Machine learning-based prediction of anatomical outcome after idiopathic macular hole surgery

Yu Xiao[1#], Yijun Hu[1,2,3#], Wuxiu Quan[4#], Bin Zhang[4], Yuqing Wu[5], Qiaowei Wu[1], Baoyi Liu[1], Xiaomin Zeng[1], Zhanjie Lin[1], Ying Fang[1], Yu Hu[6], Songfu Feng[7], Ling Yuan[6], Hongmin Cai[4], Honghua Yu[1], Tao Li[5]

[1]Guangdong Eye Institute, Department of Ophthalmology, Guangdong Provincial People's Hospital; Guangdong Academy of Medical Sciences; the Second School of Clinical Medicine, Southern Medical University, Guangzhou, China; [2]Aier Institute of Refractive Surgery, Refractive Surgery Center, Guangzhou Aier Eye Hospital, Guangzhou, China; [3]Aier School of Ophthalmology, Central South University, Changsha, China; [4]School of Computer Science and Engineering, South China University of Technology, Guangzhou, China; [5]State Key Laboratory of Ophthalmology, Zhongshan Ophthalmic Center, Sun Yat-sen University, Guangzhou, China; [6]Department of Ophthalmology, the First Affiliated Hospital of Kunming Medical University, Kunming, China; [7]Department of Ophthalmology, Zhujiang Hospital of Southern Medical University, Guangzhou, China

Contributions: (I) Conception and design: H Yu, T Li, H Cai, Y Xiao, Y Hu; (II) Administrative support: H Yu, Y Hu, H Cai; (III) Provision of study materials or patients: H Yu, H Cai, T Li, L Yuan, S Feng; (IV) Collection and assembly of data: Y Xiao, Y Fang, Y Wu, Y Hu, S Feng, X Zeng, Z Lin; (V) Data analysis and interpretation: Y Xiao, W Quan, B Zhang, Y Hu, Q Wu, B Liu; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

[#]These authors contributed equally to this work.

Correspondence to: Honghua Yu, MD, PhD. Guangdong Eye Institute, Department of Ophthalmology, Guangdong Provincial People's Hospital; Guangdong Academy of Medical Sciences; the Second School of Clinical Medicine, Southern Medical University, No. 106 Zhongshan Er Road, Yuexiu District, Guangzhou 510080, China. Email: yuhonghua@gdph.org.cn; Tao Li, MD, PhD. State Key Laboratory of Ophthalmology, Zhongshan Ophthalmic Center, Sun Yat-sen University, 7 Jinsui Road, Guangzhou 510060, China. Email: litao2@mail.sysu.edu.cn; Hongmin Cai, PhD. School of Computer Science and Engineering, South China University of Technology, Higher Education Mega Center, Panyu District, Guangzhou 510006, China. Email: hmcai@scut.edu.cn.

Background: To develop a machine learning (ML) model for the prediction of the idiopathic macular hole (IMH) status at 1 month after vitrectomy and internal limiting membrane peeling (VILMP) surgery.

Methods: A total of 288 IMH eyes from four ophthalmic centers were enrolled. All eyes underwent optical coherence tomography (OCT) examinations upon admission and one month after VILMP. First, 1,792 preoperative macular OCT parameters and 768 clinical variables of 256 eyes from two ophthalmic centers were used to train and internally validate ML models. Second, 224 preoperative macular OCT parameters and 96 clinical variables of 32 eyes from the other two centers were utilized for external validation. To fulfill the purpose of predicting postoperative IMH status (i.e., closed or open), five ML algorithms were trained and internally validated by the ten-fold cross-validation method, while the best-performing algorithm was further tested by an external validation set.

Results: In the internal validation, the mean area under the receiver operating characteristic curves (AUCs) of the five ML algorithms were 0.882–0.951. The AUC, accuracy, sensitivity, and specificity of the best-performing algorithm (i.e., random forest, RF) were 0.951, 0.892, 0.973, and 0.904, respectively. In the external validation, the AUC of RF was 0.940, with an accuracy of 0.875, a specificity of 0.875, and a sensitivity of 0.958.

Conclusions: Based on the preoperative OCT parameters and clinical variables, our ML model achieved remarkable accuracy in predicting IMH status after VILMP. Therefore, ML models may help optimize surgical planning for IMH patients in the future.

Keywords: Anatomical outcomes; idiopathic macular hole (IMH); machine learning (ML); prediction model

Page 2 of 10

Xiao et al. Predict postoperative MH status using ML

Introduction

As a leading cause of central vision loss, idiopathic macular hole (IMH) is defined as a full-thickness defect of the neurosensory retina at the fovea (1). Pathological vitreoretinal traction at the central macula has been considered the most critical mechanism of IMH (2,3). The incidence of IMH ranges from 0.1–0.8% in adults aged over 44 years, and nearly two-thirds of the IMH population are females (4-7). Older age, female sex, and pre-existing IMH in the other eye are considered the risk factors of IMH formation (3,4,8). Currently, vitrectomy and internal limiting membrane peeling (VILMP) has become the first-line treatment for full-thickness macular holes (FTMH), with an anatomical closure rate varying from 80% to 95% (9-11). A standard VILMP surgery includes a three-port vitrectomy, internal limiting membrane (ILM) peeling, and an air tamponade. In contrast to the promising overall success rates, the failure rate of large macular holes cannot be neglected, since up to 44% of them remain open after the first surgery (12,13). In such cases, more radical surgeries, such as an inverted ILM flap and an autologous ILM transplantation, may help improve the outcomes (12,14). In addition, previous studies have suggested that older age, longer hole duration, a larger base diameter (BASE), and a longer minimum linear dimension are risk factors of surgical failures, among which the anatomical parameters are the most significant for success rate evaluation (8,15,16).

Optical coherence tomography (OCT) has been extensively applied to the diagnosis and prognosis of IMH due to its noninvasive nature and accurate visualization of the retinal microstructure. Previously, there have been few attempts to analyze and identify various prognostic indicators for the surgical outcomes of IMH, including its minimum diameter (MIN), BASE, hole form factor (HFF), MH index (MHI), tractional hole index (THI) and diameter hole index (DHI) (17-20). However, as previous studies mainly focused on the predictive power of one OCT parameter only, the generalizability of such prediction could be highly limited.

Machine learning (ML) is a subfield of artificial intelligence that has been increasingly applied to ophthalmic practice as it could combine enormous predictors that interact with each other in a non-linear and highly interactive way (21). For instance, a neural network-based ML algorithm has been developed to automatically predict visual outcomes after ranibizumab treatment for diabetic macular edema (DME) (22). However, to the best of the authors' knowledge, no ML model has been developed to predict the anatomical outcomes of IMH after a standard VILMP. If available, this ML model could help vitreoretinal surgeons accurately identify patients at a higher risk of surgical failure, and therefore they can consider applying more advanced surgical techniques.

The aim of this study was to develop a ML model capable of automatically predicting IMH status (closed or open) after VILMP based on preoperative OCT parameters and clinical variables from a multi-center population. We present the following article in accordance with the TRIPOD reporting checklist (available at http://dx.doi.org/10.21037/atm-20-8065).

Methods

Participants

Full-thickness IMH eyes with at least one-month follow-up were retrospectively recruited, while those with macular holes secondary to trauma, high myopia, macular edema, epiretinal membrane, retinal detachment, or retinoschisis were excluded. Age, sex, and duration of symptoms were extracted from electronic medical records (EMR). All eyes underwent ophthalmic examinations, including slit-lamp biomicroscopy on both the anterior segment and the fundus, and spectral domain-OCT scanning (SD-OCT, Spectralis; Heidelberg Engineering, Heidelberg, Germany) upon admission and one month after VILMP. The study was conducted according to the Declaration of Helsinki (as revised in 2013) and was approved by the Institutional Review Board of the Guangdong Provincial People's Hospital (GPPH, No. GDREC2020067H). Informed consent was obtained from all patients.

OCT parameter measurement

OCT examinations were performed by experienced
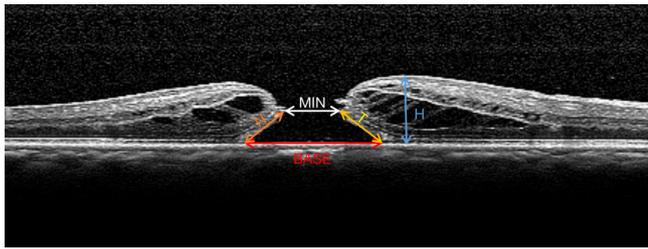
Figure 1 Manual measurement of optical coherence tomography parameters. MIN, minimum diameter of idiopathic macular hole (IMH); BASE, the extent of IMH at the retinal pigment epithelium; T, temporal arm length of IMH; N, nasal arm length of IMH; H, height of IMH.

technicians. A 20°×20° volume acquisition protocol was used to acquire a set of 25 horizontal and vertical cross-sectional B-scan images, each of which was composed of 512 A-scans. The image through the center of the IMH was determined by the simultaneous evaluation of the red-free image on the computer monitor of the OCT scanner, and only the horizontal scan showing the center of the IMH was exported for the OCT parameters measurement (16).

The preoperative macular OCT parameters [i.e., MIN, BASE, height of hole (H), temporal and nasal arm length of hole (T and N)] were manually measured by two independent retinal specialists (X.Z., Y.X.) using ImageJ software (version 1.8.0). The two retinal specialists were unaware of postoperative IMH status when the measurements were performed. The mean values of the measurements were used for training and validating the ML models (*Figure 1*). To measure the repeatability between the two retinal specialists, the intra-class coefficient (ICC) was calculated. MIN was defined as the minimum extent of IMH, while BASE was defined as the extent of IMH at the retinal pigment epithelium (RPE) (17,18). H was defined as the distance from RPE to the innermost aspect of the hole margins, while T and N were defined as the distances between the temporal endpoints or the nasal endpoints of MIN and BASE (16). After measuring all preoperative macular OCT parameters, four indices were calculated: HFF [defined as (T + N)/BASE], MHI (defined as H/BASE), DHI (defined as MIN/BASE), and THI (defined as H/MIN) (18-20).

Postoperative MH status labeling

All the VILMP surgeries were performed by senior

vitreoretinal specialists (H.Y., T.L., L.Y.). IMH closure was defined as the restoration of the continuity among neurosensory retina at the central fovea detected in all the postoperative OCT scans (16). The postoperative status of IMH was classified and labeled as "closed" or "open" by two independent retinal specialists (X.Z., Y.X.) based on the postoperative OCT images taken one month after VILMP. If there was a discordance between the two retinal specialists, arbitration was performed by a senior retinal specialist (H.Y.) to generate the final label.

Training and validation of the ML models

A total of 1,792 preoperative macular OCT parameters and 768 clinical variables of 256 eyes from both the Department of Ophthalmology, GPPH (378 preoperative OCT parameters and 162 clinical variables of 54 eyes), and Zhongshan Ophthalmic Center (ZOC, 1,414 preoperative OCT parameters and 606 clinical variables of 202 eyes) were used for ML training and internal validation. For external validation, 224 preoperative macular OCT parameters and 96 clinical variables of 32 eyes from the Department of Ophthalmology, Zhujiang Hospital of Southern Medical University (ZHSMU, 35 preoperative OCT parameters and 15 clinical variables of five eyes), and the Department of Ophthalmology, the First Affiliated Hospital of Kunming Medical University (FAHKMU, 189 preoperative OCT parameters and 81 clinical variables of 27 eyes), were included.

To develop a ML model for the accurate prediction of postoperative IMH status, five ML algorithms were trained: random forest (RF) (23), decision tree (24), support vector machine (SVM) (25), k-Nearest Neighbor (KNN) (26), and least absolute shrinkage and selection operator (LASSO) (27). For these algorithms, the inputs were OCT parameters (BASE, MIN, H, MHI, THI, THI, and HFF) and clinical variables (age, sex, and duration of symptoms), while the output was postoperative IMH status (i.e., closed or open).

To optimize the performance of the ML algorithms, the commonly hyperparametric search method, Grid Search algorithm, was used to find the best combination of parameters. By traversing the values of all hyperparameters and using ten-fold cross-validation, the group of parameters with the highest cross-validation accuracy of the training set is selected as the corresponding optimal solution. As for the ten-fold cross-validation, the dataset was first randomly split into 10 independent portions. In each run, nine portions

Page 4 of 10

Xiao et al. Predict postoperative MH status using ML
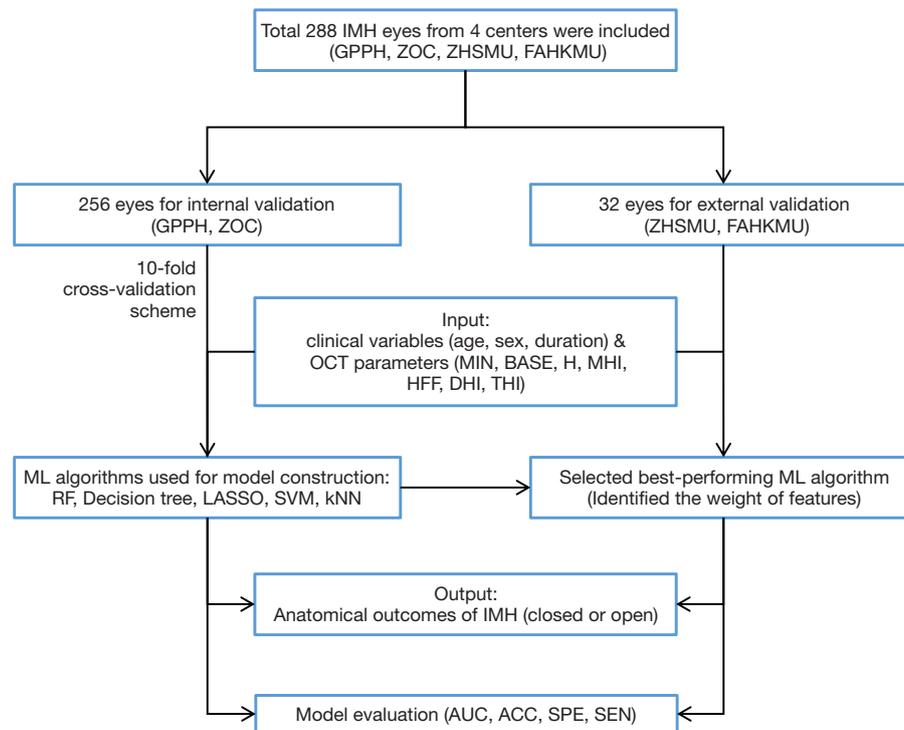
Figure 2 The work-flow diagram. IMH, idiopathic macular hole; ZOC, Zhongshan Ophthalmic Center; ZHSMU, Zhujiang Hospital of Southern Medical University; FAHKMU, First Affiliated Hospital of Kunming Medical University; GPPH, Guangdong Provincial People's Hospital; OCT, optical coherence tomography; MIN, minimum diameter of idiopathic macular hole (IMH); BASE, the extent of IMH at the retinal pigment epithelium (RPE); T, temporal arm length of IMH; N, nasal arm length of IMH; H, height of IMH; THI, tractional hole index; MHI, macular hole index; DHI, diameter hole index; HFF, hole form factor; ML, machine learning; RF, random forest; SVM, support vector machine; KNN, k-Nearest Neighbor; LASSO, least absolute shrinkage and selection operator; AUC, the area under the receiver operating characteristic curve; ACC, accuracy; SPE, specificity: SEN, sensitivity.

of the dataset were employed to train the ML model, while the last portion was used for model testing to facilitate parameter selection and model tuning.

After fine tuning in training and internal validation, the hyperparameters of the ML models were as follows:

RF: fitcensemble function, Learners = templateTree (MaxNumSplits =10), NumLearningCycles =12, LearnRate =0.24.

Decision tree: fitctree function, minleaf =10, Level =0.

SVM: LibLinear package, s =5, B =1, c =6.

KNN: fitcknn function, NumNeighbors =19.

LASSO: Alpha =1, CV =10.

Statistical analysis

In the internal validation, the area under the receiver operating characteristic (ROC) curve (AUC), accuracy, specificity, and sensitivity were generated after each run of the ten-fold CV. The mean values of the ten runs were recorded to measure the overall performance of our ML models. The predictive power of the best-performing ML algorithm was determined in the internal validation and was further verified by our independent external validation dataset. The flowchart is shown in *Figure 2*.

A series of true positive rates (TPR) and false positive rates (FPR) were obtained to form the ROC curves. TPR is also known as sensitivity, while FPR is defined as subtracting the specificity value from 1. The AUC was calculated as the area between the ROC and the FPR axis. The optimal cut-off point was obtained using the highest Youden's index (sensitivity + specificity – 1), and the corresponding optimal sensitivity and specificity values were recorded. The overall accuracy and AUC in internal validation were presented as the mean and the 95% confidence interval (CI). The ML

Table 1 Baseline characteristics of IMH eyes

Characteristics	All eyes (n=288)	Internal validation (n=256)	External validation (n=32)	P values
Age, years, mean (SD)	60.30 (10.80)	60.35 (10.31)	59.90 (14.50)	0.868
Sex, females, n (%)	183 (63.54)	165 (64.45)	18 (56.25)	–
Duration of symptoms, months, mean (SD)	7.02 (11.97)	7.11 (11.98)	6.21 (12.11)	0.691
MIN, µm, mean (SD)	445.70 (211.63)	448.94 (212.59)	419.81 (205.22)	0.464
BASE, µm, mean (SD)	914.82 (348.86)	925.05 (346.77)	832.94 (360.30)	0.159
Height, µm, mean (SD)	443.42 (92.74)	446.94 (91.40)	415.25 (99.91)	0.068
Temporal, µm, mean (SD)	344.17 (130.59)	349.62 (128.64)	300.59 (139.80)	0.067
Nasal, µm, mean (SD)	336.95 (130.01)	340.97 (129.38)	304.81 (132.65)	0.138
MHI, mean (SD)	0.61 (0.50)	0.61 (0.50)	0.66 (0.49)	0.569
THI, mean (SD)	1.47 (1.46)	1.48 (1.51)	1.39 (0.99)	0.675
DHI, mean (SD)	0.49 (0.16)	0.49 (0.16)	0.52 (0.16)	0.265
HFF, mean (SD)	0.80 (0.28)	0.80 (0.29)	0.76 (0.21)	0.434

IMH, idiopathic macular hole; SD, standard deviation; MIN, minimal diameter of IMH; BASE, base diameter of IMH; Temporal, temporal arm length of IMH; Nasal, nasal arm length of IMH; THI, tractional hole index; MHI, macular hole index; DHI, diameter hole index; HFF, hole form factor.

models were established and evaluated by MATLAB (version 9.5.0.9; MathWorks).

In particular, the RF algorithm (i.e., one of the five algorithms in this study) can be used to evaluate the importance of predictors in a dataset. It estimates the importance of each predictor by using the predictorImportance function in MATLAB. This function sums changes in the mean-squared error due to splits on every predictor and then divides the sum by the number of branch nodes. As a result, the function generates the weight of each predictor and provides a graphical display for assessment.

Results

The demographics of the eyes are summarized in *Table 1*. Among 288 eyes, 183 were from female patients (63.54%). The mean age was 60.30±10.80 years, and the mean duration of symptoms was 7.02±11.97 months. There were 208 eyes with a closed IMH (72.22%) and 80 eyes with an open IMH (27.78%) after VILMP surgery, as we intentionally included more IMH-open eyes to statistically balance the cases regardless of the real-world incidence. For the reproducibility of OCT measurements, the inter-observer ICC was 0.958 for MIN, 0.847 for BASE, 0.910 for H, 0.923 for T, and 0.933 for N, suggesting good reproducibility between the two retinal specialists (X.Z.,

Y.X.). There was no significant difference in each baseline parameter between the internal validation set and external validation set, suggesting small selection bias and good homogeneity of the two datasets.
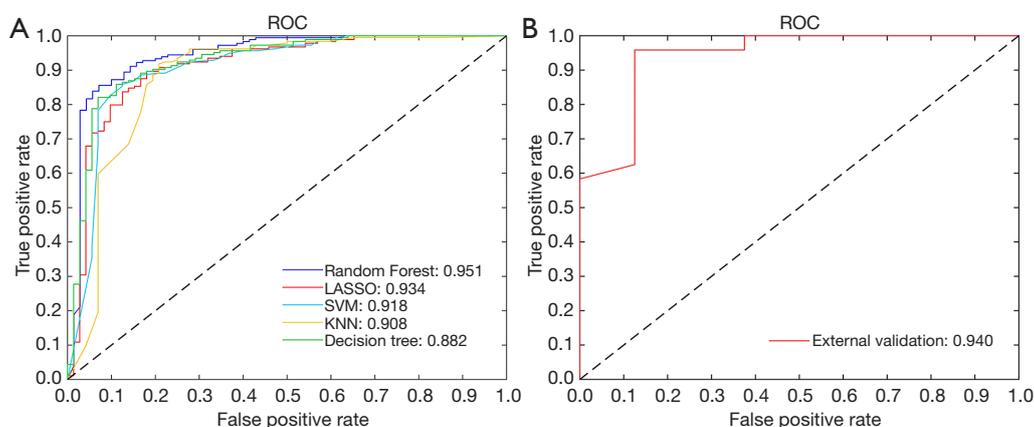
In the internal validation, the mean AUCs of the five ML algorithms were 0.882–0.951 (95% CI: 0.789–0.993), with mean accuracies of 0.857–0.892 (95% CI: 0.812–0.940), mean sensitivities of 0.865–0.934 (95% CI: 0.805–0.970), and mean specificities of 0.804–0.973 (95% CI: 0.659–1.000). Among these ML algorithms, RF achieved the best performance. The mean AUC of RF was 0.951 (95% CI: 0.908–0.993), with a mean accuracy of 0.892 (95% CI: 0.844–0.940), a mean sensitivity of 0.904 (95% CI: 0.856–0.952), and a mean specificity of 0.973 (95% CI: 0.938–1.000). The performance metrics of the ML algorithms are shown in *Table 2*, and the comparison among ROC curves is presented in *Figure 3A*.

In the external validation, the AUC obtained by the RF algorithm was 0.940, with an accuracy of 0.875, a sensitivity of 0.875, and a specificity of 0.958 (*Table 2* and *Figure 3B*). Furthermore, our best-performing model (i.e., RF algorithm) identified BASE as the most critical parameter for the accurate prediction of postoperative IMH status (*Figure 4*). To better illustrate the importance of BASE, binary logistic regression analysis was derived for prediction of IMH status using BASE in internal validation set. The ROC curve of

Page 6 of 10

Xiao et al. Predict postoperative MH status using ML

Table 2 The performance evaluation of ML models

ML models	AUC	ACC	SPE	SEN
Internal validation				
Random forest (95% CI)	0.951* (0.908–0.993)	0.892* (0.844–0.940)	0.973* (0.938–1.000)	0.904 (0.856–0.952)
LASSO (95% CI)	0.934 (0.897–0.972)	0.862 (0.828–0.897)	0.934 (0.864–1.000)	0.874 (0.805–0.943)
SVM (95% CI)	0.918 (0.879–0.956)	0.868 (0.826–0.910)	0.932 (0.888–0.977)	0.865 (0.826–0.904)
KNN (95% CI)	0.908 (0.846–0.971)	0.857 (0.812–0.901)	0.920 (0.849–0.990)	0.881 (0.833–0.930)
Decision tree (95% CI)	0.882 (0.789–0.976)	0.879 (0.831–0.926)	0.804 (0.659–0.949)	0.934* (0.899–0.970)
External validation				
Random forest	0.940	0.875	0.875	0.958

All mean prediction values were given with their 95% CI. The best results among the internal validation set were marked by * in each column. An ML Model with random forest algorithm was selected as the best-performing model and validated externally. ML, machine learning; 95% CI, 95% confidence interval; LASSO, least absolute shrinkage and selection operator; SVM, support vector machine; KNN, k-Nearest Neighbor; AUC, The area under the receiver operating characteristic curve; ACC, accuracy; SPE, specificity; SEN, sensitivity.

Figure 3 The ROC curves of the machine learning models. A graph is plotted for each prediction task, including the prediction task in the internal validation set (A) and the external validation set (B). ROC, receiver operating characteristic; SVM, support vector machine; KNN, k-Nearest Neighbor; LASSO, least absolute shrinkage and selection operator.

BASE was showed in *Figure 5*, with an AUC of 0.911.

Discussion

In the present study, five ML algorithms were trained and evaluated to predict IMH status after standard VILMP surgery. The RF algorithm (i.e., the best-performing algorithm) was further validated using an independent external validation set. In our study, the best-performing ML algorithm provided promising predictions with an AUC of 0.951 in the training and internal validation set and an AUC of 0.940 in the external validation set, suggesting its encouraging capability of accurately predicting IMH status (i.e., closed or open) after VILMP based on preoperative OCT features and clinical variables. Moreover, we identified that BASE is the most critical feature for an accurate prediction.

As it is well-acknowledged that the development of IMH is mainly caused by anterior and tangential vitreoretinal traction to the fovea, VILMP has become the first-line treatment for FTMH with a success rate of 80–95% (9-11). Despite the high success rate, IMH can remain open in some cases (13). After the initial surgery that failed to achieve IMH closure, a second surgery is needed more often than not, which is usually associated with higher medical costs and less favorable visual outcomes compared
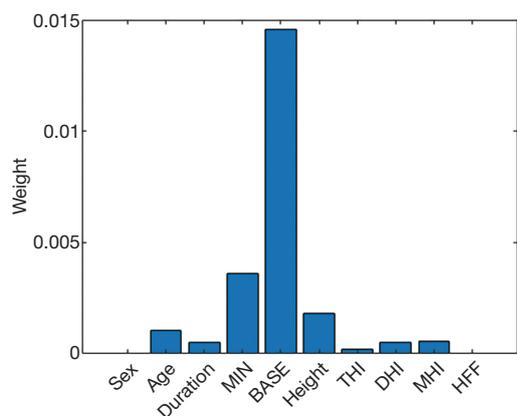
Figure 4 The weight of features. Our best-performing machine learning model (random forest scheme) identified the weight of different features after running all the samples. The blue bars indicate the weight of the features. The higher the blue bar, the more important the corresponding feature is for the prediction task. Duration, duration of symptoms; MIN, minimum diameter of IMH; IMH, idiopathic macular hole; BASE, the extent of IMH at the RPE; RPE, retinal pigment epithelium; T, temporal arm length of IMH; N, nasal arm length of IMH; Height, height of IMH; THI, tractional hole index; MHI, macular hole index; DHI, diameter hole index; HFF, hole form factor.
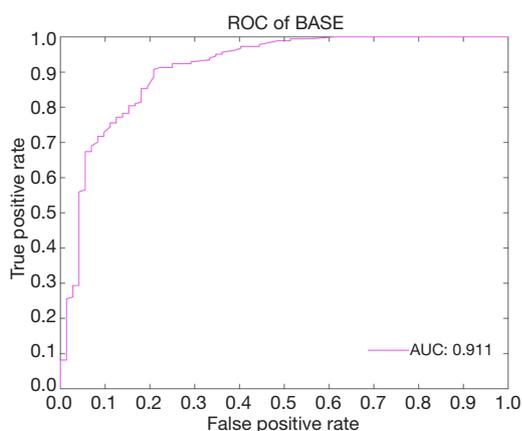
Figure 5 The ROC curve of BASE. A binary logistic regression analyse was derived for BASE using internal validation set. ROC, receiver operating characteristic; BASE, the extent of IMH at the retinal pigment epithelium.

to primary closure (9,13). Thus, it is of clinical importance to identify IMH patients susceptible to surgical failure after a standard VILMP. The prediction of the postoperative IMH status could alleviate anxieties of patients and could

help surgeons choose more radical surgical techniques, such as an inverted ILM flap and an autologous ILM transplantation, for patients with an unfavorable prognosis and refractory IMH (12,14).

Older age, larger BASE and MIN, and longer hole duration are commonly considered risk factors of the surgical failure of IMH (8,15,16). Previously, studies have focused on using preoperative OCT parameters to predict the anatomical outcomes of IMH. It was reported that IMHs with a horizontal diameter of more than 400 μm had a lower closure rate (56%) compared to those with diameters less than 400 μm (92%) (17). In addition, IMHs were closed after surgery in all the eyes with an HFF >0.9, while the mean anatomical success rate was only 67% in eyes with an HFF <0.5 (18). Recently, a study for which a binary logistic regression analysis was used to predict the anatomical outcomes of the IMH base on MHI and THI achieved an AUC of 0.909 and 0.708, respectively (16). However, these types of studies may be limited because they only analyze the predictive power of a single parameter. Moreover, the logistic regression-based prediction could also be rather unreliable when dealing with non-linear data, which could exist in this prediction. Taken together, it is particularly beneficial and meaningful to develop a ML model to automatically identify IMH patients at a higher risk of surgery failure after a standard VILMP.

Within the broad category of artificial intelligence, ML models are particularly suited to predictions based on existing data (28). First, ML algorithms can adjust hyperparameters and have a better orientation of the mission. Second, the evaluation performance of the ensemble ML method is stable, robust, and comprehensive. Lastly, compared to a logistic regression, a collinearity diagnosis of independent variables is not required for ML algorithms. ML algorithms have been intensively used to accurately predict therapeutic outcomes of various ophthalmic diseases, including age-related macular degeneration (AMD), macular edema associated with retinal vein occlusion (RVO-ME), and DME (29-32). For instance, a ML algorithm (LASSO protocol) was shown to have a reliable prediction of visual acuity (VA) in neovascular AMD patients after three consecutive anti-vascular endothelial growth factor (VEGF) injections. Base on EMR (e.g., baseline VA) and the measurement parameters from OCT (e.g., central retinal thickness), this ML model achieved a mean absolute error of 5.5 and 13 letters in the 3-month and 12-month VA forecast, respectively (29). Based on the

Page 8 of 10

Xiao et al. Predict postoperative MH status using ML

retinal thickness features from OCT images, a ML model using Extra Trees achieved AUCs of 0.76–0.83 and 0.76–0.79 in predicting the recurrence of branch retinal vein occlusion or central retinal vein occlusion within one year, respectively (32). However, to date, there is no ML model available for the prediction of IMH surgery outcomes.

In addition, our best-performing ML model (using RF algorithm) identified that BASE is the most critical feature for the prediction of postoperative IMH status, and a binary logistic regression analyse derived for BASE also presented a well-performed result, with an AUC of 0.911. These findings suggest that the size of IMH may be the most important parameter to determine the postoperative status. Most investigators have assumed that tangential traction to the fovea plays a crucial role in the development of IMH and that the success of IMH surgery is highly dependent on the relief of such traction. BASE, the maximum diameter of IMH at the RPE layer, directly reflects the size of retinal defects in the foveal neurosensory retina. In addition, it serves as an indicator of the transverse traction force to the fovea. Therefore a smaller BASE often means less tractional force, which is more likely to be removed by a standard VILMP surgery. Similarly, a prior study used a binary logistic regression analysis to evaluate the prognosis predictive power of several OCT parameters (e.g., MIN, BASE, MHI), which also showed that BASE is the strongest indicator in predicting anatomical success after IMH surgery (16).

From a computer science perspective, the RF algorithm, established by constructing a multitude of decision trees at training time, can be used to avoid overfitting and to help reduce the variance of the algorithm by selecting features based on their prediction values. As the best-performing algorithm in our study, although dealing with the multiple features and the unbalanced samples from our multi-center dataset, the RF algorithm still showed an excellent performance due to its robustness and stability. On the contrary, as a simple algorithm based on the distance function, a poor performance can be expected from KNN, while the SVM algorithm has difficulty obtaining satisfactory results when dealing with multiple features and unbalanced samples. As a simple linear model, the LASSO algorithm also underperformed compared to complex ensemble-based approaches (e.g., RF) in this study. Likewise, a previous study compared the performance of RF and SVMs with different kernel functions in the prediction of the treatment response of DME after an anti-vascular endothelial growth factor injection. It was shown that RF had the best predictive ability among five ML models with an AUC of 0.951, inferring that RF has stronger anti-noise and anti-overfit capabilities due to its random operations.

There are several limitations of this study. Firstly, the presented model relies on the manual measurement of preoperative macular OCT parameters. Nevertheless, the acceptable repeatability and reproducibility of these manual measurements on OCT have been demonstrated in previous studies (33,34). Secondly, since 3D-OCT images may be more useful than 2D-OCT images in IMH surgery, future investigation should be performed to validate the IMH outcome prediction using 3D-OCT images. Another limitation is that the number of patients included in the dataset was limited, especially in the external validation set, because only eight patients without primary closure after surgery were available; however, the cases in the training set and the external validation set were enrolled from different ophthalmic centers, and the promising performance in external validation also suggests the excellent adaptability of our ML model. Furthermore, to improve the generalizability of our ML model, it is important to recruit IMH with different characteristics (such as age and myopia severity) and collect images obtained from different OCT devices for development and validation of the ML model in future work.

In conclusion, our results suggest that it is possible to develop a ML model for the automatic and accurate prediction of IMH status after a standard vitreoretinal surgery. Predictions by the ML model could alleviate the anxieties of patients at low risk of surgery failure and could motivate both the surgeons and patients to pursue more progressive surgical methods in more challenging cases where patients are prone to an unfavorable prognosis.

Footnote

Reporting Checklist: The authors have completed the TRIPOD reporting checklist. Available at http://dx.doi.org/10.21037/atm-20-8065

Data Sharing Statement: Available at http://dx.doi.org/10.21037/atm-20-8065

Peer Review File: Available at http://dx.doi.org/10.21037/atm-20-8065

Conflicts of Interest: All authors have completed the ICMJE uniform disclosure form (available at http://dx.doi.org/10.21037/atm-20-8065). The authors have no conflicts of interest to declare.

Ethical Statement: The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. The study was conducted according to the Declaration of Helsinki (as revised in 2013) and was approved by the Institutional Review Board of Guangdong Provincial People's Hospital (No. GDREC2020067H). Informed consent was obtained from all patients.

References

1. Takahashi A, Yoshida A, Nagaoka T, et al. Idiopathic full-thickness macular holes and the vitreomacular interface: a high-resolution spectral-domain optical coherence tomography study. Am J Ophthalmol 2012;154:881-92.e2.

2. Duker JS, Kaiser PK, Binder S, et al. The International Vitreomacular Traction Study Group Classification of Vitreomacular Adhesion, Traction, and Macular Hole. Ophthalmology 2013;120:2611-9.

3. Chew EY, Sperduto RD, Hiller R, et al. Clinical course of macular holes: the Eye Disease Case-Control Study. Arch Ophthalmol 1999;117:242-6.

4. McCannel CA, Ensminger JL, Diehl NN, et al. Population-based incidence of macular holes. Ophthalmology 2009;116:1366-9.

5. Thapa SS, Thapa R, Paudyal I, et al. Prevalence and pattern of vitreo-retinal diseases in Nepal: the Bhaktapur glaucoma study. BMC Ophthalmol 2013;13:9.

6. García-Layana A, García-Arumí J, Ruiz-Moreno JM, et al. A review of current management of vitreomacular traction and macular hole. J Ophthalmol 2015;2015:809640.

7. Wang S, Xu L, Jonas JB. Prevalence of full-thickness macular holes in urban and rural adult Chinese: the Beijing Eye Study. Am J Ophthalmol 2006;141:589-91.

8. Kim Y, Kim ES, Yu SY, et al. Age-related clinical outcome after macular hole surgery. Retina 2017;37:80-7.

9. Gross JG. Late reopening and spontaneous closure of previously repaired macular holes. Am J Ophthalmol 2005;140:556-8.

10. Yao Y, Qu J, Dong C, et al. The impact of extent of internal limiting membrane peeling on anatomical outcomes of macular hole surgery: results of a 54-week randomized clinical trial. Acta Ophthalmol 2019;97:303-12.

11. Elhusseiny AM, Schwartz SG, Flynn HW Jr, et al. Long-Term Outcomes after Macular Hole Surgery. Ophthalmol Retina 2020;4:369-76.

12. Michalewska Z, Michalewski J, Adelman RA, et al. Inverted internal limiting membrane flap technique for large macular holes. Ophthalmology 2010;117:2018-25.

13. Abbey AM, Van Laere L, Shah AR, et al. Recurrent macular holes in the era of small-gauge vitrectomy: A Review of Incidence, Risk Factors, and Outcomes. Retina 2017;37:921-4.

14. Morizane Y, Shiraga F, Kimura S, et al. Autologous transplantation of the internal limiting membrane for refractory macular holes. Am J Ophthalmol 2014;157:861-9.e1.

Page 10 of 10

Xiao et al. Predict postoperative MH status using ML

15. Essex RW, Kingston ZS, Moreno-Betancur M, et al. The Effect of Postoperative Face-Down Positioning and of Long- versus Short-Acting Gas in Macular Hole Surgery: Results of a Registry-Based Study. Ophthalmology 2016;123:1129-36.

16. Wakely L, Rahman R, Stephenson J. A comparison of several methods of macular hole measurement using optical coherence tomography, and their value in predicting anatomical and visual outcomes. Br J Ophthalmol 2012;96:1003-7.

17. Ip MS, Baker BJ, Duker JS, et al. Anatomical outcomes of surgery for idiopathic macular hole as determined by optical coherence tomography. Arch Ophthalmol 2002;120:29-35.

18. Ullrich S, Haritoglou C, Gass C, et al. Macular hole size as a prognostic factor in macular hole surgery. Br J Ophthalmol 2002;86:390-3.

19. Kusuhara S, Escaño MFT, Fujii S, et al. Prediction of postoperative visual outcome based on hole configuration by optical coherence tomography in eyes with idiopathic macular holes. Am J Ophthalmol 2004;138:709-16.

20. Ruiz-Moreno JM, Staicu C, Pinero DP, et al. Optical coherence tomography predictive factors for macular hole surgery outcome. Br J Ophthalmol 2008;92:640-4.

21. Obermeyer Z, Emanuel EJ. Predicting the Future - Big Data, Machine Learning, and Clinical Medicine. N Engl J Med 2016;375:1216-9.

22. Chen SC, Chiu HW, Chen CC, et al. A novel machine learning algorithm to automatically predict visual outcomes in intravitreal ranibizumab-treated patients with diabetic macular edema. J Clin Med 2018;7:475.

23. Breiman L. Random forests. Mach Learn 2001;45:5-32.

24. Safavian SR, Landgrebe D. A survey of decision tree classifier methodology. IEEE Trans Syst Man Cybern Syst 1991;21:660-74.

25. Noble WS. What is a support vector machine? Nat Biotechnol 2006;24:1565-7.

26. Cover T, Hart P. Nearest neighbor pattern classification. IEEE Trans Inf Theory 1967;13:21-7.

27. Tibshirani R. Regression shrinkage and selection via the lasso: a retrospective. J R Stat Soc Ser B-Stat Methodol 2011;73:267-88.

28. Chen JH, Asch SM. Machine Learning and Prediction in Medicine - Beyond the Peak of Inflated Expectations. N Engl J Med 2017;376:2507-9.

29. 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;125:1028-36.

30. Liu B, Zhang B, Hu Y, et al. Automatic prediction of treatment outcomes in patients with diabetic macular edema using ensemble machine learning. Ann Transl Med 2021;9:43.

31. Wu Q, Zhang B, Hu Y, et al. Detection of morphologic patterns of diabetic macular edema using a deep learning approach based on optical coherence tomography images. Retina 2021;41:1110-7.

32. Vogl WD, Waldstein SM, Gerendas BS, et al. Predicting Macular Edema Recurrence from Spatio-Temporal Signatures in Optical Coherence Tomography Images. IEEE Trans Med Imaging 2017;36:1773-83.

33. Rahman W, Chen FK, Yeoh J, et al. Repeatability of manual subfoveal choroidal thickness measurements in healthy subjects using the technique of enhanced depth imaging optical coherence tomography. Invest Ophthalmol Vis Sci 2011;52:2267-71.

34. Chhablani J, Barteselli G, Wang H, et al. Repeatability and reproducibility of manual choroidal volume measurements using enhanced depth imaging optical coherence tomography. Invest Ophthalmol Vis Sci 2012;53:2274-80.