Performance of Drusen Detection by Spectral-Domain Optical Coherence Tomography

Ferdinand G. Schlanitz, Christian Ahlers, Stefan Sacu, Christopher Schütze, Marcos Rodriguez, Sabine Schriefl, Isabelle Golbaz, Tobias Spalek, Geraldine Stock, and Ursula Schmidt-Erfurth

PURPOSE. To evaluate the performance of automated analyses integrated in three spectral-domain optical coherence tomography (SD-OCT) devices to identify drusen in eyes with early (i.e., nonatrophic and nonneovascular) age-related macular degeneration (AMD).

METHODS. Twelve eyes of 12 AMD patients, classified as AREDS 2 and 3 and having a mean count of 113 drusen were examined with three clinical SD-OCT devices (Cirrus [Carl Zeiss Meditec, Dublin CA], 3DOCT-1000 [Topcon, Tokyo, Japan], and Spectralis [Heidelberg Engineering, GmbH, Heidelberg, Germany]) and five different scan patterns. After standard automated segmentation of the RPE was performed, every druse in each B-scan was identified and graded by two independent expert graders. Errors in the segmentation performance were classified as negligible, moderate, or severe. Correlations were based on the diameter and height of the druse and its automated segmentation. The overall drusen pattern identified by experts’ detailed delineation was plotted with a custom-made computer program to compare automated to manual identification outcomes.

RESULTS. A total of 1356 drusen were analyzed. The automated segmentation of the retinal pigment epithelium (RPE) by Cirrus made significantly fewer errors in detecting drusen than did the 3DOCT-1000 (P < 0.001). The Cirrus 200 × 200 scan pattern detected 30% of the drusen with negligible errors. Spectralis did not offer a true RPE segmentation. The drusen were identified by expert graders were significantly higher in the scans than in the standard fundus photographs (P < 0.05).

CONCLUSIONS. SD-OCT imaging proved an excellent performance in visualizing drusen-related RPE disease. However, the available automated segmentation algorithms showed distinct limitations to reliable identification of the amount of drusen, particularly smaller drusen, and the actual size. (Invest Ophthalmol Vis Sci. 2010;51:6715–6721) DOI:10.1167/iovs.10-5288

Age-related macular degeneration (AMD) is the leading cause of irreversible and severe vision loss in people aged 50 years or older in the developed world.1–2 One of its clinical characteristics and, in most cases, the first clinical finding is the presence of drusen.3 Drusen are focal deposits of extracellular material located between the basal lamina of the retinal pigment epithelium (RPE) and the inner collagenous layer of Bruch’s membrane.4 Clinically, drusen appear as yellowish white deposits underneath the reddish appearance of the retina in biomicroscopy. Drusen can be divided roughly into two main morphologic groups, hard and soft, depending on the appearance of their marginal zones.5 The presence of small, hard drusen with a diameter of <63 μm is a common finding in all age groups and therefore is not seen as a risk factor for the development of AMD.6–9 On the other hand, the occurrence of soft drusen is one of the characteristics of age-related maculopathy and is widely viewed as a prestage of AMD.10 Consequently, the drusen diameter and area, but not their amount, have been identified as a risk factor for progression to advanced AMD.11–13 However, the Age-related Eye Disease Study (AREDS) showed that the agreement between graders examining these two parameters was only moderate.14 Furthermore, it is obvious that a time-consuming manual analysis identifying and counting individual drusen on photographs is inadequate for clinical practice.

Increasing availability of spectral-domain optical coherence tomography (SD-OCT) devices, which are able to obtain a three-dimensional dataset of the retina and retinal pigment epithelium (RPE) including most of its pathologic alterations like drusen, has significantly changed clinical practice. This is especially true of the evaluation of advanced forms of AMD, such as the exudative type.15–17 To be able to use SD-OCT datasets in clinical practice, various compartments that can be observed in the single B-scans contained in a complete dataset of the retina are segmented by computerized algorithms in a first step. Data are then interpolated and presented in three dimensions to allow for realistic quantitative analysis in the whole macular region. The purpose of this is to identify anatomic structures and to measure retinal parameters relevant for diagnosis, therapy decisions, and monitoring.18–20

As drusen play a substantial role in the progression to advanced AMD,21 clinicians have begun to look for useful segmentation programs that can determine drusen size and area automatically, to clearly identify patients with a high-risk profile according to the AREDS classification and to detect progression rates reliably. The introduction of a prophylactic therapy for patients with significantly greater individual risk would markedly enhance the clinical and economic benefit of intervention and would help to reduce potential side effects of antiangiogenic treatment in advanced disease. Clearly, preventive strategies can contribute to an improved and long-lasting conservation of the patient’s visual acuity, since a relative risk reduction of approximately 25% over a period of 5 years was shown for the prophylactic substitution of vitamins in the AREDS study.22
The diagnostic performance of OCT imaging systems depends on various factors, such as imaging quality, the individual anatomy of the eye and the pathologic state of the ocular structures.\textsuperscript{25,24}

In this study, we therefore evaluated the performance of segmentation procedures integrated in current SD-OCT devices in detail to identify drusen in patients with early, (i.e., nonatrophic and nonneovascular) AMD. The evaluation of the current status of automated SD-OCT drusen analysis is of significant clinical relevance since nonneovascular AMD has increasingly become an important field of research with direct implications for current clinical care.

**Methods**

**Patients**

To avoid a possible impact of pathologic alterations other than focal drusen on the segmentation procedure, patients were selected according to the standard AREDS classification. Consecutive patients with drusen classified as AREDS 2 or 3 were included. Category 2 is defined by the presence of extensive small drusen (<63 μm), nonextensive intermediate drusen (≥63, <125 μm), or pigment epithelial abnormalities in at least one eye. Category 3 includes large drusen (≥125 μm), extensive intermediate drusen, or noncentral geographic atrophy in at least one eye.\textsuperscript{25} The eyes were examined ophthalmoscopically by an experienced retinologist.

Patients with diseases that could potentially influence the scan quality, such as clouding of the ocular media by cornea, lens, or presence of vitreal opacities, were excluded. Patients with history of ocular trauma, any kind of macular edema or surgery other than uncomplicated cataract surgery and patients with refractive error greater than +3.0 D or less than −7.0 D were also excluded from the study before the study examinations.

**Data Acquisition**

After giving informed consent, patients meeting the protocol criteria were included in the study and underwent a standardized examination procedure according to a protocol that was approved by the local ethics committee and was conducted according to the guidelines in the Declaration of Helsinki. The bestcorrected visual acuity (BCVA) of the patient was obtained and mydriatic eye drops were administered. At maximum mydriasis, a standard three-field digital monoscopic photography of the fundus (FF 450plus; Carl Zeiss Meditec, Inc., Dublin, CA) was taken, and the patients underwent examination with three clinical SD-OCT devices, the Cirrus (Carl Zeiss Meditec, Inc., Dublin, CA), the 3DOCT-1000 (Topcon, Tokyo, Japan), and the HRA Spectralis (Heidelberg Engineering GmbH, Heidelberg, Germany). With Cirrus, a 512 × 128 scan as well as a 200 × 200 scan was performed; with 3DOCT-1000 the 512 × 128 and the 256 × 256 scan pattern was used; and with Spectralis, eyes underwent a volume scan with 37 images composed of 30 frames at a scan angle of 20° and a 1024-pixel axial resolution. To determine the scan quality, we used the scales given by the acquisition software of each device: the signal strength of Cirrus (range, 1–10), the quality scale of Spectralis (in dB), and the imaging quality of 3DOCT-1000 (range, 1–100). Scans that were assessed as being of unacceptable quality (e.g., motion artifacts) were repeated until satisfactory results could be obtained. To compare the scan quality of the different devices, we subdivided the quality scales into low (Cirrus, 1–4; Spectralis, 1–19 dB; Topcon, 1–30), medium (Cirrus, 5–8; Spectralis, 20–25 dB; Topcon, 31–40), and high (Cirrus, 9–10; Spectralis, 26+ dB; Topcon, 41–100) quality.

For a precise delineation of drusen, segmentation of the RPE was chosen, as a subtle focal bending of this reflective layer is one of the characteristic morphologic findings in OCT imaging of drusen.\textsuperscript{26,27}

The review software of Cirrus (ver. 3.0.0.64) and 3DOCT-1000 (ver. 2.0.0) automatically performed a complete segmentation of the RPE in each B-scan by plotting an interpolated line at the area of highest intensity of the constituent A-scans. Because the latest Spectralis software (Heidelberg Eye Explorer, ver. 1.6.1.0, Acquisition Software 4.0.0.0) does not offer an integrated RPE segmentation program, the automated identification of Bruch’s membrane was substituted.

**Data Analysis**

After the processing of each volume scan, two experienced examiners identified each druse on every B-scan, delineating the true contour of the same structure.\textsuperscript{22,28} Each localized drusenoid pigment epithelial detachment was called a druse.

The beginning, ending (in pixels) and the maximum height (in micrometers) of every druse were documented, as well as its identification by the automated segmentation line. Errors in the delineation were graded in their relation to the drusen’s true dimensions by a standardized scheme described in Table 1 for all patients (Fig. 1).

Errors of grade 1 were labeled as negligible and 2 to 5 as moderate. According to the histologic definition of drusen and other studies in which drusen were identified on OCT images,\textsuperscript{6,29} each localized drusenoid pigment epithelial detachment was called a druse.

The beginning, ending (in pixels) and the maximum height (in micrometers) of every druse was documented, as well as its identification by the automated segmentation line. Errors in the delineation were graded in their relation to the drusen’s true dimensions by a standardized scheme described in Table 1 for all patients (Fig. 1).

Errors of grade 1 were labeled as negligible and 2 to 5 as moderate. According to the histologic definition of drusen and other studies in which drusen were identified on OCT images,\textsuperscript{6,29} each localized drusenoid pigment epithelial detachment was called a druse.

### Table 1. Error Classification of the Automatic Delineation Errors in Their Relation to the Drusen’s True Height versus True Diameter, as Identified by Experienced Examiners

<table>
<thead>
<tr>
<th>Error Classifications and Grade</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Negligible</td>
<td>[h - TH] &lt; 1/3TH and</td>
</tr>
<tr>
<td>Grade 1</td>
<td></td>
</tr>
<tr>
<td>Grade 2</td>
<td>[h - TH] &lt; 1/3TH and 1/5TD &lt;</td>
</tr>
<tr>
<td>Grade 3</td>
<td></td>
</tr>
<tr>
<td>Moderate</td>
<td>Two or more drusen delineated as one (i.e., no interruption in the continuity of the segmentation line)</td>
</tr>
<tr>
<td>Grade 4</td>
<td>1/3TH &lt;</td>
</tr>
<tr>
<td>Moderate if diameter of druse &lt;63 μm</td>
<td>[h - TH] &gt; 1/2TH and/or</td>
</tr>
<tr>
<td>Grade 6</td>
<td></td>
</tr>
<tr>
<td>Severe if diameter of druse &gt;63 μm</td>
<td></td>
</tr>
</tbody>
</table>

The error by d is mostly larger, as it is usually two-sided. h, delineated height; d, delineated diameter; TH, true height; TD, true diameter.

The statistical methods and interpretation of data were performed with SPSS, version 16.0. The x² test, the t-test, the Wilcoxon test, the one-way-ANOVA, and the post-hoc test of Duncan were used in the analyses (SPSS, ver. 16.0, 2010, Vol. 51, No. 12).
Out of each drusen coordinate (width, height, and number of the B-scan), as depicted in manual records, a set of points was built and allocated in a matrix. A custom-made software program (MatLab) displayed this matrix as an image connatural to a pseudo-SLO image, using a height-correlated color map to sufficiently visualize the drusen for clinical analysis (Fig 2).

Drusen in the photographs were counted by enlarging the digital images on a 19-in. TFT monitor. Each yellowish spot was carefully marked within the area scanned by OCT in a subsequent step.

RESULTS

Twelve eyes of 12 consecutive patients were examined. Six patients were female, and the mean age was 74.7 years (SD 8.9; range, 62–90). The mean BCVA was 20/25 Snellen (0.94 metric) (SD 0.23; range, 20/40 –20/15). Two patients were pseudophakic. The mean number of the drusen counted in standard photographs of the fundus was 113 (SD 45; range, 40 –195); a total of 1356 individual drusen were analyzed. The observed frequency of the scan quality grades can be seen in Table 2 for each OCT-device and scan pattern.

The interpolated segmentation of the RPE performed by Cirrus made significantly fewer errors in detecting drusen than did the 3DOCT-1000 (P < 0.001). The Cirrus 200 × 200-scan algorithm detected 30% of the identified drusen with negligible errors (18% in the 512 × 128 scan), whereas the 3DOCT-1000 512 × 128 scan algorithm detected only 8.6% of the drusen with negligible errors (3% in the 256 × 256 scan; Fig. 3).

The proportion of severe errors was least with both Cirrus scan rasters, with optimal performance obtained with the 200 × 200 pattern. The 3DOCT-1000 device demonstrated a less accurate performance, with a mean 29.2% rate of severe segmentation errors for both scan patterns.

The Spectralis instrument performed poorly, as its software does not offer true segmentation of the RPE. The substituted analysis of the instrument’s identification of Bruch’s membrane proved to be unsuitable for identifying drusen in the retina. However, due to the high resolution of the B-scans, the overall amount of small drusen identified by the expert examiners in the individual Spectralis B-scans was significantly higher than in the scans of the other devices (P < 0.0001). Of the drusen identified in the Spectralis data-sets, 31.4% were considered to be small versus 15.1% of those detected by Cirrus and 3DOCT-1000. The high potential of Spectralis for detecting small retinal alterations explains the percentage of 42.6% of moderate errors, as we graded a missed detection of a small druse only as a moderate error. As shown in Figure 4, the nondetection of small drusen was nearly completely responsible for the proportion of moderate errors by Spectralis.

For Cirrus and 3DOCT-1000, both grades 4 and 5, which were applied if the segmentation algorithm overlapped two or more drusen and delineated them as one (grade 4) or if the delineated druse was approximately double the diameter and half the height of the real druse (grade 5; 40.0% versus 27.5% over all devices, excluding Spectralis), accounted for most of the moderate errors, followed by missed detection of small drusen (21.9%).

Table 2. Observed Frequency of the Scan Quality Grades for Each OCT Device and Scan Pattern

<table>
<thead>
<tr>
<th>Scan Quality</th>
<th>High</th>
<th>Medium</th>
<th>Low</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cirrus 512 × 128</td>
<td>45.5</td>
<td>54.5</td>
<td>0.0</td>
</tr>
<tr>
<td>Cirrus 200 × 200</td>
<td>18.2</td>
<td>63.6</td>
<td>18.2</td>
</tr>
<tr>
<td>3DOCT-1000 512 × 128</td>
<td>33.3</td>
<td>50.0</td>
<td>16.7</td>
</tr>
<tr>
<td>3DOCT-1000 256 × 256</td>
<td>25.0</td>
<td>25.0</td>
<td>50.0</td>
</tr>
<tr>
<td>Spectralis 1027</td>
<td>25.0</td>
<td>75.0</td>
<td>0.0</td>
</tr>
</tbody>
</table>

Data are percentages of scans in each quality category.

FIGURE 1. Examples for negligible (grade 1), moderate (grades 4 and 5), and severe errors (grade 6, if druse >63 μm).

FIGURE 2. Development of a drusen map. (A) Fundus photograph of a 74-year-old man, showing the area mapped in a 3DOCT-1000 512 × 128 scan. On each B-scan (B), the beginning, end, and height of each druse was documented and plotted on a corresponding line in the map (C).
Furthermore, the calculated ratio of the height to diameter was significantly higher in delineation errors classified as moderate (0.25 vs. 0.16 for negligible and 0.19 for severe errors; \( P < 0.001 \)). With decreasing scan quality (i.e., decreasing of illumination and/or increasing of noise), there was a slight, but significant, deterioration in the performance of all devices, including Spectralis, for which there was a decrease in the number of identified small drusen (\( P < 0.05 \)).

By Cirrus, the median diameter of a successfully detected druse was 270 \( \mu \text{m} \) (SD \( \pm 192.4 \); range, 30–1470) in the 200 \( \times \) 200 scan pattern and 351.6 \( \mu \text{m} \) in the 512 \( \times \) 128 scan pattern (SD \( \pm 209.76 \); range, 47–1455). With the 3D-OCT-1000 device, the values were 340 \( \mu \text{m} \) for the 512 \( \times \) 128 scan pattern (SD \( \pm 55.4.26 \); range, 30–1450) versus 353 \( \mu \text{m} \) in the 256 \( \times \) 256 scan pattern (SD \( \pm 434.49 \mu \text{m} \); range, 80–1440; Table 3). These differences proved to be significant (\( P < 0.001 \)).

Overall counts of drusen plotted on the maps generated by manual delineation were significantly higher (mean, 179 \( \pm 59 \); range, 44–230) than in the standard photographs (mean, 113 \( \pm 45 \); range, 40–195; \( P < 0.05 \)). In 4 of 12 eyes, the total number of drusen identified was nearly twice as many as were counted in the standard fundus photographs.

**DISCUSSION**

In this study, we examined the performance of the automated segmentation procedures integrated in three commercially available SD-OCT devices and different scan patterns, to identify drusen in patients with early AMD (i.e., no atrophic or neovascular features).

As a result, the current intensity-based delineation algorithms for the RPE showed significant limitations in performing this task. As seen in Figures 5D–F, mostly large drusen \( >300 \mu \text{m} \) in diameter were delineated correctly. Given that drusen down to 30 \( \mu \text{m} \) in diameter are clinically detectable,\(^{29}\) this finding is surprising.

One main reason for this failure seems to be the interpolation procedures that are applied by the delineation algorithms identifying the RPE layer. Even though such an interpolation is understandable, as it gives the algorithms robustness, especially in advanced forms of AMD, it seems to result in lower detection rates of relevant pathologic features in early AMD, due to its more subtle alterations within the RPE. As the high proportion of nondetected small drusen and the substantial rate of moderate errors due to height-to-diameter ratio indicate, the interpolated RPE line may be too...
The analysis of the plotted maps derived from the manual recorded data revealed significant higher total amounts of drusen detectable by SD-OCT than by analysis of fundus photographs. As shown in Figures 5 and 6, this result is for the most part based on the better differentiation between neighboring drusen and the more accurate detection of small drusen. In a third of our patients, the total number of drusen was nearly twice as high as those identified by fundus photograph alone, especially in cases in which the photograph’s quality was compromised due to discrete media opacities. The fact that small, hard drusen account for most of the overall drusen number is known from histologic studies. Epidemio-
logic studies confirmed the presence of small, hard drusen as a common finding in all age groups. Therefore, they are not seen as a risk factor for the development of AMD today. However, as Sarks et al. assumed in their clinicopathologic studies in which they investigated the nature of small drusen, large amounts of small drusen may indeed represent pathologic changes underneath the retina. Accordingly, they might be seen as a risk factor for progression toward AMD, which is ignored in the established grading schemes. The fact that SD-OCT, especially in high-resolution mode, allows for an accurate manual detection of small drusen may indicate that SD-OCT is a useful tool for further clinical investigation, especially when considering the nature and risk of large quantities of small drusen.

As the beginning and end of each druse was mostly clearly definable, particularly in the high-resolution scans, the calculated total area of drusen appears to be more reliable than the estimation derived from fundus photography, as recommended by the AREDS. Studies examining these differences between the total drusen area quantification with different modalities are in progress.

As long as current automated segmentation programs fail to detect drusen without significant error, this potential advantage of SD-OCT imaging cannot be carried over to clinical practice. Our study clearly showed that drusen detection by standard fundus photography cannot be replaced by SD-OCT imaging in its current form.

In conclusion, SD-OCT imaging showed promise for precise and objective detection and measurement of drusen. However, the current segmentation algorithms often failed
to delineate the druse’s characteristics accurately, even though the druse was clearly visualized in the individual B-scans. Currently, a great effort is under way to bring automated OCT segmentation solutions to perfection. Topcon has introduced new review software (ver. 2.2) that may perform better in detecting retinal layers. However, we want to encourage software developers to place emphasis on more accurate RPE delineation, to improve the performance of drusen detection in the future. Such an advance would enable clinicians to assess drusen area more realistically, could be useful in estimating the individual prognosis of a patient more reliably, and is important since an objective base for upcoming preventive strategies in AMD is clearly needed to improve clinical care for our patients.

References


