



Reproducibility of qualitative assessment of drusen volume in eyes with age related macular degeneration

Federico Corvi ^{1,2,3} · Sowmya Srinivas^{1,2} · Muneeswar Gupta Nittala¹ · Giulia Corradetti ^{1,2} · Swetha B. Velaga ¹ · Dwight Stambolian⁴ · Jonathan Haines⁵ · Margaret A. Pericak-Vance ⁶ · Srinivas R. Sadda^{1,2}

Received: 6 June 2020 / Revised: 5 October 2020 / Accepted: 5 November 2020 / Published online: 19 November 2020
© The Author(s), under exclusive licence to The Royal College of Ophthalmologists 2020

Abstract

Background Although an optical coherence tomography (OCT)-derived central drusen volume $\geq 0.03 \text{ mm}^3$ has been found to be a risk factor for progression to late age-related macular degeneration (AMD), this parameter is not currently available on most OCT devices or acquisition protocols. The purpose of this study was to evaluate the ability of human graders to qualitatively assess drusen volume by inspection of OCT B-scans.

Methods 100 subjects (200 eyes) from the Amish Eye Study diagnosed with early or intermediate AMD underwent OCT imaging with both Cirrus OCT and Spectralis OCT. Drusen volume was automatically computed from the Cirrus OCT volumes using the Cirrus Advanced RPE Analysis software. Spectralis volume scans were reviewed by two independent, masked graders who were asked to determine whether the central drusen volume was $\geq 0.03 \text{ mm}^3$. Cohen's kappa coefficients were computed to assess the agreement.

Results After excluding 11 eyes with poor image quality and 5 eyes used for training of the graders, the remaining 184 eyes were included in this analysis. The agreement between the graders and the automated evaluation of drusen volume by the Cirrus OCT was excellent with $K = 0.88$ for grader 1 and $K = 0.82$ for grader 2. The agreement between graders was also excellent with a $K = 0.88$.

Conclusions The presence of a high central drusen volume can be assessed reliably by qualitative inspection of OCT B-scans. This approach may be useful in the assessment of risk for progression to late AMD.

Introduction

Age-related macular degeneration (AMD) is the most common cause of legal blindness among patients older than 50 years of age in the developed countries [1–3]. A number of studies and research groups have evaluated risk factors for progression from early to late AMD, defined as the development of macular neovascularization (MNV) or geographic atrophy (GA) [4–10]. Initial studies identified risk factors from colour fundus photos, and these biomarkers have included features such as drusen area, presence of large drusen, hyperpigmentation and depigmentation [8, 9]. Optical coherence tomography (OCT), however, has largely supplanted colour photography as the dominant imaging technology in clinical practice and clinical trials. A number of groups, including ours, have now identified risk factors for AMD progression on OCT scans. In a longitudinal study of individual drusenoid lesions followed over time, intraretinal hyperreflective foci (IRHF) and hyporefective

✉ Srinivas R. Sadda
ssadda@doheny.org

¹ Doheny Image Reading Center, Doheny Eye Institute, Los Angeles, CA, USA

² Department of Ophthalmology, David Geffen School of Medicine at UCLA, Los Angeles, CA, USA

³ Eye Clinic, Department of Biomedical and Clinical Science “Luigi Sacco”, Sacco Hospital, University of Milan, Milan, Italy

⁴ Department of Ophthalmology, University of Pennsylvania, Perelman School of Medicine, Philadelphia, PA, USA

⁵ Department of Population and Quantitative Health Sciences, Case Western Reserve University, Cleveland, OH, USA

⁶ John P. Hussman Institute for Human Genomics, University of Miami Miller School of Medicine, Miami, FL, USA

drusen cores (hDC) were found to predict which drusen went on to develop atrophy or MNV at that location. Reticular pseudodrusen (RPD), which have now been shown to correspond to subretinal drusenoid deposits (SDD), have also been shown to increase the risk for development of atrophy as well as MNV (especially Type 3). Using drusen volumes automatically computed by Advanced RPE analysis software on the Cirrus OCT, Abdelfattah and colleagues were able to show that a drusen volume $\geq 0.03 \text{ mm}^3$ within the central 3 mm of the macula was associated with a higher risk of progression to late AMD, and in particular GA [11]. The importance of high central drusen volume (based on this cutpoint of 0.03 mm^3) was confirmed in subsequent larger studies [12–14].

Lei and colleagues incorporated high central drusen volume and these other factors (SDD, IRHF, hDC) into a simple OCT-based staging system that was patterned after the AREDS simple scale, and was intended to be used for prognostication in clinical practice or identifying high-risk patients for enrolment into early intervention clinical trials [12]. Drusen volume measurements, however, are not automatically computed by many OCT devices, and even if 3rd party software were to be used for drusen quantification, this would generally require that clinicians had obtained high-density volume scans to allow volumetric analysis. These limitations hinder the broad applicability of these OCT-based risk stratification systems.

To address this limitation, this study evaluated whether subjective inspection of OCT B-scans alone could be used to determine if a patient's eye contained a drusen load above or below the high-risk threshold of 0.03 mm^3 .

Materials and methods

This study was approved by the Institutional Review Boards of the University of California Los Angeles, the University of Miami, and the University of Pennsylvania. All procedures were conducted in accordance with the Declaration of Helsinki and informed consent was obtained by all subjects before being enrolled.

We randomly selected 100 subjects (200 eyes) enrolled in the Amish Eye Study who were diagnosed with early or intermediate AMD in both eyes based on review of colour fundus photographs and OCT. The Amish Eye Study is an ongoing NEI-supported clinical research investigation aimed at identifying novel phenotypic and genetic biomarkers for evaluating AMD in an elderly population (individuals over the age of 50). The baseline features, including the OCT characteristics, were described in detail in a previous report [13].

Exclusion criteria for the study were evidence on OCT or colour photos of late AMD defined by the presence of MNV or RPE atrophy, the presence of any other retinal disease other than AMD, poor OCT image quality due to significant media opacity and signal strength < 6 which is the minimum level specified by the manufacturer in order to produce reliable drusen analysis. All subjects were imaged on the same day with two SD-OCT devices: (A) Cirrus OCT (Carl Zeiss Meditec, Inc., Dublin, CA) with a $6 \times 6 \text{ mm}$ macular cube protocol of 128 horizontal B-scans and 512 A-scans per B-scan and (B) Heidelberg Spectralis HRA + OCT (Heidelberg Engineering, Heidelberg, Germany) with a $6 \times 6 \text{ mm}$ macular cube protocol of 97 horizontal B-scans and 1024 A-scans per B-scan.

Grading procedure

Optical coherence tomography data for each subject were exported in a coded de-identified fashion for subsequent analysis at the Doheny Image Reading Center (DIRC). After exportation and transmission to DIRC, OCT images were reviewed for qualitative features and analysed for quantitative measures by certified OCT graders using the Cirrus Advanced RPE Analysis software (Version 6.0). The method used by the automated Advanced RPE Analysis software to quantify drusen volume has been described in previous reports [15, 16]. Briefly, the algorithm segments the RPE band following the RPE contour and also estimates the original RPE position (termed the RPE fit) in the absence of an RPE elevation. The region between the RPE band and the RPE fit is then quantified as the drusen volume. Automatic drusen volumes within the 3-mm circle (centred on the fovea) of the Cirrus grid was recorded by a senior independent grader (MN). As drusen volumes were calculated by automated FDA-cleared instrument software, and the accuracy and reliability of this software have been assessed in prior reports [15, 16], the reproducibility of this “gold-standard” metric was not re-assessed in this study.

A drusen volume training set (five cases) was chosen by the senior grader (MN) from the cases and had both Spectralis and Cirrus OCT data with known drusen volume values from the Cirrus. The training set was reviewed by two other OCT graders (FC and SS) to establish subjective recognition for the appearance on Spectralis OCT B-scans of a case which met the threshold drusen volume on the Cirrus OCT of “high central drusen volume” (Fig. 1). After this training, the two graders reviewed the cases collected for this study in an independent, masked fashion. The estimation of drusen volume was performed evaluating whether the central drusen volume (within the central 3 mm) was either < 0.03 or $\geq 0.03 \text{ mm}^3$. Thus, to assess central drusen volume, the graders only considered those B-

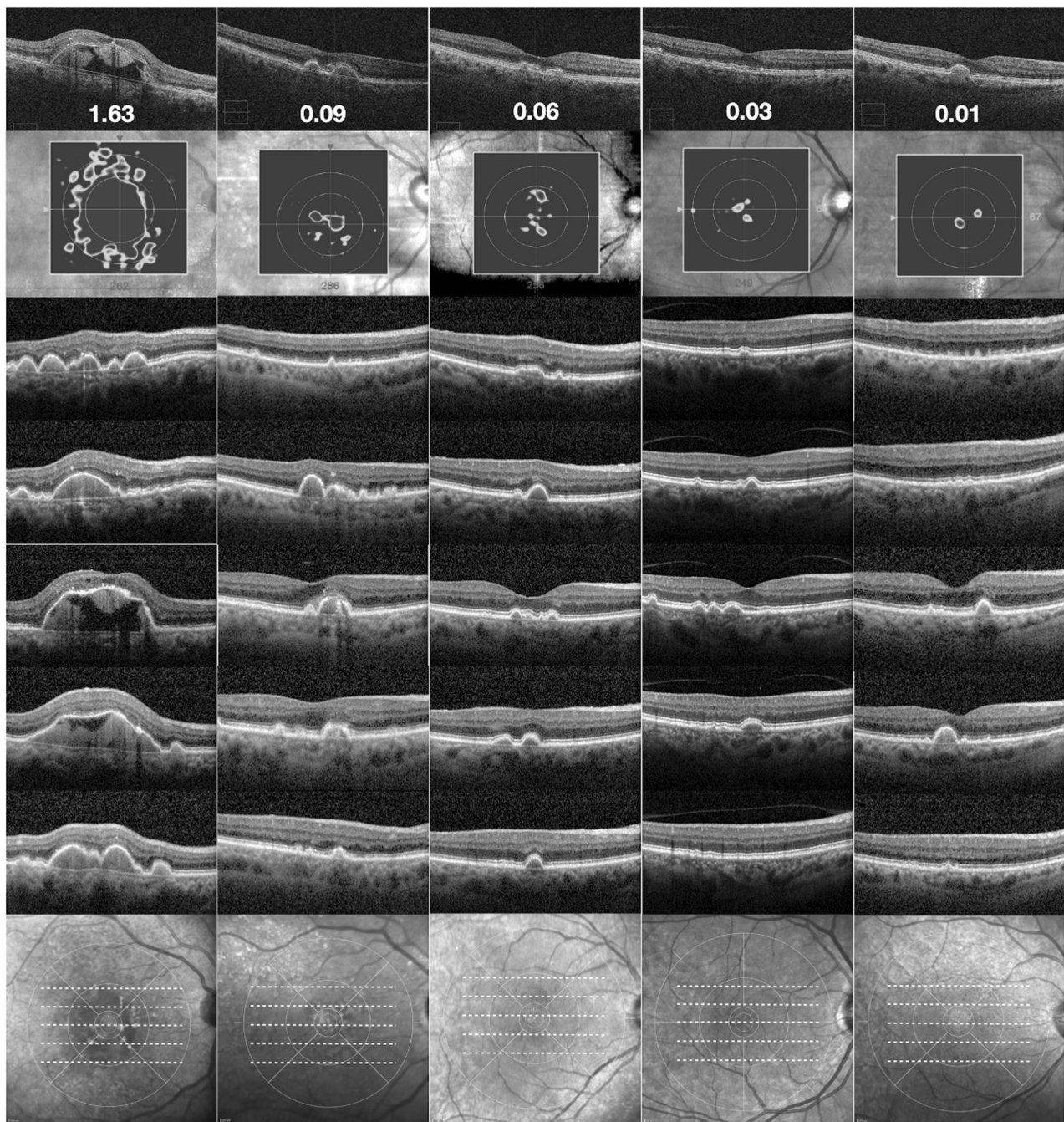


Fig. 1 Five cases of intermediate AMD used for training graders for drusen volume assessment. The first row shows a structural OCT B-scan from the Cirrus OCT volume showing the inner and outer segmentation boundaries at the level of RPE and RPE fit (estimated Bruch's membrane position), respectively. The drusen volume measurement (in mm^3) as calculated by the Cirrus Advanced RPE Analysis is shown (ranging from 1.63 to 0.01 mm^3). The second row shows the corresponding drusen volume maps produced by the Cirrus Advanced

scans (or portions of B-scans) which they deemed to fall within the central 3 mm (based on the middle ring of the Early Treatment of Diabetic Retinopathy Study (ETDRS) grid) of the macula.

RPE Analysis software, highlighting the drusen which contributed to the drusen volume. The third through seventh row shows representative B-scans from the Spectralis OCT within the central 3 mm of the ETDRS grid. The bottom row shows the Spectralis infrared reflectance images with locations of the selected B-scans shown as dashed lines. AMD age-related macular degeneration, OCT optical coherence tomography, RPE retinal pigment epithelium, ETDRS Early Treatment Diabetic Retinopathy Study.

Statistical analysis

Statistical analysis was performed with IBM SPSS Statistics for Windows, Version 26.0 (IBM Corporation, Armonk,

NY, USA). Drusen volume was measured in millimetre cubed, and the results were expressed as mean ± SD. Cohen’s kappa (κ) was calculated to measure the agreement between two graders and the agreement between graders and automatic drusen volumes.

Results

A total of 200 eyes from 100 subjects with early or intermediate AMD were included in this study. Fifty-eight (58%) were women with a mean age of the population of 77.1 years (SD, 9.5), with a median age of 79 years (range, 53–99 years). Among these 200 eyes, 11 were excluded due to poor image quality, 5 were used for initial training of the graders, and thus ultimately 184 eyes were included in the analysis.

The mean drusen volume was $0.05 \pm 0.15 \text{ mm}^3$, and 67 eyes (36%) were measured to have a drusen volume $\geq 0.03 \text{ mm}^3$ and 117 (64%) had a volume lower than 0.03 mm^3 . The distribution of drusen volumes among all cases is shown in the histogram in Fig. 2. The agreement between the graders’ subjective inspection of Spectralis B-scans and the automated Cirrus Advanced RPE Analysis for the presence of high central drusen volume ($\geq 0.03 \text{ mm}^3$) was excellent with a $K = 0.92$ (95% confidence intervals (CI) = 0.86–0.98) for grader 1 and $K = 0.84$ (95% CI = 0.76–0.93) for grader 2 (Table 1). The agreement between graders for determination of high central drusen volume was excellent with a $K = 0.88$ (exact 95% confidence intervals, 0.81–0.95). Most cases where there was either a disagreement between graders or a disagreement between both

graders and the automated drusen volume measurement were close to the threshold drusen volume (0.03 mm^3 ; Fig. 2).

Discussion

In this study, we evaluated the ability of two independent graders to qualitatively determine whether central drusen volume was $\geq 0.03 \text{ mm}^3$ based on rapid inspection of the Spectralis B-scans alone. We found an excellent agreement between the two graders and the drusen volume automatically computed by the Cirrus OCT using the Cirrus Advanced RPE Analysis software. These findings suggest that a qualitative determination of high central drusen volume may be sufficient and could potentially be used in the estimation of the risk for progression to advanced AMD.

The importance of assessing drusen volume and in particular determining the presence of high central drusen volume has been established by several studies. Baseline drusen volume appears to be an important predictor for the development of late AMD within 2 years, suggesting that OCT-derived drusen volume measurements may be a useful biomarker to identify eyes at the highest risk for progression to late AMD. In particular, Abdelfattah et al. found in patients with AMD, that a drusen volume $>0.03 \text{ mm}^3$ was associated with 4-fold increased risk for developing late AMD compared with those with lower values [11]. Nassisi et al. investigated the relationship between several OCT features, including drusen volume, and progression to late AMD in the fellow eyes of patients enrolled in the Study of Ranibizumab Administered Monthly or on an As-needed Basis in Patients with Subfoveal Neovascular AMD (HARBOR trial), finding a hazard ratio of 1.46 for high central drusen volume [14].

Drusen volume in these previous studies was based on an automatic analysis performed by the Cirrus OCT using the Advanced RPE Analysis software which has been cleared by the FDA. The automated software algorithm relies on the detection of the retinal pigment epithelium (RPE) surface and an estimated fit of the original or expected RPE assuming no elevation within the 3- and 5-mm circles centred on the fovea. Previous studies have demonstrated that OCT-based automated drusen quantification is

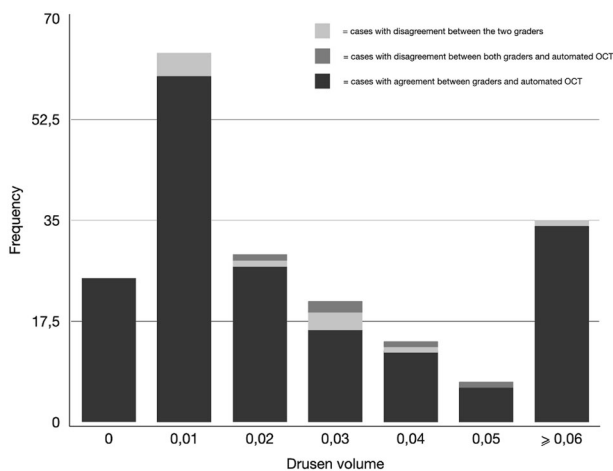


Fig. 2 Distribution of drusen volumes in study cohort. The frequency of drusen volumes (binned in 0.001 mm^3 intervals) is plotted in this histogram. The numbers of cases where the graders disagreed with the automated drusen volume analysis or the graders disagreed between themselves is shown. Note that most cases of disagreements were near the 0.03 mm^3 threshold.

Table 1 Agreement between graders.

| | Grader 1–Cirrus | Grader 2–Cirrus | Grader 1–Grader 2 |
|----------------|-----------------|-----------------|-------------------|
| Cohen’s kappa | 0.92 | 0.84 | 0.88 |
| 95% CI | 0.86–0.98 | 0.76–0.93 | 0.81–0.95 |
| Standard error | 0.03 | 0.04 | 0.03 |

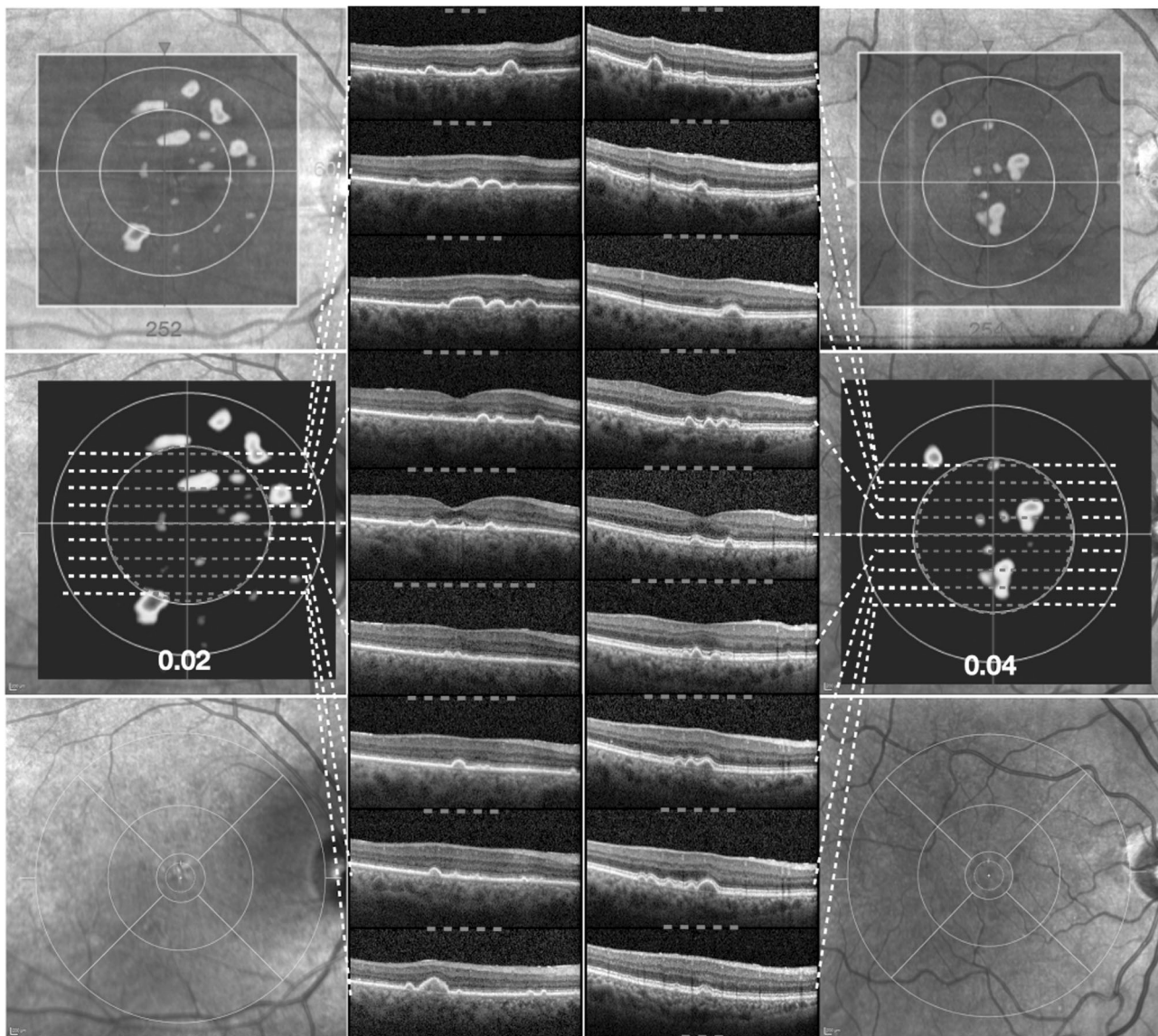


Fig. 3 Two representative cases of disagreement between graders and automated software. On the left, a case where both graders agreed that a high central the drusen volume was present but was not confirmed by the automated analysis, as the value was 0.02 mm^3 . On the right, a case where both graders agreed that the drusen volume was below the 0.03 mm^3 threshold, but the automated analysis computed a value of 0.04 mm^3 . The top right and left images show the Cirrus

infrared reflectance image with superimposed drusen volume maps from the Cirrus OCT. The middle right and left images show the superimposed drusen volume maps on the Spectralis infrared reflectance image with the red and white dashed lines indicating the locations of the Spectralis B-scans for the two cases which are shown in the central two columns.

reproducible [16–18]. A challenge for relying on automated drusen volume measurements for use in multi-centre clinical trials, clinical research or clinical practice, however, is that it is limited to the few OCT devices that perform this measurement. Most OCT devices do not automatically generate volume and even off-line image processing software tools will generally require a dense volume scan. Our study, however, shows that after a very brief training period (review of only 5 cases), human graders are able to reliably determine the presence of high central drusen volume. This is important as it may also allow this risk factor to be

assessed in many retrospective/historical OCT datasets where automated drusen volume computation may not be possible.

A post hoc review of cases in which there were disagreements in drusen volume assessment was conducted to determine if there were opportunities to further improve the subjective evaluation. There was only one case where both graders agreed that a high central drusen volume was present but was not confirmed by the automated analysis, which measured a volume of only 0.02 mm^3 (Fig. 3). In four cases, both graders agreed the drusen volume was below the

threshold but the automated software measured a value $\geq 0.03 \text{ mm}^3$ —the mean automated drusen volume in these cases was 0.037 (range: 0.03–0.05). In 10 cases, the graders disagreed on their assessment of the presence of high central drusen volume. Overall, among these 15 cases in which there was either disagreement with the automated measurement or between the two graders, the mean automated drusen volume in these cases was 0.029. As 12 of these 15 cases had a drusen volume close to the 0.03 mm^3 threshold value (range between 0.02 and 0.04 mm^3), we suspect that this reflects the limits of the precision of the human grader to subjectively discriminate drusen volumes. On the other hand, the exact 0.03 mm^3 cutpoint is somewhat arbitrary, so missing a few cases close to threshold may be acceptable for general clinical use. Overall, in the cohort, there were a total of 54 cases (30%) with drusen volume between 0.02 and 0.04 mm^3 .

Our study is not without limitations, which should be considered when assessing our results. First, graders were trained before starting the drusen volume assessments. On the other hand, the training was brief and based on only five representative cases, which suggests that most clinicians should be able to learn to do this quickly. Second, although this approach appears to be useful for determining whether a particular case meets the threshold for a pre-defined high central drusen volume, it does not provide a precise quantitative measurement. While the categorical determination is sufficient for use of the current OCT-based simple scale for AMD severity, future more precise AMD severity scales may utilize drusen volume as continuous measure.

In summary, we observed that the presence of high central drusen volume could be determined reliably by a qualitative inspection of OCT B-scans. This approach may be feasible for use in clinical practice and in clinical trials to assess the risk for progression to late AMD.

Summary

What was known before

- Although an optical coherence tomography (OCT)-derived central drusen volume $\geq 0.03 \text{ mm}^3$ has been found to be a risk factor for progression to late age-related macular degeneration (AMD), this parameter is not currently available on most OCT devices or acquisition protocols.

What this study adds

- The purpose of this study was to evaluate the ability of human graders to qualitatively assess drusen volume by inspection of OCT B-scans. The presence of a high

central drusen volume can be assessed reliably by qualitative inspection of OCT B-scans. This approach may be useful in the assessment of risk for progression to late AMD.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

References

1. Schmitz-Valckenberg S, Steinberg JS, Fleckenstein M, Visvalingam S, Brinkmann CK, Holz FG. Combined confocal scanning laser ophthalmoscopy and spectral-domain optical coherence tomography imaging of reticular drusen associated with age-related macular degeneration. *Ophthalmology*. 2010;117:1169–76.
2. Abdelsalam A, Del Priore L, Zarbin MA. Drusen in age-related macular degeneration: pathogenesis, natural course, and laser photocoagulation-induced regression. *Surv Ophthalmol*. 1999;44:1–29.
3. Schlanitz FG, Ahlers C, Sacu S, Schütze C, Rodriguez M, Schriefl S, et al. Performance of drusen detection by spectral-domain optical coherence tomography. *Invest Ophthalmol Vis Sci*. 2010;51:6715–21.
4. Bressler SB, Maguire MG, Bressler NM, Fine SL. Relationship of drusen and abnormalities of the retinal pigment epithelium to the prognosis of neovascular macular degeneration. The Macular Photocoagulation Study Group. *Arch Ophthalmol*. 1990;108:1442–7.
5. Pauleikhoff D, Barondes MJ, Minassian D, Chisholm IH, Bird AC. Drusen as risk factors in age-related macular disease. *Am J Ophthalmol*. 1990;109:38–43.
6. Ferris FL, Davis MD, Clemons TE, Lee LY, Chew EY, Lindblad AS, et al. A simplified severity scale for age-related macular degeneration: AREDS Report No. 18. *Arch Ophthalmol*. 2005;123:1570–4.
7. The Age-Related Eye Disease Study Research Group. The Age-Related Eye Disease Study severity scale for age-related macular degeneration: AREDS Report No. 17. *Arch Ophthalmol*. 2005;123:1484–98.
8. Klein R, Klein BEK, Knudtson MD, Meuer SM, Swift M, Gangnon RE. Fifteen-year cumulative incidence of age-related macular degeneration: the Beaver Dam Eye Study. *Ophthalmology*. 2007;114:253–62.
9. The Age-Related Eye Disease Study Research Group. The Age-Related Eye Disease Study system for classifying age-related macular degeneration from stereoscopic color fundus photographs: AREDS Report Number 6. *Am J Ophthalmol*. 2001;132:668–81.
10. Lamin A, El Nokrashy A, Chandra S, Sivaprasad S. Association of longitudinal changes in drusen characteristics and retinal layer volumes with subsequent subtype of choroidal neovascularisation. *Ophthalmic Res*. 2020;63:375–82.
11. Abdelfattah NS, Zhang H, Boyer DS, Rosenfeld PJ, Feuer WJ, Gregori G, et al. Drusen volume as a predictor of disease progression in patients with late age-related macular degeneration in the fellow eye. *Invest Ophthalmol Vis Sci*. 2016;57:1839–46.

12. Lei J, Balasubramanian S, Abdelfattah NS, Nittala MG, Sadda SR. Proposal of a simple optical coherence tomography-based scoring system for progression of age-related macular degeneration. *Graefes Arch Clin Exp Ophthalmol*. 2017;255:1551–8.
13. Nittala MG, Song YE, Sardell R, Adams LD, Pan S, Velaga SB, et al. AMISH EYE STUDY: baseline spectral domain optical coherence tomography characteristics of age-related macular degeneration. *Retina*. 2019;39:1540–50.
14. Nassisi M, Lei J, Abdelfattah NS, Karamat A, Balasubramanian S, Fan W, et al. OCT risk factors for development of late age-related macular degeneration in the fellow eyes of patients enrolled in the HARBOR study. *Ophthalmology*. 2019;126:1667–74.
15. Gregori G, Wang F, Rosenfeld PJ, Yehoshua Z, Gregori NZ, Lujan BJ, et al. Spectral domain optical coherence tomography imaging of drusen in nonexudative age-related macular degeneration. *Ophthalmology*. 2011;118:1373–9.
16. Nittala MG, Ruiz-Garcia H, Sadda SR. Accuracy and reproducibility of automated drusen segmentation in eyes with non-neovascular age-related macular degeneration. *Invest Ophthalmol Vis Sci*. 2012;53:8319–24.
17. Freeman SR, Kozak I, Cheng L, Bartsch DU, Mojana F, Nigam N, et al. Optical coherence tomography-raster scanning and manual segmentation in determining drusen volume in age-related macular degeneration. *Retina*. 2010;30:431–5.
18. Chiu SJ, Izatt JA, O'Connell RV, Winter KP, Toth CA, Farsiu S. Validated automatic segmentation of AMD pathology including drusen and geographic atrophy in SD-OCT images. *Invest Ophthalmol Vis Sci*. 2012;53:53–61.