when comparing Spectralis CAMS (Table 4) and Cirrus CAMS (Supplementary Table 1) to Spectralis FMS. Median interscanner agreement indices (Supplementary Table 2) for mRNFL thicknesses were suboptimal (71.07–88.22%), while median interscanner agreement indices for all other layer thicknesses were excellent (91.72–98.45%) for all comparisons of manual segmentation methods on both devices. These results support excellent agreement across the cohort for these measures. Bland-Altman analyses also revealed that, in general, segmentation of scans from the same OCT device (Spectralis), though with different techniques (CAMS and FMS), yielded the narrowest limits of agreement (LOA) (an indicator of agreement at an individual subject level) between measured thicknesses (LOA range in μm: GCIP = 8.46; INL + OPL = 10.58; ONL + PR = 9.76; RPE+ = 8.39). Considering the average thicknesses of these layers, GCIP and ONL + PR agreed well at the individual level, and also maintained reasonable agreement at the individual level when comparing Cirrus CAMS to Spectralis CAMS and Cirrus CAMS to Spectralis FMS. Bland-Altman analyses comparing thickness measures derived from different manual segmentation techniques on scans from different OCT devices (i.e., CAMS on Cirrus and FMS on Spectralis) in general showed the highest mean differences

Table 2 Bland-Altman analysis: Spectralis CAMS versus Cirrus CAMS

Layer (*n* = 82)	Mean difference (95% CI)	Lower LOA (95% CI)	Upper LOA (95% CI)	Range
mRNFL	1.94 (1.50 to 2.38)	−2.04 (−2.81 to −1.28)	5.93 (5.16 to 6.69)	7.97
GCIP	0.76 (0.18 to 1.35)	−4.54 (−5.55 to −3.53)	6.06 (5.05 to 7.08)	10.60
INL	−1.31 (−2.15 to −0.46)	−8.94 (−10.40 to −7.48)	6.33 (4.87 to 7.79)	15.28
INL + OPL	0.36 (−0.29 to 1.02)	−5.54 (−6.67 to −4.41)	6.27 (5.14 to 7.40)	11.81
ONL	−0.03 (−0.65 to 0.60)	−5.68 (−6.76 to −4.60)	5.63 (4.55 to 6.71)	11.31
ONL + PR	−1.29 (−2.09 to −0.50)	−8.52 (−9.91 to −7.14)	5.94 (4.55 to 7.32)	14.46
IS	−1.44 (−1.76 to −1.12)	−4.32 (−4.88 to −3.77)	1.44 (0.89 to 1.99)	5.76
RPE+	0.17 (−0.60 to 0.95)	−6.85 (−8.19 to −5.50)	7.20 (5.85 to 8.54)	14.04

CAMS computer-aided manual segmentation, *CI* confidence interval, *LOA* limit of agreement, *mRNFL* macular retinal nerve fiber layer, *GCIP* ganglion cell + inner plexiform layer, *INL* inner nuclear layer, *OPL* outer plexiform layer, *ONL* outer nuclear layer, *PR* photoreceptor segments, *IS* inner photoreceptor segments, *RPE* retinal pigment epithelium

Table 3 Mean retinal layer thicknesses, µm

	Spectralis CAMS	Cirrus CAMS	Spectralis FMS	Cirrus FAS
mRNFL				
HCs (SD)	19.1 (2.8)	16.7 (2.1)	22.5 (2.2)	32.7 (3.5)
MS (SD)	17.2 (2.6)	15.5 (2.1)	20.4 (2.2)	28.9 (4.4)
GCIP				
HCs (SD)	70.7 (5.9)	69.6 (5.9)	68.6 (6.3)	81.5 (7.1)
MS (SD)	62.0 (8.9)	61.4 (8.7)	59.3 (8.1)	71.3 (7.9)
INL				
HCs (SD)	33.8 (3.4)	34.8 (3.7)	33.8 (3.5)	–
MS (SD)	32.9 (3.4)	34.3 (3.9)	34.0 (4.5)	–
INL + OPL				
HCs (SD)	54.3 (3.7)	54.2 (4.5)	54.5 (4.4)	63.8 (4.6)
MS (SD)	54.2 (4.5)	53.7 (5.3)	55.2 (4.7)	64.3 (4.5)
ONL				
HCs (SD)	82.3 (7.2)	82.3 (7.3)	80.7 (7.0)	–
MS (SD)	76.5 (8.8)	76.6 (8.6)	75.2 (8.7)	–
ONL + PR				
HCs (SD)	165.6 (8.4)	168.6 (9.1)	163.3 (8.8)	120.8 (10.1)
MS (SD)	158.6 (10.3)	158.9 (11.0)	156.4 (10.2)	116.5 (9.0)
IS				
HCs (SD)	27.5 (1.5)	29.2 (1.5)	27.4 (1.8)	–
MS (SD)	26.6 (1.3)	27.9 (1.6)	26.6 (1.9)	–
RPE+				
HCs (SD)	55.7 (2.9)	57.0 (4.4)	55.4 (2.7)	–
MS (SD)	55.4 (3.4)	54.4 (3.5)	54.5 (2.8)	–

CAMS computer-aided manual segmentation, *FMS* fully manual segmentation, *FAS* fully automated segmentation, *SD* standard deviation, *mRNFL* macular retinal nerve fiber layer, *GCIP* ganglion cell + inner plexiform layer, *INL* inner nuclear layer, *OPL* outer plexiform layer, *ONL* outer nuclear layer, *PR* photoreceptor segments, *IS* inner photoreceptor segments, *RPE* retinal pigment epithelium

Table 4 Bland–Altman analysis: Spectralis CAMS versus Spectralis FMS

Layer ($n = 82$)	Mean difference (95% CI)	Lower LOA (95% CI)	Upper LOA (95% CI)	Range
mRNFL	−3.30 (−3.77 to −2.82)	−7.60 (−8.42 to −6.78)	1.00 (0.18 to 1.83)	8.61
GCIP	2.43 (1.96 to 2.90)	−1.80 (−2.61 to −0.99)	6.66 (5.85 to 7.47)	8.46
INL	−0.87 (−1.66 to −0.08)	−8.05 (−9.43 to −6.68)	6.31 (4.94 to 7.69)	14.37
INL + OPL	−0.68 (−1.26 to −0.09)	−5.97 (−6.98 to −4.95)	4.62 (3.60 to 5.63)	10.58
ONL	1.43 (0.77 to 2.09)	−4.53 (−5.66 to −3.39)	7.38 (6.24 to 8.52)	11.91
ONL + PR	2.22 (1.68 to 2.76)	−2.66 (−3.59 to −1.72)	7.10 (6.17 to 8.04)	9.76
IS	0.06 (−0.27 to 0.40)	−2.97 (−3.55 to −2.39)	3.09 (2.51 to 3.67)	6.06
RPE+	0.73 (0.27 to 1.20)	−3.46 (−4.26 to −2.66)	4.93 (4.13 to 5.73)	8.39

CAMS computer-aided manual segmentation, *FMS* fully manual segmentation, *CI* confidence interval, *LOA* limit of agreement, *mRNFL* macular retinal nerve fiber layer, *GCIP* ganglion cell + inner plexiform layer, *INL* inner nuclear layer, *OPL* outer plexiform layer, *ONL* outer nuclear layer, *PR* photoreceptor segments, *IS* inner photoreceptor segments, *RPE* retinal pigment epithelium

and widest LOA. Of additional note, for all comparisons, LOAs for the INL were among the widest, despite being one of the thinnest layers defined, suggesting poor reproducibility of this segmentation parameter. Analysis of the composite measurement, INL + OPL, resulted in narrower LOA and improved interscanner agreement indices, relative to the INL alone. Bland-Altman analyses and interscanner agreement indices were similar for HCs and MS patients. Bland-Altman plots for the GCIP and the mRNFL are shown in Fig. 2.

Due to greater macular sampling by FAS than CAMS or FMS, Pearson's correlations with FAS were assessed rather than Bland-Altman analyses. Cirrus FAS-derived measurements correlated with thickness measurements derived from Spectralis CAMS (r values: mRNFL = 0.64; GCIP = 0.87; INL + OPL = 0.68; ONL + PR = 0.73; $p < 0.0001$ for all), Cirrus CAMS (mRNFL = 0.65; GCIP = 0.84; INL + OPL = 0.67; ONL + PR = 0.75; $p < 0.0001$ for all), and Spectralis FMS (mRNFL = 0.64;
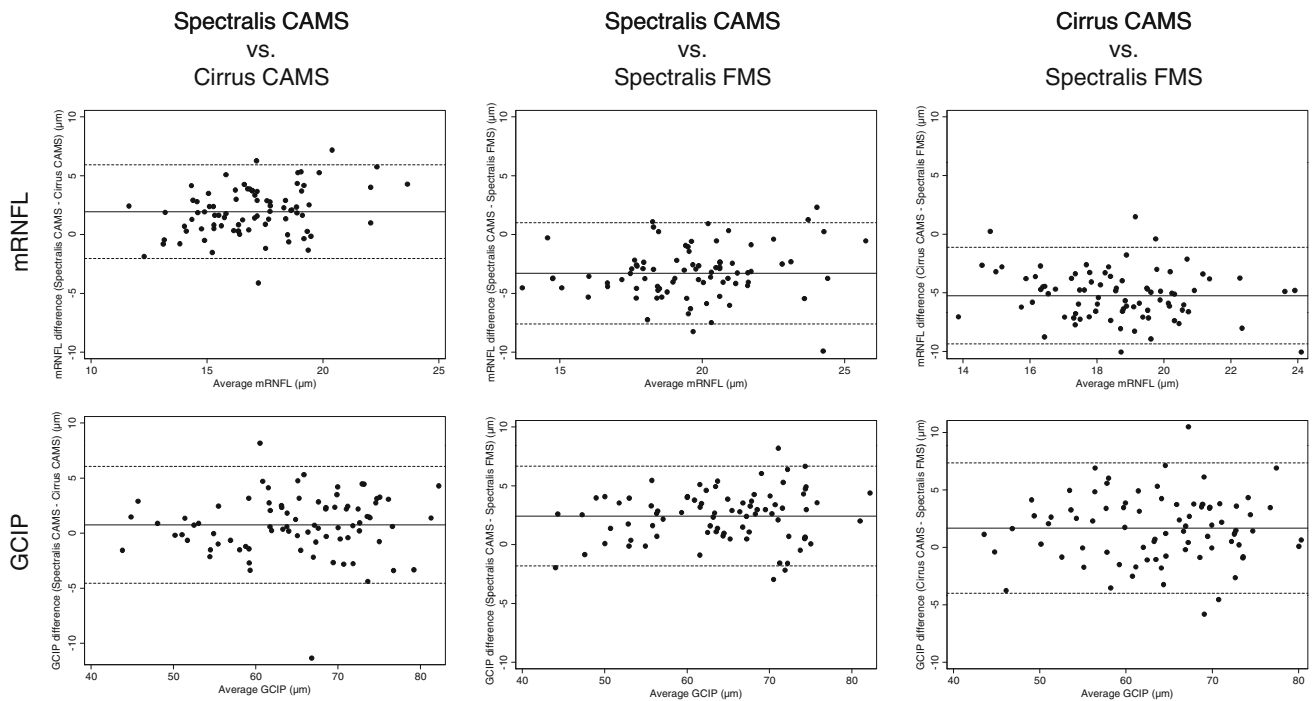
Fig. 2 Bland-Altman plots. *CAMS* computer-aided manual segmentation, *FMS* fully manual segmentation, *GCIP* ganglion cell + inner plexiform layer, *mRNFL* macular retinal nerve fiber layer

GCIP = 0.85; INL + OPL = 0.60; ONL + PR = 0.73; $p < 0.0001$ for all).

To determine the inter-rater reproducibility of Spectralis FMS, two individuals segmented 30 HC and 52 MS scans, independently and in random order. Intraclass correlation coefficients (ICCs) were high for GCIP (0.948), ONL (0.935), and ONL + PR (0.954) thickness measures, consistent with excellent inter-rater reproducibility. ICCs were low for mRNFL (0.635), INL (0.510), and IS (0.476) thickness measures, indicating poor inter-rater reproducibility. The ICC for the composite measure INL + OPL was higher (0.712) than that for the INL alone. Both segmenters also independently re-segmented 16 randomly chosen scans (8 HCs, 8 MS) to determine the intra-rater reproducibility of Spectralis FMS-derived thickness measures. The intra-rater ICCs of both segmenters were similar to the inter-rater ICCs. Intra-rater and inter-rater ICCs were similar for HCs and MS patients. ICCs are summarized in Table 5.

Exploratory comparisons of OCT segmentation measures between MS subjects and healthy controls

mRNFL, GCIP, and ONL + PR thicknesses were significantly decreased in MS compared to HCs across segmentation techniques and devices (Table 6). Additionally, Cirrus CAMS, Spectralis CAMS, and Spectralis FMS-derived ONL measures were also significantly reduced in

MS compared to HCs. Subgroup comparisons by MS disease subtype were not performed because of the small number of SPMS and PPMS patients in the study. These individual and composite layer thickness measures remained significantly lower in MS compared to HCs, when patients with ON history ($n = 16$) were excluded from analyses (Supplementary Table 3).

Additional exploratory comparisons were also performed between MS eyes, with and without ON history (Supplementary Table 4). Although ON history was associated with lower GCIP thicknesses, this was not significant for all segmentation methods studied (Cirrus FAS: $p = 0.02$; Cirrus CAMS: $p = 0.056$; Spectralis CAMS: $p = 0.024$; Spectralis FMS: $p = 0.052$). Also, mRNFL thinning was significant in patients with ON history for the Cirrus FAS-derived mRNFL measure ($p = 0.008$). It is important to interpret these results with caution given the low number of eyes with and without ON history compared.

Exploratory visual function correlations

As monocular letter acuity at 100, 2.5, and 1.25% contrast was only recorded in 20 HCs and 42 MS patients, it is important to interpret these results with caution; 2.5 and 1.25% low-contrast letter-acuity scores for MS eyes with ON history were significantly lower than those for HC eyes (2.5%: $p = 0.037$; 1.25%: $p = 0.024$). Within the MS group, Spearman's rank correlations between visual acuity

Table 5 Inter- and intra-rater intraclass correlation coefficients for Spectralis FMS

Layer	Segmenter A versus Segmenter B[a] (95% CI)	Segmenter A versus Segmenter A[b] (95% CI)	Segmenter B versus Segmenter B[b] (95% CI)
mRNFL	0.635 (0.486 to 0.748)	0.744 (0.417 to 0.902)	0.885 (0.701 to 0.958)
GCIP	0.948 (0.921 to 0.966)	0.973 (0.925 to 0.990)	0.910 (0.751 to 0.968)
INL + OPL	0.712 (0.586 to 0.804)	0.892 (0.724 to 0.961)	0.700 (0.282 to 0.888)
INL	0.510 (0.324 to 0.657)	0.412 (−0.030 to 0.736)	0.402 (−0.119 to 0.744)
ONL	0.935 (0.900 to 0.957)	0.902 (0.749 to 0.964)	0.970 (0.912 to 0.989)
ONL + PR	0.954 (0.927 to 0.971)	0.961 (0.892 to 0.986)	0.978 (0.939 to 0.992)
IS	0.476 (0.076 to 0.703)	0.606 (0.159 to 0.843)	0.765 (0.380 to 0.916)
RPE+	0.891 (0.829 to 0.931)	0.852 (0.566 to 0.949)	0.907 (0.759 to 0.966)

FMS fully manual segmentation, *CI* confidence interval, *mRNFL* macular retinal nerve fiber layer, *GCIP* ganglion cell + inner plexiform layer, *INL* inner nuclear layer, *OPL* outer plexiform layer, *ONL* outer nuclear layer, *PR* photoreceptor segments, *IS* inner photoreceptor segments, *RPE* retinal pigment epithelium

[a] $n = 82$ (52 MS, 30 HCs)

[b] $n = 16$ (8 MS, 8 HCs)

and segmentation measures were assessed (Table 7). mRNFL thicknesses correlated significantly with visual scores at all contrast levels across segmentation methods. GCIP thickness correlations with 2.5 and 1.25% low-contrast letter-acuity were significant for all segmentation methods, but with 100% high-contrast letter-acuity, only Cirrus FAS GCIP measurements achieved significant correlation. Consistent with recent studies [14], correlation coefficients were much higher between GCIP thicknesses and 2.5 and 1.25% low-contrast letter acuity than pRNFL thicknesses. Also, correlations were stronger between GCIP thicknesses and 2.5 and 1.25% low-contrast letter acuity than mRNFL thicknesses, except for the Cirrus FAS-derived mRNFL measure, which had comparable correlation coefficients to the GCIP. Cirrus FAS-derived GCIP and mRNFL thicknesses exhibited the strongest correlations with visual function among the GCIP and mRNFL segmentation measures acquired. Other retinal layers did not correlate with visual function.

Discussion

Our results indicate that GCIP, INL + OPL, ONL, and ONL + PR thickness measures of Cirrus and Spectralis macular scans agree excellently across MS and HC cohorts, irrespective of the manual OCT segmentation technique utilized. Moreover, GCIP and ONL + PR thickness measures appear to agree well at the individual level. In addition, Spectralis FMS-derived GCIP, ONL, and ONL + PR thickness measures in this study demonstrated excellent intra-rater and inter-rater reproducibility, whereas reproducibility of mRFNL measures was poor with this technique. Since superiority of a particular OCT segmentation method is difficult to establish due to lack of a "gold

standard" to which to compare, the reliability of the segmentation method used is of utmost importance. CAMS and FAS have also been previously shown to have excellent reproducibility, although for both techniques, a more limited number of retinal layers was assessed (FAS ICCs: 0.91–0.99; CAMS concordance correlation coefficients: 0.98–0.99) [13, 22]. Furthermore, exploratory analyses from this study suggest visual dysfunction in MS may correlate better with thickness measures of GCIP derived from all segmentation techniques utilized (on both OCT devices) than pRNFL, consistent with prior studies of Cirrus FAS in MS [14]. All of these findings add further support for the potential use of GCIP thickness as an outcome measure in trials of putative neuroprotective and/or remyelinating agents in MS.

Across segmentation techniques and devices, exploratory analyses revealed mRNFL, GCIP, and ONL + PR thicknesses were significantly reduced in MS compared to HCs, irrespective of ON history. Similarly, measures of the ONL (where the photoreceptor cell-bodies are located) were significantly reduced across manual OCT segmentation techniques in MS (this measure is not currently calculated by Cirrus FAS). Prior studies of Cirrus FAS demonstrate prominent GCIP thinning in MS eyes [13, 14], with and without ON history, in accordance with post-mortem data [16, 27]. Results of the current study suggest this thinning in MS is detectable across OCT devices and OCT segmentation techniques. Moreover, prior studies of Cirrus FAS demonstrate ONL + PR thinning in MS [13, 14]. Although ONL/ONL + PR pathology has not been demonstrated pathologically in MS eyes, electroretinographic studies suggest that ONL pathology occurs in MS [17, 18]. Results of the current study strongly support structural abnormalities of the photoreceptor segments and their cell bodies in MS, with exploratory analyses

Table 6 Comparison of retinal layer thickness measures of MS (n = 52) and HCs (n = 30)

Layer	Spectralis CAMS mean difference, µm (95% CI)	P value	Cirrus CAMS mean difference, µm (95% CI)	P value	Spectralis FMS mean difference, µm (95% CI)	P value	Cirrus FAS mean difference, µm (95% CI)	P value
mRNFL	−2.1 (−3.3 to −0.9)	0.001	−1.4 (−2.4 to −0.5)	0.003	−2.1 (−3.0 to −1.1)	<0.001	−3.8 (−5.7 to −1.9)	<0.001
GCIP	−9.2 (−12.7 to −5.7)	<0.001	−8.8 (−12.1 to −5.4)	<0.001	−10.0 (−13.2 to −6.8)	<0.001	−10.8 (−14.3 to −7.4)	<0.001
INL	−1.1 (−2.6 to 0.3)	0.120	−0.7 (−2.4 to −1.0)	0.430	0.4 (−1.4 to 2.2)	0.650	–	–
INL + OPL	−0.1 (−2.0 to 1.8)	0.920	−0.47 (−2.8 to −1.8)	0.690	0.6 (−1.5 to 2.6)	0.590	−0.05 (−2.2 to 2.1)	0.960
ONL	−5.9 (−9.5 to −2.3)	0.002	−5.7 (−9.3 to −2.2)	0.002	−5.6 (−9.2 to −1.9)	0.002	–	–
ONL + PR	−7.1 (−11.4 to −2.8)	0.002	−9.7 (−14.3 to −5.2)	<0.001	−7.0 (−11.3 to −2.7)	<0.001	−4.4 (−8.4 to −0.3)	0.035
RPE+	−0.3 (−1.8 to 1.2)	0.670	−2.7 (−4.5 to −0.9)	0.003	−0.8 (−2.1 to 0.5)	0.220	–	–

Multivariate linear regression was used to adjust for age and sex. Presented mean differences are adjusted for age and sex

CAMS computer-aided manual segmentation, *FMS* fully manual segmentation, *FAS* fully automated segmentation, *CI* confidence interval, *mRNFL* macular retinal nerve fiber layer, *GCIP* ganglion cell + inner plexiform layer, *INL* inner nuclear layer, *OPL* outer plexiform layer, *ONL* outer nuclear layer, *PR* photoreceptor segments, *IS* inner photoreceptor segments, *RPE* retinal pigment epithelium

Table 7 Correlation of letter acuity (100, 2.5, and 1.25% contrast) and retinal layer thickness measures

	100% LA rho	P value	2.5% LA rho	P value	1.25% LA rho	P value
GCIP						
Cirrus FAS	0.34	0.03	0.62	<0.001	0.62	<0.001
Cirrus CAMS	0.22	0.16	0.48	<0.001	0.50	<0.001
Spectralis CAMS	0.27	0.08	0.50	<0.001	0.52	<0.001
Spectralis FMS	0.28	0.07	0.53	<0.001	0.54	<0.001
mRNFL						
Cirrus FAS	0.39	0.01	0.63	<0.001	0.64	<0.001
Cirrus CAMS	0.31	0.045	0.41	0.008	0.36	0.02
Spectralis CAMS	0.36	0.02	0.33	0.03	0.33	0.03
Spectralis FMS	0.39	0.01	0.38	0.01	0.33	0.03
pRNFL						
Cirrus HD-OCT	0.09	0.57	0.24	0.13	0.27	0.08
Spectralis OCT	0.20	0.24	0.31	0.06	0.33	0.04

Spearman's rank correlation was used to determine correlations of retinal layer thickness measures and letter acuity in MS patients

FAS fully automated segmentation, *CAMS* computer-aided manual segmentation, *FMS* fully manual manual segmentation, *GCIP* ganglion cell + inner plexiform layer, *mRNFL* macular retinal nerve fiber layer, *pRNFL* peripapillary retinal nerve fiber layer, *LA* letter acuity

demonstrating quantitative abnormalities of this layer across OCT segmentation techniques and devices.

Bland-Altman analysis is used to determine agreement between different measurement methods [23]. The mean difference between methods indicates agreement on average across the cohort, whereas LOA indicates agreement at an individual level. Interpretation of Bland-Altman analyses requires taking into account the magnitude of expected differences between clinical groups, as well as the size of the mean differences and LOA ranges relative to examined measures. Although the mRNFL was the thinnest layer identified, it had high mean differences and wide LOA across manual segmentation techniques and devices, indicating poor agreement across the cohort and at an individual level. The GCIP, despite being on average more than three times thicker than the mRNFL, had narrow LOA and low mean differences, suggesting GCIP thickness measures agree well at the cohort and individual levels. Interscanner agreement indices provide additional means to assess

average agreement between measures and corroborated these findings. Our results indicate the OCT segmentation method may be as important as the OCT platform from which segmented scans are acquired. One might anticipate Spectralis FMS and Spectralis CAMS to agree best, and while LOA for GCIP, INL + OPL, and ONL + PR were narrowest for this comparison, mean differences for the majority of measures were actually lowest when the same manual segmentation technique (CAMS) was used on Cirrus and Spectralis. In general, our results suggest that future trials could potentially utilize both Cirrus and Spectralis devices for the determination of GCIP thicknesses (similar to the co-utilization of different MRI platforms in MS trials), but that patients should be scanned on the same device throughout the study, and ideally the same segmentation technique should be used across devices.

Exploratory analyses revealed GCIP thinning across OCT segmentation techniques and devices, irrespective of ON history. Furthermore, GCIP thickness measures exhibited stronger correlation than pRNFL with low-contrast letter acuity, a validated clinical measure that has been used as a clinical trial outcome measure in MS and is currently being employed in ongoing clinical trials [28, 29]. Also, the GCIP had better structure–function correlations (with low-contrast vision) than the mRNFL, with the exception of the Cirrus FAS-derived mRNFL measure, which seemed to be comparable to the GCIP. This may relate to the poorer reproducibility of manual segmentation-derived mRNFL measures observed in this study. Additionally, similar to GCIP, mRNFL measures exhibited stronger correlations with low-contrast letter acuity than the conventional pRNFL. This may be because the mRNFL is a more specific measure of macular axons crucial for visual function, whereas the pRNFL is a global measure of retinal axonal integrity. Cirrus FAS-derived GCIP and mRNFL measures exhibited the strongest correlations with low-contrast letter acuity, as compared to the manual segmentation-derived GCIP and mRNFL measures. A plausible explanation for this may be that manual segmentation was performed on only a single central horizontal scan traversing the fovea (the RNFL and GCL are devoid at the fovea) [30], whereas FAS measures the GCIP and mRNFL thicknesses of the neuronally enriched area surrounding the fovea, consequently sampling a larger area with a higher concentration of ganglion cells and their axons. Only a single horizontal scan line was segmented by both manual segmentation techniques (FMS and CAMS) due to the time-intensiveness of these procedures. All the FMS and CAMS procedures performed in this study took over 150 h. Although preferable, averaging manual segmentation-derived measures (FMS and CAMS) of multiple macular scans of the same eye, as well as manually segmenting further scans than those included in the study for the purpose of increasing our

sample size was considered impractical. For these reasons we consider our comparisons of OCT segmentation measures between MS subjects and HCs, as well as the correlations between OCT segmentation measures and visual function as exploratory, with the need to interpret these particular results with caution. The time-intensive nature of manual OCT segmentation techniques highlights the necessity for accurate automated OCT segmentation techniques.

This study highlights important limitations of available OCT segmentation protocols. We were unable to visually discriminate the GCL from the IPL on Spectralis and Cirrus scans, and we were thus unable to manually segment this particular border. Therefore, the composite measure of the GCL + IPL was used to represent GCL integrity. The automated segmentation method utilized in this study is also subject to this limitation, and additionally is unable to delineate the INL–OPL border. While segmentation of the INL from the OPL was performed manually, LOA and interscanner agreement indices for INL thicknesses were among the poorest. Additionally, Spectralis FMS reproducibility for INL measures was poor. Utilization of the INL + OPL composite measurement resolved these problems to some extent, suggesting that assessment of the INL + OPL may be preferable to assessment of the INL alone. *Post-mortem* studies of MS eyes demonstrate qualitative INL atrophy in 40% of MS eyes [16]. Failure to detect INL and/or INL + OPL thinning in our MS cohort may relate to the combination of poorer reproducibility for these measures and low sample size. Development of higher resolution OCT technology and more sophisticated segmentation techniques may overcome these limitations, enabling more precise and specific assessment of all retinal layers. Irrespective of manual segmentation technique or scan acquisition device, thickness measures for GCIP, INL + OPL, and ONL + PR correlated well with comparable thickness measures calculated by Cirrus FAS. However, as discussed, only a single horizontal scan from eyes was manually segmented, potentially representing limited and inadequate sampling of the macula.

With MS trials of remyelinating agents already initiated, the development of quantitative and reproducible outcome measures is crucial. While conventional OCT-derived pRNFL thickness measures have been shown to correlate with visual dysfunction and disease disability [31, 32], recent studies show that GCIP thickness measures (estimated with Cirrus FAS) may correlate better with disability and visual dysfunction in MS [14], suggesting GCIP may be a more suitable outcome measure in neuroprotective and/or remyelinating MS trials than conventional OCT measures. The current study demonstrates that GCIP thicknesses determined by all segmentation techniques on both Cirrus and Spectralis agree well, are reproducible, and may

correlate better with visual dysfunction in MS than pRNFL thicknesses. Results of this study also highlight that the segmentation technique used may be as important as the OCT device, and indicate the need for future automated segmentation algorithms developed for OCT devices (which may be potentially compatible in studies) to be consistent in terms of borders identified, layer thicknesses measured, and the area over which they are measured. Without this consistency, the potential for OCT segmentation in MS may be more limited. Results of this study also suggest that mRNFL, if measured with reproducible techniques such as the fully automated technique utilized in this study, may be comparable to GCIP measures and also have stronger structure–function correlations than the pRNFL in MS. Our results also indicate photoreceptor pathology in MS is detectable across OCT segmentation techniques and devices. This pathology warrants further study as it remains unclear if this represents retrograde trans-synaptic degeneration or primary retinal neuronopathy. Future studies of OCT segmentation in MS may shed further light on the neuroretinal pathobiology of this complex disorder.

References

1. Huang D, Swanson EA, Lin CP, Schuman JS, Stinson WG, Chang W, Hee MR, Flotte T, Gregory K, Puliafito CA et al (1991) Optical coherence tomography. Science 254(5035):1178–1181
2. Frohman EM, Fujimoto JG, Frohman TC, Calabresi PA, Cutter G, Balcer LJ (2008) Optical coherence tomography: a window into the mechanisms of multiple sclerosis. Nat Clin Pract Neurol 4(12):664–675
3. Cettomai D, Pulicken M, Gordon-Lipkin E, Salter A, Frohman TC, Conger A, Zhang X, Cutter G, Balcer LJ, Frohman EM, Calabresi PA (2008) Reproducibility of optical coherence tomography in multiple sclerosis. Arch Neurol 65(9):1218–1222
4. Syc SB, Warner CV, Hiremath GS, Farrell SK, Ratchford JN, Conger A, Frohman T, Cutter G, Balcer LJ, Frohman EM, Calabresi PA (2010) Reproducibility of high-resolution optical coherence tomography in multiple sclerosis. Mult Scler 16(7):829–839
5. Chen TC, Cense B, Pierce MC, Nassif N, Park BH, Yun SH, White BR, Bouma BE, Tearney GJ, de Boer JF (2005) Spectral domain optical coherence tomography: ultra-high speed, ultra-high resolution ophthalmic imaging. Arch Ophthalmol 123(12):1715–1720
6. Wolf-Schnurrbusch UE, Ceklic L, Brinkmann CK, Iliev ME, Frey M, Rothenbuehler SP, Enzmann V, Wolf S (2009) Macular thickness measurements in healthy eyes using six different optical coherence tomography instruments. Invest Ophthalmol Vis Sci 50(7):3432–3437
7. Balcer LJ (2006) Clinical practice. Optic neuritis. N Engl J Med 354(12):1273–1280
8. Sisto D, Trojano M, Vetrugno M, Trabucco T, Iliceto G, Sborgia C (2005) Subclinical visual involvement in multiple sclerosis: a study by MRI, VEPs, frequency-doubling perimetry, standard perimetry, and contrast sensitivity. Invest Ophthalmol Vis Sci 46(4):1264–1268
9. Toussaint D, Perier O, Verstappen A, Bervoets S (1983) Clinicopathological study of the visual pathways, eyes, and cerebral hemispheres in 32 cases of disseminated sclerosis. J Clin Neuroophthalmol 3(3):211–220
10. Ikuta F, Zimmerman HM (1976) Distribution of plaques in seventy autopsy cases of multiple sclerosis in the United States. Neurology 26(6 PT 2):26–28
11. Shindler KS, Ventura E, Dutt M, Rostami A (2008) Inflammatory demyelination induces axonal injury and retinal ganglion cell apoptosis in experimental optic neuritis. Exp Eye Res 87(3):208–213
12. Saidha S, Eckstein C, Ratchford JN (2010) Optical coherence tomography as a marker of axonal damage in multiple sclerosis. CML Mult Scler 2(2):33–43
13. Saidha S, Syc SB, Ibrahim MA, Eckstein C, Warner CV, Farrell SK, Oakley JD, Durbin MK, Meyer SA, Balcer LJ, Frohman EM, Rosenzweig JM, Newsome SD, Ratchford JN, Nguyen QD, Calabresi PA (2011) Primary retinal pathology in multiple sclerosis as detected by optical coherence tomography. Brain 134(Pt 2):518–533
14. Saidha S, Syc SB, Durbin MK, Eckstein C, Oakley JD, Meyer SA, Conger A, Frohman TC, Newsome S, Ratchford JN, Frohman EM, Calabresi PA (2011) Visual dysfunction in multiple sclerosis correlates better with optical coherence tomography derived estimates of macular ganglion cell layer thickness than peripapillary retinal nerve fiber layer thickness. Mult Scler J 17(12):1449–1463
15. Davies EC, Galetta KM, Sackel DJ, Talman LS, Frohman EM, Calabresi PA, Galetta SL, Balcer LJ (2011) Retinal ganglion cell layer volumetric assessment by spectral-domain optical coherence tomography in multiple sclerosis: application of a high-precision manual estimation technique. J Neuroophthalmol 31:260–264
16. Green AJ, McQuaid S, Hauser SL, Allen IV, Lyness R (2010) Ocular pathology in multiple sclerosis: retinal atrophy and inflammation irrespective of disease duration. Brain 133(Pt 6):1591–1601
17. Papakostopoulos D, Fotiou F, Hart JC, Banerji NK (1989) The electroretinogram in multiple sclerosis and demyelinating optic neuritis. Electroencephalogr Clin Neurophysiol 74(1):1–10
18. Forooghian F, Sproule M, Westall C, Gordon L, Jirawuthiworavong G, Shimazaki K, O'Connor P (2006) Electroretinographic abnormalities in multiple sclerosis: possible role for retinal autoantibodies. Doc Ophthalmol 113(2):123–132

19. Polman CH, Reingold SC, Edan G, Filippi M, Hartung HP, Kappos L, Lublin FD, Metz LM, McFarland HF, O'Connor PW, Sandberg-Wollheim M, Thompson AJ, Weinshenker BG, Wolinsky JS (2005) Diagnostic criteria for multiple sclerosis: 2005 revisions to the "McDonald Criteria". Ann Neurol 58(6):840–846

20. Warner CV, Syc SB, Stankiewicz AM, Hiremath G, Farrell SK, Crainiceanu CM, Conger A, Frohman TC, Bisker ER, Balcer LJ, Frohman EM, Calabresi PA, Saidha S (2011) The impact of utilizing different optical coherence tomography devices for clinical purposes and in multiple sclerosis trials. PLoS One 6(8): e22947

21. Hood DC, Lin CE, Lazow MA, Locke KG, Zhang X, Birch DG (2009) Thickness of receptor and post-receptor retinal layers in patients with retinitis pigmentosa measured with frequency-domain optical coherence tomography. Invest Ophthalmol Vis Sci 50(5):2328–2336

22. Hood DC, Cho J, Raza AS, Dale EA, Wang M (2011) Reliability of a computer-aided manual procedure for segmenting optical coherence tomography scans. Optom Vis Sci 88(1):113–123

23. Bland JM, Altman DG (1986) Statistical methods for assessing agreement between two methods of clinical measurement. Lancet 1(8476):307–310

24. Filippi M, van Waesberghe JH, Horsfield MA, Bressi S, Gasperini C, Yousry TA, Gawne-Cain ML, Morrissey SP, Rocca MA, Barkhof F, Lycklama a Nijeholt GJ, Bastianello S, Miller DH (1997) Interscanner variation in brain MRI lesion load measurements in MS: implications for clinical trials. Neurology 49(2):371–377

25. Filippi M, Rocca MA, Gasperini C, Sormani MP, Bastianello S, Horsfield MA, Pozzilli C, Comi G (1999) Interscanner variation in brain MR lesion load measurements in multiple sclerosis using conventional spin-echo, rapid relaxation-enhanced, and fast-FLAIR sequences. Am J Neuroradiol 20(1):133–137

26. Lee J, Koh D, Ong CN (1989) Statistical evaluation of agreement between two methods for measuring a quantitative variable. Comput Biol Med 19(1):61–70

27. Kerrison JB, Flynn T, Green WR (1994) Retinal pathologic changes in multiple sclerosis. Retina 14(5):445–451

28. Balcer LJ, Galetta SL, Calabresi PA, Confavreux C, Giovannoni G, Havrdova E, Hutchinson M, Kappos L, Lublin FD, Miller DH, O'Connor PW, Phillips JT, Polman CH, Radue EW, Rudick RA, Stuart WH, Wajgt A, Weinstock-Guttman B, Wynn DR, Lynn F, Panzara MA (2007) Natalizumab reduces visual loss in patients with relapsing multiple sclerosis. Neurology 68(16):1299–1304

29. Balcer LJ, Frohman EM (2010) Evaluating loss of visual function in multiple sclerosis as measured by low-contrast letter acuity. Neurology 74(Suppl 3):S16–S23

30. Hogan MJ, Alvarado JA, Weddell JE (1971) Histology of the human eye. W. B. Saunders, Philadelphia

31. Talman LS, Bisker ER, Sackel DJ, Long DA Jr, Galetta KM, Ratchford JN, Lile DJ, Farrell SK, Loguidice MJ, Remington G, Conger A, Frohman TC, Jacobs DA, Markowitz CE, Cutter GR, Ying GS, Dai Y, Maguire MG, Galetta SL, Frohman EM, Calabresi PA, Balcer LJ (2010) Longitudinal study of vision and retinal nerve fiber layer thickness in multiple sclerosis. Ann Neurol 67(6):749–760

32. Gordon-Lipkin E, Chodkowski B, Reich DS, Smith SA, Pulicken M, Balcer LJ, Frohman EM, Cutter G, Calabresi PA (2007) Retinal nerve fiber layer is associated with brain atrophy in multiple sclerosis. Neurology 69(16):1603–1609