• **Clinical Research** •

High interpretable machine learning classifier for early glaucoma diagnosis

Carlos Salvador Fernandez Escamez[1,2]*, Elena Martin Giral*[1]*, Susana Perucho Martinez*[1]*, Nicolas Toledano Fernandez*[1]

[1]Ophthalmology Department, Hospital de Fuenlabrada, Madrid 28942, Spain

[2]Doctorate Program in Health Sciences, Universidad Rey Juan Carlos, Alcorcon 28922, Madrid, Spain

Correspondence to: Carlos Salvador Fernandez Escamez. Servicio de Oftalmologia, Hospital de Fuenlabrada, C/ Camino del Molino S/N, Fuenlabrada, Madrid 28942, Spain. carlossalvador.fernandez@salud.madrid.org

Abstract

● **AIM:** To develop a classifier for differentiating between healthy and early stage glaucoma eyes based on peripapillary retinal nerve fiber layer (RNFL) thicknesses measured with optical coherence tomography (OCT), using machine learning algorithms with a high interpretability.

● **METHODS:** Ninety patients with early glaucoma and 85 healthy eyes were included. Early glaucoma eyes showed a visual field (VF) defect with mean deviation >-6.00 dB and characteristic glaucomatous morphology. RNFL thickness in every quadrant, clock-hour and average thickness were used to feed machine learning algorithms. Cluster analysis was conducted to detect and exclude outliers. Tree gradient boosting algorithms were used to calculate the importance of parameters on the classifier and to check the relation between their values and its impact on the classifier. Parameters with the lowest importance were excluded and a weighted decision tree analysis was applied to obtain an interpretable classifier. Area under the ROC curve (AUC), accuracy and generalization ability of the model were estimated using cross validation techniques.

● **RESULTS:** Average and 7 clock-hour RNFL thicknesses were the parameters with the highest importance. Correlation between parameter values and impact on classification displayed a stepped pattern for average thickness. Decision tree model revealed that average thickness lower than 82 µm was a high predictor for early glaucoma. Model scores had AUC of 0.953 (95%CI: 0.903-0998), with an accuracy of 89%.

● **CONCLUSION:** Gradient boosting methods provide accurate and highly interpretable classifiers to discriminate between early glaucoma and healthy eyes. Average and 7-hour RNFL thicknesses have the best discriminant power.

● **KEYWORDS:** machine learning; glaucoma; diagnosis; optical coherence tomography

INTRODUCTION

Glaucoma is one of the principal causes of preventable irreversible visual loss. Its diagnosis and early treatment are critically important in order to preserve visual function[1]. However, detection of glaucoma in its early stages can be challenging due to the wide variation in the normal appearance of the optic nerve[2].

The newest optical coherence tomography (OCT) devices have enabled the detection of subtle defects in the retinal nerve fiber layer (RNFL) that were previously undetectable through clinical examination[3]. The OCT yields a huge amount of numerical data, whose combined analysis can be performed using machine learning methods. Among these methods, some have the property of interpretability which allow the comparison of the results with prior knowledge, as well as discovering how the algorithm works to obtain the result[4].

The goal of the present study was to develop and validate an interpretable classifier able to distinguish between healthy eyes and early glaucoma by means of the peripapillary RNFL thickness measured with OCT. The newest tree gradient boosting (GB) methods were used, for which there exists a GB explainer extension that allowed to achieve a high level of interpretability of the model.

SUBJECTS AND METHODS

Ethical Approval The study had the approval of the Institutional Review Board of the hospital with approval number APR 18/16 and adhered to the principles of the Declaration of Helsinki. Informed consent was required to all subjects. All participants underwent a comprehensive ophthalmic examination.

Study Design and Participants This was an observational cross-sectional study of consecutive glaucoma patients at Fuenlabrada Hospital institution (Madrid, Spain) during the years 2018 to 2019. Normal subjects were recruited from the population who came to clinic for regular eye examination. Glaucoma patients were selected from those who came to clinic as a first diagnosis or for successive examination and showed only mild visual field (VF) damage.

Spectral domain OCT examination was performed on all subjects using a commercially available equipment (Topcon 3D OCT 2000 MarkII, FastMap v. 3.40, Topcon Corp., Japan). Tomographic images of the peripapillary area were obtained with the three-dimensional disc scan on a measured circle 3.46 mm in diameter. Scans were obtained after pupil dilation with 1% tropicamide. Automatic perimetry with 24-2 Swedish Interactive Threshold Algorithm of the Humphrey Visual Field Analyzer (Carl Zeiss Meditec Inc.) was conducted on all subjects.

All tests were completed within a week in order to avoid changes in glaucoma stage while collecting the data.

Examination of the subjects was carried out using E-Snellen visual acuity test and conventional slit-lamp and funduscopic techniques to rule out anomalies of the anterior or posterior segment. Intraocular pressure (IOP) was measured using contact Goldman tonometry. Subjects accepted for research had best corrected visual acuity of 20/40 or better, spherical refraction between ±5.0 diopters and cylinder correction less than 3.0 diopters. They did not have any previous ocular surgery, except uncomplicated cataract extraction and intraocular lens implantation, and had no other anterior or posterior segment disease.

Eligibility Criteria One eye was randomly selected if both eyes were eligible for the study. Inclusion criteria for the patients with early stage glaucoma included a basal IOP ≥21 mm Hg, gonioscopic open angle, and a reproducible VF defect in the absence of any other neurological or ocular defect that may cause an abnormal VF test result. VF test was conducted at least three times. Based on Hoddap-Parrish-Anderson criteria, early VF damage was defined as an arc-like perimetric defect with mean deviation ≥ -6.00 dB, glaucoma hemifield test result outside normal limit or corrected pattern standard deviation (CPSD)/pattern standard deviation (PSD) significant at $P<0.05$[5]. Superior, inferior or both hemifields were affected. Reliability was calculated on a false negative rate <15%, false positive rate <10% and fixation losses <10%. Exclusion criteria for early glaucoma group included eyes with close or anomalous angle, IOP<21 mm Hg, unreliable VFs or anomalous disc aspect such as tilted discs.

Inclusion criteria for controls were those with no history of eye disease, no family history of glaucoma, IOP<21 mm Hg, no sign of glaucoma optic disc damage, and a normal result on three followed reliable VF tests. Eyes with VF defects, IOP>21 mm Hg or any anomalous appearance of the optic nerve were excluded in the control group.

OCT scans with segmentation errors or signal-to-noise ratio index lower than 45 were rejected. All examinations were reviewed by an experienced ophthalmologist and discarded if there was any suspicion of being decentered or not being completely reliable.

Analysis Algorithms All statistical and machine learning algorithms were carried out using R-free software environment for statistical computing (version 3.6.2 for Windows).

Unsupervised k-medoid cluster analysis was taken with the unlabeled dataset, using all the peripapillary RNFL measurements obtained with OCT. Points located further than the sum of the mean distance of the points in each cluster to its center plus two times the standard deviation were considered as outliers and excluded of the dataset for further analysis[6]. Tree GB methods were applied to the processed data in order to produce a prediction model in the form of an ensemble of decision trees. These algorithms improve the accuracy of predictions, at the expense of loss in interpretation of the model[7]. Relative importance of the parameters in GB trees was calculated, and those with the greatest importance were selected. GB explainer extension was used to interpret the generated GB model, exploring the correlations between the numeric value of each parameter and its impact on classification.

In order to minimize overfitting, only parameters with the highest importance on the GB model were used as input data to build a pruned decision tree. A decision tree was chosen as the preferred classification algorithm due to its high level of interpretability, allowing the results obtained in our model to be checked against previous knowledge regarding RNFL damage in glaucoma. In order to minimize the number of early glaucoma eyes misclassified as normal eyes, a weight matrix was introduced in the model, and a fitted tree classifier was constructed.

The outcome of our study will be a decision tree classifier. Detection of parameters with the greatest relative importance will be assessed. The ability to distinguish between glaucoma and controls will be evaluated using ten-fold cross validation, estimating the accuracy rate, AUC curves and sensibility and specificity.

RESULTS

In this study we recruited 90 eyes with early stage glaucoma and 85 controls, who were included in the study. All participants were Caucasian race.

Baseline characteristics of the study population are summarized in Table 1. Optic disc size difference was not significant among the groups.

Table 1 Clinical characteristics of the control and early glaucoma groups

Parameters	Healthy (*n*=85)	Early glaucoma (*n*=90)	*P*
Age (y)	61.07±8.93	65.61±10.20	0.005[a]
Male (%)	55.2	43.3	0.141[b]
Left eye (%)	44.7	60.0	0.051[b]
VA (logMAR)	0.02±0.05	0.07±0.09	<0.001[a]
SE (D)	0.21±1.33	0.45±1.60	0.332[a]
IOP (mm Hg)	17.21±2.01	22.56±3.6	<0.001[a]
VFI (%)	98.54±1.47	86.90±8.12	<0.001[c]
SAP MD (dB)	-0.64±1.36	-47.74±1.76	<0.001[c]
SAP PSD (dB)	1.85±0.50	6.09±2.90	<0.001[c]

VA: Visual acuity; SE: Spherical equivalent; D: Diopters; IOP: Intraocular pressure; VFI: Visual field index; SAP: Standard automated perimetry; MD: Mean deviation; PSD: Pattern standard deviation. [a]Unpaired *t*-test; [b]χ^2 test; [c]Krusal-Wallis test.

Kolmogorov-Smirnov one-sample test showed Gaussian distribution of the variables. Unpaired *t*-test analysis demonstrated significant difference for all RNFL values except for 9 clock-hour RNFL thickness.

Unsupervised clustering revealed four eyes in the early glaucoma group as outliers which were excluded from the dataset for further analysis.

GB analysis displayed average thickness as the parameter with the highest importance in the classification model, followed by 7 clock-hour thickness (Figure 1).

Scatter-plot graphic correlating thickness values for the different RNFL parameters and the impact on classification, estimated through the log-odd of classification, showed a stair-like pattern for average thickness, with a plateau at the values <80 μm, >90 μm and between 80 and 90 μm (Figure 2). Similar correlation graphs for the rest of the RFNL values exhibited a flat pattern with no step depending on thickness value.

The pruned decision tree built after excluding all those parameters which GB proved to have a low impact on classification is shown in Figure 3. Eyes with an average thickness inferior to 82 μm are glaucomatous eyes in 98% of cases. Eyes with an average thickness >90 μm and 7 clock-hour thickness >99 μm have a 95% predicted probability of being healthy. The overall classification accuracy evaluated through ten-fold cross validation was 89%. Estimated false negative rate was 13% and AUC of the model was 0.953 (95%CI: 0.903-0.998). Sensibility reached 0.89 with specificity 0.885.

A modified tree constructed after adding a weight matrix in order to minimize misclassification of glaucoma eyes included average and inferior quadrant thickness as the splitting parameters, with different cleavage values than those in the non-weighted tree (Figure 4). However, classification accuracy
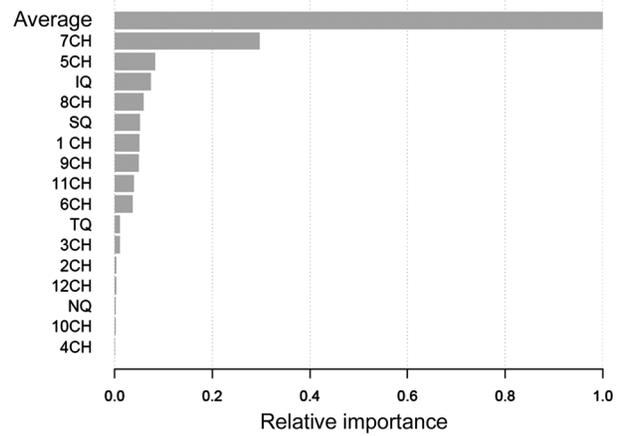

Figure 1 Relative importance of average, quadrants, and clock-hour segments of peripapillary retinal nerve fiber layer thickness on the tree gradient boosting model CH: Clock-hour; IQ: Inferior quadrant; SQ: Superior quadrant; TQ: Temporal quadrant; NQ: Nasal quadrant.
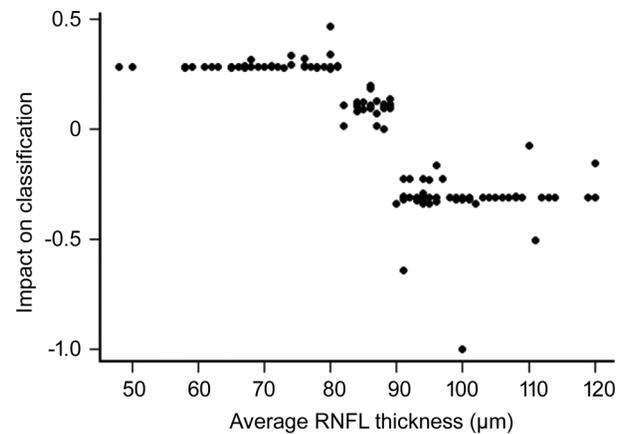

Figure 2 Impact on classification represented against average RNFL thickness Impact on classification is expressed as the log-odd of classification. A positive impact value is associated with a greater probability of being classified as early glaucoma. RNFL: Retinal nerve fiber layer.
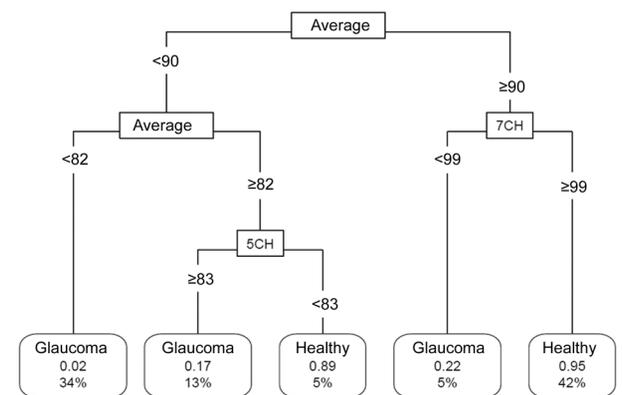

Figure 3 Pruned classification tree constructed after excluding outliers and including features which the greatest relative importance in gradient boosting model Connecting lines include the splitting value for the parameter. Leaf nodes show the class of the node, the predicted probability of being healthy and the percentage of observations in the node. Average: Average RNFL thickness; CH: Clock-hour thickness.

Figure 4 Modified classification tree built adding a loss matrix to minimize misclassification of glaucoma as healthy eyes Connecting lines include the splitting value for the parameter. Leaf nodes show the label of the node, the predicted probability of being healthy and the percentage of observations in the node. Average: Average RNFL thickness; CH: Clock-hour thickness.
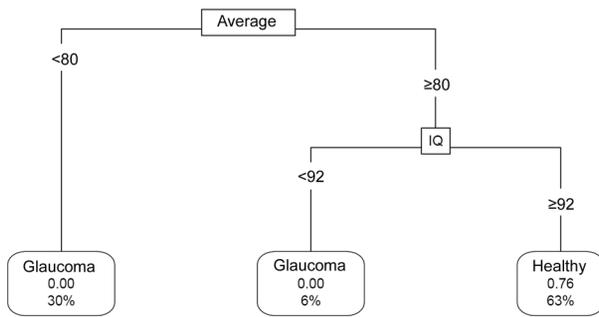
of the model drops to 84% as estimated with ten-fold cross validation. AUC of the model was 0.85 (95%CI: 0.752-0.958) and sensitivity was 0.745 with specificity of 0.956.

DISCUSSION

Early diagnosis in glaucoma constitutes one of the biggest challenges in ophthalmology. The importance of its early detection is justified by Mansberger: it is a significant public health problem; it is treatable and is able to be detected in early stages[8].

Clinical evaluation has been, until recent times, the gold standard for early detection[2]. However, VF defects are often difficult to evaluate due to the high variance in patients with low reproducibility, especially in elderly patients with cognitive impairment[9]. OCT technology provides an objective and reproducible analysis of the optic nerve and peripapillary RNFL. Many studies have shown evidence that structural damage often precedes the detection of functional changes, which supports the validity of the analysis of the RNFL with OCT for early diagnosis in glaucoma[1].

Several studies have reported the correlation between RNFL defects and VF damage[10-14]. Some authors have described macular ganglion complex cell (GCC) thinning in glaucoma. However, numerous studies support that RNFL parameters are superior to GCC for glaucoma diagnosis with OCT[15-17]. Vascular changes at the optic disc might provide useful information for diagnosis of glaucoma. Bojikian et al[18] have shown significant correlation between peripapilary blood flux and VF parameters and structural biometrics in primary open-angle glaucoma. In any case, OCT provides a huge amount of information that is not easily interpreted with traditional statistical techniques.

Machine learning offer tools to deal with datasets which exhibit a large size and high multidimensionality. Using these techniques, we can find hidden patterns that would be undetected with classical statistical analysis. The number of these methods increases day by day, making it extremely

complicated to decide which is the best method to apply. Classification and regression trees, linear discriminant analysis, Naïve Bayes, artificial neural networks, deep learning algorithms, random forest and other machine learning methods have been widely used to predict, classify or detect ocular diseases[19-25]. These methods obtain a high level of accuracy and can achieve a nearly perfect classification rate. However, they often lack one of the fundamental aspects of science: interpretability.

Interpretability is defined as the capacity to gain insight into how a data mining method gets to the result[4]. It provides the possibility to link the results obtained using data mining techniques with previous knowledge. Tree based methods are highly interpretable and straightforward for visual understanding. However, they suffer from a high variance depending on the dataset used to train the model[26].

Gradient boosting is a new algorithm based on the construction of an ensemble of models[7]. It builds multiple models and combines their results to increase performance. GB also provides a rate of importance on the model for the different parameters that compounds the input, which have demonstrated a high local stability and consistency, opposite to other inconsistent techniques such as random forest and other machine learning methods[27]. Moreover, GB has an explainer extension containing a method to clarify why each example is classified in each group.

GB analysis applied to our data revealed that average RNFL thickness is the critical parameter to discriminate between early glaucoma and normal eyes, with a relative importance that is three times higher than the next most important parameter. Average thickness values display a stair-like graphic when plotted against its impact on the classification model, with steps at about 80 and 90 μm, indicating that inside the segments of <80 μm, 80 to 90 μm and >90 μm the influence on the classification model remains stable.

RNFL thickness at 7 clock-hour is the second most important parameter in our classifier. When plotted against its degree of impact on the model a flat horizontal graph is obtained, with no significant step, representing a linear relation between 7 clock-hour thickness and impact on the result of the model.

Using a decision tree method, Huang found that the inferior quadrant was the parameter in RNFL with the best classification importance in glaucoma eyes. Using univariate analysis, this article revealed average, inferior quadrant and 6 and 5 clock-hour thicknesses as the parameters with the best AUC[28]. In Medeiros et al[29] the highest discrimination power of individual OCT parameters was from the inferior quadrant thickness. Baskaran et al[30] reports that average and inferior quadrant RNFL thicknesses discriminate glaucoma better than the other parameters. Some authors describe that combining

OCT parameters by means of a multivariable predictive model outperforms univariable models in terms of AUC and classification rates[31-32]. Multiple studies combine several OCT features in a single parameter which contain information of its components, in order to obtain better classification models[33-34]. Nonetheless, this combination makes interpretation of the model completely unreachable. All these univariable and multivariable studies find disparate results regarding the parameter with the greatest discrimination power, though the general consensus is that average thickness and inferior hemisphere segments have the best AUC, similar to the importance rate we describe in our model.

Our classification tree constructed including the best performing discriminating parameters reveals that average thickness remains as the key parameter. Eyes with an average thickness <82 µm are mostly predicted as glaucoma (probability: 98%). If average thickness >90 µm and RNFL at 7 clock-hour >99 µm, the eye is classified as normal (predicted probability: 95%). These cut-off points for average thickness values are closely related to those found in our GB model.

Using ten-fold cross validation, the calculated overall accuracy of the model was 89%, with an AUC of 0.953, which is better than AUC for any univariable model. Specificity of the model is 0.89 for a sensibility of 0.88.

Classifying a glaucoma eye as normal is a much more serious error than labeling a normal eye as glaucoma, since undiagnosed glaucoma can lead to irreversible visual loss. This consideration was included in our tree through a loss matrix, in order to minimize the false negative rate. This modified tree revealed average thickness as the principal parameter for classification. Eyes with an average thickness <80 µm are all glaucoma eyes. Average thickness ≥80 µm with inferior quadrant thickness ≥92 µm suggests a highly probable normal eye (predicted probability: 76%).

The estimated accuracy for the weighted model was 84% of correctly classified eyes, which is considered a good classifier. Moreover, the lower accuracy is balanced with a minimum false negative rate. AUC of the model was 0.85, which is lower than AUC in the non-weight model, but specificity increased to 0.956.

Limitations to our study include the lack of a test group. Instead, we have made use of cross validation as a verified re-sampling method to evaluate the performing of our model. Overfitting has been controlled by means of outliers detection, selecting those parameters with the best relative importance in GB model and pruning of the tree. We have only included eyes with early damage and IOP>21 mm Hg. Therefore, our results may not be applicable to more advanced glaucoma eyes or normal-tension glaucoma.

As a conclusion, GB methods based on tree structures selected RNFL average thickness as the parameter with the greatest power to distinguish early glaucoma eyes. Thickness at the 7 clock-hour is the second most important parameter for classification. The fact that our classifier depends only on RNFL thickness favors its use in patients with unreliable visual field test, expanding the number of cases where our model can be valuable.

Although the number of subjects included in the present study is limited, we strongly believe that the development of an interconnected digital data warehouse coming from the electronic devices of multiple ophthalmology institutions will offer the foundations for future decision and classification models if adequate interpretable machine learning methods are applied.

REFERENCES

1 Lucy KA, Wollstein G. Structural and functional evaluations for the early detection of glaucoma. *Expert Rev Ophthalmol* 2016;11(5):367-376.

2 Tatham AJ, Medeiros FA, Zangwill LM, Weinreb RN. Strategies to improve early diagnosis in glaucoma. *Prog Brain Res* 2015;221:103-133.

3 Kansal V, Armstrong JJ, Pintwala R, Hutnik C. Optical coherence tomography for glaucoma diagnosis: an evidence based meta-analysis. *PLoS One* 2018;13(1):e0190621.

4 Doshi-Velez F, Kim B. Towards A rigorous science of interpretable machine learning. 2017: arXiv:1702.08608[stat.ML]. https://arxiv.org/abs/1702.08608

5 Hodapp E, Parrish IIRK, Anderson DR. *Clinical decisions in glaucoma.* St. Louis: CV Mosby comp. 1993.

6 Kaliyaperumal S, Kuppusamy M. Clustering Algorithm Based Outlier Detection in Data Mining. *Int J Comput Appl* 2016;4(6):9-19.

7 Chen TQ, Guestrin C. XGBoost: a scalable tree boosting system. 2016: arXiv:1603.02754[cs.LG]. https://arxiv.org/abs/1603.02754

8 Mansberger SL, Demirel S. Early detection of glaucomatous visual field loss: why, what, where, and how. *Ophthalmol Clin North Am* 2005;18(3):365-373,v-vi.

9 Rao HL, Yadav RK, Begum VU, Addepalli UK, Choudhari NS, Senthil S, Garudadri CS. Role of visual field reliability indices in ruling out glaucoma. *JAMA Ophthalmol* 2015;133(1):40.

10 Leung CKS, Cheung CYL, Weinreb RN, Qiu QL, Liu S, Li HT, Xu GH, Fan N, Huang L, Pang CP, Lam DSC. Retinal nerve fiber layer imaging with spectral-domain optical coherence tomography: a variability and diagnostic performance study. *Ophthalmology* 2009;116(7):1257-1263.e2.

11 Park SB, Sung KR, Kang SY, Kim KR, Kook MS. Comparison of glaucoma diagnostic capabilities of Cirrus HD and Stratus optical coherence tomography. *Arch Ophthalmol* 2009;127(12):1603-1609.

12 Hwang YH, Kim YY. Glaucoma diagnostic ability of quadrant and clock-hour neuroretinal rim assessment using cirrus HD optical coherence tomography. *Invest Ophthalmol Vis Sci* 2012;53(4):2226-2234.

13 Wollstein G, Schuman JS, Price LL, Aydin A, Beaton SA, Stark PC, Fujimoto JG, Ishikawa H. Optical coherence tomography (OCT) macular and peripapillary retinal nerve fiber layer measurements and automated visual fields. *Am J Ophthalmol* 2004;138(2):218-225.

14 Mwanza JC, Oakley JD, Budenz DL, Anderson DR, Cirrus Optical Coherence Tomography Normative Database Study Group. Ability of cirrus HD-OCT optic nerve head parameters to discriminate normal from glaucomatous eyes. *Ophthalmology* 2011;118(2):241-248.e1.

15 Elbendary AM, Abd El-Latef MH, Elsorogy HI, Enaam KM. Diagnostic accuracy of ganglion cell complex substructures in different stages of primary open-angle glaucoma. *Can J Ophthalmol* 2017;52(4):355-360.

16 Barua N, Sitaraman C, Goel S, Chakraborti C, Mukherjee S, Parashar H. Comparison of diagnostic capability of macular ganglion cell complex and retinal nerve fiber layer among primary open angle glaucoma, ocular hypertension, and normal population using Fourier-domain optical coherence tomography and determining their functional correlation in Indian population. *Indian J Ophthalmol* 2016;64(4):296-302.

17 Vidinova CN, Gouguchkova PT, Vidinov KN. Ganglion cell complex map for detecting early damage in high tension and normal tension glaucoma. *Klin Monbl Augenheilkd* 2016;233(1):72-78.

18 Bojikian KD, Chen CL, Wen JC, Zhang QQ, Xin C, Gupta D, Mudumbai RC, Johnstone MA, Wang RK, Chen PP. Optic disc perfusion in primary open angle and normal tension glaucoma eyes using optical coherence tomography-based microangiography. *PLoS One* 2016;11(5):e0154691.

19 Hothorn T, Lausen B. Bagging tree classifiers for laser scanning images: a data- and simulation-based strategy. *Artif Intell Med* 2003;27(1):65-79.

20 Quintana JM, Arostegui I, Alberdi T, Escobar A, Perea E, Navarro G, Elizalde B, Andradas E. Decision trees for indication of cataract surgery based on changes in visual acuity. *Ophthalmology* 2010;117(8):1471-1478.e3.

21 Sacchetti M, Lambiase A, Mantelli F, Deligianni V, Leonardi A, Bonini S. Tailored approach to the treatment of vernal keratoconjunctivitis. *Ophthalmology* 2010;117(7):1294-1299.

22 Gulshan V, Peng L, Coram M, Stumpe MC, Wu D, Narayanaswamy A, Venugopalan S, Widner K, Madams T, Cuadros J, Kim R, Raman R, Nelson PC, Mega JL, Webster DR. Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. *JAMA* 2016;316(22):2402.

23 Rohm M, Tresp V, Müller M, Kern C, Manakov I, Weiss M, Sim DA, Priglinger S, Keane PA, Kortuem K. Predicting visual acuity by using machine learning in patients treated for neovascular age-related macular degeneration. *Ophthalmology* 2018;125(7):1028-1036.

24 Sugimoto K, Murata H, Hirasawa H, Aihara M, Mayama C, Asaoka R. Cross-sectional study: Does combining optical coherence tomography measurements using the 'Random Forest' decision tree classifier improve the prediction of the presence of perimetric deterioration in glaucoma suspects? *BMJ Open* 2013;3(10):e003114.

25 Zhang K, Liu XY, Jiang JW, Li WT, Wang S, Liu L, Zhou XJ, Wang LM. Prediction of postoperative complications of pediatric cataract patients using data mining. *J Transl Med* 2019;17:2.

26 James G, Witten D, Hastie T, Tibshirani R. *An introduction to statistical learning*. New York, NY: Springer New York, 2013.

27 Lundberg SM, Lee SI. Consistent feature attribution for tree ensembles. 2017: arXiv:1706.06060[cs.AI]. https://arxiv.org/abs/1706.06060

28 Huang ML, Chen HY, Lin JC. Rule extraction for glaucoma detection with summary data from StratusOCT. *Invest Ophthalmol Vis Sci* 2007;48(1):244-250.

29 Medeiros FA, Zangwill LM, Bowd C, Vessani RM, Susanna R Jr, Weinreb RN. Evaluation of retinal nerve fiber layer, optic nerve head, and macular thickness measurements for glaucoma detection using optical coherence tomography. *Am J Ophthalmol* 2005;139(1):44-55.

30 Baskaran M, Ong EL, Li JL, Cheung CY, Chen D, Perera SA, Ho CL, Zheng YF, Aung T. Classification algorithms enhance the discrimination of glaucoma from normal eyes using high-definition optical coherence tomography. *Invest Ophthalmol Vis Sci* 2012;53(4):2314-2320.

31 Mwanza JC, Warren JL, Budenz DL, Ganglion Cell Analysis Study Group. Combining spectral domain optical coherence tomography structural parameters for the diagnosis of glaucoma with early visual field loss. *Invest Ophthalmol Vis Sci* 2013;54(13):8393-8400.

32 Mwanza JC, Lee G, Budenz DL, Warren JL, Wall M, Artes PH, Callan TM, Flanagan JG. Validation of the UNC OCT index for the diagnosis of early glaucoma. *Trans Vis Sci Tech* 2018;7(2):16.

33 Kim SJ, Cho KJ, Oh S. Development of machine learning models for diagnosis of glaucoma. *PLoS One* 2017;12(5):e0177726.

34 Bizios D, Heijl A, Hougaard JL, Bengtsson B. Machine learning classifiers for glaucoma diagnosis based on classification of retinal nerve fibre layer thickness parameters measured by Stratus OCT. *Acta Ophthalmol* 2010;88(1):44-52.