

Probabilistic Forecasting of Anti-VEGF Treatment Frequency in Neovascular Age-Related Macular Degeneration

Maximilian Pfau^{1,2}, Soumya Sahu^{3,4}, Rawan Allozi Rupnow^{3,4}, Kathleen Romond³, Desiree Millet², Frank G. Holz², Steffen Schmitz-Valckenberg^{2,5}, Monika Fleckenstein⁵, Jennifer I. Lim³, Luis de Sisternes⁶, Theodore Leng⁷, Daniel L. Rubin¹, and Joelle A. Hallak³

¹ Department of Biomedical Data Science, Stanford University, Palo Alto, CA, USA

² Department of Ophthalmology, University of Bonn, Bonn, Germany

³ Department of Ophthalmology and Visual Sciences, University of Illinois at Chicago, Chicago, IL, USA

⁴ Department of Epidemiology and Biostatistics, School of Public Health, University of Illinois at Chicago, Chicago, IL, USA

⁵ John A. Moran Eye Center, University of Utah, Salt Lake City, UT, USA

⁶ Carl Zeiss Meditec AG, Dublin, CA, USA

⁷ Byers Eye Institute at Stanford, Stanford University School of Medicine, Palo Alto, CA, USA

Correspondence: Joelle A. Hallak, Department of Ophthalmology & Visual Sciences, University of Illinois at Chicago, 1855 W. Taylor Street, M/C 648, 60612 Chicago IL, USA. e-mail: joelle@uic.edu

Received: November 25, 2020

Accepted: May 19, 2021

Published: June 29, 2021

Keywords: personalized medicine; machine-learning; age-related macular degeneration; anti-VEGF treatment; probabilistic prediction

Citation: Pfau M, Sahu S, Rupnow RA, Romond K, Millet D, Holz FG, Schmitz-Valckenberg S, Fleckenstein M, Lim JI, de Sisternes L, Leng T, Rubin DL, Hallak JA. Probabilistic forecasting of anti-VEGF treatment frequency in neovascular age-related macular degeneration. *Transl Vis Sci Technol.* 2021;10(7):30. <https://doi.org/10.1167/tvst.10.7.30>

Purpose: To probabilistically forecast needed anti-vascular endothelial growth factor (anti-VEGF) treatment frequency using volumetric spectral domain–optical coherence tomography (SD-OCT) biomarkers in neovascular age-related macular degeneration from real-world settings.

Methods: SD-OCT volume scans were segmented with a custom deep-learning-based analysis pipeline. Retinal thickness and reflectivity values were extracted for the central and the four inner Early Treatment Diabetic Retinopathy Study (ETDRS) subfields for six retinal layers (inner retina, outer nuclear layer, inner segments [IS], outer segments [OS], retinal pigment epithelium–drusen complex [RPEDC] and the choroid). Machine-learning models were probed to predict the anti-VEGF treatment frequency within the next 12 months. Probabilistic forecasting was performed using natural gradient boosting (NGBoost), which outputs a full probability distribution. The mean absolute error (MAE) between the predicted versus actual anti-VEGF treatment frequency was the primary outcome measure.

Results: In a total of 138 visits of 99 eyes with neovascular AMD (96 patients) from two clinical centers, the prediction of future anti-VEGF treatment frequency was observed with an accuracy (MAE [95% confidence interval]) of 2.60 injections/year [2.25–2.96] ($R^2 = 0.390$) using random forest regression and 2.66 injections/year [2.31–3.01] ($R^2 = 0.094$) using NGBoost, respectively. Prediction intervals were well calibrated and reflected the true uncertainty of NGBoost-based predictions. Standard deviation of RPEDC-thickness in the central ETDRS-subfield constituted an important predictor across models.

Conclusions: The proposed, fully automated pipeline enables probabilistic forecasting of future anti-VEGF treatment frequency in real-world settings.

Translational Relevance: Prediction of a probability distribution allows the physician to inspect the underlying uncertainty. Predictive uncertainty estimates are essential to highlight cases where human-inspection and/or reversion to a fallback alternative is warranted.

Introduction

Age-related macular degeneration (AMD) is the leading cause of legal blindness in industrialized countries.¹⁻³ Anti-vascular endothelial growth factor (anti-VEGF) therapy may halt or significantly delay vision loss in eyes with neovascular AMD^{4,5}; however, a plethora of studies highlight treatment variability, undertreatment, and loss of initially gained visual acuity improvement over time.⁶⁻⁸

Undertreatment may be related to (1) misdiagnosis and conservative re-treatment criteria⁹ and (2) to the overall burden of treatment, which includes the burden on patients, as well as the complexity of scheduling visits for elderly patients in need of a companion.¹⁰ Recent innovative therapeutic approaches such as long-acting anti-VEGF inhibitors,¹¹ or administration of therapeutics via a port delivery system,¹² have the potential to reduce treatment burden in neovascular AMD. However, because of potential side effects, preselection of patients in need of frequent injections is required.¹¹⁻¹³ Machine learning (ML) models can potentially be applied to assist in patient screening.

Recently, ML and deep learning (DL) applications have shown great promise for AMD. Applications include screening of neovascular AMD,^{14,15} predicting future neovascular conversion of AMD,^{16,17} and imputing, from imaging data, visual function in terms of the best-corrected visual acuity,¹⁸ retinal light sensitivity,^{19,20} and vision-related quality of life.²¹ By extension, ML and DL approaches may help to identify “poor responders” to anti-VEGF therapy in need of a high injection frequency.²² Bogunović et al.²² were able to classify patients with low and high treatment frequency with an area under the receiver operating characteristic curve of 0.7 and 0.77 in the setting of a standardized clinical trial data set. However, the applied prediction model only provided the prediction point for each patient without a full probability distribution over the entire outcome space. For clinical decision making, an estimate of the uncertainty for each single prediction constitutes a prerequisite. Predictions with high certainty (i.e., with a narrow probability distribution) would allow physicians to use the prediction for clinical decision making about upcoming anti-VEGF therapy approaches. In contrast, uncertain predictions would enable physicians to opt for a manual fallback alternative (e.g., conventional “pro re nata” [PRN] or “treat and extend” [T&E] protocol with established anti-VEGF agents). Recently, multiple algorithms have been proposed to allow for probabilistic forecast-

ing.^{23,24} These models may particularly be useful for medical applications.

The purpose of our study was to evaluate the ability of a novel probabilistic forecasting model to predict future anti-VEGF treatment frequency with an estimate of uncertainty using real-world clinical data. This innovative probabilistic forecasting model provided a measure of predictive uncertainty for each individual prediction. Specifically, we extended the previously developed NGBoost algorithm²³ with the addition of a negative binomial distribution as a probability distribution to adequately reflect the needed anti-VEGF injection frequency.

Methods

Patients

Imaging and injection frequency data were collected from two tertiary centers: the Department of Ophthalmology and Visual Sciences, University of Illinois at Chicago, Chicago, IL, USA, and the University Eye Hospital Bonn, University of Bonn, Germany. Data were exported in an anonymized manner, and this study adhered to the tenets of the Declaration of Helsinki. The study was approved by the local institutional review boards at the University of Illinois at Chicago and at Bonn University.

The inclusion criteria included patients older than 55 years of age, and choroidal neovascularization secondary to AMD in at least one eye.²⁵ The diagnosis of AMD was based on the presence of drusen and pigmentary changes. Exclusion criteria included history of vitreoretinal surgery, laser photocoagulation, and ocular radiation therapy or other retinal diseases in the study eye such as retinal vascular diseases.²⁵ Patients were treated on the basis of a PRN²⁶ or T&E protocol with a conventional anti-VEGF inhibitor (bevacizumab, ranibizumab or aflibercept). Patients could be either treatment naïve at the time of the spectral domain–optical coherence tomography (SD-OCT) scan (beginning of the one-year interval) or pretreated. Thus the overall variability of number of injections per year was high in this cohort, and the assessment of model performance constitutes a conservative estimate. Because the primary aim of this study was to demonstrate the applicability of probabilistic forecasting, rather than the refined assessment of the efficacy of anti-VEGF agents, all data were pooled. This approximation is reasonable given that large-scale, real-world data showed that the difference in the required number of injections between agents tends to be small.^{27,28}

Imaging Protocol

After pupillary dilation with 0.5% tropicamide and 2.5% phenylephrine, patients underwent dilated slit lamp and indirect ophthalmoscopy and $20^\circ \times 15^\circ$ SD-OCT imaging (19 B-scans, ART 8) using a Spectralis HRA+OCT device (Heidelberg Engineering, Heidelberg, Germany).

Deep-Learning-Based SD-OCT Segmentation

The SD-OCT B-scans were segmented as previously described.²⁹ A convolutional neural network (CNN [Deeplabv3 model with a ResNet-50 backbone]) was trained and validated with a large data-set comprising of 9680 B-scans from 80 patients with late AMD (4840 B-scans from 40 patients [$30^\circ \times 25^\circ$ raster scan with 121 B-scans] with Macular neovascularization (MNV) and Geographic Atrophy (GA), respectively). The CNN was trained from scratch. A random subset of 20% of the patients served as validation data. The model was trained for 30 epochs (Adam optimization algorithm, initial learning rate of 0.001, learning rate decay [γ] of 0.1 every 10 epochs, sum of the Dice loss and cross entropy as loss function). The final model was selected on the basis of the optimal validation loss to avoid overfitting. The segmented layers included the inner retina, Outer Nuclear Layer (ONL), IS, OS, and retinal pigment epithelium-drusen complex (RPEDC) and choroid as in previous studies.²⁰ Importantly, Henle's fiber layer and hyporeflexive wedge-shaped bands were counted toward the ONL in this study. Subretinal fluid was counted toward the OS layer/compartments per definition, whereas subretinal hyperreflexive material, as well as type 1 and type 2 neovascular membranes, were counted toward the RPEDC (cf. Fig. 1).

Feature Extraction

En face thickness maps and en face mean-, minimum- and maximum-intensity projections (i.e., four maps) were generated for each retinal layer (six layers) using a custom software written in Python. Layer thickness maps constituted two-dimensional maps showing the axial layer thickness along each A-scan. The mean-, minimum-, and maximum-intensity projections represent the intensity values for a given layer along a given A-scan. In addition, the sub-RPE mean projection, sub-RPE restricted summed-area projection (RSAP),³⁰ and full retinal mean projection were added to the en-face stack (total of 27 maps [$4 \times 6 + 3$]). Subsequently, the average and standard

deviation of the thickness and of the intensity in the mean/minimum/maximum-intensity projections were extracted for the central and four inner Early Treatment Diabetic Retinopathy Study (ETDRS) subfields (i.e., 270 imaging features were extracted: $2 [\text{mean/SD}] \times 5 [\text{ETDRS subfields}] \times 27$, Fig. 1).

Machine-Learning

As the outermost loop, we generate 10 unique data subsets with only one visit per patient using random seeds. This step ensured that only one visit of one eye per patient was included in the model fitting. Embedded in this outermost loop, all models were fitted using nested cross-validation to obtain an unbiased estimate for the model accuracy (outer resampling) and tune hyperparameters (nested inner resampling) simultaneously. Specifically, we applied outer leave-one out cross-validation to assess the model accuracy. Nested within each outer training split, we applied nested inner 10-fold cross-validation to optimize the hyperparameters of the respective model.

Three conventional machine-learning algorithms, which are applicable in the setting of correlated predictors were probed: LASSO regression (using the R package *glmnet*),³¹ principal component regression (using the R package *pls*) and random forest regression (using the R package *randomForest*).³² For LASSO regression and principal component regression,³³ predictors were centered and scaled. These pre-processing transformations were estimated (both for the outer and inner resampling) only from the training data and then applied to the respective data set to avoid subtle information leakage. As part of the inner resampling, the following hyperparameters were tuned: lambda (100 values from 0.0001 to 1.5) for LASSO regression, number of components (1 to 20) for principal component regression and mtry (20, 40, 80, 160) for random forest regression. The optimal hyperparameter in the nested inner 10-fold cross validation was selected based on the mean absolute error (MAE) between observations and predictions.

Probabilistic Forecasting

For probabilistic forecasting, we implemented for the first time, to the best of our knowledge, NGBoost with a negative binomial distribution as the probability distribution. Specifically, in this method we assumed that the conditional distribution of the response variable (treatment frequency for 12 months), Y , conditioned on the predictor variables (Image Features), X , is a negative binomial distribution with param-

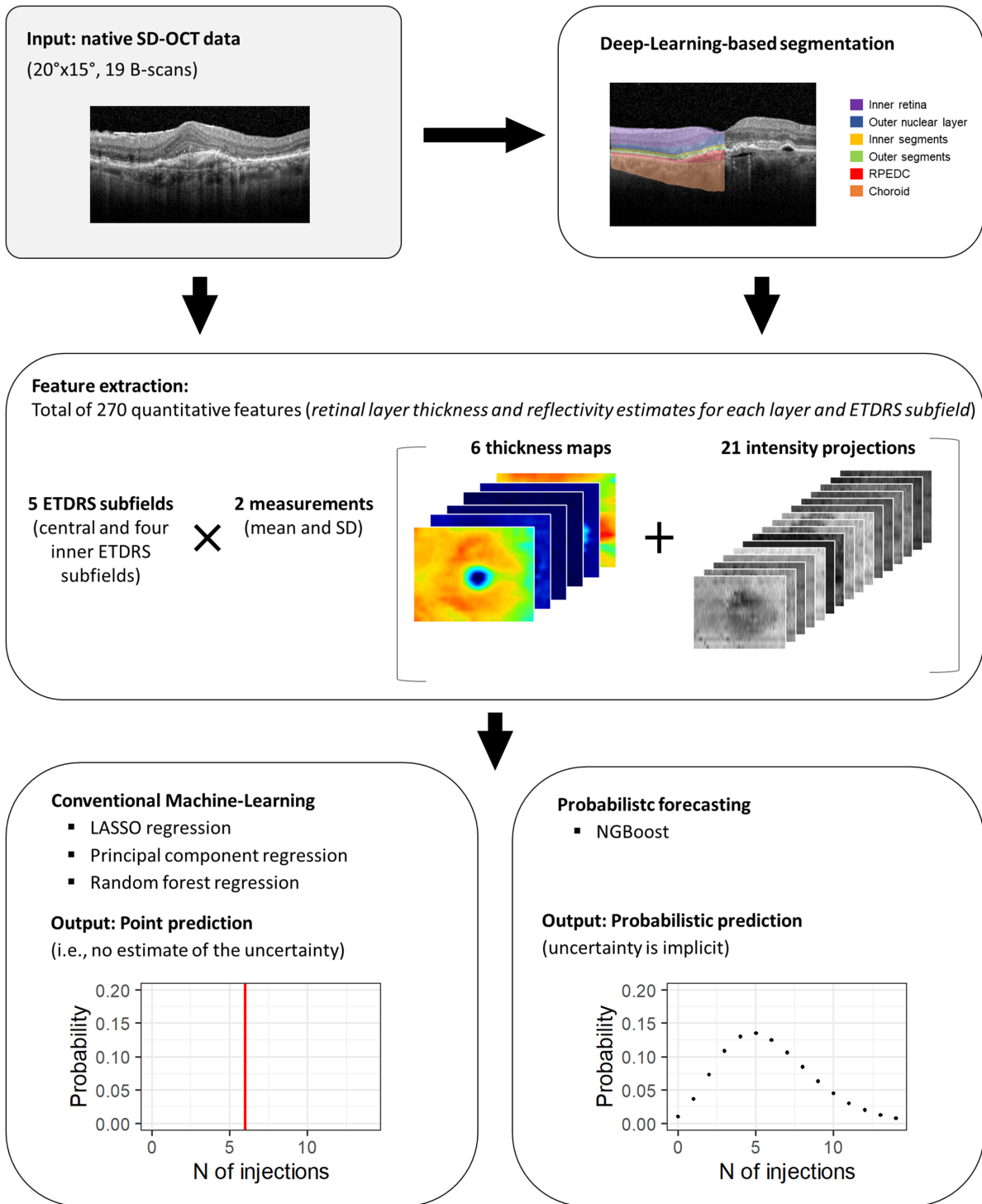


Figure 1. Image analysis pipeline. For this study, images were segmented using a previously validated, deep-learning-based pipeline.²⁹ Subsequently, the average (mean) and variability (standard deviation) of the layer thickness and layer reflectivity (minimum-, mean-, maximum-intensity projections) were extracted for each ETDRS subfield. These imaging biomarkers were then used to predict the future anti-VEGF treatment frequency, using conventional machine-learning approaches, as well as probabilistic forecasting.

eters $\mu(X)$ (mean of the distribution) and $n(X)$, both of which are functions of X . We assumed there is a true parent distribution of $Y|X$, so we defined a loss function, which is the Kullback Leibler divergence between our assumed Negative Binomial distribution and the true distribution. $\mu(X)$ and $n(X)$ were estimated by minimizing this loss function. Minimization of this loss function is equivalent to minimizing the score function (negative of log-likelihood). Gradient boosting has been used in this purpose where we used natural gradient instead of the traditional gradient.²³ We have used regression trees as the base learner. A small learning rate with a large number of weak learners (decision trees) was used. It takes about 10 minutes for model training. Once we estimated $\mu(X)$ and $n(X)$, we were able to estimate the distribution of $Y|X$. In this way for given set of image features, we can predict the probability distribution of injection frequency for some given X . The number of estimators in gradient boosting and depth of the regression tree were selected based on cross-validation. SHAP (SHapley Additive exPlanations) feature importance values were computed for the NGBoost model to provide an intuitive interpretation of the model.

Statistical Analyses

All estimates for the model accuracy were obtained from the outer resampling. Since predictions for multiple visits and for multiple seeds (from the outermost loop) were available, mixed effects models were applied to compute the MAE estimates and coefficient of determination (R^2) as performance metrics. For the NGBoost model, the MAE was computed by using the mean of the forecasted distribution as the prediction. Eye nested in patient and the iteration of the outermost loop were considered as random effect terms. Herein, the (marginal) R^2 describes the proportion of variance explained by the predictions alone (i.e., without patient-specific random factors). In addition, the cross-validated predictions were compared to the true frequency of injections using Bland-Altman plots. For the probabilistic forecasting, the percentage of true observations, which fell into the cross-validated 20%, 40%, 60%, and 80% interval predictions was analyzed to confirm the calibration of the interval predictions. The model specific indicators of feature importance were summarized across folds using boxplots. Specifically, we analyzed the absolute coefficient of for LASSO regression, the weighted sum of absolute coefficients for principal component regression, the permutation importance for random forest regression

and the Gini importance for NGBoost-based probabilistic forecasting.

Results

Cohort

A total of 148 visits of 99 eyes with neovascular AMD from 96 patients were included in this study. Specifically, 40 visits of 40 eyes of 37 patients from the first clinical site, and 108 visits of 59 eyes of 59 patients from the second clinical site were available. The one-year intervals for the repeated visits from the second clinical site did not overlap (i.e., baseline to month 12 and month 12 to month 24).

Prediction of One-Year Anti-VEGF Treatment Frequency Using Conventional Machine Learning

Three common machine-learning algorithms, which allow for conventional point estimation, were applied to predict future anti-VEGF treatment frequency (Fig. 2, Table). LASSO regression and principal component regression allowed for prediction of the anti-VEGF treatment frequency with an accuracy of (MAE [95% CI]) 2.76 injections/year [2.39–3.14] ($R^2 = 0.038$), and of 2.74 injections/year [2.38–3.11] ($R^2 = 0.173$), respectively. Random forest regression, which also allows for modeling of nonlinear relationships and interaction effects, provided a prediction accuracy of 2.60 injections/y [2.25–2.96] ($R^2 = 0.390$).

The probabilistic prediction of the anti-VEGF treatment frequency was similarly accurate. Specifically, NGBoost allowed for the prediction of the injection frequency with a MAE of 2.66 injections/year [2.31–3.01] ($R^2 = 0.094$). The probabilistic predictions were well calibrated in terms of the interval prediction. Specifically, the cross-validated 20%, 40%, 60%, and 80% interval predictions encompassed the true value in 22.8%, 38.6%, 52.4%, and 77.2%, respectively. This highlights the validity of the intervals.

Feature Importance

The importance of the imaging features was overall similar across all models (Fig. 3). For example, the standard deviation of the RPEDC thickness in the central ETDRS subfield thickness was the top ranked feature for LASSO regression and among the top five features across all models (with a median [rank] coefficient of 0.34 [no. 1] for LASSO regression, weighted sum of coefficients of 0.17 [no. 4] in principal

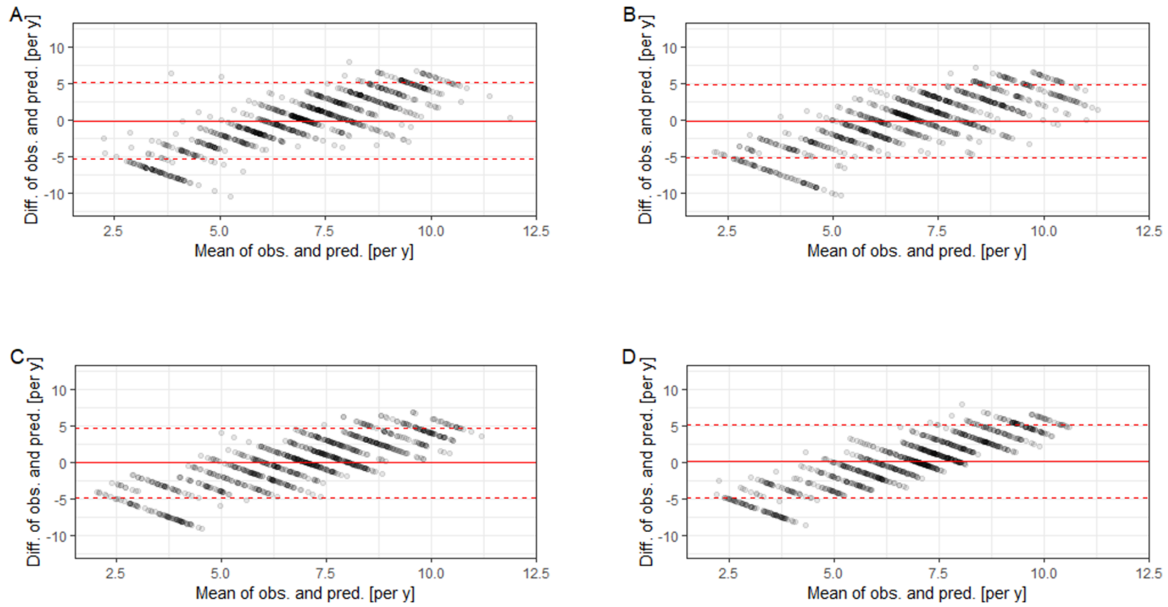


Figure 2. Prediction accuracy. The Bland-Altman plots show the difference between the observed and predicted frequency of anti-VEGF injections based on LASSO regression (A), principal component regression (B), random forest regression (C) as well as for NGBoost (D). All points were plotted semitransparent to avoid overplotting. The dashed lines indicate the 95% limits of agreement and the solid line the mean difference (both calculated considering the hierarchical nature of the data [eye nested in patient]). Notably, all model tended to slightly overestimate the required injections frequency in patients with few injections and underestimate the required injection frequency for patients with a high number of injections.

Table. Model Performance (Test Performance, I.E., Outer-Loop of the Nested Cross Validation)

Model	Mean Absolute Error (Injections/Year) [95% CI]	Cross-Validated R ²	ROC-AUC for Low Treatment Requirement*	ROC-AUC for High Treatment Requirement †
LASSO regression	2.76 [2.39–3.14]	0.038	0.61	0.63
Principal component regression	2.74 [2.38–3.11]	0.173	0.63	0.7
Random forest regression	2.60 [2.25–2.96]	0.390	0.68	0.7
NGBoost	2.66 [2.31–3.01]	0.094	0.68	0.69

*Receiver operating characteristic area under the curve (ROC-AUC) was calculated for the identification of eyes with an injection requirement of 4 injections/y or less.²²

†ROC-AUC for 10 injections/y or more.²²

component regression, a permutation importance of 1.36 % Inc. MSE [no. 3] for random forest regression and a Gini importance of 1.53 [no. 4] for NGBoost-based probabilistic forecasting). The second most important feature for LASSO regression constituted the standard deviation of the IS thickness in the nasal inner ETDRS subfield (coefficient of 0.19 [no. 2] for LASSO regression, weighted sum of coefficients of 0.15 [no. 4] in principal component regression, a permutation importance of 1.40 %Inc. MSE [no. 2] for random forest regression, and a Gini importance of 5.08 [no. 2] for NGBoost-based probabilistic forecasting).

Figure 4 shows probabilistic forecasting for two representative patients. For these patients, the forecasted distribution is very much coherent with the clinical imaging characteristics.

Discussion

In this study, we applied probabilistic forecasting to predict future anti-VEGF treatment frequency in patients with neovascular AMD. This work highlights the potential of ML- and DL-based algorithms to

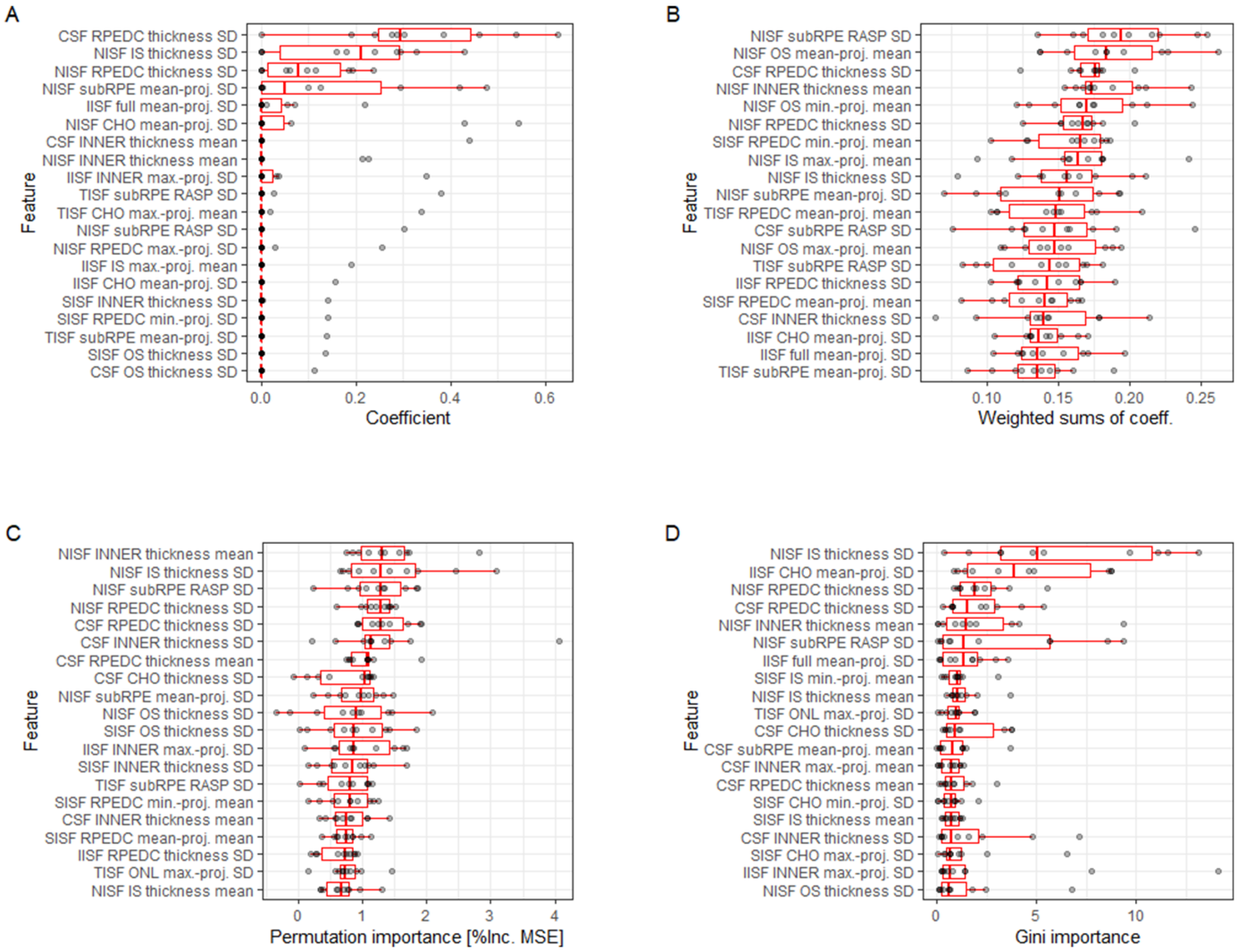


Figure 3. Feature importance. The panels show the feature importance for the prediction of the anti-VEGF injection frequency for 12 months for LASSO regression (A, unit: coefficient), principal component regression (B, unit: weighted sum of the absolute coefficients), random forest regression (C, unit: percentage of increase in mean squared error [%Inc MSE]), and NGBoost (D, unit: Gini importance). The results from the 10 outermost repeats for the analysis (random seeds) are shown as dots. The boxplots summarize the results. Notably, LASSO regression results in variable selection. Thus the coefficient is sometimes zero.

inform clinical practice, facilitate patient scheduling, and identify patients who may benefit from long-acting treatment modalities.

Our approach is comparable to the performance of previous work (Table),²² despite our use of real-world clinical data versus clinical trial study data. The results of this study constitute a step towards image-guided predictions of treatment frequency that may significantly enhance the definition of treatment intervals in the management of neovascular AMD. Specifically, the currently used PRN and T&E protocols have disadvantages. PRN is somewhat unfavorable because of the high frequency of visits, whereas visit scheduling for T&E treatment is only possible one visit at a time.

Forecasting of treatment requirements for the next 12 months could have significant impact on compliance and treatment effectiveness because it may allow the implementation of “augmented T&E” protocols—for example, for patients with predicted high treatment requirements, the interval extension could be limited to one instead of two weeks. However, clinical implementation of such a new protocol would obviously require prospective comparisons to standard-of-care T&E treatments.

In addition to direct implementation into clinical practice, probabilistic forecasting of future anti-VEGF treatment frequency would be helpful for patient counseling, scheduling of visits, and for informed

**Cross-validated interval prediction
for the upcoming 12 months**

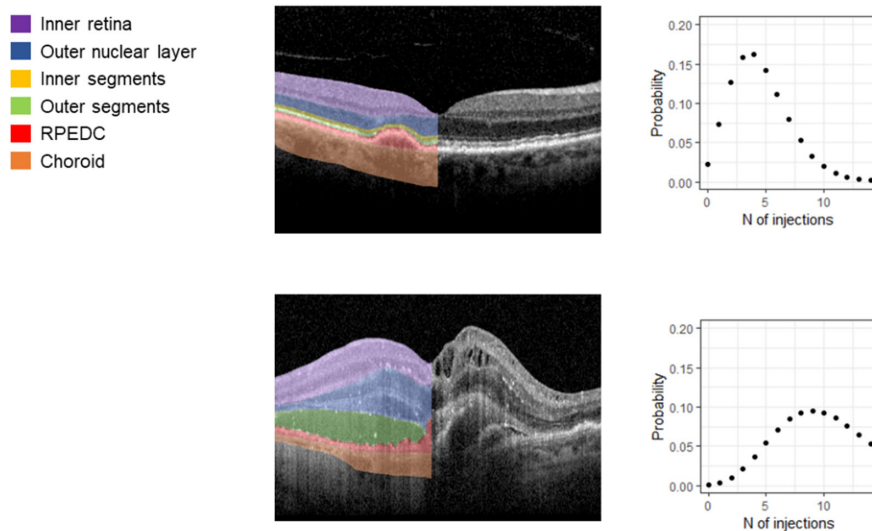


Figure 4. Exemplary patients. The figure shows the central spectral-domain optical coherence tomography B-scan of two patients and the probabilistic forecast for the upcoming 12 months. The upper patient shows a type 1 choroidal neovascularization with no intra-retinal fluid and only subtle subretinal fluid (in neighboring B-scans). The predictive model predicts three to four injections/year for this eye (true number of required injections = 2). In contrast, the model predicts seven to eight injections/year for the eye of the lower patient, which is characterized by marked intraretinal and subretinal and a type 2 neovascular membrane (true number of required injections = 10).

decision making regarding upcoming long-term treatment modalities. Applications of interval predictions to highlight model uncertainty (instead of point predictions) may be helpful in a clinical setting, given the manifold factors that may lead to invalid point estimates. Such factors include out-of-distribution samples (e.g., MNV subtypes like polypoidal choroidal vasculopathy) and poor image quality. A plethora of innovative therapeutic approaches have been introduced or are in the final stages of clinical development. This includes long-acting anti-VEGF inhibitors^{11,34,35} or administration of therapeutic anti-VEGF agents via a port delivery system,¹² as well as gene therapy.³⁶ Multiple studies have shown that only a small group of patients require monthly treatment.^{37,38} Thus clear-cut criteria, or ideally a fully automated pipeline, as presented here, to identify eligible patients would be valuable. Other than risk/benefit ratio considerations, probabilistic forecasting may help to estimate the cost-effectiveness of the aforementioned therapeutics on the level of an individual patient. These data may be beneficial for health coverage approvals, if necessary. Furthermore, anti-VEGF therapy in routine clinical practice tends to result in undertreatment and suboptimal visual acuity outcomes compared to clinical trials as highlighted by the multi-country AURA

and LUMIERE studies.^{7,39} Although several factors may contribute to the tendency to undertreat, a realistic forecast for the next 12 months may help ensure treatment compliance with required office visits.¹⁰

In terms of feature importance, standard deviation of the central ETDRS subfield thickness of the RPEDC constituted the most informative predictor across models. Specifically, the SHAP plot revealed that higher values for the standard deviation of the central ETDRS subfield thickness of the RPEDC were associated with a higher need of injections (Supplementary Figure S1). From a biological perspective, this is plausible because the irregularity of the RPEDC and the thickness and volume of MNV complex in this area are likely reflective of the severity of the underlying neovascular membrane and the need for more intensive anti-VEGF dosing.

This study is limited by the overall sample size. Moreover, (out-of-domain) disease phenotypes, which were not included in the training data, may lead to inadequate predictions. This includes, for example, aneurysmal type 1 neovascularization (polypoidal choroidal vasculopathy) as commonly found in Asian patients.^{40,41} Future work will be needed to validate that the uncertainty estimates for NGBoost truly highlight out-of-distribution samples, especially in

conjunction with external data sets. Based on the currently included features and sample size, it is evident that all models tend to overestimate the number of injections for patients with low treatment requirements and underestimate the number of injections for patients with high treatment requirements. This proportional bias indicates that a larger training set and the addition of further informative features may improve model performance. Pooling of data from multiple physicians from two clinical sites and the inclusion of patients treated with PRN, as well as T&E, most likely increased the variability of our target variable (annual injection frequency). Hence, the reported prediction accuracies must be considered as conservative estimates that would likely improve if the training data are not only increased, but also tailored to a specific treatment regimen. Of note, some retinal features in this study were indirectly encoded. For example, intraretinal fluid would be indicated by a combination of ONL thickening in conjunction with reduced ONL reflectivity. For clinical applications, direct segmentation of features such as intraretinal and subretinal fluid or hyperreflective foci would be preferable to further enhance the interpretability of the relationship between model inputs and predictions.^{16,42} Furthermore, refinement (post-processing) of the segmentation results from the CNN-based segmentation could possibly enhance the prediction accuracies of the ML models. For future work, the predictive value of multimodal imaging data, for example optical coherence tomography angiography-based biomarkers for neovascular activity, needs to be investigated. In addition, electronic health record (EHR)-based predictors including the absence or number of prior injections would likely improve the prediction accuracy. Moreover, DL-based prediction models may provide a better prediction accuracy, but at the cost of interpretability. Last, prospective clinical validation of models is needed to substantiate possible benefits in clinical practice.

In summary, both canonical ML-models and a novel probabilistic prediction model (NGBoost with a negative binomial probability distribution) allow forecasting of future anti-VEGF treatment frequency in neovascular AMD with moderate accuracy. Specifically, probability forecasting is clinically advantageous, given that the probability interval can help physicians consider the model output in a thoughtful manner or use standard treatment as a fallback alternative. These models may help physicians select patients for whom long-term treatment options beyond conventional anti-VEGF injections may provide a favorable risk/benefit ratio.

Acknowledgments

Supported by the research award 2020 of the Association of Rhine-Westphalian Ophthalmologists (RWA) to MP, the German Research Foundation (DFG) grant PF950/1-1 to MP; a BrightFocus Foundation grant M2019155 to J.A.H., an Unrestricted Grant for Research to Prevent Blindness and a P30 Core Grant for Vision Research (2P30EY001792), Department of Ophthalmology and Visual Sciences, University of Illinois at Chicago, Chicago, IL (J.A.H., J.I.L.); and in part by Novartis Pharma GmbH, Germany (M.F.) and Unrestricted Grant from Research to Prevent Blindness, New York, NY, to the Department of Ophthalmology & Visual Sciences, University of Utah; and in part by an Unrestricted Grant from Research to Prevent Blindness, New York, NY, to the Department of Ophthalmology, Byers Eye Institute at Stanford, Stanford University School of Medicine and NIH grant P30-EY026877.

Disclosure: **M. Pfau**, Heidelberg Engineering (F), Optos (F), Carl Zeiss Meditec (F), CenterVue (F); **S. Sahu**, None; **R.A. Rupnow**, None; **K. Romond**, None; **D. Millet**, Heidelberg Engineering (F), Optos (F), Carl Zeiss Meditec (F), CenterVue (F); **F.G. Holz**, Acucela (R), Allergan (R), Apellis (R), Bayer (R), Bioeq/Formycon (R), CenterVue (R), Ellex (R), Roche/Genentech (R), Geuder (R), Kanghong (R), NightStarx (R), Novartis (R), Optos (R), Zeiss (R), Acucela (C), Allergan (C), Apellis (C), Bayer (C), Boehringer-Ingelheim (C), Roche/Genentech (C), Geuder (C), Grayburg Vision (C), LinBioscience (C), Kanghong (C), Novartis (C), Pixium Vision (C), Oxurion (C), Stealth BioTherapeutics (C), Zeiss (C); **S. Schmitz-Valckenberg**, Acucela/ Kubota Vision (F), Apellis (C), Novartis (C, F, R), Allergan (C, F, R), Bayer (F, R), Bioeq/Formycon (F, C), Carl Zeiss Meditec (F, R), CenterVue (F), Galimedix (C), Roche (F, R), Heidelberg Engineering (F), Katairo (F), Optos (F), Oxurion (C), Roche/Genentech (F,C); **M. Fleckenstein**, Heidelberg Engineering (F, C), Optos (F), Carl Zeiss Meditec (F), Genentech/Roche (F, C), CenterVue (F), Novartis (F, C) US20140303013 (P); **J.I. Lim**, Genentech (F), Regeneron (F), Allergan (R), Alcon (R), Opthea (C), Santen (C), Quark (C), Aura Biosciences (C), Iveric Bio (R), Cognition (C), Aldeyra (F), Novartis (C), Graybug (F), Stealth (F), Chengdu (F); **L. de Sisternes**, Carl Zeiss Meditec (E); **T. Leng**, Kodiak (F), 3T (F), Genentech (C), Regeneron (C), Verana (C); **D.L. Rubin**, None; **J.A. Hallak**, BrightFocus Foundation Grant (R)

References

- Cheung LK, Eaton A. Age-related macular degeneration. *Pharmacotherapy*. 2013;33:838–855.
- Holz FG, Schmitz-Valckenberg S, Fleckenstein M. Recent developments in the treatment of age-related macular degeneration. *J Clin Invest*. 2014;124:1430–1438.
- Rasmussen A, Bloch SB, Fuchs J, et al. A 4-year longitudinal study of 555 patients treated with ranibizumab for neovascular age-related macular degeneration. *Ophthalmology*. 2013;120:2630–2636.
- Bakri SJ, Thorne JE, Ho AC, et al. Safety and efficacy of anti-vascular endothelial growth factor therapies for neovascular age-related macular degeneration: a report by the American Academy of Ophthalmology. *Ophthalmology*. 2019;126:55–63.
- Munk MR, Kiss C, Huf W, et al. One year follow-up of functional recovery in neovascular AMD during monthly anti-VEGF treatment. *Am J Ophthalmol*. 2013;156:633–643.
- Rofagha S, Bhisitkul RB, Boyer DS, et al. Seven-year outcomes in ranibizumab-treated patients in ANCHOR, MARINA, and HORIZON: a multicenter cohort study (SEVEN-UP). *Ophthalmology*. 2013;120:2292–2299.
- Holz FG, Tadayoni R, Beatty S, et al. Multi-country real-life experience of anti-vascular endothelial growth factor therapy for wet age-related macular degeneration. *Br J Ophthalmol*. 2015;99:220–226.
- Gillies MC, Campain A, Barthelmes D, et al. Long-term outcomes of treatment of neovascular age-related macular degeneration: data from an observational study. *Ophthalmology*. 2015;122:1837–1845.
- Liakopoulos S, Spital G, Brinkmann CK, et al. ORCA study: real-world versus reading centre assessment of disease activity of neovascular age-related macular degeneration (nAMD). *Br J Ophthalmol*. 2020;104:1573–1578.
- Wintergerst MWM, Bouws J, Loss J, et al. Reasons for delayed and discontinued therapy in age-related macular degeneration. *Ophthalmologie*. 2018;115:1035–1041.
- Dugel PU, Koh A, Ogura Y, et al. HAWK and HARRIER: Phase 3, multicenter, randomized, double-masked trials of brolucizumab for neovascular age-related macular degeneration. *Ophthalmology*. 2020;127:72–84.
- Campochiaro PA, Marcus DM, Awh CC, et al. The port delivery system with ranibizumab for neovascular age-related macular degeneration: results from the randomized phase 2 Ladder Clinical Trial. *Ophthalmology*. 2019;126:1141–1154.
- Rosenfeld PJ, Browning DJ. Is This a 737 Max Moment for Brolucizumab? *Am J Ophthalmol*. 2020;216:A7–A8.
- Keremany DS, Goldbaum M, Cai W, et al. Identifying medical diagnoses and treatable diseases by image-based deep learning. *Cell*. 2018;172:1122–1131.e9.
- De Fauw J, Ledsam JR, Romera-Paredes B, et al. Clinically applicable deep learning for diagnosis and referral in retinal disease. *Nat Med*. 2018;24:1342–1350.
- Yim J, Chopra R, Spitz T, et al. Predicting conversion to wet age-related macular degeneration using deep learning. *Nat Med*. 2020;26:892–899.
- Hallak JA, De Sisternes L, Osborne A, et al. Imaging, genetic, and demographic factors associated with conversion to neovascular age-related macular degeneration: secondary analysis of a randomized clinical trial. *JAMA Ophthalmol*. 2019;137:738–744.
- Rohm M, Tresp V, Müller M, et al. Predicting visual acuity by using machine learning in patients treated for neovascular age-related macular degeneration. *Ophthalmology*. 2018;125:1028–1036.
- Pfau M, von der Emde L, Dysli C, et al. Determinants of cone- and rod-function in geographic atrophy: AI-based structure-function correlation. *Am J Ophthalmol*. 2020;217:162–173.
- von der Emde L, Pfau M, Dysli C, et al. Artificial intelligence for morphology-based function prediction in neovascular age-related macular degeneration. *Sci Rep*. 2019;9:11132.
- Künzel SH, Möller PT, Lindner M, et al. Determinants of quality of life in geographic atrophy secondary to age-related macular degeneration. *Invest Ophthalmol Vis Sci*. 2020;61:63.
- 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.
- Duan T, Avati A, Ding DY, et al. NGBoost: natural gradient boosting for probabilistic prediction. arXiv Prepr arXiv 191003225. 2019.
- Marz A. *XGBoostLSS – An extension of XGBoost to probabilistic forecasting*. arXiv preprint arXiv:1907.03178. 2019.

25. Fleckenstein M, Grassmann F, Lindner M, et al. Distinct genetic risk profile of the rapidly progressing diffuse-trickling subtype of geographic atrophy in age-related macular degeneration (AMD). *Investig Ophthalmol Vis Sci*. 2016;57:2463–2471.
26. Lalwani GA, Rosenfeld PJ, Fung AE, et al. A variable-dosing regimen with intravitreal ranibizumab for neovascular age-related macular degeneration: year 2 of the PrONTO Study. *Am J Ophthalmol*. 2009;148:43–58.e1.
27. Ferreira A, Sagkriotis A, Olson M, et al. Treatment frequency and dosing interval of ranibizumab and aflibercept for neovascular age-related macular degeneration in routine clinical practice in the USA. *PLoS One*. 2015;10:e0133968.
28. Lee AY, Lee CS, Egan CA, et al. UK AMD/DR EMR REPORT IX: comparative effectiveness of predominantly as needed (PRN) ranibizumab versus continuous aflibercept in UK clinical practice. *Br J Ophthalmol*. 2017;101:1683–1688.
29. Pfau M, von der Emde L, de Sisternes L, et al. Progression of Photoreceptor Degeneration in Geographic Atrophy Secondary to Age-related Macular Degeneration. *JAMA Ophthalmol*. 2020;138:1026–1034.
30. Chen Q, Niu S, Shen H, et al. Restricted Summed-Area Projection for Geographic Atrophy Visualization in SD-OCT Images. *Transl Vis Sci Technol*. 2015;4:2.
31. Friedman J, Hastie T, Tibshirani R. Regularization paths for generalized linear models via coordinate descent. *J Stat Softw*. 2010;33:1–22.
32. Breiman L. Random forests. *Mach Learn*. 2001;45:5–32.
33. Mevik BH, Wehrens R, Liland KH. {pls}: Partial Least Squares and Principal Component regression. R package version. 2011;2(3).
34. Khanani AM, Patel SS, Ferrone PJ, et al. Efficacy of every four monthly and quarterly dosing of faricimab vs ranibizumab in neovascular age-related macular degeneration: the STAIRWAY phase 2 randomized clinical trial. *JAMA Ophthalmol*. 2020;138:964–972.
35. Puliafito CA, Wykoff CC. New frontiers in retina: highlights of the 2020 angiogenesis, exudation and degeneration symposium. *Int J Retin Vitreol*. 2020;6:18.
36. Rakoczy EP, Magno AL, Lai CM, et al. Three-year follow-up of phase 1 and 2a rAAV.sFLT-1 Subretinal Gene Therapy Trials for Exudative Age-Related Macular Degeneration. *Am J Ophthalmol*. 2019;204:113–123.
37. Ying G, 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;122:2523–31.e1.
38. 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*. 2015;19:1–298.
39. Cohen SY, Mimoun G, Oubraham H, et al. Changes in visual acuity in patients with wet age-related macular degeneration treated with intravitreal ranibizumab in daily clinical practice: the LUMIERE study. *Retina*. 2013;33:474–481.
40. Dansingani KK, Gal-Or O, Sadda SR, et al. Understanding aneurysmal type 1 neovascularization (polypoidal choroidal vasculopathy): a lesson in the taxonomy of ‘expanded spectra’—a review. *Clin Exp Ophthalmol*. 2018;46:189–200.
41. Kokame GT, Liu K, Kokame KA, et al. Clinical characteristics of polypoidal choroidal vasculopathy and anti-vascular endothelial growth factor treatment response in Caucasians. *Int J Ophthalmol Zeitschrift fur Augenheilkd*. 2020;243:178–186.
42. Keenan TDL, Chakravarthy U, Loewenstein A, et al. Automated quantitative assessment of retinal fluid volumes as important biomarkers in neovascular age-related macular degeneration. *Am J Ophthalmol*. 2021;224:267–281.