Validation of an Automated Quantification of Relative Ellipsoid Zone Reflectivity on Spectral Domain-Optical Coherence Tomography Images

Sarah Thiele1,2, Ben Isselmann1, Maximilian Pfau1–3, Frank G. Holz1,2, Steffen Schmitz-Valckenberg1,2, Zhichao Wu4,5, Robyn H. Guymer4,5, and Chi D. Luu4,5

1 Department of Ophthalmology, University of Bonn, Bonn, Germany
2 GRADEReading Center, University of Bonn, Bonn, Germany
3 Department of Biomedical Data Science, Stanford University, Stanford, California, USA
4 Centre for Eye Research Australia, Royal Victorian Eye and Ear Hospital, East Melbourne, Australia
5 Ophthalmology, Department of Surgery, University of Melbourne, Melbourne, Australia

Correspondence: Sarah Thiele, FEBO, Department of Ophthalmology, University of Bonn, Venusberg-Campus 1, Gebäude 04, 53127 Bonn, Germany. e-mail: sarah.thiele@ukbonn.de

Received: July 2, 2020
Accepted: September 24, 2020
Published: October 14, 2020

Keywords: reflectivity; imaging; biomarker


Purpose: Relative ellipsoid zone reflectivity (rEZR) represents a potential biomarker of photoreceptor health on spectral-domain optical coherence tomography (SD-OCT). Because manual quantification of rEZR is laborious and lacks of spatial resolution, automated quantification of the rEZR would be beneficial. The purpose of this study was to evaluate the reliability and reproducibility of an automated rEZR quantification method.

Methods: The rEZR was acquired using a manual and an automated approach in eyes with age-related macular degeneration (AMD) and healthy controls. The rEZR obtained from both methods was compared and the agreement between the methods and their reproducibility assessed.

Results: Forty eyes of 40 participants with a mean (± standard deviation) age of 65.2 ± 7.8 years were included. Both the manual and automated method showed that control eyes exhibit a greater rEZR than AMD eyes (P < 0.001). Overall, the limits of agreement between the manual and automated method were −7.5 to 7.3 arbitrary units (AU) and 95% of the data points had a difference in the rEZR between the methods of ±8.2%. An expected perfect reproducibility was observed for the automated method, whereas the manual method had a coefficient of repeatability of 6.3 arbitrary units.

Conclusions: The automated quantification of rEZR method is reliable and reproducible. Further studies of the rEZR as a novel biomarker for AMD severity and progression are warranted.

Translational Relevance: Automated quantification of SD-OCT–based rEZR allows for its comprehensive and longitudinal characterization evaluating its relevance as an in vivo biomarker of photoreceptor function and its prognostic value for AMD progression.

Introduction

Currently, clinical classification systems for early and intermediate age-related macular degeneration (AMD) are based on drusen size and the presence of retinal pigmentary changes.1–3 These features have also been used to determine the individual’s risk of progression toward advanced AMD. However, these features are not sensitive enough to monitor disease progression or predict risk of progression over a short period of time. Thus, current clinical trials of new treatments for early stages of AMD require a large cohort of participants and long follow-up period to have sufficient power to demonstrate whether the treatment is efficacious. A biomarker that allows better monitoring of AMD progression during the early stages would facilitate and expedite the development and evaluation of novel therapies targeting early disease stages.
With the introduction of spectral-domain optical coherence tomography (SD-OCT), a high-resolution retinal ultrastructure can be imaged in vivo, allowing a more detailed study of the pathologic changes in the macular area offering a more refined diagnosis, classification, risk assessment, and monitoring of AMD.\textsuperscript{4–6} In AMD, features of interest on SD-OCT scans are the four hyperreflective zones in the outer retina and retinal pigment epithelium (RPE).\textsuperscript{7} Of the four hyperreflective outer retinal zones, the signal of the second hyperreflective band, termed the ellipsoid zone (EZ), is thought to derive from light scattering mitochondria, which are tightly packed in the photoreceptor inner segments and that have a crucial role for photoreceptor metabolic function, structural integrity and health.\textsuperscript{8–10} Furthermore, it has been shown that light scattering properties of the mitochondria are dependent upon the function, morphology and density of the photoreceptors.\textsuperscript{10–13} Thus, it is believed that changes in the EZ reflectivity on SD-OCT scans might reflect the pathologic alterations of the outer retina. Hence, EZ reflectivity from SD-OCT images is of great research interest and could potentially represent a novel biomarker for photoreceptor integrity and act as a surrogate marker of photoreceptor function.

Wu et al.\textsuperscript{14} have previously quantified the relative EZ reflectivity (rEZr) and showed that the rEZr was strongly correlated with retinal function assessed by multifocal electroretinography. The same group also showed that the rEZr was significantly reduced in eyes with intermediate (i) AMD, particularly in eyes with reticular pseudodrusen and pigmentary changes, compared with the normal control eyes.\textsuperscript{15,16} These findings suggest that rEZr might have a role in AMD characterization and monitoring disease progression. However, the rEZr data in those studies were obtained manually from a limited number of locations on a single OCT line scans (B-scan) through the fovea. This process is not only very time consuming, but also lacks spatial resolution. To characterize the prognostic relevance of the rEZr for disease progression, a high-resolution rEZr two-dimensional map derived from longitudinal volume OCT scans is needed, which furthermore would improve the sensitivity to detect localized rEZr changes. A method to automatically calculate rEZr from a three-dimensional OCT volume scan would overcome these limitations, and we thus developed an algorithm to automatically quantify the rEZr for this purpose.

The aim of this study was to determine the feasibility and reliability of the automated method by examining the agreement between the manual and automated method for quantifying the rEZr, and the reproducibility of the two methods.

### Methods

#### Study Cohort

Participants were recruited as part of a natural history cohort of intermediate AMD (iAMD) at the Center for Eye Research Australia (CERA). The study was approved by the Human Research Ethics Committee of the Royal Victorian Eye and Ear Hospital in Melbourne, Australia, and conducted according to the Declaration of Helsinki. All study participants provided written informed consent before enrollment.

Eligible AMD participants were required to be at least 50 years of age and were diagnosed with iAMD, defined as having at least one druse larger than 125 μm in diameter within 1500 μm from the fovea in both eyes.\textsuperscript{1} Spouses and friends of the AMD participants were recruited as control participants and included only if both eyes did not exhibit signs of AMD, although drusen 63 μm or smaller (druplets) were allowed (normal aging changes). All participants were required to have a best-corrected visual acuity of 20/40 or better in both eyes. The exclusion criteria for any participant were the presence of any late-stage AMD (including geographic atrophy, or OCT-defined atrophy or nascent geographic atrophy), choroidal neovascularization, significant cataract, glaucoma, amblyopia, refractive error with spherical equivalent of 3.0 diopters or greater, or any corneal pathology that could comprise vision in either eye. Only data from one eye (study eye) per participant were included in this study. The eye with lower visual acuity was chosen as study eye. If the visual acuity is the same for both eyes, the right eye was selected as the study eye.

#### Imaging

High-resolution combined and simultaneous confocal scanning laser ophthalmoscopy +SD-OCT imaging was performed with a Spectralis HRA+OCT device (Heidelberg Engineering, Heidelberg, Germany). Although both a single horizontal scan and volume scans were collected, only data from the single horizontal scan were used in this validation study. This was because it is not feasible to perform manual quantification of the rEZr on the volume scans for all the participants. The single horizontal line scan through the fovea was acquired for each study eye using the high-resolution (1538 pixel along the x-axis) setting and an automated real-time mode of 100 frames.
Image Preprocessing

An automated segmentation of retinal layers was performed, and if required, manual corrections were applied using the Heidelberg Eye Explorer software (HEYEX, software version 1.10.4.0, Heidelberg Engineering). Once correct retinal layer segmentation was confirmed the segmentation data were exported via XML-files for each OCT line scan.

All quantitative analysis of pixel reflectivities were performed on OCT raw images using the export function for raw image data in the HEYEX software. The raw OCT (linear displayed) images provide reflectivity data of the actual non-transformed signals, which is different to the reflectivity signals of logarithmic displayed OCT images typically seen on the HEYEX reviewing software. The signals in logarithmic displayed OCT images are transformed using a proprietary algorithm to enhance the perception of contrast toward the lower end of the dynamic range (i.e., hyporeflective structures). This image transformation leads to misrepresentation of real differences in reflectivity and loss of information given the 8-bit grayscale images (i.e., 256 values for intensity) of exported B-scans. To account for real differences in reflectivities, only the actual and therefore linearly displayed reflectivity information were determined in the “native” raw image data using MatLab (The MathWorks, Version 9.5 Natick, MA) (Figs. 1A and 1B).

Segmentation coordinates (exported via XML-files) of retinal layers were superimposed to the raw OCT images. The OCT B-scans were then straightened along the coordinates of the RPE (Fig. 1C) to allow for accurate determination of the rEZR even in eyes with pronounced curvatures of the posterior pole. Images of left eyes were horizontally flipped so that the data from each coordinate of the left eye were matched to that of the right eye.

Determination of the rEZR

On each OCT B-scan, seven predefined and equally distributed regions of interest (ROIs) up to an eccentricity of 8.9° from the fovea were generated (Fig. 1D). This strategy allowed us to obtain the rEZR data from seven different retinal locations on each OCT scan. The width of each ROI was set at 20 pixels (approximately 120 μm in high-resolution Spectralis OCT imaging) along the image x-axis. The mean pixel intensity (dynamic range of grey values: 0–1 [arbitrary units (AU)]) for the EZ band and external limiting membrane (ELM) band within the ROI was then determined and the rEZR was calculated as the ratio of the EZ and ELM peak intensity.

The EZ and EML intensity was determined by both the manual and automated methods. For the conventional manual quantification approach, both the EZ and the ELM peak was visually identified from the reflectivity profile in each predetermined ROI and their peak intensity values were obtained manually. The manual method was performed by a single examiner (S.T.). For the automated method, the EZ and the ELM peak intensity was determined using an algorithm developed in MatLab software. The algorithm automatically marked the predetermined ROIs and identified the EZ and ELM peaks at each ROI. Similar to the manual method, the width of each ROI was set at 20 pixels along the image x-axis. To identify the EZ and ELM peak, the system automatically marked the subregions for EZ and ELM determination on the reflectivity profile of each ROI. The positions of the subregions were determined based on the 95% prediction interval using the following functions.
A representative reflectivity profile of the outer retina (for corresponding OCT image, see Fig. 1) indicating the distance between the peak reflectivity of the RPE (first peak from left) and the EZ (second peak), determined as the mean [pixels] of the “Area of determination of EZ” and the RPE and the ELM (third peak) determined as the mean [pixels] the area of determination of ELM.

Specifically, the distances (in pixels) between the RPE peak and the EZ peak (RPE-EZ distance) and the RPE and the ELM peaks (RPE-ELM distance; Fig. 2) was predetermined in the algorithm based on the data obtained from a separate group of 19 healthy controls (37 eyes) with a mean age of 64.3 ± 7.8 years. To measure these distances, we manually identified the peaks of the three hyper-reflective bands (RPE, EZ, and ELM) from the reflectivity profile at each ROI. The average and standard deviation of these distances were computed for each ROI to take into account the spatial variation in these distances (e.g., a wider RPE-EZ distance at the fovea compared with the periphery). Once the EZ and ELM peak were identified, the system captured the peak intensity values within the area of determination and stored them in an Excel file.

Subregion for EZ determination = mean RPE-EZ distance ± 2 standard deviation [pixels].
Subregion for ELM determination = mean RPE-ELM distance ± 2 standard deviation [pixels].

The outcome parameter was the rEZR (AU) determined at each ROI. For the purpose of this study, the rEZR at all ROI areas, including areas with drusen, were included in the agreement analysis. Although structural changes within the outer retina, like drusen, are known to impact the localized reflectivity profiles, the aim of our study was to determine the agreement between the manual and automated method in determining the rEZR, regardless of whether there was a presence or absence of pathology. Therefore, we did not exclude the rEZR data in areas with drusen in this study. By including all the data points, a large range of rEZR data was collected, which allowed a better detection on potential relationships between the difference and the mean rEZR as part of the Bland–Altman analysis. The average rEZR for the total study cohort and subgroups was calculated. Comparison of the rEZR between control and AMD eyes was performed using a linear mixed effect model, adjusted for age and ROI location. To determine accuracy, the peak detection rate of the automated method, correct and incorrect peak identification were identified by visual inspection of the underlying reflectivity plot for each of the 280 ROIs, where the specific subregion for EZ and for ELM determination as well as the peaks were graphically highlighted. The scatter and Bland–Altman plots were used to inspect the agreement characteristics between the manual and automated method and the limits of agreement between the two methods were determined. The cumulative percentage of difference in rEZR between the two methods was also plotted to examine the distribution and magnitude of the discrepancy between the rEZR quantification methods. The percentage of difference in rEZR was calculated by the formula ((difference rEZR ÷ average rEZR) × 100).

Forty eyes of 40 participants were included in this study. A summary of the demographics of these participants is given in the Table. In brief, there were 20 healthy control participants without any AMD signs and 20 participants with iAMD. There was no difference in mean (± standard deviation) age between the iAMD (67.2 ± 8.0 years) and control group (63.1 ± 7.2 years, P = 0.0982). The unadjusted average rEZR
of the control group was significantly greater (manual = 66.6 AU, automated = 66.7 AU) than that of the iAMD group (41.9 vs. 42.0 AU), regardless whether it was measured with a manual or automated method ($P < 0.001$ for both methods, Table). The difference in rEZR between the control and AMD group remained significant after adjusted for age and ROI locations using a linear mixed effect model (marginal means were 69.3 vs. 40.9 AU for the manual method and 68.0 vs. 39.7 AU for the automated method, $P < 0.001$ for both methods). There was no significant difference in the average rEZR between the methods for both the control (mean difference, $-0.07$; $P = 0.992$) and iAMD group (mean difference, $-0.07$; $P = 0.987$, Table).

Using the automated method, the EZ and ELM peaks were correctly detected in 271/280 ROIs, giving the accuracy peak detection rate for the automated method of approximately 96.8%. Of the nine ROIs with incorrectly identified EZ and ELM peaks, two ROIs had incorrectly placed peaks and the other seven ROIs (from one control and six AMD eyes), the algorithm could not detect the EZ or EML peak since these peaks were outside the area of determination. These nine data points were subsequently not included for further data analysis on rEZR and for evaluating the agreement between the manual and automated method. Representative cases of correct and incorrect determination of the peaks are shown in Figure 3.

### Interapproach Agreement

The agreement between the manual and automated method is demonstrated in Figure 4. The scatter and Bland–Altman plots showed a characteristic linear relationship between the manual and automated method and there was no obvious relationship between the magnitude of the average and the difference of the two methods. Overall, the Bland–Altman plots showed that approximately 95% of all test points had a difference in rEZR between the methods of less than 10 AU, though the difference in rEZR in the control group was greater than that of the iAMD group. The limits of agreement between the two methods were $-7.5$ to $7.3$ for all eyes, $-4.8$ to $3.5$ for AMD eyes and $-9.7$ to $9.5$ for control eyes. The cumulative distribution plots showed that 90% of all test points were within 10% variation between the two methods.

### Repeatability

To examine the repeatability of the manual and automated approach, the same dataset was analyzed twice by both methods, being masked to the initial grading results. The agreement between the first and second rEZR measurement is shown in Figure 5. The automated method showed an expected perfect agreement between the first and the second measurement as demonstrated by the scatter and Bland-Altman plot. The manual method exhibited a small variation between the measurements with a coefficient of repeatability of 6.3 AU.

### Discussion

In this study, we investigated an analysis algorithm for automated extraction of the rEZR from SD-OCT images. We found that the automated method not only achieved strong agreement with the manual approach in determining the rEZR ($<10\%$ of measurement variation), it also had a better reproducibility compared with the manual approach. Consistent with previous studies, we found that the rEZR was decreased in iAMD eyes compared with control eyes.15,16

Because it might be difficult to appreciate the degree of agreement between the two methods using the rEZR AU, we have calculated the percentage of the difference
in the rEZR between the methods. Our data showed that 95% of the data points had a difference in rEZR between the methods of ±8.2%. This result means that the rEZR value obtained by the automated method is equal to at least 92% of the rEZR value obtained by the manual method in the majority (95%) of the measurements, indicating a high level of agreement between the two methods. In addition, by using the percentage of the difference in rEZR, measurements for the control and AMD group were also converted to a common scale, which allowed for a direct comparison in measurement variability between the groups. This strategy is used to overcome the potential incompatibility comparison associated with different rEZR ranges between the control and AMD group.

To our knowledge, there has been only one other study on automated quantification of the rEZR. Ha et al.\textsuperscript{23} recently reported the application of automated rEZR quantification in healthy and glaucoma eyes. However, it is inappropriate to make a direct comparison in the performance of the method used in the study Ha et al. with ours owing to several
Quantification of Ellipsoid Zone Reflectivity

Figure 4. Scatter plots and Bland–Altman plots of rEZR showed a strong agreement between the manual and automated method for the total participants, as well as separately for control and iAMD subjects. Cumulative distribution plots showed that approximately 90% of all test points had a rEZR difference between the two methods of less than 10% of its magnitude. The diagonal black lines on the scatter plots represent the expected perfect agreement of the two measurements (line of unity). The grey areas over the ±2 standard deviation lines on the Bland–Altman plots represent the 95% confidence intervals for the upper and lower limit of agreement.

differences in the methodology and study design. First, the study by Ha et al. recruited glaucoma patients with the outer retinal morphology presumably intact, whereas we examined patients with iAMD, all with large drusen and other disease-associated structural changes in the outer retinal layers. Because the EZ and ELM reside in the outer retina, any morphologic changes to the outer retina imposes a challenge in the correct identification and quantification of the intensity of the EZ and ELM. Despite that, our automated method demonstrated an excellent performance with a total failure rate of only 3.2%. With further development of the algorithm we believe that this error rate could be reduced further. Second, our approach processes raw data of the SD-OCT images whereas that of Ha et al. used logarithmic transformed OCT images. The raw OCT images are scaled linearly and, thus, allow obtaining the actual quantitative pixel’s intensity information. Although logarithmic transformed SD-OCT images allow a better visualization of the retinal structures clinically, this contrast enhancement leads to a misrepresentation of real reflectivity signals and must be viewed cautiously in the context of quantitative reflectivity analyses. To optimize peak assessment from the reflectivity profiles, we also implemented a process to ensure the correct alignment of each axial OCT A-scan based on RPE segmentation coordinates to obtain straightened OCT B-scans. This process facilitates not only an accurate EZ and ELM peak analysis in the reflectivity plots by removing any redundant image noise, but also enables rEZR determination in eyes with different morphologies such as eyes with more curved posterior poles.
Robust biomarkers that can identify disease progression in the early stages of AMD are still lacking. The rEZR has been postulated as a surrogate measure of photoreceptor function and a potential indicator for AMD progression. However, to date rEZR has been assessed manually at a few locations on a SD-OCT single line scans. The manual method clearly has numerous constrains and, given the required time to assess the rEZR manually, it is not feasible to apply this method to a large and longitudinal dataset. In this study, it took approximately 10 minutes to manually obtain the rEZR from seven ROIs on a single OCT B-scan. If this manual approach was applied to a volume OCT scan with 49 B-scans and 51 ROIs in each B-scan, it would take approximately 60 hours to complete one volume OCT scan. Owing to feasibility issues associated with the manual method, it has not yet been possible to develop a more refined characterization of the rEZR, including a topographical analysis across the whole macular area, as an innovative AMD biomarker and evaluation of its prognostic value regarding AMD staging and disease progression. As shown in this study, our automated method outperformed the manual approach in terms of efficiency and reproducibility. For any parameters to be used for monitoring disease progression, a high level of reproducibility is of particular important as it allows early detection of any subtle changes. The reliability, efficiency, and reproducibility of the automated rEZR quantification approach support further development and evaluation of the method for monitoring changes in rEZR and the associated disease progression.

There were some limitations associated with the study. First, the sample size of the study was relatively small with 40 participants. However, there were seven discrete data points collected from each study eye, generating a total of 280 data points, which was sufficient for examining the agreement and reproducibility of the two methods. Second, the study only examined the rEZR on single line scan. As mentioned elsewhere in this article, this work was an initial study to test our automated algorithm and it was not feasible to perform manual quantification of the rEZR in such a large volume scans. Because the data from this study showed a high agreement between the automated and manual methods, our next step will be to expand this research to examine the rEZR of the entire volume scan. We believe that the adaptation of the automated algorithm on volumetric OCT data is easy to implement. Because the process of analyzing volumetric data is exactly the same as that used in a single OCT line scan, except that it is repeated on multiple B-scans,
we do not expect any difference in the performance of the automatic method when applied to volumetric OCT data. Further studies are warranted to evaluate the performance of our proposed automated method in determining the rEZR on OCT images acquired from other systems and on other retinal conditions. Third, the coordinates of the segmentation lines were generated semiautomatically in this study; thus, the entire process of obtaining the rEZR was not fully automated because we still needed to perform manual review and correction of OCT segmentation. However, the actual detection of the peaks and measurement of the reflectivity was an automated process. This step was our first in the development toward a fully automated approach. We are working toward implementing artificial intelligence in the process to perform the segmentation automatically in the future. We believe that the implantation of artificial intelligence not only improves the efficiency of the rEZR determination process, but also allows us to obtain the rEZR data from three-dimensional OCT volume scans and in context of large and longitudinal AMD studies. Finally, in this study, the RPE was flattened; the current automated method required a relatively straight horizontal reference line for an accurate detection of the reflectivity peaks. In our experience, this reference line does not have to be absolutely flat for the algorithm to detect the peaks, as demonstrated by a high peak detection rate in this study. The process of flattening the RPE layer inevitably also altered the contour of the retinal layers above it. Although changes in the retinal contour did not seem to affect the rEZR quantification, to overcome this limitation and further improve the accuracy of the peak detection we are working on a new algorithm to detect the coordinates of the EZ and ELM without the need to flatten the RPE layer.

In conclusion, we presented an innovative method for automated determination of the rEZR based on digital image analysis and applicable to AMD affected eyes. The automated method did not only demonstrate a strong agreement with the manual approach, it also showed superior reproducibility. Given this approach’s applicability to longitudinal as well as three-dimensional SD-OCT data, further studies are warranted to precisely evaluate the rEZR as an innovative in vivo biomarker for AMD progression.

Acknowledgments

Supported by the National Health & Medical Research Council of Australia (Project Grant no.: APP1027624 [RHG, CDL]; and fellowship grant nos.: GNT1103013 [RHG], Bupa Health Foundation, Australia (RHG, CDL), BONFOR GEROK Program, Faculty of Medicine, University of Bonn, Grant No O-137.0026 to ST; by the Maria-von-Linden Program, University of Bonn to ST; by the Werner Jackstädt Nachwuchspreis of the German Retina Society to ST and by the BM-AXIS program of the University of Bonn and University of Melbourne to ST. The Centre for Eye Research Australia receives operational infrastructure support from the Victorian Government.

Disclosure: S. Thiele, Heidelberg Engineering (R, F); Optos (F), Zeiss (F), CenterVue (F); Novartis (R); Bayer (R); B. Isselmann, None; M. Pfau, Heidelberg Engineering (F); Optos (F), Zeiss (F), CenterVue (F); F.G. Holz, Acucela (C, F, R), Allergan (F, R), Apellis (C, R), Bayer (C, F, R), Boehringer-Ingelheim (C), Bioeq/Formycon (F, C), CenterVue (F), Ellex (R), Roche/Genentech (C, F, R), Geuder (C), Grayburg Vision (C, R), Heidelberg Engineering (C, F, R), Kanghong (C, F), LinBioscience (C, R), NightStarX (F), Novartis (C, F, R), Optos (F), Pixium Vision (C, F, R), Oxurion (C, R), Stealth BioTherapeutics (C, R), Zeiss (F, R); S. Schmitz-Valckenberg, Acucela/Kubota Vision (F), Novartis (C, F, R), Allergan (C, F, R); Bayer (F, R), Bioeq/Formycon (F, C); Carl Zeiss Meditec (F, R); CenterVue (F); Galimedix (C), Genentech/Roche (F, R); Heidelberg Engineering (F); Katairo (F), Optos (F), Oxurion (C), Roche/Genentech (F, C); Z. Wu, None; R.H. Guymer, Novartis (S), Bayer (S), Roche Genentech (S), Appelis (S); C.D. Luu, None

References


