Predicting visual acuity by using machine learning in patients treated for neovascular age-related macular degeneration

Markus Rohm[1, 2], Volker Tresp[2], Michael Müller[1], Christoph Kern[1], Ilja Manakov[1,2], Maximilian Weiß[1], Dawn A. Sim[3], Siegfried Priglinger[1], Pearse A. Keane[3], Karsten Kortuem[1, 3]

Abbreviations:

AMD	Age-related macular degeneration
CRT	Central retinal thickness
EMR	Electronic medical records
ETL	Extract, transform, load
DW	Data warehouse
GEID	Generic eye identifier
ICD-10	International classification codes of diseases version 10
MAE	Mean absolute error
nAMD	neovascular age-related macular degeneration
OCT	Optical coherence tomography
PRN	Pro re nata treatment scheme
RMSE	Root mean square error
VA	Visual acuity
XML	Extensible Markup Language

No conflicting relationship exists for any author

Presented at the American Academy of Ophthalmology Annual Meeting, 2016, Chicago

Financial Support: None

Running head: Predicting visual acuity in wet AMD patients by using machine-learning

Corresponding author:
Karsten Kortuem, FEBO, MHBA
University Eye Hospital Munich
Mathildenstr. 8
80336 Munich, Germany
karsten.kortuem@med.uni-muenchen.de

[1]Department of Ophthalmology, Ludwig-Maximilians-University Munich, Germany
[2]Department of Computer Science, Ludwig-Maximilians-University Munich, Germany
[3]Moorfields Eye Hospital, London, United Kingdom

Abstract

Objective:
To predict, by using machine learning, visual acuity (VA) at 3 and 12 months in patients with neovascular age-related macular degeneration (nAMD) after initial upload of 3 Anti-VEGF (vascular endothelial growth factor) injections.

Design:
Database study

Subjects:
For the 3-month VA forecast, 653 patients (379 females) with 738 eyes and an average age of 74.1 years were included. The baseline VA before the first injection was 0.54 logMAR (+/- 0.39). 456 of these patients (270 females, 508 eyes, average age: 74.2 years) had sufficient follow-up data to be included for a 12-month VA prediction. The baseline VA before the first injection was 0.56 logMAR (+/- 0.42).

Methods:
Five different machine learning algorithms (AdaBoost.R2, Gradient Boosting, Random Forests, Extremely Randomized Trees, and Lasso) were used to predict VA in patients with nAMD after treatment with 3 Anti-VEGF injections. Clinical data features came from a data warehouse (DW) containing electronic medical records (41 features, e.g. VA) and measurements features from optical coherence tomography (OCT) (124 features, e.g. central retinal thickness). VA of patient eyes excluded from machine learning was predicted and compared with the ground truth, namely the actual VA of these patients as recorded in the DW.

Main Outcome Measure:
Difference in logMAR VA after 3 and 12 months after upload phase between prediction and ground truth as defined above.

Results:
For the 3-month VA forecast, the difference between the prediction and ground truth was between 0.11 logMAR (5.5 letters) mean absolute error (MAE)/0.14 logMAR (7 letters) root mean square error (RMSE) and 0.18 logMAR (9 letters) MAE/0.2 logMAR (10 letters) RMSE. For the 12-month VA forecast, the difference between the prediction and ground truth was between 0.16 logMAR (8 letters) MAE/ 0.2 logMAR (10 letters) RMSE and 0.22 logMAR (11 letters) MAE/0.26 logMAR (13 letters) RMSE.
The best performing algorithm with regard to forecasts was the Lasso protocol.

Conclusions:
Machine-learning allowed VA to be predicted for three months after three anti-VEGF injections, with a comparable result to VA measurement reliability. For a forecast after 12 months of therapy, VA prediction may help to encourage patients adhering to intravitreal therapy.

Introduction

In computer science, machine learning is a term that covers diverse approaches to artificial intelligence in computers and is nowadays quite a common phenomenon. The original aim, which was stated as early as 1959, was for computers to have the ability to learn without being explicitly programmed [1]. Many methods were tried over the years, but it took until the 1990s when machines started to beat humans. Such stories were then broadcast in the mass media and gained a great deal of publicity. In 1996 and 1997, Garry Kasparov lost a chess match against the IBM super-computer "Deep Blue"; this was the first time that a long-standing world champion (1985-2000) was defeated by a computer. Another well-publicized milestone was reached when IBM Watson beat two expert human players in the television show "Jeopardy" in 2011 [2].

In the medical sciences involving vision, several machine learning techniques have been applied in various sub-specialties [3]. Mostly, glaucoma and retinal imaging-related problems have been addressed. As the number of patients with diabetic retinopathy is continuing to rise [4], efforts have recently been made to detect early forms of diabetic retinopathy by means of machine learning [5-7]. These programs have been shown to provide a highly sensitive tool for detecting such disease complications [8]. First approaches have been published for the automatic imaging analysis of optical coherence tomography (OCT) by using machine learning [9, 10]. Additional studies are underway to automate the detection and classification of pathological features in eye imaging including fundus photography and OCT[11].

Age-related macular degeneration (AMD) is still one of the leading causes of legally defined blindness in industrial countries [12]. However, the use of anti-VEGF agents such as bevacizumab, ranibizumab, or aflibercept in AMD and other retinal diseases such as diabetic macular edema or venous occlusions have allowed, for the first time, the sustainable improvement, stabilization, or slow-down of disease progression [13-20]. Nevertheless, even after several years of experience, real-life results in patients vary greatly from results obtained in clinical trials [21, 22].

To improve insights into our own clinical results and to facilitate clinical research, a data warehouse (DW) has been set up in our institution [23]. It comprises data, such as diagnoses, medication, and undergone surgery, from electronic medical records (EMR), as developed in our institution based on the hospital-wide EMR system "SAP i.s.h.med", from

over 320,000 patients. It also contains measurement data from OCT (e.g., central retinal thickness, macular volume) and other measurement devices such as IOLMaster and Pentacam.

For this study, we have used five modern machine learning algorithms to predict the outcome of visual acuity (VA) after one year of anti-VEGF treatment in newly diagnosed neovascular AMD patients. The base for the prediction was our DW with its clinical and measurement data. The forecasting of VA might help to reduce the psychological pressure on patients, as many fear sight loss as a consequence of the diagnosis of wet AMD [24]. In particular, before the first injections, patients are extremely anxious and nervous, as they do not know what the outcome of the therapy will be and whether it is worth pursuing[25].

Material and Methods

Data warehouse as repository

To estimate VA after an initial upload of 3 anti-VEGF injections in patients with neovascular AMD after 3 and 12 months of intravitreal therapy, we applied machine learning algorithms on real-life data from our DW. This is updated every night with new and altered information from the EMR (i.s.h.med, Cerner AG, Erlangen, Germany). The nightly transfer from EMR includes diagnoses, clinical data such as VA and intraocular pressure, intravitreal injections, medications, appointments, and surgical operations. Up until December 2016, it incorporated data from 330,801 patients including 44,134 recorded intravitreal injections and 402,001 VA measurements. Our Ethics Committee ruled that approval was not required for this study. This study adheres to the tenets of the Declaration of Helsinki.

In addition to clinical data, 75,750 Spectralis OCT (Heidelberg Engineering, Heidelberg, Germany) measurements, but no images, were extracted from the software “Heidelberg Eye Explorer”. It provides an XML (Extensible Markup Language) export, which is a structured file format of measurement values, as in the company’s software. This export contains data like central retinal thickness (CRT), retinal volume and measurements in nine zones grid over the macular region (Figure 1). The XML files have been incorporated into the DW with a custom-made extract, transform, load (ETL) script, which is also programmed in Java. The DW itself was built on a “Microsoft SQL” database server running in a database cluster in the hospital’s datacenter. The PC used for the machine learning algorithms was a quad core Intel i5 (3.30 GHz) PC with 32 gigabytes of RAM, running on Ubuntu 16.04 LTS and Python (3.6.0). In

Python, the following libraries were used: scipy (Version: 0.18.1), numpy (Version: 1.12.0), pandas (Version: 0.19.2), jupyter (Version: 1.0.0), scikit-learn (Version: 0.18.1), and Theano (Version 0.8.2).

Data Pre-processing

For our test, we identified all patient eyes in the DW that, at any stage, received the international classification codes of diseases version 10 (ICD-10) "H35.3", which stands for age-related macular degeneration, and which received at least 3 injections of anti-VEGF medication (aflibercept, bevacizumab or ranibizumab) as noted in the surgical report. The standard operating procedure for treating 153 nAMD in our institution requires monthly injections for the first three months. After the initial upload of three injections, patients were treated according to the most recent treatment recommendations of the German Ophthalmological Society (Deutsche Ophthalmologische Gesellschaft – DOG)[26, 27]. Accordingly, a pro re nata scheme (PRN) was recommended until 2015. From 2015 onwards, a treat and extend scheme was recommended alongside the PRN regime. This led to 456 patients with 508 eyes for the long-term prediction goal (365 days) and 653 patients with 738 eyes for the short-term prediction goal (90 days); these were the patients that remained after the removal of all patients with invalid entries. The patients used for the long-term prediction represent a subgroup of the short-term prediction group.

Before algorithms can be trained on these patients, data need to be pre-processed and joined into a single dataset. The data came from the EMR and the OCT measurements. Most of the machine learning algorithms expect a dataset with no missing values; this is difficult to achieve in real-life datasets such as ours. Therefore, we centered all the other values around VA measurements. Consequently, we associated VA values with the most recently available measurements from previous (but not future) records in order to rule out potential data mix-ups shortly before and after an intravitreal medication application.

In the data collection process, we treated every eye as a separate patient and associated it with a generic eye identifier (GEID). The newly created dataset included not only the most recent measurement values (e.g., VA and CRT), but also aggregated versions with different timeframes (3 months and 12 months) and different aggregation functions (mean, variance, minimum and maximum). These synthetic features were designed in order to capture trends in the temporal progression of the disease. After all preparations, there were 165 features, which were used for training. 41 features were from the EMR or calculated based on EMR

data, like days since first injection. 124 features are based on OCT measurement data. Table 1 (available online only) gives an overview of all features used for training.

Individual patient eye histories were now built from the dataset created to date. These histories consisted of a specified number of previous temporal consecutive visual events up to a certain time point. It was essentially a condensed version of the past visits of the patient until the third injection was given, by using the resulting dataset, as traditional machine learning algorithms cannot use temporal information directly.

Algorithms used for Visual Acuity Prediction

To predict the VA outcome of nAMD patients, we applied five different machine learning algorithms, all of which are freely available online and form part of the previously mentioned libraries. In detail, they were:

- AdaBoost.R2 [28]: This belongs to the so-called boosting methods, which try to reduce the bias of combined estimators by specially built base models (often weak learners). The basic idea of the popular ensemble algorithm AdaBoost is to train an initial weak learner (in our case a simple decision tree) on the original dataset. Afterwards, multiple copies (49 in our case) are trained on the same dataset, but with more focus on examples where the performance was previously poor (i.e., the ones that were difficult to learn). The final prediction output is created by a majority vote of all weak learners.
- Gradient Boosting [29]: This approach, being the second boosting method, is a generalization of boosting to arbitrary differentiable loss functions. At every optimization step, a regression tree is fit to the negative gradient of the loss function. These models are very popular, powerful, and robust against outliers.
- Random Forests [30]: This belongs to the so-called averaging methods, where the basic idea is to build multiple independent models and to average their predictions to retrieve the final prediction. Under the assumption that the errors made by the individual models are also independent, the averaging of their predictions will likely reduce the error. In effect, Random Forests are an ensemble of simple decision trees that are built from a bootstrap sample (i.e., a sample drawn with replacement). In addition, each of the (100) decision trees is trained on a random subset of all available features. Despite this being a very simple model, it has been shown to be extremely powerful in many practical applications.

- Extremely Randomized Trees [31]: These are special versions of Random Forests, but instead of choosing the thresholds for the randomly picked features by how discriminative they are, these thresholds are also selected at random. This may enhance the variance of the ensemble at the price of higher bias.
- Lasso [32]: A linear model trained with L1 prior as a regularizer and coordinate descent for training. It estimates sparse model coefficients (in contrast to L2 prior), i.e., it leads to compact models, where zero is assigned to most of the low gain features. Therefore, it is often used for feature selection.

Evaluation of Models

To evaluate the performance of our predictive models, we used 10-fold cross validation. We randomly split the available GEIDs ($\mathbb{D}$) into 10 approximately equally sized, mutually exclusive subsets ($\mathbb{D}_1 0, \mathbb{D}_1 1, \ldots, \mathbb{D}_1 9$). In every iteration i of the 10 total iterations, we used all subsets without the i-th ($\mathbb{D} \setminus \mathbb{D}_1 i$) for training (or validation) and the i-th for testing. Therefore, the tests of the model were performed on patient eyes that the models did not previously see. For evaluation, either the data (e.g. VA) of the last one, two, three or four visits prior to the third injection were available. Resulting accuracy was stratified by the number of previous visits. The computing time for the training and testing of the models was about 8 hours. To make the test results more comparable, we used fixed but initially randomly generated seeds for shuffling the GEIDs before splitting. To provide a quality measure for our predictions per algorithm, which shows how close they are to the ground truth, we indicate the root of the mean squared error (RMSE= $\sqrt{\frac{1}{N}\sum_{i=1}^{N}(\tilde{y}_i - y_i)^2}$ for N=number of predictions per fold) and the mean absolute error (MAE= $\frac{1}{N}\sum_{i=1}^{N}|\tilde{y}_i - y_i|$) for 90 and 365 days' prediction as compared with the ground truth, which is the actual patient VA as measured in clinics. When describing accuracy of machine learning algorithms, usually both values are indicated. RMSE does penalize outliers more, thereby allowing you to choose the more robust algorithm. This is especially helpful, if the MAE is comparable between tested algorithms. MAEs are usually used in medical publications only.

Results

The demographics of all patients/eyes included in the study are summarized in Table 2. For the long-term prediction, out of 456 patients, 270 were female (59.2%). For the shorter prediction period, out of 653 patients, 379 were female (58%). The average of patients was slightly above 74 years in both groups. The VA at first injection was 0.54 (+/- 0.39) logMAR for the long-term prediction cohort and 0.56 (+/- 0.42) logMAR for the short-term group. Tables 3 and 4 show the accuracy of the applied machine learning algorithms used to predict VA (in logMAR scale) as compared with the ground truth (i.e., the outcome as recorded during the patient's visit in our clinic) after 90 days and 365 days, respectively. In each table, the different numbers of previous visits, e.g., condensed histories of the patient's eye with varying complexity, were used to predict the VA after 90 or 365 days. Predicted values were compared with the closest VA measurement available. The best learner in all cases was the L1 regularized linear model Lasso. If only data of one previous visit before the third injection was taken into account for validation, the RMSE of VA over the ground truth was 0.14 logMAR (equals 7 letters) for the short-term prediction (90 days) and 0.23 logMAR (equals 11.5 letters) for the long-term prediction. The MAE of VA over the ground truth was 0.11 logMAR (5.5 letters) for the short-term prediction and 0.16 logMAR (8 letters) for the long-term prediction. The RMSE was lowest (0.2 logMAR – 10 letters), if data of 4 previous visits before the third injection were taken into account. For the MAE there was no improvement, if 4 previous visits were included in validation. Figures 2 and 3 show the difference in visual acuity between prediction and the ground truth, which are the VA values captured in our clinic. Figure 4 shows the weights of features for the visual acuity prediction for 90 days (Figure 4a) and for one year (Figure 4b). Figure 5 shows the change of VA between the true three-month baseline and predicted VA after 3 months and one year.

Discussion

Several approaches have been made to predict progression in patients with AMD. Some of them use genetic information to predict progression to an advanced form of wet AMD [33]. Others are based on imaging information from OCT and assess, for example, the presence and size of drusen [10, 34, 35]. Based on the familiar Comparison of AMD Treatments Trials (CATT) trial, two studies evaluated predictors for vision outcomes after one respectively two years of treatment[36, 37]. Several rules of thumb could be identified: For example, older age,

better baseline VA and greater total fovea thickness, were independently associated with less improvement in VA after at least one year of therapy. These predictors are applicable in many patients, but it can't be used in individual patients to predict his or her vision outcome after a certain time. In this study, we have been able to predict VA in individual patients with neovascular AMD undergoing anti-VEGF therapy after a further 90 or 365 days by using machine learning based on real-life data from an EMR and OCT measurement data, but no OCT images. Figure 4 shows that also in our study, previous VA values have great influence on our prediction model. We also noted (figure 5), that predicted VA values, both short-term as long-term, were different from the ground-truth at the 3-month baseline suggesting that after month 3 changes in regard to VA during active therapy can be observed.

As expected, VA was best predicted for the short-term goal of 90 days. The best machine learning algorithm was able to predict this with a RMSE of 7 letters (+/- 1.5 letters) or a MAE of 5.5 letters (+/- 1 letter). The prediction was within the variability of measurements of VA in a large eye clinic in which a real change can be detected if it is greater than 0.15 logMAR [38]. The best VA prediction (RMSE: 0.20 logMAR (+/- 0.08), MAE: 0.16 logMAR (+/- 0.04)) for 365 days after further therapy was also close to this value of 0.15 logMAR. As mentioned before, the RMSE is more sensitive to outliers, as larger errors are punished more severely. This can be also observed in Figures 2 and 3. One would assume that the prediction is better the more previous encounters are included, but for four previous visits, this was not applicable. One reason might be that too few patients were included in the testing.

VA is, in most medical retina clinical trials, the primary endpoint, although it has not only scientific value, but it gives an indication as to how patients are affected in their daily lives. As is well known, VA is closely correlated with patient quality of life [22]. A recently published review article analyzing 14 studies that assess the psychological impact of anti-VEGF treatments of wet AMD showed that patient often feared persisting vision loss and inability to lead an independent life[24]. A reduction in VA also leads to a reduction in quality of life and the ability to perform daily duties. One of the most precious activity of daily life for many is the ability to drive a car, which is often a necessity for leading an independent life in more rural areas. Approximately 90% of patients have been shown to meet the requirements to drive a car after 24 months of intravitreal anti-VEGF therapy [39]. The prediction of their VA outcome after a defined time alleviates the anxieties of patients and motivated them to pursue Anti-VEGF therapy, which requires frequent consultations with an unknown number

of planned injections. On the other hand, mental support for patients with an unfavorable prognosis needs also to be provided. This helps to reduce the incidence of depression, if initiated at a suitable time [40].

Further refinement can be achieved by the use of patient genetic data. Genetic factors are well known as being involved in disease progression and are more precise in conjunction with further data. It has led to the creation of the "AMD risk calculator", which can be accessed online and for which various entries need to be made (e.g., age, sex, smoking status, body dimensions, and different AMD gene mutations) [41]. A combination of these approaches might help to build a robust decision support tool for AMD, which might help to deliver evidence-based personalized treatment for wet AMD patients.

Major information about the disease state and reaction to therapy is included in OCT images. In this study, we have used numerical measurement information such as CRT from the vendor's software. However, this provides limited information concerning anatomical aspects of the retina during the examination. Future approaches could make use of the OCT scans itself and learn directly from the image by using deep learning. The use of this more advanced technology could provide major performance gains. Initial promising studies involving the application of deep learning to OCT scans have been published. For example, use of deep learning technology has made it possible to classify normal versus age-related macular degeneration OCT scans[9]. In another study, drusen progression over time at the level of a single druse could be predicted over a mean follow-up of 37.8 months, as drusen volume is known to be predictor for progression to late AMD[10]. Further innovations in this field are expected and will likely contribute to more precise VA forecasts.

Interestingly, from a computer science perspective, a simple linear model such as Lasso outperforms complex ensemble-based approaches such as Gradient Boosting Machines (GBM), which usually achieve better results. Moreover, bagging with 10 estimators does not really help to increase the performance indicating that the bagged estimators are not able to learn various aspects of the data and make independent errors. Even training with Lasso's selected features (all non-zero model variables) did not enhance other learners. A linear model thus seems to grasp the internal relationship in the given data relatively well. In this case, the greater capacity of the other learners seems to lead to slight overfitting.

However, more work is necessary to improve real-life data preparation and cleaning, as data input errors occur in daily routine work. An approach with regard to this limitation

might be the creation of national data registries, as more data should improve prediction quality [42]. The presented model relied on data of previous visits before VA prediction after the third injection could be done. This limitation might be overcome in the future, if more clinical data will be available and also OCT image recognition using deep learning is integrated in a model. This would enable instant VA prediction on day of diagnosis or e.g. in second opinion cases. Also, randomized clinical trials must be set up to test the reliability and accuracy of decision support software.

In conclusion, the prediction of VA by using real-life clinical data is a valuable first step to developing a clinical decision support software. It demonstrates the value of well-structured EMRs. It can be used to optimize anti-VEGF treatment in neovascular AMD patients and the allocation of resources to those who need it most.

Figure legends

Figure 1: This figure shows nine sectors, which are placed over the macula within the Heidelberg Eye Explorer software, C0 being the fovea. Data measurements in the XML files were given for the whole grid as well as for each induvial sector. Therefore, features of an individual sector in Table 1 were count nine times. The letters T (temporal), I (inferior), N (nasal) and S (superior) indicate the anatomic position of the sector.

Figure 2: Plots showing differences between visual acuity prediction (red line) and ground truth (green line) for a prediction for 90 further days of treatment. On the x-axis, a consecutive number for each patient eye in the test set is shown. Positive values on y-axis mean improvement of visual acuity, negative values mean deterioration. A graph is plotted for every prediction depending on previously included visits (see headline for the exact number). VA=visual acuity (logMAR).

Figure 3: Plots showing differences between visual acuity prediction (red line) and ground truth (green line) for a prediction for 365 further days of treatment. On the x-axis, a consecutive number for each patient eye in the test set is shown. Positive values on y-axis mean improvement of visual acuity, negative values mean deterioration. A graph is plotted for every prediction depending on previously included visits (see headline for the exact number). VA=visual acuity (logMAR).

Figure 4: The plot shows the weight of the different features for the visual acuity prediction task over 3 months (4a) and 1 year (4b). The light blue bar indicates how important the feature was on average for the model on the different test runs (10 in total). The dark blue bar gives an intuition of how instable the feature was in the different test runs in terms of an importance vs. variance ratio. The more of the light blue bar is overlayed with the dark blue bar, the more instable the feature was during the test runs. T2=Outer temporal measurement section according to the ETDRS grid as placed on the macula, E11.30=ICD-10 code for diabetic type 2 eye complications, I2=outer inferior section of the ETDRS grid, VPP=volume in a grid sector of the ETDRS grid, C0=central location of the ETDRS grid.

Figure 5: This figures shows the difference between true three-month visual acuity and predicted visual acuity after 90 days (figure 5b)/365 days (figure 5b). All values are given in logMAR units. The vertical axis indicates the relative frequency of each VA delta value.

References

1. Samuel AL. Some studies in machine learning using the game of checkers. IBM Journal of research and development 1959;3(3):210-29.
2. Ferrucci DA. Introduction to “this is watson”. IBM Journal of Research and Development 2012;56(3.4):1: -: 15.
3. Caixinha M, Nunes S. Machine Learning Techniques in Clinical Vision Sciences. Current Eye Research 2016:1-15.
4. Thomas RL, Dunstan FD, Luzio SD, et al. Prevalence of diabetic retinopathy within a national diabetic retinopathy screening service. British Journal of Ophthalmology 2015;99(1):64-8.
5. Gulshan V, Peng L, Coram M, et al. Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. JAMA 2016;316(22):2402-10.
6. Oliveira CM, Cristóvão LM, Ribeiro ML, Abreu J. Improved automated screening of diabetic retinopathy. Ophthalmologica 2011;226(4):191-7.
7. Sim DA, Keane PA, Tufail A, et al. Automated retinal image analysis for diabetic retinopathy in telemedicine. Current diabetes reports 2015;15(3):1-9.
8. Haritoglou C, Kernt M, Neubauer A, et al. Microaneurysm formation rate as a predictive marker for progression to clinically significant macular edema in nonproliferative diabetic retinopathy. Retina 2014;34(1):157-64.
9. Lee CS, Baughman DM, Lee AY. Deep Learning Is Effective for Classifying Normal versus Age-Related Macular Degeneration Optical Coherence Tomography Images. Ophthalmology Retina 2017.
10. Bogunović H, Montuoro A, Baratsits M, et al. Machine Learning of the Progression of Intermediate Age-Related Macular Degeneration Based on OCT ImagingMachine Learning of the Progression of AMD Based on OCT. Investigative Ophthalmology & Visual Science 2017;58(6):BIO141-BIO50.
11. De Fauw J, Keane P, Tomasev N, et al. Automated analysis of retinal imaging using machine learning techniques for computer vision. F1000Research 2016;5:1573.
12. Quartilho A, Simkiss P, Zekite A, et al. Leading causes of certifiable visual loss in England and Wales during the year ending 31 March 2013. Eye (Lond) 2016;30(4):602-7.
13. Ferrara N, Hillan KJ, Gerber H-P, Novotny W. Discovery and development of bevacizumab, an anti-VEGF antibody for treating cancer. Nat Rev Drug Discov 2004;3(5):391-400.
14. Group TCR. Ranibizumab and Bevacizumab for Neovascular Age-Related Macular Degeneration. New England Journal of Medicine 2011;364(20):1897-908.
15. Massin P, Bandello F, Garweg JG, et al. Safety and Efficacy of Ranibizumab in Diabetic Macular Edema (RESOLVE Study) A 12-month, randomized, controlled, double-masked, multicenter phase II study. Diabetes care 2010;33(11):2399-405.
16. Regillo CD, Brown DM, Abraham P, et al. Randomized, double-masked, sham-controlled trial of ranibizumab for neovascular age-related macular degeneration: PIER Study year 1. American journal of ophthalmology 2008;145(2):239-48. e5.
17. Brown DM, Campochiaro PA, Singh RP, et al. Ranibizumab for macular edema following central retinal vein occlusion: six-month primary end point results of a phase III study. Ophthalmology 2010;117(6):1124-33. e1.
18. Heier JS, Brown DM, Chong V, et al. Intravitreal aflibercept (VEGF trap-eye) in wet age-related macular degeneration. Ophthalmology 2012;119(12):2537-48.

19. Korobelnik J-F, Do DV, Schmidt-Erfurth U, et al. Intravitreal aflibercept for diabetic macular edema. Ophthalmology 2014;121(11):2247-54.
20. Brown DM, Heier JS, Clark WL, et al. Intravitreal aflibercept injection for macular edema secondary to central retinal vein occlusion: 1-year results from the phase 3 COPERNICUS study. American journal of ophthalmology 2013;155(3):429-37. e7.
21. 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. British Journal of Ophthalmology 2014:bjophthalmol-2014-305327.
22. Finger RP, Wiedemann P, Blumhagen F, et al. 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 ophthalmologica 2013;91(6):540-6.
23. Kortüm KU, Müller M, Kern C, et al. Using electronic health records to build an ophthalmological data warehouse and visualize patients' data. American Journal of Ophthalmology 2017.
24. Senra H, Ali Z, Balaskas K, Aslam T. Psychological impact of anti-VEGF treatments for wet macular degeneration—a review. Graefe's Archive for Clinical and Experimental Ophthalmology 2016;254(10):1873-80.
25. Thetford C, Hodge S, Harding S, et al. Living with age-related macular degeneration treatment: Patient experiences of being treated with ranibizumab (Lucentis)(R) intravitreal injections. British Journal of Visual Impairment 2013;31(2):89-101.
26. Redaktionskommittee, Pauleikhoff D, Bertram B, et al. Die Anti-VEGF-Therapie bei der neovaskulären altersabhängigen Makuladegeneration: Therapeutische Strategien Stand Dezember 2012. Klin Monatsbl Augenheilkd 2013;230(02):170-7.
27. Gesellschaft DO. Die Anti-VEGF-Therapie bei der neovaskulären altersabhängigen Makuladegeneration – therapeutische Strategien. Der Ophthalmologe 2015;112(3):237-45.
28. Drucker H. Improving regressors using boosting techniques. ICML1997; v. 97.
29. Chen T, Guestrin C. XGBoost: A Scalable Tree Boosting System. Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. San Francisco, California, USA: ACM, 2016.
30. Breiman L. Random forests. Machine learning 2001;45(1):5-32.
31. Geurts P, Ernst D, Wehenkel L. Extremely randomized trees. Machine learning 2006;63(1):3-42.
32. Tibshirani R. Regression shrinkage and selection via the lasso. Journal of the Royal Statistical Society Series B (Methodological) 1996:267-88.
33. Ding Y, Liu Y, Yan Q, et al. Bivariate Analysis of Age-Related Macular Degeneration Progression Using Genetic Risk Scores. Genetics 2017;206(1):119-33.
34. de Sisternes L, Simon N, Tibshirani R, et al. Quantitative SD-OCT Imaging Biomarkers as Indicators of Age-Related Macular Degeneration ProgressionPredicting AMD Progression Using SD-OCT Features. Investigative Ophthalmology & Visual Science 2014;55(11):7093-103.
35. Chiu C-J, Mitchell P, Klein R, et al. A Risk Score for the Prediction of Advanced Age-Related Macular Degeneration. Ophthalmology 2014;121(7):1421-7.
36. Ying G-s, 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(12):2523-31.e1.

37. Ying G-s, Huang J, Maguire MG, et al. Baseline Predictors for One-Year Visual Outcomes with Ranibizumab or Bevacizumab for Neovascular Age-related Macular Degeneration. Ophthalmology 2013;120(1):122-9.
38. Siderov J, Tiu AL. Variability of measurements of visual acuity in a large eye clinic. Acta Ophthalmologica Scandinavica 1999;77(6):673-6.
39. Bressler NM, Chang TS, Varma R, et al. Driving Ability Reported by Neovascular Age-related Macular Degeneration Patients after Treatment with Ranibizumab. Ophthalmology 2013;120(1):160-8.
40. Rovner BW, Casten RJ, Hegel MT, et al. Low Vision Depression Prevention Trial in Age-Related Macular Degeneration. Ophthalmology 2014;121(11):2204-11.
41. Seddon JM, Silver RE, Kwong M, Rosner B. Risk Prediction for Progression of Macular Degeneration: 10 Common and Rare Genetic Variants, Demographic, Environmental, and Macular CovariatesRisk Prediction for Progression of AMD. Investigative Ophthalmology & Visual Science 2015;56(4):2192-202.
42. Parke Ii DW, Lum F, Rich WL. The IRIS(R) Registry : Purpose and perspectives. Ophthalmologe 2017;114(Suppl 1):1-6.