PLOS ONE

RESEARCH ARTICLE

Reliability of Intra-Retinal Layer Thickness Estimates

Timm Oberwahrenbrock[1]*, Maria Weinhold[2], Janine Mikolajczak[1], Hanna Zimmermann[1], Friedemann Paul[1,3], Ingeborg Beckers[2], Alexander U. Brandt[1]

1 NeuroCure Clinical Research Center, Charité–Universitätsmedizin Berlin, Berlin, Germany, **2** Optics and Laser Technology Laboratory, Beuth University of Applied Sciences, Berlin, Germany, **3** Department of Neurology, Charité–Universitätsmedizin, Berlin, Germany

* timm.oberwahrenbrock@charite.de

CrossMark
click for updates

OPEN ACCESS

Citation: Oberwahrenbrock T, Weinhold M, Mikolajczak J, Zimmermann H, Paul F, Beckers I, et al. (2015) Reliability of Intra-Retinal Layer Thickness Estimates. PLoS ONE 10(9): e0137316. doi:10.1371/journal.pone.0137316

Editor: Ralf Andreas Linker, Friedrich-Alexander University Erlangen, GERMANY

Received: December 19, 2014

Accepted: August 15, 2015

Published: September 8, 2015

Copyright: © 2015 Oberwahrenbrock et al. This is an open access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All relevant data are within the paper and its Supporting Information files.

Funding: This study was supported by a grant from the Berlin Institute for Applied Sciences (IFAF) to IB and a grant from the German Research Society (DFG, Exc. 257) to FP. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: Friedemann Paul is a PLOS ONE Editorial Board member. This does not alter the authors' adherence to all the PLOS ONE policies on

Abstract

Purpose

Measurement of intra-retinal layer thickness using optical coherence tomography (OCT) has become increasingly prominent in multiple sclerosis (MS) research. Nevertheless, the approaches used for determining the mean layer thicknesses vary greatly. Insufficient data exist on the reliability of different thickness estimates, which is crucial for their application in clinical studies. This study addresses this lack by evaluating the repeatability of different thickness estimates.

Methods

Studies that used intra-retinal layer segmentation of macular OCT scans in patients with MS were retrieved from PubMed. To investigate the repeatability of previously applied layer estimation approaches, we generated datasets of repeating measurements of 15 healthy subjects and 13 multiple sclerosis patients using two OCT devices (Cirrus HD-OCT and Spectralis SD-OCT). We calculated each thickness estimate in each repeated session and analyzed repeatability using intra-class correlation coefficients and coefficients of repeatability.

Results

We identified 27 articles, eleven of them used the Spectralis SD-OCT, nine Cirrus HD-OCT, two studies used both devices and two studies applied RTVue-100. Topcon OCT-1000, Stratus OCT and a research device were used in one study each. In the studies that used the Spectralis, ten different thickness estimates were identified, while thickness estimates of the Cirrus OCT were based on two different scan settings. In the simulation dataset, thickness estimates averaging larger areas showed an excellent repeatability for all retinal layers except the outer plexiform layer (OPL).

Conclusions

Given the good reliability, the thickness estimate of the 6mm-diameter area around the fovea should be favored when OCT is used in clinical research. Assessment of the OPL

sharing data and materials. The other authors have declared that no competing interests exist.

was weak in general and needs further investigation before OPL thickness can be used as a reliable parameter.

Introduction

Optical coherence tomography (OCT) is a widely established means of acquiring high-resolution retinal images [1]. The technique is based on low coherence interferometry, in which near-infrared light is used to generate spatial images of biological target tissue. Over the years, OCT evolved into a number of different approaches. A significant breakthrough in development was the transition from *time domain* OCT (TD-OCT), which is hampered in terms of its application in medicine by low acquisition speed and limited resolution, to *frequency-* (or *spectral-*) *domain* OCT (SD-OCT), which greatly increased imaging speed while simultaneously improving the signal-to-noise ratio [2].

Today's SD-OCT devices allow high-resolution, 3D imaging of retinal tissue. They expand the use of OCT from macroscopic qualitative interpretation towards quantitative applications. One emerging field for this quantitative OCT is the measurement of neurodegeneration and neuroinflammation in neurological disorders. As part of the central nervous system, the retina is an excellent site for monitoring neurological pathologies, from well-known and economically significant diseases like multiple sclerosis (MS) [3], Parkinson's disease [4] or Alzheimer's disease [5], to less common diseases like neuromyelitis optica [6], Friedreich's ataxia [7], Susac syndrome [8] or the spinocerebellar ataxias [9].

Recently, quantitative OCT imaging of the retina has shifted from the peripapillary retinal nerve fiber layer (RNFL) thickness as a marker of retinal neurodegeneration to distinct macular intra-retinal layer thicknesses. However, quantification of intra-retinal layer thickness requires the segmentation of OCT scans into the respective layers. Currently, this task is performed either manually, by experienced graders, semi-automatically, with manual inspectors correcting automated results, or fully automatically. Moreover, the algorithms used to automate the segmentation have been developed independently by the device manufacturers. Consequently, several post-processing algorithms exist that aim to meet this requirement as built-in or add-on software, but the lack of standardization between these limits their prospects for diagnosis and disease monitoring. Until this lack is redressed, OCT, promising as it may be, cannot be used as a reliable marker for neurodegeneration in either therapeutic trials or clinical care, as previously suggested [10].

The necessary reliability standards can be met if two current deficiencies are redressed: Firstly, the current algorithms have to be improved to ensure that all techniques discriminate reliably between different layers [11]. Secondly, the approaches used to calculate the final outcome values for the mean layer thickness differ greatly in area and amount of single measurements included for averaging. These *layer thickness estimates* need to be standardized and investigated for reliability and comparability [12]. Especially when semi-automatic segmentation is performed, which is standard practice today, researchers tend towards smaller thickness estimates with a limited number of included B-scans in order to decrease workload for manual inspection. The goal of the present study was to investigate the precision of different layer thickness estimates for intra-retinal OCT segmentation measurements under repeatability conditions. We were especially interested in studies in MS, since here the most studies using quantitative OCT have been published.

Methods

Literature Review

All research articles that included intra-retinal segmentation of macular OCT scans in MS were retrieved from the PubMed database using the following search term: "optical coherence tomography multiple sclerosis" (search performed on 01/24/2014). Only articles written in English that discussed thickness or volume estimates of distinct retinal layers or combined layers were included. Exclusion criteria were the use of retinal segmentation in OCT scans other than macular scans (i.e. scans of the optic nerve head) and if the segmentation was solely performed to determine the thickness of the whole retina (from inner limiting membrane to Bruch's membrane or retinal pigment epithelium).

Subjects

To test the repeatability of different layer thickness estimates under various conditions we generated three datasets: two control cohorts of 15 healthy subjects each underwent three consecutive examinations with Spectralis SD-OCT (measured by MW) or Cirrus HD-OCT (measured by TO) and 13 MS/CIS patients receiving two consecutive examinations with both OCT devices (measured by JM). A detailed demographic overview of the investigated cohorts can be found in Table A in S1 File. Subject inclusion criteria were best corrected visual acuity of at least 0.8 according to ETDRS charts, using the Optec 6500P measurement device (Stereo Optical Inc., Chicago, Illinois, USA). Exclusion criteria were any history of neurological (except MS) or ophthalmological diseases and a refractive error of more than ±5.0 diopters. None of the healthy subjects showed any retinal alterations that affected OCT imaging. Patients were not matched to controls but were chosen because of known retinal affections related to MS or ON in case of CIS patients. All participants gave informed, written consent.

Ethics

The study was approved by the ethics committee of the Charité–Universitätsmedizin Berlin and was conducted in accordance with the Declaration of Helsinki in its current version and applicable German laws.

Optical Coherence Tomography

OCT examinations were performed with the Spectralis SD-OCT (Heidelberg Engineering, Heidelberg, Germany; Heidelberg eye explorer version 5.7.5.0) and the Cirrus HD-OCT 4000 (Carl Zeiss Meditec, Dublin, USA; instrument software version 6.5.0.772). Both eyes of every subject were measured and included in the analysis. We determined five different Spectralis scan designs replicating the settings used in the published studies that met the inclusion criteria (Fig 1). Previous studies with Cirrus OCT made use of the two standardly implemented macula volume scans (Fig 1).

For our repeatability datasets all scan protocols were applied with at least 5 minutes between consecutive examinations. The test subjects were asked to leave the room between examinations and the OCT devices were reset to an initial position. All measurements were performed with activated eye tracker, but without the device's follow-up function.

Intra-Retinal Segmentation

To analyze repeatability of different Spectralis layer thickness estimates, initial intra-retinal segmentation was performed using the Heidelberg Eye Explorer software (Version 6.0.0.2), after which an operator (MW for controls and TO for patients) reviewed the results and occasionally

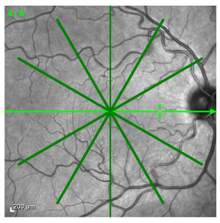
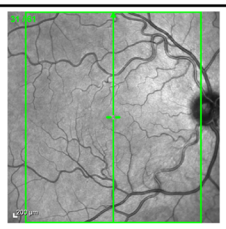
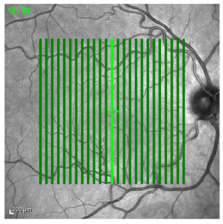
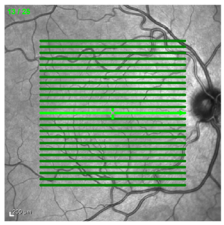
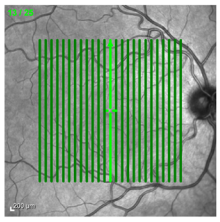
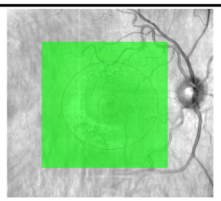
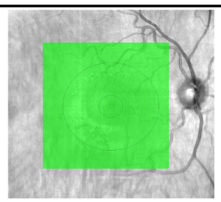
OCT device	Scan abbreviation	Fundus image with B-scans	B-scan direction	No. of B-scans	No. of A-scans per B-scan	Scanning angle	Automatic realtime
Spectralis SD-OCT (Heidelberg Engineering)	R-6-1536		Radial	6	1536 (for HC) 768 (for MS)	30°	100
Spectralis SD-OCT (Heidelberg Engineering)	V-61-768		Vertical	61	768	30° x 25°	13
Spectralis SD-OCT (Heidelberg Engineering)	V-25-512		Vertical	25	512	20° x 20°	9
Spectralis SD-OCT (Heidelberg Engineering)	H-25-512		Horizontal	25	512	20° x 20°	9
Spectralis SD-OCT (Heidelberg Engineering)	V-25-1024		Vertical	25	1024	20° x 20°	49
Cirrus HD-OCT (Carl Zeiss Meditec)	H-128-512		Horizontal	128	512	6 mm x 6 mm	NA
Cirrus HD-OCT (Carl Zeiss Meditec)	H-200-200		Horizontal	200	200	6 mm x 6 mm	NA

Fig 1. OCT scan settings used for simulation of the repeatability of different thickness estimates. The first column indicates the OCT device used to produce the scan in each row. The following column names each set of scanning parameter as a concatenation of scan direction, the number of B-scans and the number of A-scans (see columns 4–6). The third column shows for each scan design the generated OCT scans as lines or boxes on exemplary fundus images. The heights and widths of the scan areas are given in column 7, followed by the number of averaging frames in automatic real-time mode in the last column (only applicable for the Spectralis SD-OCT scan settings). NA = not applicable.

doi:10.1371/journal.pone.0137316.g001

corrected the segmentation lines where necessary (semi-automatic segmentation). Segmentation line correction was performed individually for each scan without using any other scan for reference. Fig 2A shows an exemplary B-scan from a Spectralis SD-OCT including the segmentation lines used to delineate the following retinal layers: RNFL, ganglion cell layer (GCL), inner plexiform layer (IPL), inner nuclear layer (INL), outer plexiform layer (OPL) and outer nuclear layer (ONL). As previous works have utilized the combined ganglion cell and inner plexiform layers (GCIP), we also included this parameter in our analysis. The positions of the segmentation lines were then exported to a spreadsheet, which was imported into Matlab (Version 2012b, MathWorks, Munich, Germany) to calculate mean layer thicknesses with different thickness estimates (Fig 3, thickness estimates A—J). For thickness estimates based on a single B-scan only, we extracted the corresponding B-scans from an appropriate volume scan (Fig 3, thickness estimates G—J).

The Cirrus volume scans were automatically segmented with the Cirrus review software (6.5.0.772) with no corrections of segmentation lines possible in this software. The mean

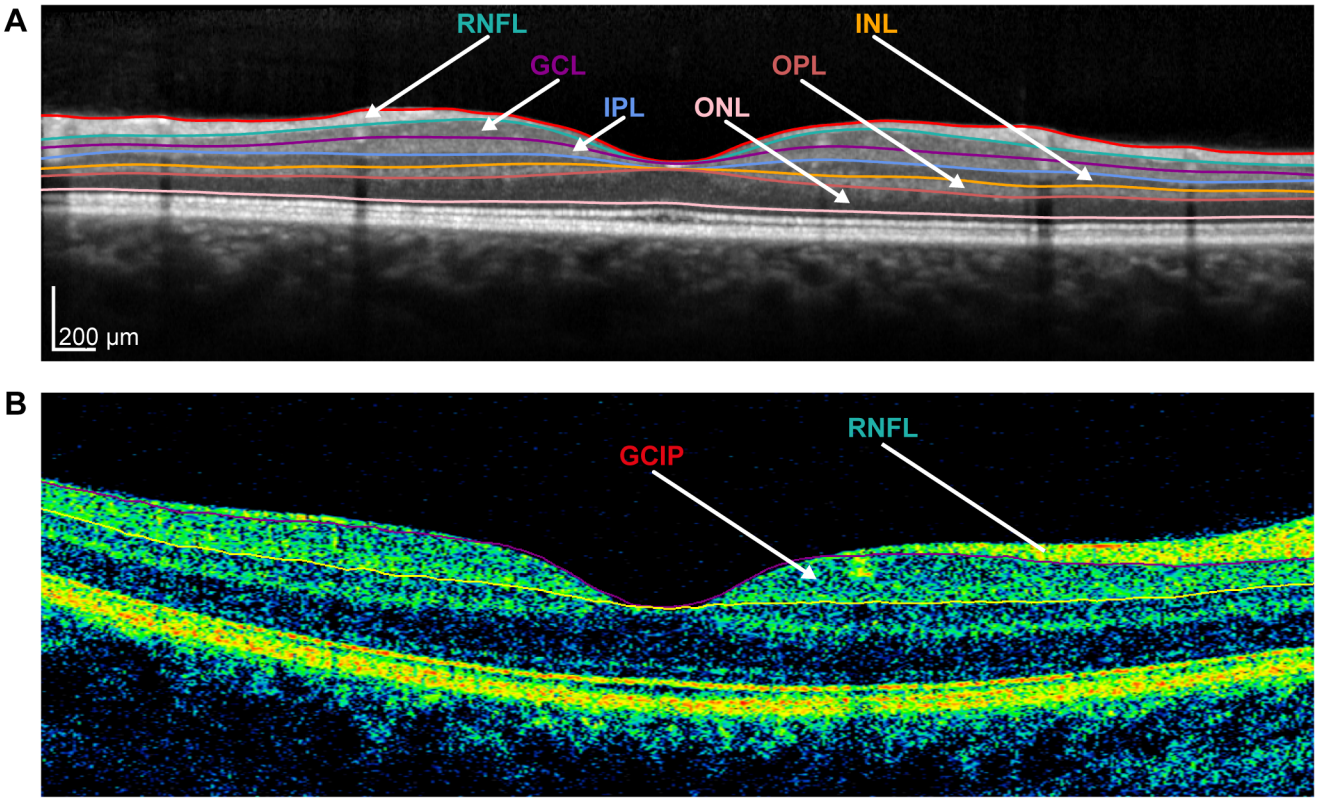
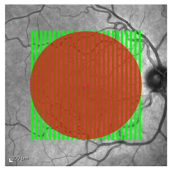
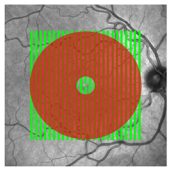
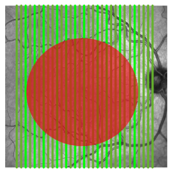
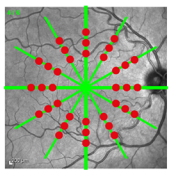
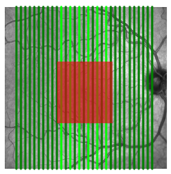
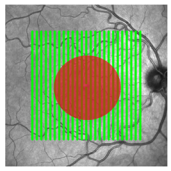


Fig 2. Intra-retinal layer segmentation of OCT B-scans. (A) Exemplary B-scan of a Spectralis SD-OCT with manually reviewed and corrected segmentation into the following retinal layers: retinal nerve fiber layer (RNFL), ganglion cell layer (GCL), inner plexiform layer (IPL), inner nuclear layer (INL), outer plexiform layer (OPL) and outer nuclear layer (ONL). (B) Exemplary Cirrus HD-OCT B-scan with segmentation in the RNFL and combined ganglion cell and inner plexiform (GCIP) layer.

doi:10.1371/journal.pone.0137316.g002

Thickness estimate		Analysis description	Ref.
A		• Spectralis SD-OCT. • Acquisition scan: V-25-512. • All 25 B-scans. • Mean of all values within 6 mm diameter around the fovea. • Heidelberg Eye Explorer.	20
B		• Spectralis SD-OCT. • Acquisition scan: V-25-1024 • All 25 B-scans. • Area for thickness estimation: Mean of all values within 6 mm diameter around the fovea (excluding central 1 mm diameter cirle area). • Heidelberg Eye Explorer	34
C		• Spectralis SD-OCT. • Acquisition scan: V-61-768 • Central B-scans and 6 (every fourth) B-scan in each direction (totaling 13 B-scans) • Mean of all values within 6 mm diameter around the fovea. • Custom Matlab program.	35, 37
D		• Spectralis SD-OCT. • Acquisition scan: R-6-1536 • All 6 B-scans • Mean of 36 seperate points (3 for each radial line). • Custom Matlab program.	36
E		• Spectralis SD-OCT. • Acquisition scan: V-61-768 • Central B-scans and 3 (every fourth) B-scan in each direction (totaling 7 B-scans) • Mean of all values of the central box (side length 1/3 of the B-scan length). • Custom Matlab program.	6
F		• Spectralis SD-OCT. • Acquisition scan: V-25-1024 • All 25 B-scans. • Mean of all values within 3 mm diameter around the fovea. • Heidelberg Eye Explorer.	39

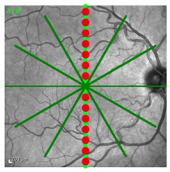
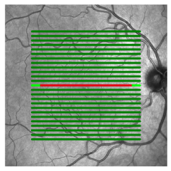
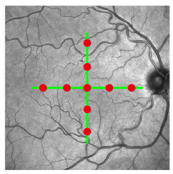
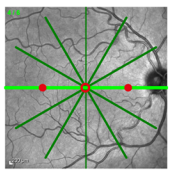
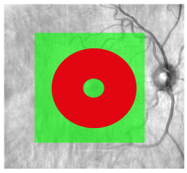
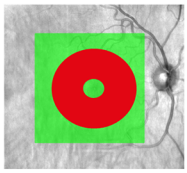
Thickness estimate		Analysis description	Ref.
G		• Spectralis SD-OCT. • Acquisition scan: R-6-1536 • Vertical B-scan. • Mean of 7 points (equally distributed between 1.75° and 7°) for each hemifield (totaling 14 points). • Cutstom Matlab program.	24
H		• Spectralis SD-OCT. • Acquisition scan: H-25-512 • Central B-scan. • Mean of macular width of 5 mm. • Cutstom Matlab program.	12, 26
I		• Spectralis SD-OCT. • Acquisition scans: V-25-512 & H-25-512 • Both central B-scans of each volume scan. • Mean values at 9 locations. • Custom Matlab program.	40
J		• Spectralis SD-OCT. • Acquisition scan: R-6-1536 • Central horizontal B-scan. • Mean of the 2 values at the thickest point nasally and temporally (ONL measured at the single central position). • Custom Matlab program.	12, 27
K		• Cirrus HD-OCT. • Acquisition scan: H-128-512. • All scans. • Mean of the values within an elliptic annulus with outer radi 2.4 and 2.0mm and inner radi 1.2 and 1.0mm. • Cirrus Review Software.	12, 18, 21, 23, 26, 28, 29, 30, 32, 33
L		• Cirrus HD-OCT. • Acquisition scan: H-200-200. • All scans. • Mean of the values within an elliptic annulus with outer radi 2.4 and 2.0mm and inner radi 1.2 and 1.0mm. • Cirrus Review Software.	32

Fig 3. Explanation of different thickness estimates used for the simulation of repeatability. The red areas or points on the fundus images indicate the values that were averaged to generate the layer thickness estimates. The text to the right of each image refers to (top-to-bottom): 1) the used OCT device 2) the applied scan as elucidated in Fig 1; 3) the (subset of) B-scans used for the thickness estimate; 4) the procedure used to calculate the thickness estimates; 5) the program used for calculating the thickness estimate. The last column indicated the article(s) in which the layer estimated was applied. For a detailed overview of the segmentation methods applied in the retrieved articles, see Table B in S1 File. Note: the number and spacing of B-scans on the fundus image for the thickness estimates C and E are not correct due to space limitations.

doi:10.1371/journal.pone.0137316.g003

thicknesses were calculated for the RNFL and the GCIP (see Fig 2B) and exported to XML files for analysis. Layer estimates for Cirrus scans (Fig 3, thickness estimates K and L) were estimated in an elliptical annulus around the fovea with outer radii of 2.4 mm (horizontal) and 2.0 mm (vertical), and inner radii of 0.6 mm (horizontal) and 0.5 mm (vertical), respectively [13].

Statistical Analysis

The thickness estimates for each layer were analyzed using modified Bland-Altman plots [14]. Here, the differences between each pair of repeated measurements were plotted against the mean of the pair. Plots were visually inspected for the occurrence of systematic bias or outliers. Next, the reliability of repeated measurements was calculated as intra-class correlation coefficients (ICC) based on the variance components of a one-way ANOVA fixed effects model with single measures (also referred to as ICC(1,1)) in R (Version 3.1.3) with the package ICC [15]. The ICC can take values from 0 to 1, with 1 representing perfect correlation and 0 no correlation. We considered ICC values above 0.9 as excellent, ICC values between 0.9 and 0.8 as moderate and less than 0.8 as insufficient, as previously suggested [16]. ICC statistics were performed for the mixed cohort of MS/CIS patients and healthy controls, as well as for each group separately. Since ICC are relative indices that are influenced by the heterogeneity of the measurement across different subjects, we also analyzed the coefficient of repeatability (CR) as an absolute reliability index, which is calculated by multiplying the within-subject standard deviation with 2.77 (= $\sqrt{2}$ x 1.96). This yields the interval of the absolute difference between repeated measurements with a probability of 95% [16]. The CRs propose a magnitude for clinically relevant changes in retinal layer mean thicknesses.

Results

Published MS Studies with Intra-Retinal Segmentation of Macular OCT Scans

The search term "optical coherence tomography multiple sclerosis" yielded 279 articles in PubMed, of which 27 met the inclusion criteria [6,12,17–41]. Of the latter, eleven applied Spectralis SD-OCT (Heidelberg Engineering, Heidelberg, Germany) [6,20,24,27,34–37,39–41]; nine used Cirrus OCT (Carl Zeiss Meditec, Dublin, CA, USA) [18,21,23,25,28–30,32,33]; two studies used both devices [12,26] and two used RTVue-100 (Optovue, Fremont, CA, USA) [19,38]. The OCT-1000 (Topcon Corp., Itabashi, Japan) [31], the Stratus OCT (Carl Zeiss Meditec, Dublin, CA, USA) [22] and a custom research device [17] were used in one study each. An analysis was performed of the OCT examinations used in each study, as reflected by the scan settings listed in Table B in S1 File and a detailed overview of the thickness estimates, the segmentation software type (automatic, semi-automatic or manual) and the investigated retinal layers in Table C in S1 File.

For Cirrus OCT two different thickness estimates were used and for both automatic segmentation software were applied (Table C in S1 File). In contrast, we identified ten different thickness estimates used for Spectralis OCT scans. In this study, we analyzed the repeatability of Spectralis and Cirrus OCT layer thickness estimates by generating repetitive OCT measurements in MS/CIS patients and healthy controls. Fig 3 summarizes the thickness estimates, how we implemented them to investigate the repeatability and in which studies they were applied.

Repeatability of OCT Layer Thickness Estimates

The intra-observer repeatability of the twelve OCT layer thickness estimates (Fig 3) and for each of the seven retinal layers (including the composite GCIP layer) are presented for the mixed cohort of patients and controls by ICC statistics in Table 1 and CR in Table 2. In summary, thickness estimates averaging larger areas (Spectralis thickness estimates A–E and estimates K and L for Cirrus) showed mostly excellent ICC values (>0.9) and low CR (1–4 μm). The deviation of thickness estimates F–J were broader across the retinal layers and some only

Table 1. Repeatability of thickness estimates and retinal layers in a mixed cohort of healthy controls and MS/CIS patients as measured by intra-class correlation coefficients (ICC) with 95%-confidence intervals.

Estimate	mRNFL	GCL	IPL	GCIP	INL	OPL	ONL
A	0.98 [0.99–0.97]	0.99 [0.99–0.98]	0.99 [0.99–0.98]	1.00 [1.00–1.00]	0.97 [0.98–0.96]	0.82 [0.88–0.72]	0.98 [0.99–0.97]
B	0.97 [0.98–0.95]	0.99 [1.00–0.99]	0.99 [0.99–0.98]	1.00 [1.00–1.00]	0.96 [0.98–0.94]	0.75 [0.84–0.63]	0.97 [0.98–0.95]
C	0.98 [0.99–0.97]	1.00 [1.00–0.99]	0.99 [1.00–0.99]	1.00 [1.00–1.00]	0.98 [0.99–0.96]	0.78 [0.86–0.68]	1.00 [1.00–0.99]
D	0.98 [0.99–0.97]	0.96 [0.98–0.94]	0.96 [0.97–0.93]	0.98 [0.99–0.97]	0.87 [0.92–0.80]	0.87 [0.92–0.80]	1.00 [1.00–0.99]
E	0.97 [0.98–0.95]	0.99 [1.00–0.99]	0.98 [0.99–0.97]	1.00 [1.00–0.99]	0.90 [0.94–0.85]	0.69 [0.80–0.56]	0.99 [1.00–0.99]
F	0.88 [0.92–0.81]	0.99 [0.99–0.99]	0.98 [0.99–0.97]	0.99 [1.00–0.99]	0.83 [0.89–0.75]	0.62 [0.75–0.47]	0.92 [0.95–0.87]
G	0.96 [0.98–0.94]	0.93 [0.95–0.88]	0.89 [0.93–0.83]	0.94 [0.96–0.91]	0.72 [0.82–0.59]	0.68 [0.79–0.54]	0.99 [0.99–0.99]
H	0.92 [0.95–0.87]	0.96 [0.98–0.94]	0.96 [0.97–0.93]	0.98 [0.99–0.97]	0.73 [0.82–0.60]	0.59 [0.72–0.43]	0.98 [0.99–0.98]
I	0.81 [0.88–0.71]	0.96 [0.97–0.93]	0.93 [0.96–0.89]	0.98 [0.99–0.97]	0.83 [0.90–0.75]	0.55 [0.69–0.38]	0.98 [0.99–0.97]
J	0.92 [0.95–0.88]	0.82 [0.89–0.73]	0.77 [0.86–0.66]	0.86 [0.91–0.79]	0.76 [0.85–0.65]	0.80 [0.87–0.70]	0.92 [0.95–0.87]
K	0.98 [0.99–0.97]	NA	NA	1.00 [1.00–1.00]	NA	NA	NA
L	0.99 [0.99–0.98]	NA	NA	1.00 [1.00–1.00]	NA	NA	NA

Abbreviations: mRNFL = macular retinal nerve fiber layer, **GCL** = ganglion cell layer, **IPL** = inner plexiform layer, **GCIP** = combined GCL and IPL, INL = inner nuclear layer, **OPL** = outer plexiform layer, **ONL** = outer nuclear layer.

doi:10.1371/journal.pone.0137316.t001

reached moderate or even insufficient ICC levels. Particularly thickness estimate J was characterized by higher degree of measurement noise (CR >8 μm for all layers).

The widths of the ICC confidence intervals (CI) as a measure of the precision were highly dependent on the ICC values: with high ICC values (>0.90) we also found very narrow CI widths (<0.1), but CI widths were wider for lower ICC values.

In addition to the mixed cohort, were performed distinct analysis of the patient and control cohorts, see Fig 4 for ICC statistics and Fig 5 for CR measures. The comparison of the repeatability results for the different thickness estimates (Fig 4A) revealed a similar distribution for both cohorts. Of note is, that both cohorts were measured, reviewed and manually corrected by different operators, which might introduce additional bias to the group comparison. Nevertheless, the distribution of ICC values indicate that our previous findings are valid independent of the health status of the measured retina.

Repeatability of Different Retinal Layers

To further investigate differences in single retinal layers, Fig 4B and Fig 5B show the repeatability indices against the distinct layers. The weakest correlation between repeated measurements was present in the OPL in which thickness estimates mainly had lower ICC values indicative for insufficient repeatability. This high measurement noise can be explained by the divergent appearance of the OPL in some of the repeated measurements: Fig 6 shows an example of the OPL in two repeated measurements of the same subject taken 16 min apart (Fig 6A and 6B). In this case, the averaged local OPL differences between both examinations reached up to 25 μm and on single B-scans the differences were up to 40 μm (see Fig 6C).

Systematic Error Analysis

Visual review of the Bland-Altman plots for individual thickness estimates and retinal layers showed no systematic errors or outliers (data not shown).

Table 2. Coefficient of repeatability (CR, in μm) of thickness estimates and retinal layers in a mixed cohort of healthy controls and MS/CIS patients.

Estimate	mRNFL	GCL	IPL	GCIP	INL	OPL	ONL
A	2.06	1.50	1.18	1.23	1.15	2.31	2.30
B	2.91	1.23	1.10	1.22	1.31	2.91	3.21
C	2.24	1.02	0.85	0.88	1.06	2.57	2.50
D	3.23	1.87	1.80	2.33	2.20	1.82	2.06
E	2.43	1.47	1.62	1.70	2.31	3.75	3.41
F	2.87	2.40	2.14	2.90	3.48	5.89	6.28
G	4.35	2.92	2.68	4.40	3.77	2.99	3.44
H	3.64	3.03	2.31	3.67	4.50	5.40	5.15
I	3.78	3.75	3.53	4.34	4.55	6.00	5.56
J	9.07	9.19	8.64	13.23	8.88	9.90	17.75
K	1.91	NA	NA	1.66	NA	NA	NA
L	1.60	NA	NA	1.38	NA	NA	NA

Abbreviations: mRNFL = macular retinal nerve fiber layer, **GCL** = ganglion cell layer, **IPL** = inner plexiform layer, **GCIP** = combined GCL and IPL, INL = inner nuclear layer, **OPL** = outer plexiform layer, **ONL** = outer nuclear layer.

doi:10.1371/journal.pone.0137316.t002

Discussion

Our investigation showed that a wide range of segmentation approaches and layer thickness estimates are currently used in intra-retinal segmentation studies in MS, unfortunately, with equally divergent reliability. As a result, readers should be careful when interpreting published studies using weak thickness estimates, identified by this study.

Spectralis SD-OCT and Cirrus HD-OCT are currently the two most common devices used for intra-retinal segmentation in MS. For the Cirrus HD-OCT the segmentation algorithm of the GCIP has been an integral part of the software since version 6.0 and its reproducibility had been investigated elsewhere [13]. Our repeatability analysis for both of the two standard Cirrus scanning protocols showed very high intra-observer reliability. While all but one of the studies using the Cirrus rely on this built-in automatic segmentation, we identified a total of ten different thickness estimates that were used in studies with the Spectralis device. Our analysis of the different Spectralis thickness estimates revealed that estimates averaging larger areas (thickness estimates A–E, Fig 3) showed excellent repeatability for all retinal layers except the OPL. In contrast, thickness estimates based on only a few single scan locations are more strongly affected by measurement noise introduced in repeated examinations and thus show less reliability. Of note is that the three thickness estimates (A—C) using the 6mm-diameter area around the fovea with different scan settings performed similarly well, independently of the transversal resolution and image averaging rate (automatic real-time, ART). Although higher scan settings, which also result in slower scanning speed, might not be beneficial in terms of repeatability, they can have an impact on the sensitivity. In this study, we did not assess the sensitivity of thickness estimates to detect changes in individual layer measurements, but instead only investigated the repeated measurement reliability of different thickness estimates. However, low CR in our study might indicate that layer measurements are potentially sensitive for tracking longitudinal changes, but this needs to be investigated further.

Studies in MS patients have revealed pronounced reduction in the ganglion cell layer and this may be the most important target for tracking neurodegeneration at least in MS patients [37,39]. Since the ganglion cell layer is thickest in the macula, thickness estimates that entirely cover the macula likely have the best sensitivity for tracking changes. The CR of the GCL

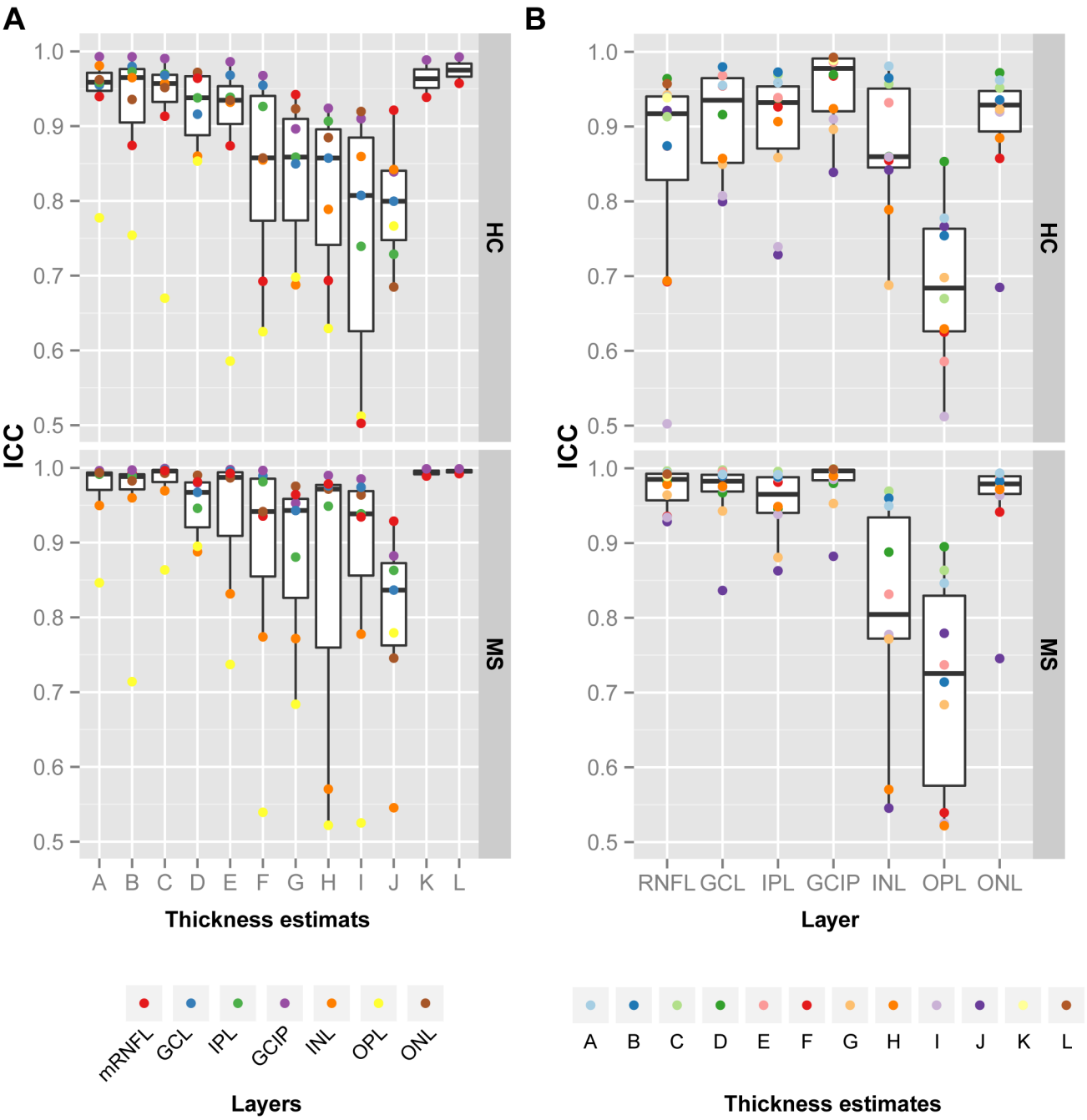


Fig 4. Intraclass correlation coefficients (ICC) for different thickness estimates and retinal layers in MS/CIS patients and healthy controls. Intraclass correlation coefficient (ICC) for the distinct cohorts of healthy controls (HC) and patients (MS) ordered for (A) the twelve thickness estimates (Fig 3) and (B) the seven retinal layers. Abbreviations: mRNFL = macular retinal nerve fiber layer, GCL = ganglion cell layer, IPL = inner plexiform layer, GICP = combined GCL and IPL, INL = inner nuclear layer, OPL = outer plexiform layer, ONL = outer nuclear layer.

doi:10.1371/journal.pone.0137316.g004

(1.02 μm), IPL (0.85 μm) and the CR of the GCIP (0.88 μm) using the thickness estimate C had the overall best performance. A recent study estimated yearly GCL/GCIP loss at 0.5 μm using layer estimate K, suggesting that SD-OCT should be able to track changes over 24 months [42]. After acute optic neuritis, average GCIP loss was reported at 9 μm in a recent study [23], making several thickness estimate good candidates to track GCL/GCIP changes after acute ON.

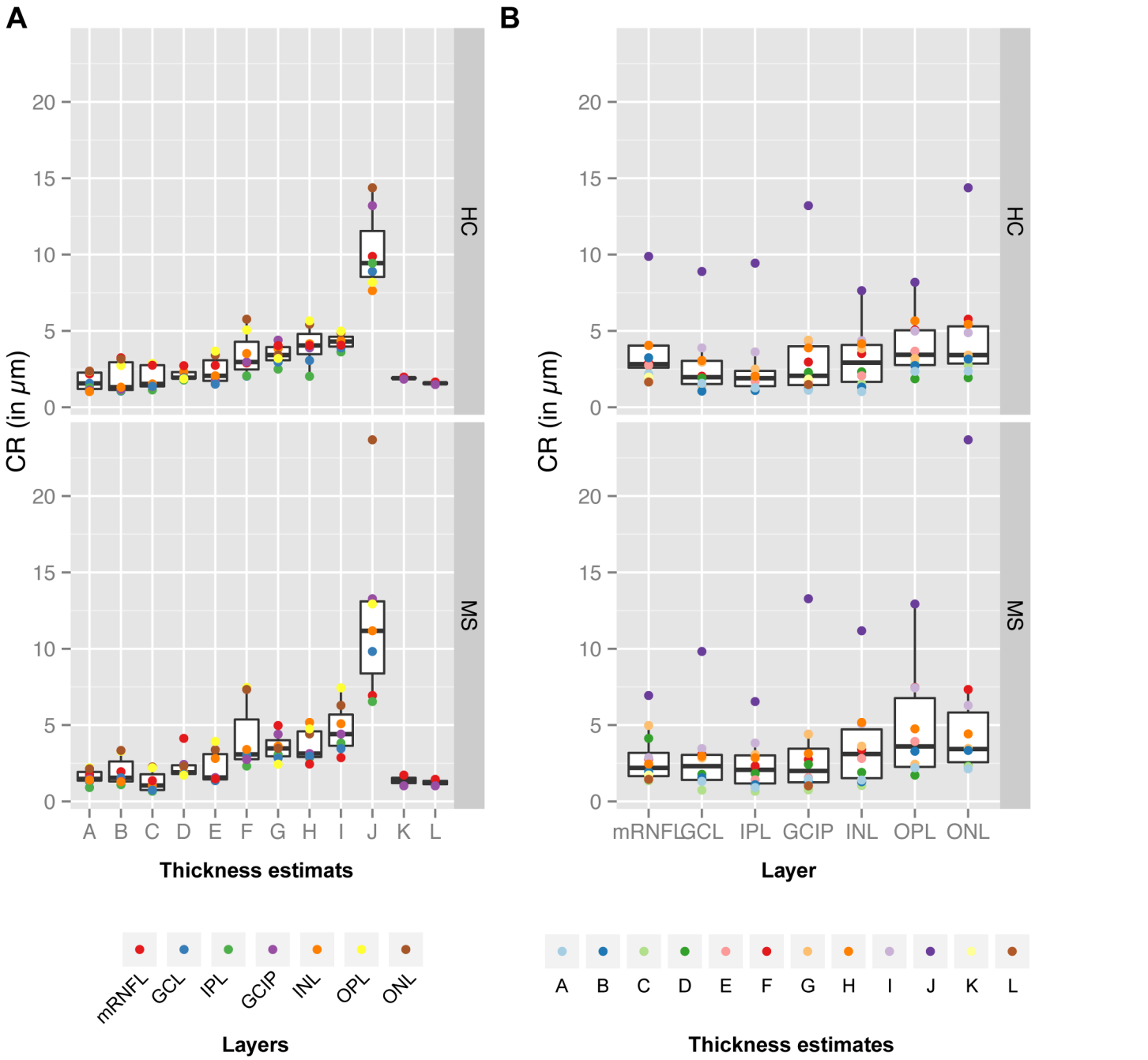


Fig 5. Coefficient of repeatability (CR) for different thickness estimates and retinal layers in MS/CIS patients and healthy controls. Coefficient of repeatability (CR) for the distinct cohorts of healthy controls (HC) and patients (MS) ordered for (A) the 12 thickness estimates (Fig 3) and (B) the 7 retinal layers. **Abbreviations**: **mRNFL** = macular retinal nerve fiber layer, **GCL** = ganglion cell layer, **IPL** = inner plexiform layer, **GICP** = combined GCL and IPL, **INL** = inner nuclear layer, **OPL** = outer plexiform layer, **ONL** = outer nuclear layer.

doi:10.1371/journal.pone.0137316.g005

Another layer of interest is the INL, in which thickening in the area of the macula has been reported in relation to optic neuritis and MS disease activity [28,35,36,43]. Here, the CR was just above 1 μm in layer estimate A—C. After acute optic neuritis, intermittent INL thickening was reported by two studies in the range of 1–4 μm [23,34]. There is currently no estimate for the magnitude of annual changes in MS outside acute ON. Considering the above CR, current

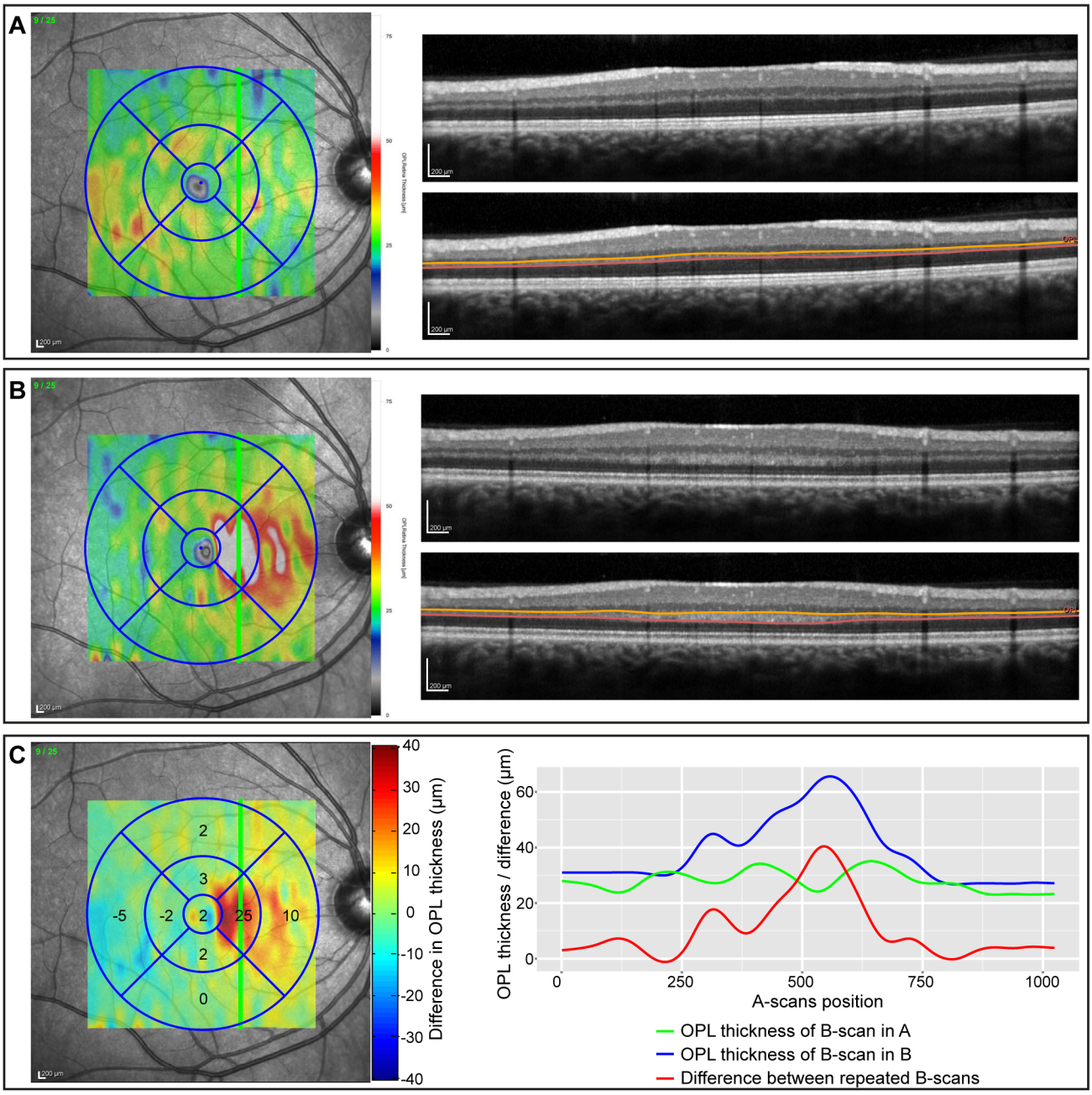


Fig 6. Differences in the outer plexiform layer (OPL) in repeated OCT measurements. (A) and (B) show two repeated OCT measurements (16 min between both examinations) of the same subject with the same scan settings (Scan V-25-1024 in Fig 1). On the left, the fundus images are given with an overlay of the OPL thickness maps generated by the Eye Explorer software (Heidelberg Engineering, Germany). The B-scans (with and without OPL segmentation lines) are shown on the right and are located at the position indicated by the green line on the fundus images. (C) The difference map on the left was calculated by subtracting the OPL thickness map in (A) from the map in (B) (custom Matlab program). The values in the grid are the mean OPL thickness differences for each sector. The right graph maps the OPL thickness of the B-scans in (A) (green line) and (B) (blue line), respectively. The red line indicates the difference between the repeated B-scans.

doi:10.1371/journal.pone.0137316.g006

SD-OCT might just fall short in reliably detecting INL changes over typical observation times both in acute ON and in MS.

The OPL results deserve special mention. Axons in the OPL of photoreceptor nuclei, commonly known as Henle's fiber layer (HFL), run obliquely in the tissue and its OCT appearance is highly dependent on the angle of the OCT's light beam relative to this layer [44]. The HFL can either appear hyperintense or hypointense, which can skew the determination of the OPL thickness (Fig 6) [45]. The OPL's weak reliability is thus a consequence of a complex, local anatomy and not *per se* noise introduced by small changes in the scan window, as with the other layers. The reliability is therefore less dependent on the thickness estimate than on an exact beam placement during measurement. When investigating OPL in clinical or research settings, an optimal OCT scan with exact beam placement is crucial to prevent uncorrectable OPL differences in the raw data. For this reason, the OSCAR-IB quality criteria, particularly criterion B ("beam placement"), should be followed rigorously [46–48].

Our study has the following limitations: data acquisition and manual correction was performed by different operators for the study cohorts, but nevertheless results were still very similar in patients and controls. However, in real multicenter settings more influencing factors are possible, like the use of different scanners or software versions or multiple operators. Further studies are necessary to tackle these issues and such studies would require the inclusion of several operators at different centers, which was beyond the scope our study. Furthermore, we did not have access to analyzable data from RT-Vue and Topcon; and the raw data of these devices could behave differently when looking at different thickness estimates. Lastly, the sample size of our study was relatively low and given its exploratory nature we were not able to perform a-priory sample size calculations. However, CI were very narrow for the excellent thickness estimates and we therefore believe that our data still support the notion of excellent repeatability of some thickness estimates and particular layers.

Several clinical trials using OCT-derived parameters as primary or secondary study outcome parameters are ongoing or are about to commence. Current and future studies should favor the most reliable thickness estimates over weaker ones. Although the thickness estimates A—E all performed well, the overall performance of estimate A–the mean of all values in the 6mm-diameter ring area centered on and including the fovea–was above average and we strongly encourage clinical researchers, as well as device manufacturers, to rely on this thickness estimate when implementing intra-retinal segmentation measurement across all macular layers. This thickness estimate should also serve as the basis for acquiring normative data from healthy controls to aid comparability between different devices and to detect measurements deviant from healthy subjects. If a proven reliable thickness estimate is not viable within the context of an individual study, scientists should explain how they selected the alternative methods and quantify the reliability of these.

Supporting Information

S1 Data. Results of OCT layer thickness estimate calculations.
(ZIP)

S1 File. Supplementary materials. This file contains overviews of the study cohorts (Table A), the scientific articles and OCT scan settings used for retinal OCT segmentation in patients with multiple sclerosis (Table B) and the applied segmentation methods used in the retrieved articles (Table C).
(DOCX)

Author Contributions

Conceived and designed the experiments: TO IB AUB. Performed the experiments: MW JM TO. Analyzed the data: TO MW HZ. Contributed reagents/materials/analysis tools: FP IB. Wrote the paper: TO FP AUB.

References

1. Huang D, Swanson EA, Lin CP, Schuman JS, Stinson WG, Chang W, et al. Optical coherence tomography. Science. 1991; 254: 1178–1181. PMID: 1957169
2. Drexler W, Morgner U, Ghanta RK, Kärtner FX, Schuman JS, Fujimoto JG. Ultrahigh-resolution ophthalmic optical coherence tomography. Nat Med. 2001; 7: 502–507. doi: 10.1038/86589 PMID: 11283681
3. Oberwahrenbrock T, Schippling S, Ringelstein M, Kaufhold F, Zimmermann H, Keser N, et al. Retinal damage in multiple sclerosis disease subtypes measured by high-resolution optical coherence tomography. Mult Scler Int. 2012; 2012: 530305. doi: 10.1155/2012/530305 PMID: 22888431
4. Roth NM, Saidha S, Zimmermann H, Brandt AU, Isensee J, Benkhellouf-Rutkowska A, et al. Photoreceptor layer thinning in idiopathic Parkinson's disease. Mov Disord. 2014; 29: 1163–1170. doi: 10.1002/mds.25896 PMID: 24789530
5. Koronyo-Hamaoui M, Koronyo Y, Ljubimov AV, Miller CA, Ko MK, Black KL, et al. Identification of amyloid plaques in retinas from Alzheimer's patients and noninvasive in vivo optical imaging of retinal plaques in a mouse model. NeuroImage. 2011; 54 Suppl 1: S204–217. doi: 10.1016/j.neuroimage.2010.06.020 PMID: 20550967
6. Schneider E, Zimmermann H, Oberwahrenbrock T, Kaufhold F, Kadas EM, Petzold A, et al. Optical Coherence Tomography Reveals Distinct Patterns of Retinal Damage in Neuromyelitis Optica and Multiple Sclerosis. PLoS ONE. 2013; 8: e66151. doi: 10.1371/journal.pone.0066151 PMID: 23805202
7. Seyer LA, Galetta K, Wilson J, Sakai R, Perlman S, Mathews K, et al. Analysis of the visual system in Friedreich ataxia. J Neurol. 2013; doi: 10.1007/s00415-013-6978-z
8. Brandt AU, Oberwahrenbrock T, Costello F, Fielden MM, Gertz KM, Kleffner I, et al. RETINAL LESION EVOLUTION IN SUSAC SYNDROME. Retina. doi: 10.1097/IAE.0000000000000700
9. Stricker S, Oberwahrenbrock T, Zimmermann H, Schroeter J, Endres M, Brandt AU, et al. Temporal Retinal Nerve Fiber Loss in Patients with Spinocerebellar Ataxia Type 1. PLoS ONE. 2011; 6: e23024. doi: 10.1371/journal.pone.0023024 PMID: 21829579
10. Barkhof F, Calabresi PA, Miller DH, Reingold SC. Imaging outcomes for neuroprotection and repair in multiple sclerosis trials. Nat Rev Neurol. 2009; 5: 256–266. doi: 10.1038/nrneurol.2009.41 PMID: 19488083
11. Seigo MA, Sotirchos ES, Newsome S, Babiarz A, Eckstein C, Ford E, et al. In vivo assessment of retinal neuronal layers in multiple sclerosis with manual and automated optical coherence tomography segmentation techniques. J Neurol. 2012; 259: 2119–2130. doi: 10.1007/s00415-012-6466-x PMID: 22418995
12. Sotirchos ES, Seigo MA, Calabresi PA, Saidha S. Comparison of Point Estimates and Average Thicknesses of Retinal Layers Measured Using Manual Optical Coherence Tomography Segmentation for Quantification of Retinal Neurodegeneration in Multiple Sclerosis. Curr Eye Res. 2013; 38: 224–228. doi: 10.3109/02713683.2012.722243 PMID: 22954302
13. Mwanza J-C, Oakley JD, Budenz DL, Chang RT, Knight OJ, Feuer WJ. Macular Ganglion Cell–Inner Plexiform Layer: Automated Detection and Thickness Reproducibility with Spectral Domain–Optical Coherence Tomography in Glaucoma. Invest Ophthalmol Vis Sci. 2011; 52: 8323–8329. doi: 10.1167/iovs.11-7962 PMID: 21917932
14. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. Lancet. 1986; 1: 307–310. PMID: 2868172
15. Wolak ME, Fairbairn DJ, Paulsen YR. Guidelines for estimating repeatability. Methods Ecol Evol. 2012; 3: 129–137. doi: 10.1111/j.2041-210X.2011.00125.x
16. Vaz S, Falkmer T, Passmore AE, Parsons R, Andreou P. The Case for Using the Repeatability Coefficient When Calculating Test–Retest Reliability. PLoS ONE. 2013; 8: e73990. doi: 10.1371/journal.pone.0073990 PMID: 24040139
17. Choi SS, Zawadzki RJ, Keltner JL, Werner JS. Changes in Cellular Structures Revealed by Ultra-high Resolution Retinal Imaging in Optic Neuropathies. Invest Ophthalmol Vis Sci. 2008; 49: 2103–2119. doi: 10.1167/iovs.07-0980 PMID: 18436843
18. Saidha S, Syc SB, Ibrahim MA, Eckstein C, Warner CV, Farrell SK, et al. Primary retinal pathology in multiple sclerosis as detected by optical coherence tomography. Brain. 2011; 134: 518–533. doi: 10.1093/brain/awq346 PMID: 21252110
19. Garas A, Simó M, Holló G. Nerve fiber layer and macular thinning measured with different imaging methods during the course of acute optic neuritis. Eur J Ophthalmol. 2011; 21: 473–483. doi: 10.5301/EJO.2010.5844 PMID: 21038310
20. Davies EC, Galetta KM, Sackel DJ, Talman LS, Frohman EM, Calabresi PA, et al. Retinal Ganglion Cell Layer Volumetric Assessment by Spectral-Domain Optical Coherence Tomography in Multiple

Sclerosis: Application of a High-Precision Manual Estimation Technique. J Neuro-Ophthalmol Sept 2011. 2011; 31: 260–264. doi: 10.1097/WNO.0b013e318221b434

21. Saidha S, Syc SB, Durbin MK, Eckstein C, Oakley JD, Meyer SA, et al. 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. 2011; 17: 1449–1463. doi: 10.1177/1352458511418630

22. Tátrai E, Simó M, Iljicsov A, Németh J, DeBuc DC, Somfai GM. In Vivo Evaluation of Retinal Neurodegeneration in Patients with Multiple Sclerosis. PLoS ONE. 2012; 7: e30922. doi: 10.1371/journal.pone.0030922 PMID: 22292076

23. Syc SB, Saidha S, Newsome SD, Ratchford JN, Levy M, Ford E, et al. Optical coherence tomography segmentation reveals ganglion cell layer pathology after optic neuritis. Brain. 2012; 135: 521–533. doi: 10.1093/brain/awr264 PMID: 22006982

24. Sriram P, Graham SL, Wang C, Yiannikas C, Garrick R, Klistorner A. Transsynaptic Retinal Degeneration in Optic Neuropathies: Optical Coherence Tomography Study. Invest Ophthalmol Vis Sci. 2012; 53: 1271–1275. doi: 10.1167/iovs.11-8732 PMID: 22222509

25. Walter SD, Ishikawa H, Galetta KM, Sakai RE, Feller DJ, Henderson SB, et al. Ganglion Cell Loss in Relation to Visual Disability in Multiple Sclerosis. Ophthalmology. 2012; 119: 1250–1257. doi: 10.1016/j.ophtha.2011.11.032 PMID: 22365058

26. Seigo MA, Sotirchos ES, Newsome S, Babiarz A, Eckstein C, Ford E, et al. In vivo assessment of retinal neuronal layers in multiple sclerosis with manual and automated optical coherence tomography segmentation techniques. J Neurol. 2012; 259: 2119–2130. doi: 10.1007/s00415-012-6466-x PMID: 22418995

27. Albrecht P, Ringelstein M, Müller AK, Keser N, Dietlein T, Lappas A, et al. Degeneration of retinal layers in multiple sclerosis subtypes quantified by optical coherence tomography. Mult Scler J. 2012; 18: 1422–1429. doi: 10.1177/1352458512439237

28. Saidha S, Sotirchos ES, Ibrahim MA, Crainiceanu CM, Gelfand JM, Sepah YJ, et al. Microcystic macular oedema, thickness of the inner nuclear layer of the retina, and disease characteristics in multiple sclerosis: a retrospective study. Lancet Neurol. 2012; 11: 963–972. doi: 10.1016/S1474-4422(12)70213-2 PMID: 23041237

29. Ratchford JN, Saidha S, Sotirchos ES, Oh JA, Seigo MA, Eckstein C, et al. Active MS is associated with accelerated retinal ganglion cell/inner plexiform layer thinning. Neurology. 2013; 80: 47–54. doi: 10.1212/WNL.0b013e31827b1a1c PMID: 23267030

30. Saidha S, Sotirchos ES, Oh J, Syc SB, Seigo MA, Shiee N, et al. Relationships between retinal axonal and neuronal measures and global central nervous system pathology in multiple sclerosis. JAMA Neurol. 2013; 70: 34–43. doi: 10.1001/jamaneurol.2013.573 PMID: 23318513

31. Fernandes DB, Raza AS, Nogueira RGF, Wang D, Callegaro D, Hood DC, et al. Evaluation of Inner Retinal Layers in Patients with Multiple Sclerosis or Neuromyelitis Optica Using Optical Coherence Tomography. Ophthalmology. 2013; 120: 387–394. doi: 10.1016/j.ophtha.2012.07.066 PMID: 23084127

32. Zimmermann H, Freing A, Kaufhold F, Gaede G, Bohn E, Bock M, et al. Optic neuritis interferes with optical coherence tomography and magnetic resonance imaging correlations. Mult Scler J. 2013; 19: 443–450. doi: 10.1177/1352458512457844

33. Sotirchos ES, Saidha S, Byraiah G, Mealy MA, Ibrahim MA, Sepah YJ, et al. In vivo identification of morphologic retinal abnormalities in neuromyelitis optica. Neurology. 2013; 80: 1406–1414. doi: 10.1212/WNL.0b013e31828c2f7a PMID: 23516321

34. Gabilondo I, Sepúlveda M, Ortiz-Perez S, Fraga-Pumar E, Martínez-Lapiscina EH, Llufriu S, et al. Retrograde retinal damage after acute optic tract lesion in MS. J Neurol Neurosurg Psychiatry. 2013; 84: 824–826. doi: 10.1136/jnnp-2012-304854 PMID: 23715916

35. Kaufhold F, Zimmermann H, Schneider E, Ruprecht K, Paul F, Oberwahrenbrock T, et al. Optic Neuritis Is Associated with Inner Nuclear Layer Thickening and Microcystic Macular Edema Independently of Multiple Sclerosis. PLoS ONE. 2013; 8: e71145. doi: 10.1371/journal.pone.0071145 PMID: 23940706

36. Kaushik M, Wang CY, Barnett MH, Garrick R, Parratt J, Graham SL, et al. Inner nuclear layer thickening is inversley proportional to retinal ganglion cell loss in optic neuritis. PloS One. 2013; 8: e78341. doi: 10.1371/journal.pone.0078341 PMID: 24098599

37. Oberwahrenbrock T, Ringelstein M, Jentschke S, Deuschle K, Klumbies K, Bellmann-Strobl J, et al. Retinal ganglion cell and inner plexiform layer thinning in clinically isolated syndrome. Mult Scler Houndmills Basingstoke Engl. 2013; 19: 1887–1895. doi: 10.1177/1352458513489757

38. Tugcu B, Soysal A, Kılıc M, Yuksel B, Kale N, Yıgıt U, et al. Assessment of structural and functional visual outcomes in relapsing remitting multiple sclerosis with visual evoked potentials and optical

coherence tomography. J Neurol Sci. 2013; 335: 182–185. doi: 10.1016/j.jns.2013.09.027 PMID: 24148562

39. Balk LJ, Tewarie P, Killestein J, Polman CH, Uitdehaag BMJ, Petzold A. Disease course heterogeneity and OCT in multiple sclerosis. Mult Scler J. 2014; 20: 1198–1206. doi: 10.1177/1352458513518626

40. Park K-A, Kim J, Oh SY. Analysis of spectral domain optical coherence tomography measurements in optic neuritis: differences in neuromyelitis optica, multiple sclerosis, isolated optic neuritis and normal healthy controls. Acta Ophthalmol (Copenh). 2014; 92: e57–e65. doi: 10.1111/aos.12215

41. Garcia-Martin E, Polo V, Larrosa JM, Marques ML, Herrero R, Martin J, et al. Retinal Layer Segmentation in Patients with Multiple Sclerosis Using Spectral Domain Optical Coherence Tomography. Ophthalmology. 2014; 121: 573–579. doi: 10.1016/j.ophtha.2013.09.035 PMID: 24268855

42. Narayanan D, Cheng H, Bonem KN, Saenz R, Tang RA, Frishman LJ. Tracking changes over time in retinal nerve fiber layer and ganglion cell-inner plexiform layer thickness in multiple sclerosis. Mult Scler Houndmills Basingstoke Engl. 2014; 20: 1331–1341. doi: 10.1177/1352458514523498

43. Brandt AU, Oberwahrenbrock T, Kadas EM, Lagrèze WA, Paul F. Dynamic formation of macular microcysts independent of vitreous traction changes. Neurology. 2014; doi: 10.1212/WNL.0000000000000545

44. Otani T, Yamaguchi Y, Kishi S. IMPROVED VISUALIZATION OF HENLE FIBER LAYER BY CHANGING THE MEASUREMENT BEAM ANGLE ON OPTICAL COHERENCE TOMOGRAPHY. Retina March 2011. 2011; 31: 497–501. doi: 10.1097/IAE.0b013e3181ed8dae PMID: 21102368

45. Ouyang Y, Walsh AC, Keane PA, Heussen FM, Pappuru RKR, Sadda SR. Different phenotypes of the appearance of the outer plexiform layer on optical coherence tomography. Graefes Arch Clin Exp Ophthalmol Albrecht Von Graefes Arch Für Klin Exp Ophthalmol. 2013; 251: 2311–2317. doi: 10.1007/s00417-013-2308-5

46. Tewarie P, Balk L, Costello F, Green A, Martin R, Schippling S, et al. The OSCAR-IB Consensus Criteria for Retinal OCT Quality Assessment. PLoS ONE. 2012; 7: e34823. doi: 10.1371/journal.pone.0034823 PMID: 22536333

47. Balk LJ, de Vries-Knoppert WAEJ, Petzold A. A simple sign for recognizing off-axis OCT measurement beam placement in the context of multicentre studies. PloS One. 2012; 7: e48222. doi: 10.1371/journal.pone.0048222 PMID: 23144857

48. Schippling S, Balk L, Costello F, Albrecht P, Balcer L, Calabresi P, et al. Quality control for retinal OCT in multiple sclerosis: validation of the OSCAR-IB criteria. Mult Scler Houndmills Basingstoke Engl. 2014; doi: 10.1177/1352458514538110