Quantitative SD-OCT Imaging Biomarkers as Indicators of Age-Related Macular Degeneration Progression

Luis de Sisternes,1 Noah Simon,2,3 Robert Tibshirani,2 Theodore Leng,4 and Daniel L. Rubin1,5

1Department of Radiology, Stanford University, Stanford, California, United States
2Department of Statistics, Stanford University, Stanford, California, United States
3Department of Biostatistics, University of Washington, Seattle, Washington, United States
4Byers Eye Institute at Stanford, Stanford University School of Medicine, Palo Alto, California, United States
5Department of Medicine (Biomedical Informatics), Stanford University, Stanford, California, United States

Correspondence: Theodore Leng, Byers Eye Institute at Stanford, 2452 Watson Court, Palo Alto, CA 94303, USA; tedleng@stanford.edu.
Daniel L. Rubin, Richard M. Lucas Center, 1201 Welch Road, P285, Stanford, CA 94305-5488, USA; dlrubin@stanford.edu.
Submitted: May 30, 2014
Accepted: September 29, 2014
Citation: de Sisternes L, Simon N, Tibshirani R, Leng T, Rubin DL. Quantitative SD-OCT imaging biomarkers as indicators of age-related macular degeneration progression. Invest Ophthalmol Vis Sci. 2014;55:7093–7103. DOI:10.1167/iovs.14-14918

Purpose. We developed a statistical model based on quantitative characteristics of drusen to estimate the likelihood of conversion from early and intermediate age-related macular degeneration (AMD) to its advanced exudative form (AMD progression) in the short term (less than 5 years), a crucial task to enable early intervention and improve outcomes.

Methods. Image features of drusen quantifying their number, morphology, and reflectivity properties, as well as the longitudinal evolution in these characteristics, were automatically extracted from 2,146 spectral-domain optical coherence tomography (SD-OCT) scans of 330 AMD eyes in 244 patients collected over a period of 5 years, with 36 eyes showing progression during clinical follow-up. We developed and evaluated a statistical model to predict the likelihood of progression at predetermined times using clinical and image features as predictors.

Results. Area, volume, height, and reflectivity of drusen were informative features distinguishing between progressing and nonprogressing cases. Discerning progression at follow-up (mean, 6.16 months) resulted in a mean area under the receiver operating characteristic curve (AUC) of 0.74 (95% confidence interval [CI], 0.58, 0.85). The maximum predictive performance was observed at 11 months after a patient’s first early AMD diagnosis, with mean AUC 0.92 (95% CI, 0.85, 0.98). Those eyes predicted to progress showed a much higher progression rate than those predicted not to progress at any given time from the initial visit.

Conclusions. Our results demonstrate the potential ability of our model to identify those AMD patients at risk of progressing to exudative AMD from an early or intermediate stage.

Keywords: age-related macular degeneration, optical coherence tomography, statistical modeling, risk assessment, prediction

Age-related macular degeneration (AMD) is the leading cause of irreversible severe vision loss in the developed world in individuals over the age of 65.1 The disease can manifest in several stages, mainly described as early, intermediate, or advanced. Eyes classified as early or intermediate present a nonexudative form, showing lesions called drusen,1 and can suddenly convert to the advanced category, either in a nonexudative form with manifestations of geographic atrophy or in an exudative form, in which abnormal blood vessel growth (choroidal neovascularization) leads to blood and protein leakage under the macula, causing irreversible damage to the photoreceptors and rapid vision loss if left untreated. Early detection and prompt intervention in advanced exudative AMD have been shown to improve visual outcomes; thus, it is crucial to identify its development at the earliest stage possible.2 Patients at early and intermediate stages can convert to advanced exudative AMD (what we will refer to herein as "AMD progression") without any previous noticeable visual changes, and this conversion often is detected once visual changes are irreversible and most treatments have suboptimal outcomes. Thus, identifying patients at high risk of imminent progression from early and intermediate AMD to its exudative form is important, but doing so is a challenging problem that is unsolved to date. Development of technology to identify accurately the risk of AMD progression at a given time would allow for optimal clinical follow-up, with more frequent screening, and potential earlier treatment and better clinical outcomes. This technology also may aid in the clinical trials of agents to prevent AMD progression by providing better biomarkers of AMD, and using them to calculate the risk of progression before and after receiving treatment.

Evaluation of drusen3 in color fundus photographs represents the current clinical practice standard for assessing early and intermediate AMD. There is positive correlation in the number, area, and extent of drusen observed in the photographs with risk of early-to-advanced progression in more than two years.4–7 These characteristics usually are estimated by visual inspection with comparison to a set of standardized circles.4,8 However, this classification has limited value for predicting AMD progression, because it is very coarse in the short term (it can only identify a subgroup of patients with a maximum risk of progression of 8.8% over two years), and even
patients classified with early AMD suddenly can have progression to the advanced exudative form. More advanced experimental predictive methods include a combination of the aforementioned drusen classification in photographs combined with genetic, demographic, and environmental factors, such as smoking or diet. While these methods show promising results, they do not exploit a variety of quantitative features of drusen that can be useful predictors of progression, such as their volumetric properties (there is no depth resolution in the photographs) or the information obtained by quantitatively evaluating feature changes as AMD progresses over time.

Optical coherence tomography (OCT) is the only in vivo imaging method capable of resolving cross-sectional retinal substructures. The technique was first commercialized by Carl Zeiss Meditec, Inc. (Dublin, CA, USA) for inner retina imaging and is considered now superior to the current standard of care for the evaluation of a number of conditions. The more recently introduced spectral-domain OCT (SD-OCT) allows very fast scanning (more than 20,000 axial scans per second) over a retinal area, with axial resolutions as low as 5 µm. The higher acquisition speed and sensitivity of SD-OCT allows the collection of three-dimensional (3D) scans with higher pixel density while minimizing artifacts due to patient movement or ocular contractions. The result is a high-density 3D image volume composed by a set of two dimensional images, called B-scans, defined in the horizontal-axial dimension, each at a different vertical positioning. While scanning patterns are different for each commercial system, some of them allow user-specific patterns, for the SD-OCT system used in this study (CirrusOCT; Carl Zeiss Meditec, Inc.), a series of 128 or 200 B-scans normally are collected per cube. Each B-scan is formed by a series of axial lines called A-scans. Figure 1a displays the typical SD-OCT nomenclature. In recent years, SD-OCT has become a key diagnostic technology in retinal diseases and potentially is valuable in providing detailed imaging characteristics of the disease phenotype that could be used for predicting AMD progression. The SD-OCT enables accurately identifying drusen (seen on the white outlines in Figs. 1b, 1c, and in a 3D representation in Fig. 1d), while its depth differentiation allows quantifying their volumetric and reflective properties. Previous studies indicate that a degenerative retinal process is associated with the height of drusen observed in SD-OCT imaging, and many other characteristics quantified via SD-OCT may be useful as disease biomarkers. Though several methods exist in the literature for the segmentation and quantification of the RPE containing drusen in SD-OCT images, and classification of healthy and intermediate AMD eyes using SD-OCT phenotypic features, to our knowledge, there is no published method that accurately identifies patients who are likely to have progression from early or intermediate nonexudative AMD to its...
Predicting AMD Progression Using SD-OCT Features

**METHODS**

**Study Dataset**

We collected retrospective data from 330 eyes in 244 patients at our institution’s Vitreoretinal Clinic, comprising a total of 2146 longitudinal SD-OCT exams obtained over a 5-year time interval. The exams were obtained from consecutive patients followed the tenets of the Declaration of Helsinki. In this study, we were interested in analyzing the conversion of eyes presenting with early or intermediate nonexudative AMD to advanced exudative AMD, which we refer to as AMD progression. Exudative AMD was detected as any manifestation of choroidal neovascularization, fibrovascular scar formation, or fibrovascular pigment epithelial detachment. Patients with an early or intermediate nonexudative AMD diagnosis in an initial visit and at least one more follow-up visit with either a similar AMD status or a conversion to advanced exudative AMD diagnosis were of interest when evaluating AMD progression (first two columns in Table 1). Group-based demographics are summarized in Table 1. We did not evaluate conversion from an early or intermediate to an advanced nonexudative state (typical of patients presenting with geographic atrophy).

All the SD-OCT scans were acquired using an instrument that produced an imaging volume with dimensions of 6 (horizontal) × 6 (vertical) × 2 (axial) mm with voxel dimensions of approximately 12, 47, and 2 µm, respectively (512, 128, and 1024 voxels in each direction). The raw data produced by the SD-OCT instrument were imported into the vendor’s proprietary software for analysis and reconstruction (Zeiss Research Browser, version 6.0.2.81; Carl Zeiss Meditec, Inc.) and later exported to files describing the reflectivity measured at each voxel location with 8-bit precision (this was imposed by proprietary software). All the data processing and methods were later implemented and carried out using Matlab (The MathWorks, Inc., Natick, MA, USA).

**Quantification and Predictive Modeling Pipeline**

We developed a fully automated pipeline for the segmentation feature extraction (de Sisternes L, et al. IOVS 2013;54:ARVO E-Abstract 4154), predictive modeling, and testing of our novel classification methods (key elements shown in Fig. 2). The data input comprised the series of longitudinal patient data. Drusen were detected automatically and segmented from the collection of the SD-OCT scans (labeled as 1 in Fig. 2) and automatically quantified by extracting a series of 22 features defining their number, extent, area, volume, shape, density, and reflectivity characteristics, as well the evolution of each of these characteristics over time (2 in Fig. 2). A prediction model was built to obtain a score

---

**Table 1.** Patient Demographics, Number of Eyes and Visits, and Those Employed in the Mixed-Effects ANOVA Analysis

<table>
<thead>
<tr>
<th>Patients, n</th>
<th>97</th>
<th>31</th>
<th>25</th>
<th>93</th>
<th>244</th>
</tr>
</thead>
<tbody>
<tr>
<td>Eyes, n</td>
<td>150</td>
<td>36</td>
<td>38</td>
<td>106</td>
<td>350</td>
</tr>
<tr>
<td>N of visits per eye, mean (SD)</td>
<td>4.35 (3.11)</td>
<td>3.92 (2.95)</td>
<td>1.00 (0.00)</td>
<td>7.35 (5.67)</td>
<td>6.43 (5.36)</td>
</tr>
<tr>
<td>Months between visits, mean (SD)</td>
<td>6.28 (5.77)</td>
<td>5.65 (4.90)</td>
<td>-</td>
<td>4.40 (4.45)</td>
<td>4.82 (4.93)</td>
</tr>
<tr>
<td>Sex, % female</td>
<td>55.01</td>
<td>68.79</td>
<td>80.00</td>
<td>65.22</td>
<td>72.03</td>
</tr>
<tr>
<td>Age at visits, y, mean (SD)</td>
<td>78.19 (9.83)</td>
<td>83.26 (5.99)</td>
<td>79.80 (11.47)</td>
<td>81.58 (7.19)</td>
<td>80.98 (8.05)</td>
</tr>
</tbody>
</table>

---

advanced exudative form using these features in a quantitative, fully automated, and reproducible manner.

In this study, we tested the hypothesis that quantitative volumetric properties of drusen observed in SD-OCT images, which can be characterized automatically, may yield a better prediction of future AMD conversion from an early or intermediate nonexudative stage to an advanced exudative stage than the current clinical practice based in visual evaluation of drusen in photographs. We identified quantitative features in SD-OCT images to characterize drusen that appear promising to predict AMD progression, and we developed and evaluated a statistical model to identify patients at high risk for AMD progression in the short term.
indicating the likelihood of a conversion from early or intermediate nonexudative AMD to advanced exudative AMD at a given time in the future, considering the quantitative features, presence of advanced exudative AMD in the fellow eye, patient demographics (age and sex), and the time elapsed between subsequent clinic visits from the same patient and eye (labeled as 3 in Fig. 2). We did not include genetic data as possible predictors, as these were not collected in this cohort of clinical patients. Evaluation of the prediction model was conducted through 10-fold cross-validation using a cohort of patients held out during model building (4 in Fig. 2) and considering their known outcome (progression versus persistence of an early or intermediate AMD stage).

**Drusen Segmentation**

Drusen were segmented automatically in the SD-OCT images using a method we described previously,32 which uses image processing algorithms (image filtering to reduce noise, image thresholding, and morphological operations) to identify the boundary of the RPE and its particular curvature, reflectivity, and topology.33,34 The result is a volumetric segmentation of the boundary of the RPE and its particular curvature, reflectivity, and topological properties.

**Description of Quantitative Imaging Features**

A set of 11 quantitative image features were extracted from the druse segmentations in each SD-OCT image cube, including characteristics of druse shape, geometry, and reflectivity; total area and volume; total number; mean area and volume per detected druse; maximum height in the axial direction; extent of retinal area affected by drusen; density in affected area; slope; and texture properties in the volume delimited by drusen (de Sisternes I, et al. IOVS 2013;54:ARVO E-Abstract 4154 and Leng T, et al. IOVS 2013;54:ARVO E-Abstract 4150).

Drusen area was measured as the area of drusen regions in an axial projection of the cube, and the extent of retinal area affected by drusen was formed by generating the convex hull of those projected drusen regions. Drusen density was detected as the ratio of total drusen area by retinal area of those projected drusen regions. Drusen density was measured as the area of drusen regions in an axial projection of the cube, and the extent of retinal area affected by drusen was formed by generating the convex hull of those projected drusen regions. Drusen density was computed as the ratio of total drusen area by retinal area of those projected drusen regions.

**Progression Model**

Time elapsed since a clinic visit with early or intermediate AMD diagnosis was taken into consideration when computing a risk of progression score, as patients have a higher chance of progression within longer time intervals. We formulated a risk score related to the chances of progression within a given future time using a L1-penalized Poisson model with logarithm of time-to-prediction as an offset, with imaging features and clinical parameters as predictors. To make the model more flexible, we allowed piecewise-linear functions by using an expanded basis for each imaging feature with hinge functions that had knots at the sample deciles. This risk score was then computed by:

\[
S(t, \bar{z}) = \left( \beta_0 + \beta_1 z_1 + \beta_2 z_2 + \ldots + \beta_K z_K \right) \cdot t,
\]

where \(\bar{z} = [z_1, z_2, \ldots, z_K]\) is a set of \(K\) features used as predictors, \(t\) is the considered time to prediction, and \(\hat{\beta} = [\hat{\beta}_0, \hat{\beta}_1, \hat{\beta}_2, \ldots, \hat{\beta}_K]\) are the set of coefficients in the model.

Let us define an observation as the pair formed by the set of extracted features from a patient visit in which early or intermediate AMD was diagnosed and the diagnosis in a future additional visit, either maintaining early or intermediate status or where a conversion to exudative AMD was detected. Given a set of \(M\) training observations, \(X = [x_1, x_2, \ldots, x_M]\), each defined with a set of \(K\) observed features \(x_m = [x_{m1}, x_{m2}, \ldots, x_{mK}]\), with associated future known outcomes \(Y = [y_1, y_2, \ldots, y_M]\) (0, maintenance of status; 1, exudative AMD event) at times \(t = [t^1, t^2, \ldots, t^M]\), we found the coefficients \(\hat{\beta}\) using generalized
TABLE 2. List of Extracted Drusen Features, and q Values From ANOVA Study Between Progressing and Nonprogressing Patients Within Each Tested Time Interval

<table>
<thead>
<tr>
<th>Feature</th>
<th>6</th>
<th>12</th>
<th>18</th>
<th>24</th>
<th>30</th>
<th>42</th>
<th>48</th>
</tr>
</thead>
<tbody>
<tr>
<td>Patient age, y</td>
<td>0.04*</td>
<td>&lt;0.01*</td>
<td>&lt;0.01*</td>
<td>&lt;0.01*</td>
<td>&lt;0.01*</td>
<td>0.03*</td>
<td>&lt;0.01*</td>
</tr>
<tr>
<td>Patient sex; 1, female; 0, male</td>
<td>0.70</td>
<td>0.70</td>
<td>0.70</td>
<td>0.70</td>
<td>0.70</td>
<td>0.70</td>
<td>0.70</td>
</tr>
<tr>
<td>N of drusen separate regions</td>
<td>0.68</td>
<td>0.68</td>
<td>0.69</td>
<td>0.66</td>
<td>0.70</td>
<td>0.67</td>
<td>0.48</td>
</tr>
<tr>
<td>Mean volume per drusen region, mm³/drusen</td>
<td>0.66</td>
<td>0.26</td>
<td>0.26</td>
<td>0.26</td>
<td>0.28</td>
<td>0.27</td>
<td>0.27</td>
</tr>
<tr>
<td>Total volume of drusen regions, mm³</td>
<td>0.38</td>
<td>0.32</td>
<td>0.16</td>
<td>0.07</td>
<td>0.01*</td>
<td>0.02*</td>
<td>0.02*</td>
</tr>
<tr>
<td>Mean area per drusen region, mm²/drusen</td>
<td>0.61</td>
<td>0.27</td>
<td>0.32</td>
<td>0.28</td>
<td>0.27</td>
<td>0.26</td>
<td>0.26</td>
</tr>
<tr>
<td>Total area of drusen regions, mm²</td>
<td>0.32</td>
<td>0.44</td>
<td>0.26</td>
<td>0.14</td>
<td>0.02*</td>
<td>0.02*</td>
<td>0.02*</td>
</tr>
<tr>
<td>Drusen density in affected area</td>
<td>0.67</td>
<td>0.52</td>
<td>0.67</td>
<td>0.51</td>
<td>0.50</td>
<td>0.51</td>
<td>0.32</td>
</tr>
<tr>
<td>Extent of drusen affected area, mm²</td>
<td>0.33</td>
<td>0.42</td>
<td>0.29</td>
<td>0.35</td>
<td>0.25</td>
<td>0.26</td>
<td>0.26</td>
</tr>
<tr>
<td>Maximum drusen height, mm</td>
<td>0.70</td>
<td>0.34</td>
<td>0.04*</td>
<td>&lt;0.01*</td>
<td>&lt;0.01*</td>
<td>&lt;0.01*</td>
<td>&lt;0.01*</td>
</tr>
<tr>
<td>Advanced exudative AMD present in contralateral eye</td>
<td>0.70</td>
<td>0.70</td>
<td>0.70</td>
<td>0.69</td>
<td>0.67</td>
<td>0.68</td>
<td>0.58</td>
</tr>
<tr>
<td>Mean drusen slope</td>
<td>0.70</td>
<td>0.69</td>
<td>0.66</td>
<td>0.42</td>
<td>0.42</td>
<td>0.53</td>
<td>0.48</td>
</tr>
<tr>
<td>Mean reflectivity inside drusen region</td>
<td>0.68</td>
<td>0.67</td>
<td>0.49</td>
<td>0.68</td>
<td>0.69</td>
<td>0.24</td>
<td>0.16</td>
</tr>
<tr>
<td>Standard of reflectivity inside drusen regions</td>
<td>0.44</td>
<td>0.26</td>
<td>0.46</td>
<td>0.27</td>
<td>0.26</td>
<td>0.12</td>
<td>0.06*</td>
</tr>
<tr>
<td>Evolution n of drusen separate regions over time, 1/mo</td>
<td>0.52</td>
<td>0.49</td>
<td>0.63</td>
<td>0.26</td>
<td>0.52</td>
<td>0.53</td>
<td>0.54</td>
</tr>
<tr>
<td>Evolution volume per drusen over time, mm³/mo</td>
<td>0.70</td>
<td>0.26</td>
<td>0.32</td>
<td>0.26</td>
<td>0.32</td>
<td>0.50</td>
<td>0.55</td>
</tr>
<tr>
<td>Evolution volume of drusen regions over time, mm³/mo</td>
<td>0.69</td>
<td>0.65</td>
<td>0.62</td>
<td>0.55</td>
<td>0.58</td>
<td>0.65</td>
<td>0.68</td>
</tr>
<tr>
<td>Evolution mean area per drusen over time, mm²/drusen-mo</td>
<td>0.70</td>
<td>0.27</td>
<td>0.35</td>
<td>0.31</td>
<td>0.35</td>
<td>0.51</td>
<td>0.57</td>
</tr>
<tr>
<td>Evolution area of drusen regions over time, mm²/mo</td>
<td>0.68</td>
<td>0.67</td>
<td>0.65</td>
<td>0.69</td>
<td>0.69</td>
<td>0.67</td>
<td>0.69</td>
</tr>
<tr>
<td>Evolution drusen density in affected area over time, 1/mo</td>
<td>0.69</td>
<td>0.68</td>
<td>0.67</td>
<td>0.20</td>
<td>0.26</td>
<td>0.32</td>
<td>0.38</td>
</tr>
<tr>
<td>Evolution extent of drusen affected area over time, mm²/mo</td>
<td>0.69</td>
<td>0.66</td>
<td>0.68</td>
<td>0.26</td>
<td>0.16</td>
<td>0.26</td>
<td>0.21</td>
</tr>
<tr>
<td>Evolution maximum drusen height over time, mm/mo</td>
<td>0.69</td>
<td>0.69</td>
<td>0.70</td>
<td>0.23</td>
<td>0.26</td>
<td>0.31</td>
<td>0.26</td>
</tr>
<tr>
<td>Evolution to advanced exudative AMD in contralateral eye over time, 1/mo</td>
<td>0.69</td>
<td>0.68</td>
<td>0.28</td>
<td>0.02*</td>
<td>0.04*</td>
<td>0.12</td>
<td>0.17</td>
</tr>
<tr>
<td>Increase of mean drusen slope over time, 1/mo</td>
<td>0.70</td>
<td>0.69</td>
<td>0.65</td>
<td>0.70</td>
<td>0.69</td>
<td>0.70</td>
<td>0.69</td>
</tr>
<tr>
<td>Evolution mean reflectivity inside drusen region over time</td>
<td>0.70</td>
<td>0.32</td>
<td>0.42</td>
<td>0.68</td>
<td>0.69</td>
<td>0.68</td>
<td>0.70</td>
</tr>
<tr>
<td>Evolution standard of reflectivity inside drusen regions over time</td>
<td>0.70</td>
<td>0.34</td>
<td>0.42</td>
<td>0.38</td>
<td>0.41</td>
<td>0.40</td>
<td>0.34</td>
</tr>
</tbody>
</table>

* q values under 0.05 were considered statistically significant.

piecewise linear regression with Lasso regularization that best fitted the equation:

\[
\ln(Y^m) = \beta_0 + \beta_1 x_1^m + \beta_2 x_2^m + \ldots + \beta_K x_K^m + \ln(t^m)
\]

for \( m = 1, \ldots, M \).

Given the score, \( S(t, \beta) \) as computed in Equation 1, we can define a threshold \( T \) that separates cases in two risk categories (high-risk or low-risk) considering their chances of progression at time \( t \):

\[
\text{category} \begin{cases} 
\text{High - risk} & \text{if } S(t, \beta) \geq T \\
\text{Low - risk} & \text{if } S(t, \beta) < T
\end{cases}
\]

**Progression Percentage Computation**

The progression percentage of each category (high-risk or low-risk) was computed as a probability rate of progressing from early or intermediate nonexudative AMD to advanced exudative AMD using the Kaplan-Meier estimator of survival in the form:

\[
PP(t) = 100 \cdot \left[1 - \frac{\text{surv}(t)}{\text{surv}(t) - \text{surv}(t - 1)}\right]
\]

where \( \text{surv}(t) \) is the survival function at time \( t \) of the category testing for progression at time \( t \), and \( \text{ND}(t; \beta) \), \( \text{NW}(t; \beta) \), and \( \text{NC}(t; \beta) \) are the number of eyes in the category testing for progression at time \( t \) that are at risk of progressing, a progression was detected, and were censored at time \( t \), respectively. Censoring of cases was due to either the patient being stopped for follow-up at time earlier than tested or an unknown AMD status at tested time (interval between a patient’s last nonexudative diagnosis and first exudative diagnosis, in which progression status is not known).

**RESULTS**

**Association of Quantitative Imaging Features With Progression Event**

We investigated the statistical differences in each of the extracted quantitative image features and patient demographics between the cohort of visits presenting AMD progression within given times (from 6–48 months at 6-month intervals) versus the cohort that remained in an early or intermediate stage using a mixed-effects ANOVA, where AMD status is a fixed effect and the particular patient is a random effect. Those visits in which there was not absolute certainty of the eye’s AMD status at the tested time either due to right-censoring (time after last known early or intermediate stage using a mixed-effects ANOVA, where AMD status is a fixed effect and the particular patient is a random effect). Those visits in which there was not absolute certainty of the eye’s AMD status at the tested time either due to right-censoring (time after last known early or intermediate stage using a mixed-effects ANOVA, where AMD status is a fixed effect and the particular patient is a random effect).
using a multiple hypothesis testing analysis by estimating the false-positive discovery rate $q$ values, using the method introduced by Storey,\textsuperscript{40} and Storey and Tibshirani\textsuperscript{41} considering all tested features at different prediction times. The distribution differences between the progressing and non-progressing cohorts were analyzed for each feature, and those presenting a $q$ value under 0.05 were considered statistically significant. In addition to the quantitative imaging features, we also analyzed clinical and demographic features (patient age, sex, and the presence of exudative AMD in the contralateral eye) as possible predictors. The $q$ values resulting from testing the distribution differences between the features extracted in those clinic visits where a progression event was recorded and those where no progression event occurred, analyzed within given time intervals, are summarized in Table 2.

**Prediction of Progression Event Accuracy**

We evaluated the accuracy of our predictive model in identifying patients with a higher risk of progression at their given follow-up times. A total number of 790 possible observations from 186 eyes of 128 patients was constructed from our retrospective dataset described above, formed by the features extracted at the clinic visits where an eye is diagnosed with early or intermediate AMD, and a known diagnosis at a subsequent visit (i.e., that the patient would either maintain an early or intermediate status or have converted to exudative AMD at the next time point). The predictive model then was trained and tested using the known outcomes at follow-up time. A histogram of the number of constructed observations as a function of the probability risk score produced by the prediction model (computed as given in Equation 1). Number of observations is displayed as the percentage from the total of the category. The dotted line indicates the threshold value that results in at least 80% sensitivity in the prediction. (C) Number of eyes with knowledge of maintaining an early or intermediate status, progressing to an advanced exudative status, and with unknown status as a function of time elapsed since first early or intermediate AMD diagnosis. (D) The AUC mean values and 95% CI for the ROC curves as a function of testing for progression at elapsed time since each eye first had an early or intermediate AMD diagnosis. (E) The ROC curves for testing for progression at 11 (top left), 16 (top right), 18 (bottom left), and 48 (bottom right) elapsed months since first early or intermediate diagnosis per eye. The AUCs and operating points of at least 80% sensitivity are indicated.
10-fold cross-validation, where all observations from the same patient were required to be in the same fold. Lasso regularization yielded an optimal $\lambda$ value of $\lambda = 0.74 \times 10^{-2}$. We assessed the prediction accuracy by constructing a receiver operating characteristic (ROC), shown in Figure 3A, and computing its area under the curve (AUC), yielding a mean value of 0.74 with 95% confidence interval (CI) of (0.58, 0.85). Following previous studies on predicting disease progression, our goal also was to identify a cutoff where sensitivity resulted at least of 80%. Patients with a risk score over the identified cutoff were labeled as high-risk cases, while cases under the threshold were labeled as low-risk cases. Assuming such scoring threshold ($T = 1.54 \times 10^{-2}$, as identified in Fig. 3B), the operating point on the ROC curve corresponded to 80.95% sensitivity and 51.24% specificity. Figure 3B displays the histograms of test observations in which there was exudative AMD in the patient’s next follow-up visit (progressors) and those that remained at an early or intermediate nonexudative stage (nonprogressors) as a function of the probability risk score produced by our prediction model.

We also evaluated the performance in making a correct prediction of AMD progression at a given future time after the first clinic visit where an eye was diagnosed with early or intermediate AMD. Those eyes with unknown status at an elapsed time after this baseline (either due to lack of patient follow-up while the patient remained in an early or intermediate stage or due to a time interval between his/her last early or intermediate AMD diagnosis and the first exudative diagnosis, in which the progression status was not known) were treated as censored cases for that particular time (right censoring and interval censoring, respectively). The total number of eyes maintaining an early or intermediate status, progressing to an exudative status, and those in which status is unknown (censoring) as a function of elapsed time since their first early or intermediate AMD diagnosis is displayed in Figure 3C.

We assessed the prediction accuracy by constructing ROC curves with the eye prediction scores, testing for progression within 1 to 60 months (at 1-month intervals) from the first diagnosis (baseline). The mean AUC and the 95% CIs were computed for testing at each elapsed time interval by bootstrapping $10^5$ samples, resampling from the set of available eyes (not censored) with replacement, and the results are shown in Figure 3D. Figure 3E displays the particular ROC curves constructed when testing at four representative elapsed time intervals, particularly at 11, 16, 18, and 48 elapsed months, with AUCs (95% CI) of 0.92 (0.83, 0.98), 0.86 (0.77, 0.95), 0.70 (0.54, 0.84), and 0.79 (0.69, 0.87), respectively. The cutoff where sensitivity resulted at least of 80% also is identified in each ROC curve, with its particular sensitivity and specificity values.

**Progression Percentage per Category in Classification**

We also computed a progression percentage at variable future times for the eyes labeled by our predictive model as high-risk and low-risk of progression, respectively. Progression percentage mean and 95% CIs were computed by bootstrapping $10^5$ samples with replacement from the set of eyes. Eyes were stratified to a time-dependent risk category (high-risk or low-risk) associated with the future time at which we are predicting progression. Figure 4A shows the number of total eyes (including censored eyes) and the number of progressing eyes as a function of elapsed time since the first early or intermediate AMD diagnosis per eye. Figure 4B shows the progression percentage mean and 95% CI for those eyes labeled as high-risk and low-risk by our prediction model as a function of time since the first early or intermediate AMD diagnosis per eye.

**DISCUSSION**

In this study we introduce a novel approach to predicting conversion from an early or intermediate nonexudative AMD stage to an advanced exudative stage in the short term based on extracting a set of quantitative features that characterize the AMD disease phenotype in SD-OCT images and using them, in combination with selected clinical features, to create a predictive model. Although the previous Age-Related Eye Disease (AREDS) studies developed a model for stratifying patients into risk categories of their chances of developing...
Predicting AMD Progression Using SD-OCT Features

Table 3. Progression Percentage Means and 95% CI at Given Times Since First Early or Intermediate AMD Diagnosis for Eyes Classified as High-Risk and Low-Risk by Our Predictive Model

<table>
<thead>
<tr>
<th>Time Elapsed Since First Early or Intermediate AMD Diagnosis, mo</th>
<th>Cross-Validated Progression Percentage, Mean (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Cases Classified as Low-Risk, %</td>
</tr>
<tr>
<td>6</td>
<td>0.00 (0.00, 0.00)</td>
</tr>
<tr>
<td>12</td>
<td>0.00 (0.00, 0.00)</td>
</tr>
<tr>
<td>18</td>
<td>2.91 (0.00, 4.97)</td>
</tr>
<tr>
<td>24</td>
<td>2.94 (0.00, 5.05)</td>
</tr>
<tr>
<td>30</td>
<td>3.96 (1.43, 6.61)</td>
</tr>
<tr>
<td>36</td>
<td>5.01 (1.56, 8.07)</td>
</tr>
<tr>
<td>42</td>
<td>5.05 (1.59, 8.18)</td>
</tr>
<tr>
<td>48</td>
<td>5.04 (1.56, 8.15)</td>
</tr>
<tr>
<td>56</td>
<td>6.28 (2.13, 9.90)</td>
</tr>
<tr>
<td>60</td>
<td>9.96 (4.52, 15.03)</td>
</tr>
</tbody>
</table>

AMD progression is summarized in Table 2, which shows that patient age is a significant discriminator for AMD progression. This work also is novel in being the first to study the evolution of drusen features over time as a predictor of AMD progression. Although these evolution characteristics had a role in our prediction model, we did not observe significant differences between progressing and nonprogressing cases for the tested set. This may be due to the small number of patients available for this study. Validation of these findings in a prospective clinical trial will be needed.

Analysis of the relevance of individual features in predicting AMD progression is summarized in Table 2, which shows that patient age is a significant discriminator for AMD progression, becoming very important when considering progression within a year from an early or intermediate AMD diagnosis. Quantitative imaging features also contributed to earlier prediction of progression; specifically, the maximum height of the detected drusen resulted in a significant discriminator for progression within 18 months, and total area of drusen had significant higher values for cases progressing within 30 months. Previous studies indicated that there is a degenerative retinal process associated with the height of drusen observed in SD-OCT imaging.27 and drusen height and width have been mentioned as possible strong predictors of progression over a 5-year interval (Dieaconescu D, et al. IOVS 2012;53:ARVO E-Abstract 2910). We also found that volumetric drusen characteristics not investigated previously were discriminators of progression; total drusen volume in the SD-OCT cube was a significant predictor for progression within 30 months. Another interesting finding is that texture properties of drusen may also have an important role when characterizing the chances of progression. The standard deviation of the reflectivity values recorded inside a drusen volume was a discriminator of progression within 48 months, with higher deviation values corresponding to higher chances of progression. Surprisingly, we did not find statistically significant differences between progressing and nonprogressing cases at the evaluated progression times in the presence of exudative AMD in the contralateral eye, while this feature is used commonly as a progression factor at longer time intervals than analyzed here.4,5 However, the evolution (or progression) of the contralateral eye from an early or intermediate stage to advanced exudative AMD was a discriminator of progression within a 24- to 30-month interval. The lack of statistical significance in the earlier feature, may be due to the distribution of the acquired data; however, the results seem to indicate that the time elapsed between the conversion of one patient’s eye and the fellow eye concentrates at approximately a 24- to 30-month interval.

The performance of our prediction model testing at a follow-up visit (0.74; 95% CI, 0.58, 0.85; with a particular operating point of 80.9% sensitivity and 51.2% specificity) seemed similar to results reported previously on models using purely genetic or a combination of genetic, phenotypic, and environmental data. However, in such prior work, image features obtained from SD-OCT were not studied. Our method also seems to provide several novel and unique advantages by providing a shorter prediction time frame compared to the prior methods. In the prior AMD progression studies, study times ranged between 4 and 5 years. In evaluating our model, follow-up visits occurred within an average of 6.16 months (5.63 SD), ranging from a minimum of 3 days to a maximum of 2.87 years (histogram of number of observations per elapsed time to a follow-up visit in Fig. 3A), which constitute much shorter ranges than any published study to our knowledge. Our method also has the novelty of being able to produce a prediction score related to the chances of progression at an arbitrary future time. When evaluating the prediction accuracy at particular given times since the first early or intermediate AMD diagnosis, the discrimination performance seemed to have the highest values in the range between 10 and 15 months (Fig. 3D), producing an AUC of 0.91 (CI, 0.78, 0.99) in the former and 0.90 (CI, 0.80, 0.97) in the later, with a maximum AUC of 0.92 (CI, 0.83, 0.98) when testing for progression within 11 months (Fig. 3E). A decrease in its performance was observed when testing for progression within 18 months (AUC, 0.70; CI, 0.54, 0.84), with seemingly increasing performance for later times. The increase of mean performance and decrease of 95% CI of the model after a 10-month range, when compared to earlier times, may be due to an insufficient number of progressing patients at an earlier time (as seen in Fig. 3C), while the decrease of performance after the 15-month range may be due to the reduced number of training observations with such larger time intervals (as shown in the histogram of observations in Fig. 3A).

In a purely genetic model, Hageman et al.9 achieved an AUC of 0.80 and 63.2% specificity at 91.7% sensitivity predicting which eyes were due to develop exudative AMD at any time in their lives. Yu et al.15 reported an AUC of 0.885 for progression
at 5 years after an early or intermediate AMD diagnosis, and 0.895 for progression at 10 years using a multistate Markov model and genetic data. Seldon and Reynolds\(^1\) demonstrated an increase of AUC from 0.757 to 0.83 when including genetic data to phenotypic and environmental data in an unconditional logistic model. In a population-based study, Buiten-dijk et al.\(^1\) recently reported an AUC of 0.60 (CI, 0.55, 0.65) in a minimal predictive model consisting only in age and sex of the patient as predictors (only demographic factors used in our model), and increase to 0.78 (CI, 0.74, 0.82) when using demographic, baseline phenotype in color fundus photograph, and environmental data (smoking, body mass index), and excluding genetic data and a higher increase to 0.87 (CI, 0.85, 0.89) when also considering an optimal set of genetic predictors for discriminating the chances of progression, in a median follow-up time of 11.1 years. Recently, Perlee et al.\(^1\) reported an impressive AUC of 0.96 through selection of the best performance Cox proportional hazard model from a combination of genotypic, phenotypic, and environmental data. However, it was stated that the genetic profile minimally differentiates risk of progression in patients with early stages of AMD. We acknowledge that genetic data are important features with predictive potential\(^1\) and ideally we would have included such data in our analyses; however, given that our data were drawn from a representative clinical practice and in a standard clinical setting, genetic data were not available. Public AMD datasets from the AREDS study include genetic data, but not SD-OCT image data. We anticipated that the inclusion of SD-OCT quantitative imaging features together with genetic, environmental, and other phenotypic factors could lead to superior results, and we plan to undertake such studies in the future to improve our model.

We also computed the progression risks for those eyes identified in high-risk and low-risk categories as the fraction of progressing eyes in the cohort within given time. The eyes identified in the high-risk category presented a much higher rate of progression than those in the low risk category at any tested time (Fig. 4, Table 3). We found that 20.61% (CI, 12.87, 28.09) of the eyes identified as high-risk of progression within 30 months actually progressed within the given time, while only 3.96% (CI, 1.43, 6.61) of the eyes in the low-risk category actually progressed. When evaluating progression at 4 years, 27.45% (CI, 18.94, 36.04) and 5.0% (CI, 1.56, 8.15) of the eyes in the high-risk category and in the low-risk category progressed, respectively.

Although the performance of our model still is limited, we believe this is due mainly to the limited data that were available for our study. Specifically, a larger prospective study where AMD patients are scheduled routinely for an SD-OCT scan at regular time intervals is needed for further evaluation, especially for the prediction of imminent AMD progression. Future work will consider the inclusion of such patient data in the development of our models, the optimization of predictors and model operative point, and testing in an independent dataset.

As quantitative imaging features are extracted from SD-OCT imaging data and used as predictors, the performance of the prediction model depends on the quality of the images as well as on the accuracy and stability of the drusen segmentations. In this work, we chose to use images acquired during clinical practice using the current practice protocols and standards so as to evaluate the performance of the method as similar as in an actual clinical setting. No exclusion criterion was considered in terms of image quality other than the images being of adequate quality to the clinician for use during a clinical encounter. Imaging of the fovea region is done typically with 128 scans in the vertical direction by Cirrus OCT instruments, and drusen that were too small to be visible by this acquisition standard would, indeed, not be detected by our methods (an evaluation of the drusen segmentation method used here has been described by Chen et al.\(^2\)). However, given these possible limitations, the prediction method showed promising results. It also is important to note that three of the patients included in this study had two sequential scans acquired within a week presenting early or intermediate AMD. Drusen are unlikely to change or form noticeably over such short time periods and feature differences between consecutive scans of such characteristics would be part of a reproducibility study. For such scans, we found that there was no significant difference in the extracted drusen features, but we believe that a larger study is needed to fully address reproducibility of our methods.

In the near future, we will report the results of a study that determined the reproducibility of the automated drusen segmentation method, and their associated extracted features and computed risk score by evaluating the differences produced over a set of eyes scanned several times on the same day with varying signal strength (quality). This enables us to assess the robustness of relying on quantitative assessment of drusen as a biomarker of disease and the risk scores computed by our prediction method.

One of the novelties of the methods presented here is the analysis of texture properties of the reflectivity inside drusen regions in SD-OCT images and their correlation with AMD progression. It is important to note the difficulty in standardizing reflectivity values within SD-OCT images. Current SD-OCT systems (such as the ones used to collect the images in this study) have internal functions to standardize the reflectivity across the cube B-scans integrated within their instruments. In our approach, we further normalized the pixel values within the cube with respect to the highest intensity pixel values recorded in the RPE layer, of highest intensity and given a value of 1 after normalization, and the lowest intensity pixel values within the vitreous region, of lowest intensity and given a value of 0 after normalization. Although these normalized values do not represent actual physical quantities of reflectivity, we considered that this approach was sufficient to analyze a mean and SD of the recorded values to differentiate highly reflective drusen, whose pixel values have similarly high reflectivity as those present in the RPE with lower variance (low SD), from drusen with lower reflectivity, whose pixel values have much lower reflectivity and have a larger range of values (high SD). Analysis of reproducibility of such measurements will be presented in the near future, as pointed out in the previous paragraph.

Our work has practical clinical value by providing clinicians with a risk score that can guide management of patients presenting an early or intermediate AMD status. Patients with high probability risk score could be recommended for more frequent follow up examination, at which time OCT could be repeated and a new risk score calculated. At present, the follow-up of AMD patients is variable, and providing a more objective approach to recommended follow up could reduce the time interval between a conversion to exudative AMD and its detection, with improvement in clinical outcomes. On the other hand, a limitation of our work is that the specificity is relatively poor at high levels of sensitivity. While this could be improved in the future with refinements to the model, a more adequate SD-OCT scan dataset with images taken at routinely shorter follow-up times, and incorporation of other relevant predictors, such as environmental and genetic factors, the downside (additional OCT scanning and more frequent follow up) is not too substantial (the cost of additional OCT imaging).

The clinical importance of our predictive method presented here and others that allow the prediction of progression from early or intermediate to advanced AMD within a variety of future time intervals is that they could enable identifying
Predicting AMD Progression Using SD-OCT Features

patients that are likely to have progression and thereby permit tailoring the clinical follow-up and treatment decision making. Our approach is innovative and may even be cost-effective, since it uses purely computational analysis of the images in addition to existing clinical data, and it may, thus, be practical to introduce into the clinical workflow. The methods described have the potential of enabling earlier treatment and better clinical outcomes in AMD patients.

Acknowledgments
The authors thank Lauren Kutzscher, Jeffrey Ma, and Vibha Mahendra, all from Stanford University, for their help collecting patient diagnosis information.

Supported by a grant from the Bio-X Interdisciplinary Initiatives Program of Stanford University, and the Spectrum-SPADA innovation grant of Stanford University. The authors alone are responsible for the content and writing of the paper.

Disclosure: L. de Sisternes, None; N. Simon, None; R. Tibshirani, None; T. Leng, None; D.L. Rubin, None

References


