

Prediction of Anti-VEGF Treatment Requirements in Neovascular AMD Using a Machine Learning Approach

Hrvoje Bogunović,¹ Sebastian M. Waldstein,¹ Thomas Schlegl,^{1,2} Georg Langs,^{1,2} Amir Sadeghipour,¹ Xuhui Liu,^{1,3} Bianca S. Gerendas,¹ Aaron Osborne,⁴ and Ursula Schmidt-Erfurth¹

¹Christian Doppler Laboratory for Ophthalmic Image Analysis, Department of Ophthalmology, Medical University of Vienna, Vienna, Austria

²Computational Imaging Research Lab, Department of Biomedical Imaging and Image-guided Therapy, Medical University of Vienna, Vienna, Austria

³Department of Ophthalmology, The First Affiliated Hospital of Zhengzhou University, Zhengzhou, China

⁴Genentech, Inc., South San Francisco, California, United States

Correspondence: Ursula Schmidt-Erfurth, Medical University of Vienna, Spitalgasse 23, 1090 Vienna, Austria; ursula.schmidt-erfurth@meduniwien.ac.at

Submitted: November 3, 2016

Accepted: May 22, 2017

Citation: Bogunović H, Waldstein SM, Schlegl T, et al. Prediction of anti-VEGF treatment requirements in neovascular AMD using a machine learning approach. *Invest Ophthalmol Vis Sci.* 2017;58:3240–3248. DOI: 10.1167/iovs.16-21053

PURPOSE. The purpose of this study was to predict low and high anti-VEGF injection requirements during a pro re nata (PRN) treatment, based on sets of optical coherence tomography (OCT) images acquired during the initiation phase in neovascular AMD.

METHODS. Two-year clinical trial data of subjects receiving PRN ranibizumab according to protocol specified criteria in the HARBOR study after three initial monthly injections were included. OCT images were analyzed at baseline, month 1, and month 2. Quantitative spatio-temporal features computed from automated segmentation of retinal layers and fluid-filled regions were used to describe the macular microstructure. In addition, best-corrected visual acuity and demographic characteristics were included. Patients were grouped into low and high treatment categories based on first and third quartile, respectively. Random forest classification was used to learn and predict treatment categories and was evaluated with cross-validation.

RESULTS. Of 317 evaluable subjects, 71 patients presented low (≤ 5), 176 medium, and 70 high (≥ 16) injection requirements during the PRN maintenance phase from month 3 to month 23. Classification of low and high treatment requirement subgroups demonstrated an area under the receiver operating characteristic curve of 0.7 and 0.77, respectively. The most relevant feature for prediction was subretinal fluid volume in the central 3 mm, with the highest predictive values at month 2.

CONCLUSIONS. We proposed and evaluated a machine learning methodology to predict anti-VEGF treatment needs from OCT scans taken during treatment initiation. The results of this pilot study are an important step toward image-guided prediction of treatment intervals in the management of neovascular AMD.

Keywords: anti-VEGF therapy, image analysis, machine learning, optical coherence tomography, choroidal neovascularization

AMD is the leading cause of irreversible vision loss in the elderly population in the developed world.¹ Furthermore, with aging in the modern population, the number of AMD patients is expected to keep growing steeply. Anti-VEGF agents are highly effective and have revolutionized the treatment of neovascular AMD,² significantly reducing AMD-associated blindness and visual impairment.^{3,4} However, the high drug cost and the need for frequent injections are placing a large socioeconomic burden on health care systems and patients. In addition, many patients do not maintain initial best-corrected visual acuity (BCVA) gains with long-term follow-up.²

AMD is a highly complex disease with a broad spectrum of pathophysiologic factors, genetic backgrounds, and morphologic features. From previous clinical trials, it is clear that interindividual treatment requirements are vastly heterogeneous, indicating that the optimal treatment should be tailored to an individual,⁵ using precision medicine instruments.

Introduction of spectral-domain optical coherence tomography (SD-OCT) initially allowed the qualitative and subsequently the quantitative examination of pathomorphologic features of the retina, becoming essential in active monitoring, treatment decisions, and patient visit scheduling on an individualized basis. The most commonly used individualized treatment regimens for the treatment of neovascular AMD, pro re nata (PRN), and treat-and-extend (TE),⁶ both rely on continued OCT imaging to inform decisions on treatment and monitoring. Nevertheless, in real-world clinical practice, both regimens often result in general undertreatment, because both hospitals and patients find it difficult to sustain frequent resource-consuming monitoring visits.² To facilitate resource management, injection decision making, and patient counseling, it is of great interest to be able to predict the extent of treatment requirements for each patient at the beginning of the therapeutic course. However, currently it is not clear what



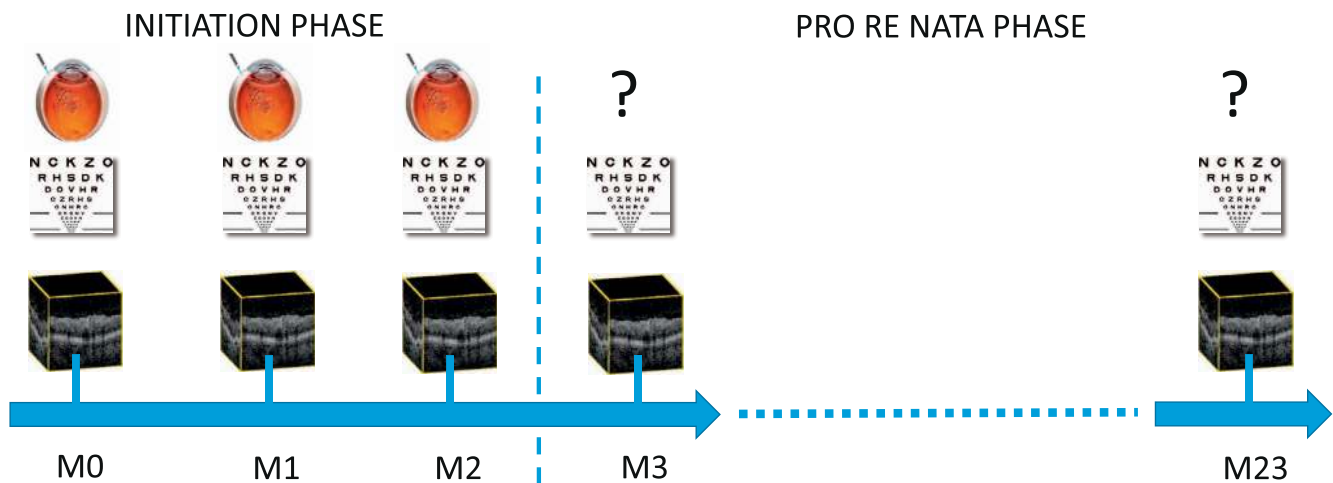


FIGURE 1. Illustration of the monitoring and treatment schedule. The initiation phase consisted of monthly injections for 3 months (M0–M2) followed by the PRN treatment regimen. PRN was based on monthly monitoring and administering 0 to 21 potential injections.

differentiates patients with low or high treatment needs, and OCT imaging biomarkers to predict these individual treatment requirements represent an unmet medical and socioeconomic need.^{7,8}

The aim of this pilot study is to predict, on an individual patient level, low and high anti-VEGF injection requirements during a PRN treatment regimen of patients with neovascular AMD. Our hypothesis suggests that these requirement categories can be predicted by observing retinal morphology and treatment response as early as during the standardized initiation phase of the treatment course. Using automated computational analysis of OCT, a set of spatiotemporal features was extracted from imaging series, characterizing the retina and its anatomic response to the initial anti-VEGF treatment. Machine learning methods were then applied to build a predictive model of the future therapeutic requirements during the PRN regimen. The model was trained and validated on 2-year data from a large-scale prospective randomized controlled trial in treatment-naïve AMD patients.

METHODS

Participants

This post hoc analysis was performed on data of patients undergoing PRN treatment within the HARBOR clinical trial (ClinicalTrials.gov number, NCT00891735). HARBOR was a 24-month, phase III, randomized, multicenter, double-masked, active treatment-controlled study with 1095 randomized patients to evaluate efficacy and safety of intravitreal ranibizumab 0.5 and 2.0 mg administered monthly or on a PRN basis in treatment-naïve patients with subfoveal neovascular AMD. Patients in the PRN groups had monthly evaluations and received ranibizumab monthly for the first three doses (Fig. 1). At the month 3 visit and thereafter, they received ranibizumab only if the retreatment criteria were met (at least five-letter decrease in BCVA from the previous visit or any evidence of disease activity on SD-OCT). For our analysis, we pooled the eyes receiving 0.5 and 2.0 mg because the trial did not report any significant differences between the two doses⁵ and the initiating monthly regimen was identical. The study was conducted in compliance with the Declaration of Helsinki, and approval for this post hoc analysis was obtained by the Ethics Committee at the Medical University of Vienna. Patients

provided written informed consent to participate in the HARBOR trial.

OCT Image Processing and Analysis

The proposed methodology is based on a fully automated image processing and analysis pipeline available at the Vienna Reading Center (VRC), Vienna, Austria. No manual corrections have been performed in this study. All images were acquired with Cirrus HD-OCT III (Carl Zeiss Meditec, Inc., Dublin, CA, USA) presenting $512 \times 128 \times 1024$ voxels, with a size of $11.7 \times 47.2 \times 2.0 \mu\text{m}^3$, covering a volume of $6 \times 6 \times 2 \text{mm}^3$.

OCT Motion Correction. Involuntary eye movements during the acquisition of OCT scans create motion artefacts that affect three-dimensional (3D) image analysis. As a preprocessing step, we reduced motion artifacts using the method of Montuoro et al.⁹ The method takes advantage of self-similarity property of the retina and simultaneously retains the retinal curvature, shape, and potential pathologies. It first corrects motion artifacts along the axial direction by shifting individual A-scans to restore the local shape symmetry of the retina. Second, maximizing pairwise phase correlation between B-scans, correction along the primary (horizontal) scan direction is obtained. It is especially advantageous to our task as it can be retrospectively applied to already acquired OCT images regardless of the scan protocol or device, whereas most other methods require special scanning patterns or multiple orthogonal acquisitions. The motion correction facilitates the subsequent layer segmentation and feature extraction, which both rely on 3D image information.

Retinal Layer Segmentation. Automated retinal layer segmentation is performed with a graph-theoretic method, part of the Iowa Reference Algorithms.^{10,11} The method transforms the problem into a multiscale 3D graph search to optimally and efficiently segment a set of surfaces according to image-based cost function and satisfying a priori hard constraints on surface smoothness and intersurface distances. As the a priori constraints are valid for healthy retinas, only a subset of layer interfaces is well segmented in neovascular AMD population. Thus, the following four principle layer thickness maps were extracted, which were empirically found to be robustly segmented: inner retina (IR), outer nuclear layer (ONL), photoreceptor outer segments with retinal pigment epithelium (OR), and total retinal thickness (TRT). An example of segmented surfaces denoting those layers is shown in Figure 2.

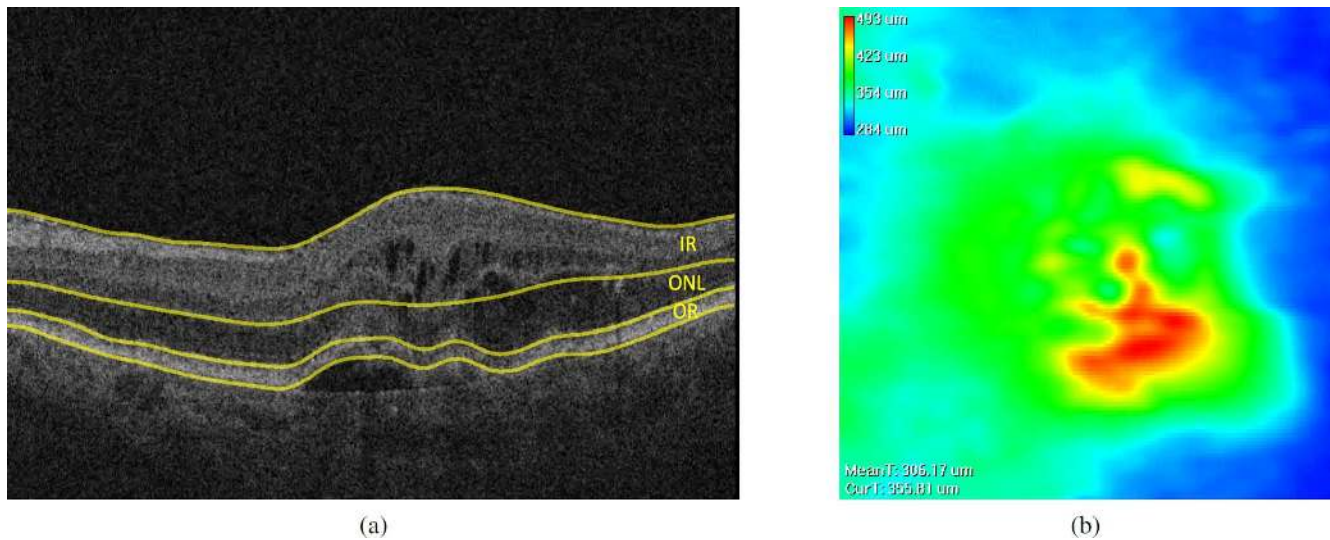


FIGURE 2. Example of (a) the automated layer segmentation result with the four principal surfaces denoted (in yellow) and (b) the total retinal thickness map.

Intraretinal and Subretinal Fluid Segmentation. Segmentation of intraretinal cystoid fluid (IRF) and subretinal fluid (SRF) was performed per B-scan using a validated segmentation algorithm based on deep learning.¹² First, based on the top and the bottom retinal layer, a mask is computed denoting the retina extending from the inner limiting membrane (ILM) to the RPE. Then, every voxel within the mask is classified with a multiscale convolutional neural network (CNN) as belonging to one of the three classes: Normal retina, IRF, or SRF (Fig. 3). The CNN had been trained in a supervised manner using a training set of 157 OCT volumes with $\approx 20,000$ manually annotated B-scans, acquired with the same OCT device model (Cirrus; Zeiss) and having the same pathology (neovascular AMD), which were disjoint from the set of images in the HARBOR trial.

Predictive Model of Treatment Requirements

For each eye, from its longitudinal series of three OCT volumes (baseline, month 1, and month 2) and the derived segmentations, we extracted a set of quantitative features characterizing the underlying retinal pathomorphology. For the imaging features to correspond across subjects, before the feature extraction, all scans of left eyes were mirrored to conform to scans of a right eye. From the image segmentations 2D maps

were computed corresponding to the thickness maps of the four layers, as well as volume and en face area maps of both IRF and SRF, resulting in eight 2D maps in total, with examples shown in Figure 4a. Analyzing data in high-dimensional OCT volumes is affected by the so-called ‘‘Curse of Dimensionality,’’ where learning is very difficult and prone to overfitting. To limit the dimensionality of the feature vector and facilitate the machine learning, we summarized the A-scan properties spatially across the regions defined by the Early Treatment Diabetic Retinopathy Study (ETDRS) grid as depicted in Figure 4b. The ETDRS grid was placed at the center of the scan, and the mean feature values per ETDRS subregions were computed. In addition to the nine ETDRS grid cells, we additionally included the central 3 mm, central 6 mm, and the rings corresponding to the parafoveal and perifoveal bands, resulting in 13 spatial regions in total. Such ETDRS-related features have the additional advantage of being easier to interpret than A-scan related ones, due to widespread use of ETDRS grid in ophthalmology. To this set of imaging features, we added the measured BCVA. To measure the rate of change of the longitudinal features, the differences between the corresponding features of the consecutive time points (month 1 – month 0 and month 2 – month 1) were further included. This resulted in the number of local spatio-temporal features being 525, computed as follows: (8 feature maps \times 13 spatial regions + 1

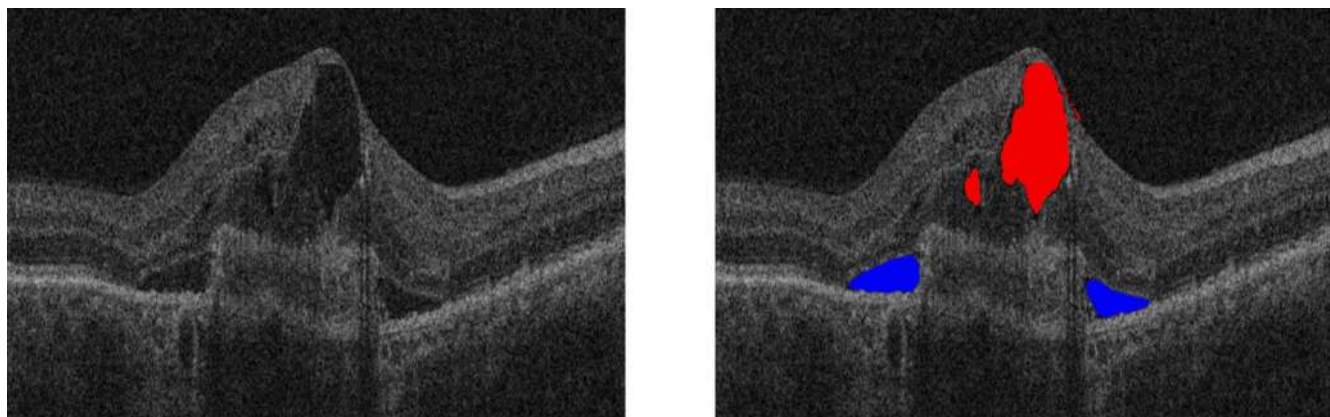
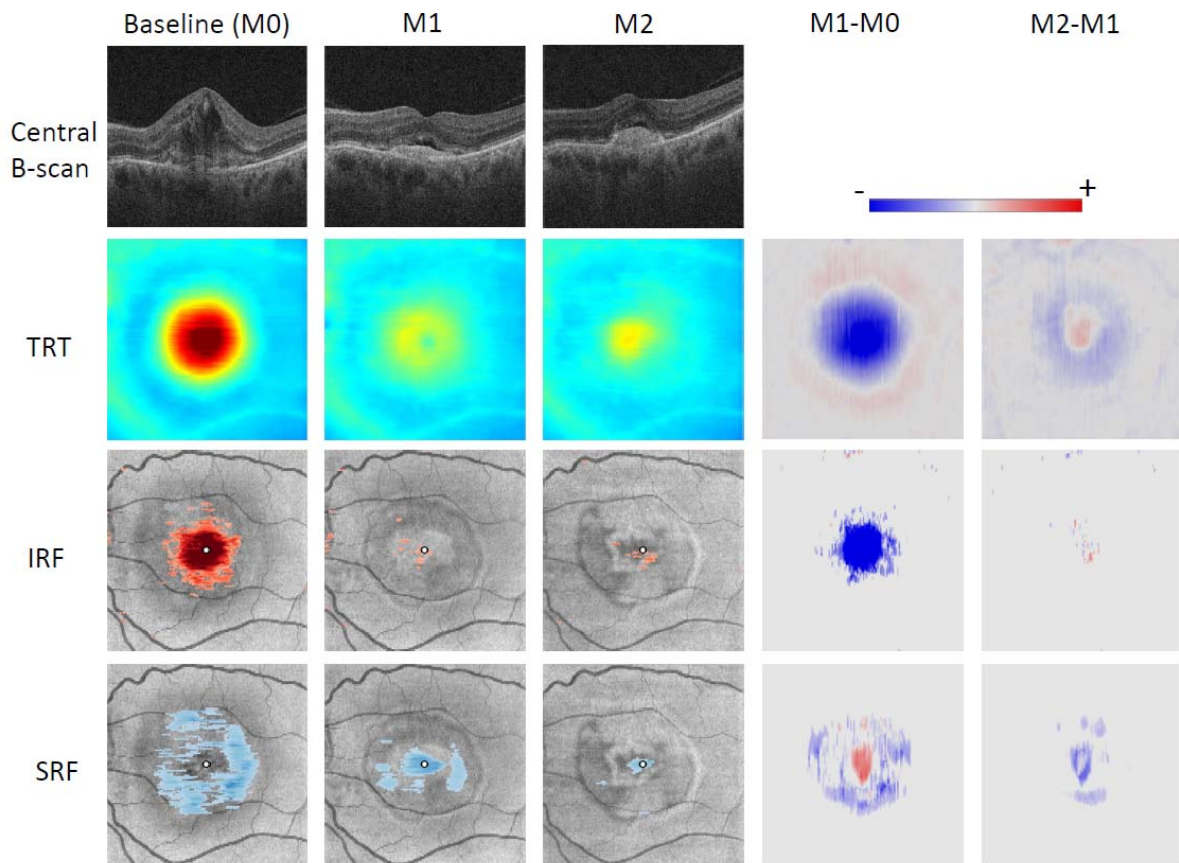
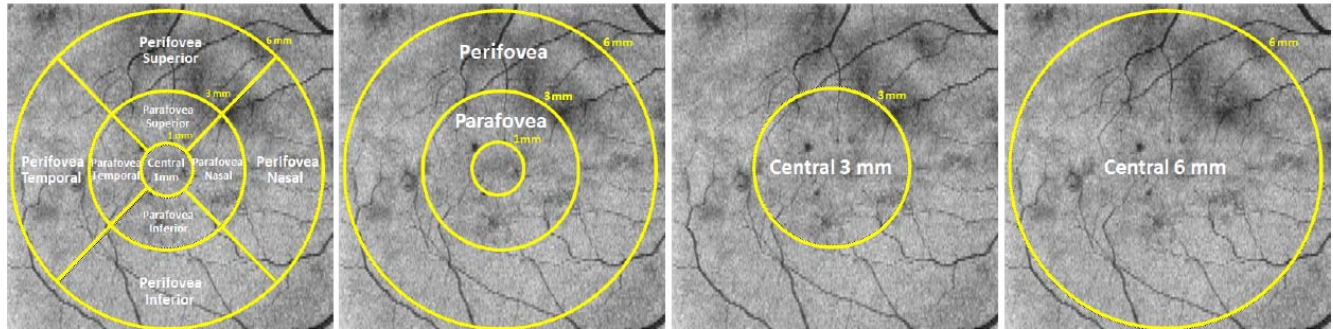


FIGURE 3. Example of the automated fluid segmentation result of intraretinal (in red) and subretinal (in blue) fluid.



(a)



(b)

FIGURE 4. (a) Example of the initiation phase OCT images and corresponding segmentations of TRT, IRF, and SRF for the five temporal elements across months (M). (b) Spatial localization of the features based on the 13 regions obtained from ETDRS grid with circle diameters of 1, 3, and 6 mm.

BCVA) \times 5 temporal elements. Last, demographic features were added: sex, race, age, and smoking status together with the fluorescein angiogram pattern type, for a total of 530 features.

The maximum number of injections during the 2-year PRN regimen is 21 (months 3 to 23). We defined the category of “low” requirements to consist of patients in lower quartile of the number of injections, which corresponded to receiving no more than five injections. Analogously, the category of “high” requirements was defined to consist of patients in upper quartile, which corresponded to receiving ≥ 16 injections. The remaining eyes in the interquartile range were assigned to the “medium” requirements category. We aim to discriminate the patients in the low requirement group from the medium and high requirement groups, and analogously, the ones in the high requirement group from the medium and low requirement

groups. Thus, we pose the problem as a multiclass one-versus-all classification.

Finally, a machine learning approach based on the random forest classifier¹³ was used to obtain a predictive model of the low and high treatment requirements from the set of the above features. Random forest was grown with 1000 trees for which the out of bag mean squared error was observed to have converged. The number of features to randomly sample as candidates at each split of a tree was chosen to be the square root of the number of features ($\sqrt{530}$), which is the default setting for a classification task.¹³

Statistical Analysis

To assess how the results would generalize to an independent data set, the performance of the predictive model was

evaluated using cross-validation (CV). The sample is randomly partitioned into 10 equal sized complementary subsamples. Then, across 10 iterations, onefold is selected as the validation set and the other nine folds serve as the training set. Thus, in each iteration, 90% of the data was used for training and 10% for validation. In such a setting, all patients in the data set participate in a validation, and each is predicted exactly once. In our study, 10-fold CV, as opposed to exhaustive or 2-fold CV, provided a very good compromise between minimizing the correlation between the learned models across the folds and the size of the training set available to a model in each fold. Because the random forest model produces a probabilistic estimate (percentage of the trees that voted positive) of belonging to a certain treatment requirement category, the quantitative performance across all the validated predictions is summarized with an area under the receiver operating characteristic (ROC) curve and presented as sensitivity and specificity at an operating point.

To put the results into perspective and create a human performance benchmark, we asked a retinal specialist (XL) to perform the grading by making an estimate of the treatment requirement category based on observing the three volumetric OCT scans from the initiation phase (i.e., the same set of scans as used by the predictive model). For comparison, the operating point of the predictive model was also reported at the level of the grader's sensitivity/specificity.

To evaluate the predictive role of features, the importance measure implemented within random forest classifier relies on permuting the values of a feature and measuring how much the permutation decreases the prediction accuracy of the model. Important features can then be detected as those where the permutation decreases the prediction accuracy the most. Finally, to evaluate if a single feature separates treatment requirement categories with statistical significance, a two-sample *t*-test was performed, and the Hochberg-Bonferroni approach was used to control the overall significance level for the multiple comparison effect.

RESULTS

Patient Characteristics

Of a total of 548 patients in the PRN arms, 375 ($\approx 70\%$) were randomly selected, whereas the remainder was kept for future use. Of these, 40 were discarded due to not having completed the 2-year study, whereas some were discarded due to missing scans (12 eyes) or image quality and segmentation issues (6 eyes) during the initiation phase. Finally, data of 317 eyes were used in our study. In this subset of the HARBOR study population, the mean \pm SD age of patients was 78 ± 8 years (range, 53 to 97 years); 57% were female, and 96% were white. The mean \pm SD baseline visual acuity (VA) was 55 ± 12 letters (range, 15 to 76 letters). Overall, 50% of patients had minimally classic choroidal neovascularization (CNV) lesions, 14% had predominantly classic lesions, and 36% had purely occult CNV.

Prediction of Treatment Requirements

The number of injections administered during the PRN period from month 3 until month 23 ranged from 0 to 21. One patient required no injections, and 16 patients required monthly injections. The distributed injection burden in the HARBOR population was as follows: 70 (22%) required low, 71 (22%) required high, and 176 (56%) required a medium number of flexible injections. A ROC of the predictive model, representing the trade-off between specificity and sensitivity, is shown in Figure 5a. We also built the predictive model in steps as the

measurements are becoming available along the progress of the initiation phase (Fig. 5b). The areas under the curve (AUCs) for predicting the categories grew monotonically, starting with 0.60 and 0.61 at baseline, 0.68 and 0.74 at month 1, and finally 0.70 and 0.77 at month 2 for the low and high requirements, respectively. Hence, an estimation after the first treatment interval was close in performance to the final prediction scores. The model relied mainly on the last measured time point, as including earlier time points did not result in the further increase in performance.

For predicting the low injection requirements, the false positives (patients wrongly predicted to have low requirements) are considered clinically more adverse than the false negatives. Thus, there the operating point was set to favor specificity over sensitivity. No such preference exists for predicting the high requirements, so there the operating point was set to maximize the specificity and the sensitivity. Using such defined operating points, the predictive model detected the low requirements patients with 71% specificity and 58% sensitivity and the high requirements patients with 71% specificity and 70% sensitivity (Fig. 5a).

A human grader has achieved sensitivities of 0.41 and 0.37 and specificities of 0.84 and 0.84 for detecting the low and high treatment requirements (Fig. 5a), respectively. Thus, the grader was conservative in assigning low and high treatment categories, and consequently, the errors were mostly false negatives, where the low/high treatment categories were graded as the medium ones. Corresponding sensitivities for the equivalent 0.84 specificity of the predictive model were 0.38 and 0.54 for the low and the high treatment requirements, respectively. Thus, the model had a comparable performance for predicting the low and almost 50% better performance in predicting the high treatment requirements.

Relevance of Retinal Features

The top 15 important features found are shown in Figure 6a. The feature importance was correlated with the course of the initiation phase, with the most important features being measured at the end of the initiation phase at month 2. The distribution of the top 50 most important features across the type and the time point measured (Fig. 6b) shows the highest number of features at month 2, especially the ones related to SRF, as well as layer thicknesses and SRF at month 1, whereas baseline features were poorly represented. The importance of differential features was low, with none appearing in the top 15, and only four such features appearing in the top 50. In comparison to the imaging data, the role of BCVA was moderate, appearing in the middle of the list of features sorted by importance. In comparison to the longitudinal data, demographic and fluorescein angiogram features were found not to be important, with only age having a moderate importance. In fact, the four features at the very bottom of the list were sex, fluorescein angiogram CNV type, smoking status, and race.

The role of TRT, IRE, and SRF area and volume, at the central 1 and 3 mm, during the three time points were further individually investigated for statistically significant differences between the retreatment categories. After correcting for multiple comparisons ($5 \times 2 \times 3 = 30$ features examined), the formal statistical significance level was set to $P < 0.002$. The following features were found to be statistically significantly different between the groups. The role of TRT at the central 1 mm is shown in Figure 7. The difference between the three groups gradually increased with time, and at the end of the initiation phase (month 2), there was a difference between the low and the medium groups ($P = 0.02$) and a significant difference between the medium and the high ($P < 0.001$)

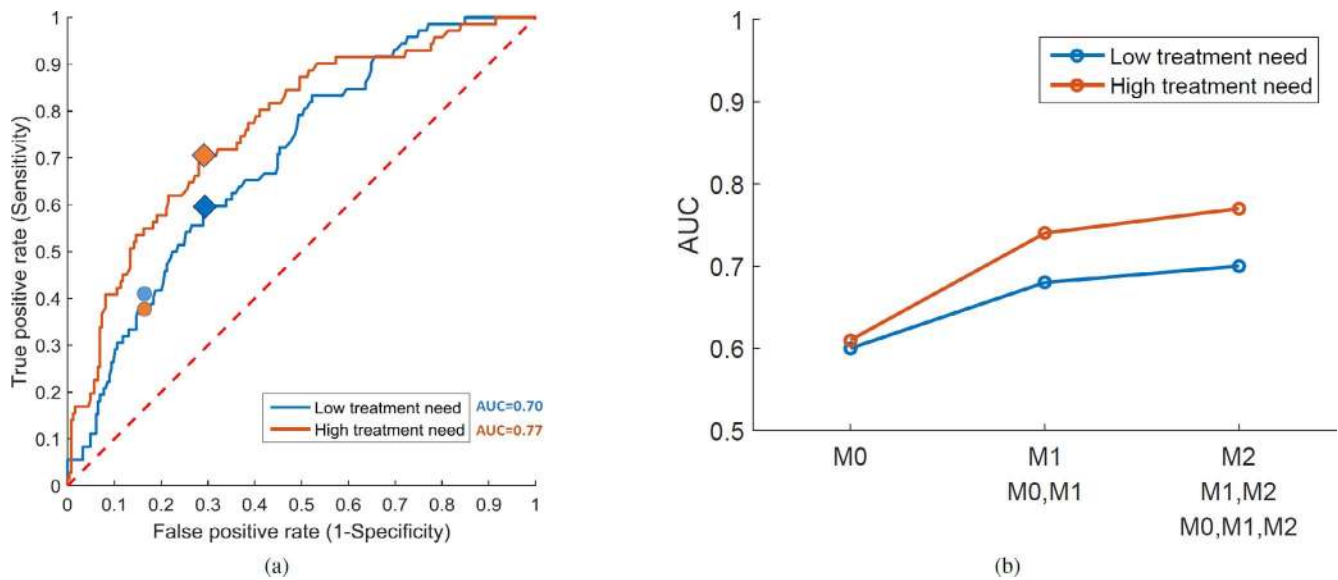


FIGURE 5. (a) ROC of the predictive model with the denoted operating points (diamond) and the operating points of the human grader (circle). (b) Progress of the AUC of the prediction along the initiation phase from month 0 (M0) to month 2 (M2).

groups. The role of IRF and SRF became evident at month 2 as well. Differences in IRF areas at the central 3 mm (Fig. 8a) were close to significant ($P = 0.003$) between the low and medium groups and strongly significant between the medium and high groups ($P < 0.001$). Regarding the role of SRF (Fig. 8b), the difference in volume in the central 3 mm was close to significant ($P = 0.006$) between the low and medium groups and was strongly significantly different between the medium and high requirement subjects ($P < 0.001$). All the above features were strongly statistically significantly different between the low and high groups ($P < 0.001$).

DISCUSSION

In this pilot study, we presented and evaluated a computer-based method to learn and predict low and high anti-VEGF treatment requirements of neovascular AMD patients from a longitudinal series of OCT scans acquired during the initiation phase. The predictions were based on fully automated image

analysis and obtained by machine learning using a random forest as a nonlinear predictive classification model. Development and validation were performed on data from the HARBOR clinical trial, which is particularly appropriate due to its large cohort size, standardized OCT imaging, and an effective PRN retreatment protocol.⁵

Our results demonstrate that a solid AUC of 70% to 80% was achieved for predicting both the low and the high treatment categories. The same performance was achieved by taking the last available time point only, because, interestingly, differential features measuring the change along the initiation phase were not found to have an important role. Thus, the state of the retina after the initial anti-VEGF doses seems more predictive of the future treatment requirements than the preceding baseline condition or the magnitude of morphologic improvement. Predicting the high requirements proved to be a more successful task, with the overall AUC performance being better than the prediction of the low treatment requirements, which is in line with clinical needs to avoid undertreatment. For both categories, similar performance was already achieved using

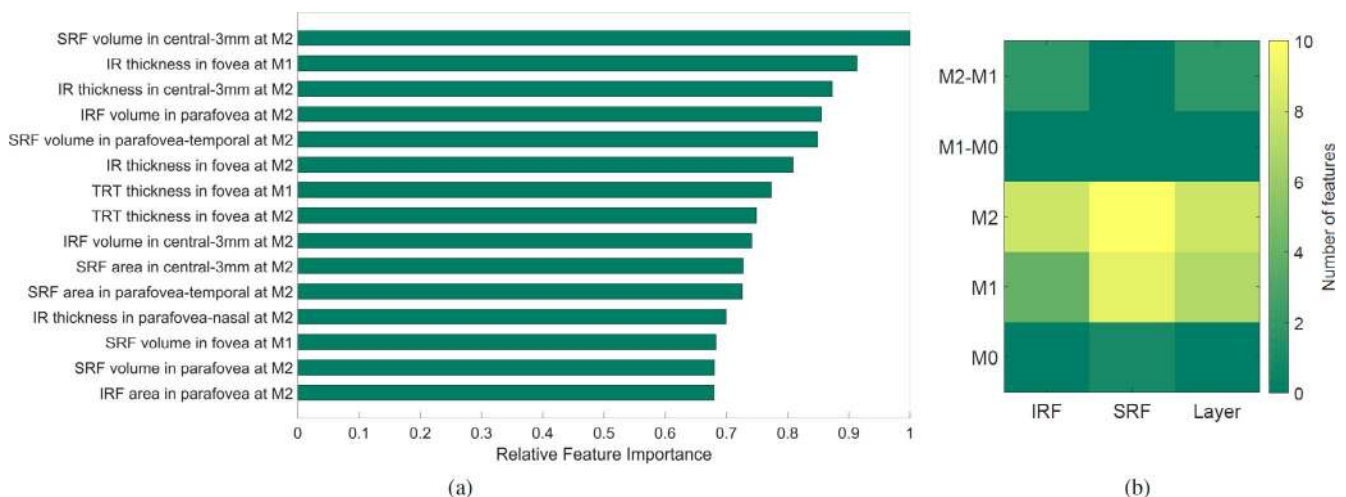


FIGURE 6. Feature importance. (a) Top 15 features sorted by the estimated feature importance. (b) Distribution of the top 50 features grouped by the type and the measured time element.

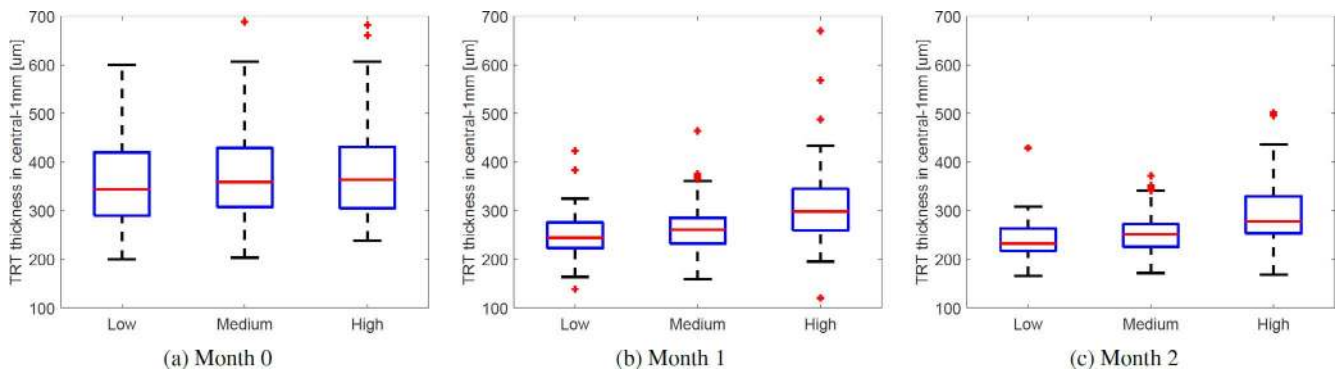


FIGURE 7. TRT at the central 1 mm between the three treatment categories during the initiation phase.

measurements from month 1, indicating that the observation of the retinal response after only one injection was already predictive of future management requirements.

The human performance evaluation revealed that this prediction task is also difficult for a human grader as well, who performed similarly across the detection of the low and high treatment categories, producing low sensitivity and high specificity. We should emphasize that such a task is not commonly done in the clinic; it is not part of ophthalmologists' training, and a large intergrader variability is expected. However, the evaluation showed that such treatment requirement prediction task is very suited for a machine learning approach where the machine could learn highly complex multidimensional patterns from high-resolution OCT scans and use the knowledge to make equal (for the low category) or even better (for the high category) predictions than human graders.

Inspecting a few individual features, the TRT in the central 1 mm, IRF area, and SRF volume in the central 3 mm were found to be discriminative. The low retreatment group had mostly dry retinas by the end of the initiation phase. Nevertheless, this was also the case with some of the subjects from the medium requirement group. Such absence of any visible exudative features on the OCT made the task of distinguishing the low from the medium requirements group more difficult, resulting in poorer sensitivity for predicting the low requirement category and encouraging clinicians to continue with a tight monitoring regimen despite a satisfactory response early on. For the high requirement group, the more persistent the exudates in the retina were during treatment initiation, the more retreatments were required.

This work follows on the recent breakthroughs in the field of artificial intelligence. In particular, machine learning fueled by large data sets is showing a great promise for its application in precision medicine. Furthermore, advances in automated 3D

image analysis are allowing streamlined, objective, and repeatable quantification of the underlying pathomorphologic properties. The combination of the two technologies has recently led to successful first steps in using longitudinal OCT imaging data for predicting disease recurrence,¹⁴ treatment responders,¹⁵ and progression to late AMD.¹⁶ To the best of our knowledge, our study is the first to attempt to predict the anti-VEGF treatment requirements in AMD, especially on such a large homogeneous cohort.

The study has several limitations. The predictive model relies on features resulting from image segmentation methods, and they therefore have to be accurately extracted (Figs. 2a, 3). We did not perform correction of the automated segmentation results because using such a large number of scans (951 OCT volumes) makes manual corrections prohibitively time consuming. However, we do expect the predictive model to be resistant to some segmentation errors due to the dimensionality reduction performed by averaging over the ETDRS grid cells and the large cohort size used that should make random forest robust to outliers. In addition, we could use only the main four retinal layers as the others were found not to be well segmented. Future improvements in automated intraretinal layers of patients with macular edema would allow us to rely on information from more layers and further boost the performance of the predictive model. Regarding the human performance evaluation, the grader relied on the OCT scans only and was not informed of the demographic variables, but we don't expect that this impacted the comparison as the predictive model did not find them to be very predictive. Finally, the fovea was assumed to be at the very center of the scan, potentially resulting in misalignments of the superimposed ETDRS grid. Nevertheless, due to feature averaging over relatively large spatial areas formed by the ETDRS grid regions, small position inaccuracies are not expected to have a large negative effect.

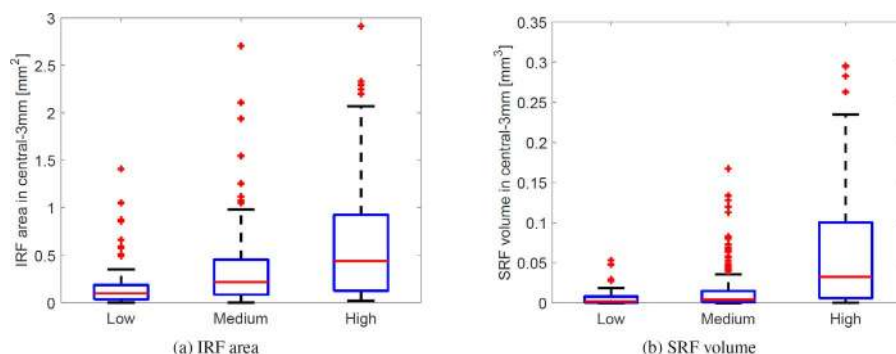


FIGURE 8. IRF area (a) and SRF volume (b) at month 2 across the three treatment categories.

Our underlying hypothesis was that longitudinal OCT images during the initiation phase contain the necessary information that can reveal future treatment requirements. The treatment categories were defined based on injection percentiles to focus on detecting patients lying on the extremes of treatment requirements. Our results show that prediction is possible, but currently it is not clear what is the upper bound of accuracy of such a prediction task (i.e., how much of treatment requirement variability can be explained by OCT imaging alone?). Also, the treatment requirements observed in our learning model were the result of specific, protocol-defined, retreatment criteria with a specific anti-VEGF treatment (ranibizumab) used in the HARBOR clinical trial and might not represent the clinically optimal treatment requirements for other anti-VEGF drugs. Nevertheless, the retreatment criteria of HARBOR appeared to be very efficient as shown by excellent outcomes in comparison with other PRN-guided trials. In addition, as the population participating in the trial was subject to strict inclusion criteria, how our predictive model would perform on the general population requires further investigation. Nevertheless, as a result of the large variability in aggressive activity in neovascular AMD, individualization of the therapeutic management is clearly an appropriate strategy. Only a minor proportion of patients require a monthly retreatment regimen (approximately 5% in our representative study). Rates were retrospectively shown to be similarly low in other PRN-guided trials such as CATT, IVAN, and GEFAL.¹⁷⁻¹⁹ The most important risk for loss of initial BCVA gains during flexible regimens is undertreatment, which has been demonstrated both in long-term clinical trials such as HORIZON²⁰ and the VIEW Open-Label Extension Study (ClinicalTrials.gov number, NCT00964795) as well as in real-world studies and registries, such as LUMINOUS and WAVE.^{21,22}

Attempts to treat all patients with a bimonthly regimen as used for aflibercept in the VIEW studies may be adequate for many patients. Averaging results for treatment efficacy and frequency in this way can lead to noninferior outcomes. However, advanced analyses have meanwhile shown that such a 2q8 regimen misses optimal efficacy for a subgroup of patients.²³ Some patients will be overtreated, with unnecessary injections and associated expenses being incurred, and most importantly, some with an aggressive disease course will be undertreated and significantly lose vision.²³ Advanced prediction analysis, however, offers an individualized method to adjust retreatment schedules to disease activity while simultaneously reducing the socioeconomic burden of AMD therapy. Also, the proposed prediction is already efficient from the first 2 months of treatment initiation, and it applies to almost half of the AMD population, including the high-risk groups of under-/overtreatment at both ends of the spectrum.

As part of future work, further understanding and interpretation of the effect of different phenotypes on the treatment requirements is needed. This should be combined with the efforts to include genetic markers and quantify additional imaging biomarkers, namely outer retinal tubulation, hyperreflective foci, subretinal hyperreflective material, and fibrous scarring.⁷ Also, retreatment indications may be modified by experience from structure/function correlation showing variable associations between different locations of fluid pooling.⁷ Finally, we plan to improve the precision of the model to predict the number of required injections per year or the mean time to retreatment.

In summary, results of our pilot study show that early response to anti-VEGF therapy for AMD is predictive of treatment requirements and indicate the potential for imaging to guide monitoring and treatment regimen. The presented precision medicine tool represents a first step toward predicting the expected treatment frequency consistent with

the level of disease activity that can ultimately lead to substantial improvement in resource management and patient counseling in a reliable way. The search for features relevant for the prognosis of management of neovascular AMD, such as IRF/SRF at different locations and timelines, will also strongly improve the insight into pathophysiologic mechanisms of disease progression.

Acknowledgments

Supported by the Austrian Federal Ministry of Science, Research and Economy, the National Foundation for Research, Technology and Development, and Genentech, Inc.

Disclosure: **H. Bogunović**, None; **S.M. Waldstein**, Bayer Healthcare AG (C), Novartis Pharma AG (C); **T. Schlegl**, None; **G. Langs**, None; **A. Sadeghipour**, None; **X. Liu**, None; **B.S. Gerendas**, None; **A. Osborne**, Genentech (E); **U. Schmidt-Erfurth**, Bayer Healthcare AG (C), Novartis Pharma AG (C), Alcon Laboratories (C), Boehringer Ingelheim GmbH (C)

References

- Jager RD, Mieler WF, Miller JW. Age-related macular degeneration. *N Engl J Med*. 2008;358:2606-2617.
- Maguire MG, Martin DF, Ying G-S, et al. Five-year outcomes with anti-vascular endothelial growth factor treatment of neovascular age-related macular degeneration: the comparison of age-related macular degeneration treatments trials. *Ophthalmology*. 2016;123:1751-1761.
- Bloch SB, Larsen M, Munch IC. Incidence of legal blindness from age-related macular degeneration in Denmark: year 2000 to 2010. *Am J Ophthalmol*. 2012;153:209-213.
- Sloan FA, Hanrahan BW. The effects of technological advances on outcomes for elderly persons with exudative age-related macular degeneration. *JAMA Ophthalmol*. 2014;132:456-463.
- Busbee BG, Ho AC, Brown DM, et al. Twelve-month efficacy and safety of 0.5 mg or 2.0 mg ranibizumab in patients with subfoveal neovascular age-related macular degeneration. *Ophthalmology*. 2013;120:1046-1056.
- Gupta OP, Shienbaum G, Patel AH, Fecarotta C, Kaiser RS, Regillo CD. A treat and extend regimen using ranibizumab for neovascular age-related macular degeneration clinical and economic impact. *Ophthalmology*. 2010;117:2134-2140.
- Schmidt-Erfurth U, Waldstein SM. A paradigm shift in imaging biomarkers in neovascular age-related macular degeneration. *Prog Retin Eye Res*. 2016;50:1-24.
- Chakravarthy U, Goldenberg D, Young G, et al. Automated identification of lesion activity in neovascular age-related macular degeneration. *Ophthalmology*. 2016;123:1731-1736.
- Montuoro A, Wu J, Waldstein S, et al. Motion artefact correction in retinal optical coherence tomography using local symmetry. In: Golland P, Hata N, Barillot C, Hornegger J, Hower R, eds. *Medical Image Computing and Computer-Assisted Intervention LNCS* Vol. 8674. Cham, Switzerland: Springer; 2014:130-137.
- Garvin MK, Abramoff MD, Wu X, Russell SR, Burns TL, Sonka M. Automated 3-D intraretinal layer segmentation of macular spectral-domain optical coherence tomography images. *IEEE Trans Med Imaging*. 2009;28:1436-1447.
- Li K, Wu X, Chen DZ, Sonka M. Optimal surface segmentation in volumetric images: a graph-theoretic approach. *IEEE Trans Pattern Anal Mach Intell*. 2006;28:119-134.
- Schlegl T, Waldstein SM, Vogl W-D, Schmidt-Erfurth U, Langs G. Predicting semantic descriptions from medical images with convolutional neural networks. In: Ourselin S, Alexander DC, Westin CF, Cardoso MJ, eds. *Information Processing in Medical Imaging*. LNCS Vol. 9123. Cham, Switzerland: Springer; 2015:437-448.

13. Breiman L. Random forests. *Mach Learn*. 2001;45:5–32.
14. Vogl W-D, Waldstein S, Gerendas B, et al. Spatio-temporal signatures to predict retinal disease recurrence. In: Ourselin S, Alexander DC, Westin CF, Cardoso MJ, eds. *Information Processing in Medical Imaging*. LNCS Vol. 9123. Cham, Switzerland: Springer; 2015:152–163.
15. Bogunović H, Abramoff MD, Zhang L, Sonka M. Prediction of treatment response from retinal OCT in patients with exudative age-related macular degeneration. In: Chen X, Garvin MK, Liu JJ, eds. *Proceedings of the Ophthalmic Medical Image Analysis First International Workshop, OMIA 2014*. Boston, MA, September 14, 2014;129–136.
16. 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.
17. Ying GS, Maguire MG, Daniel E, et al. Association of baseline characteristics and early vision response with 2-year vision outcomes in the comparison of AMD Treatments Trials (CATT). *Ophthalmology*. 2015;122:2523–2531.
18. Chakravarthy U, Harding SP, Rogers CA, et al. A randomised controlled trial to assess the clinical effectiveness and cost-effectiveness of alternative treatments to inhibit VEGF in age-related choroidal neovascularisation (IVAN). *Health Technol Assess (Rockv)*. 2015;19:1–298.
19. Kodjikian L, Souied EH, Mimoun G, et al. Ranibizumab versus bevacizumab for neovascular age-related macular degeneration: results from the GEFAL noninferiority randomized trial. *Ophthalmology*. 2013;120:2300–2309.
20. Singer MA, Awh CC, Sadda S, et al. HORIZON: an open-label extension trial of ranibizumab for choroidal neovascularization secondary to age-related macular degeneration. *Ophthalmology*. 2012;119:1175–1183.
21. Finger RP, Wiedemann P, Blumhagen F, Pohl K, Holz FG. Treatment patterns, visual acuity and quality-of-life outcomes of the WAVE study: a noninterventional study of ranibizumab treatment for neovascular age-related macular degeneration in Germany. *Acta Ophthalmol*. 2013;91:540–546.
22. Holz FG, Bandello F, Gillies M, et al. Safety of ranibizumab in routine clinical practice: 1-year retrospective pooled analysis of four European neovascular AMD registries within the LUMINOUS programme. *Br J Ophthalmol*. 2013;97:1161–1167.
23. Jaffe GJ, Kaiser PK, Thompson D, et al. Differential response to anti-VEGF regimens in age-related macular degeneration patients with early persistent retinal fluid. *Ophthalmology*. 2016;123:1856–1864.