Artificial intelligence and glaucoma progression

Ricardo Yuji Abe¹², Felipe A. Medeiros³, Vital Paulino Costa¹

¹Department of Ophthalmology, University of Campinas - UNICAMP, Campinas, Brazil; ²Hospital Oftalmológico de Brasília, Brasília, Distrito Federal, Brazil; ³Department of Ophthalmology, Duke Eye Center, Duke University, Durham, NC, USA

Abstract

Detection of progression in glaucoma is crucial to avoid visual impairment and blindness. Throughout the clinical course of the disease, glaucoma patients can present very different trajectories, as some patients may remain stable using single eye drops whereas other patients may require surgical procedures to control the disease. Thus, the decision of intensifying a treatment by adding new eyedrops or performing a glaucoma surgery need to rely on precise data of true progression of the disease. In addition, assessing the velocity of progression can help to identify rapid progressors that are more prone to develop functional impairment. In clinical practice, we use both structural (retinography and optical coherence tomography) and functional (visual field) measurements, along with clinic-demographical data to evaluate if the patient is progressing. However, in some patients the correlation between structural and functional exams makes the detection of progression a challenge. Currently we are facing a growing use of artificial intelligence in medicine with the application of complex algorithms such as deep learning models. In this review, we summarize the findings from recent studies that investigated the use of artificial intelligence in detecting glaucoma progression.

Keywords: artificial intelligence, glaucoma progression, optical coherence tomography, retinography, visual field

Correspondence: Ricardo Yuji Abe, MD, PhD, Department of Ophthalmology, University of Campinas - UNICAMP, Cidade Universitária Zeferino Vaz - Barão Geraldo, Campinas - SP, 13083-970, Brazil.
E-mail: ricardoabe85@yahoo.com.br
1. Introduction

Glaucoma is the leading cause of irreversible blindness worldwide and it is expected to affect over 118.5 million people in 2040. Patients with glaucoma can have visual impairment if the disease is not properly treated and followed, leading to worsening quality of life. Avoiding the progression of the disease requires compliance to treatment and adequate follow-up with regular consultations. In addition, procedures such as lasers and surgeries may be needed according to the severity of the disease and the rate of progression. Unfortunately, the detection of glaucoma progression can be challenging due to intrinsic variability of the devices commonly used in clinical practice, such as standard automated perimetry (SAP), retinography, and optical coherence tomography (OCT). This variability can delay or even hamper the detection of progression in some cases, increasing the risk of the patient develop functional loss and visual impairment.

In clinic, when evaluating a patient with glaucoma, we must incorporate all clinical and demographic data, along with intraocular pressure (IOP) measurements, visual field information, retinographies, and OCT scans to decide if the disease is worsening. This task can be misleading, and the patient may be overtreated due to a false diagnosis of progression or even worse, be considered as stable when there is a true progression. Therefore, the use of artificial intelligence (AI) to help the clinicians in identifying true progressors has been increasing in the past years. Since the 1990s, several authors started to apply AI in glaucoma for both diagnosis and progression.

2. Artificial intelligence and glaucoma progression

The pivotal studies performed in the 1990s have focused on AI performance to discriminate glaucomatous patients or to improve the detection of glaucoma progression using visual field parameters. None of these studies incorporated data from retinography or clinical data from the patients. Several factors have made the use of AI in glaucoma more common in recent years. First, in contrast to retinography, OCT scans offer objective numeric measurements from different optical structures (optic nerve neural rim, retinal nerve fiber layer [RNFL], and macular thickness). In addition, the widespread use of OCT and the increasing affordability of OCT devices contributed to expand the data available for research around the world. Second, electronic medical charts of thousands of patients from different centers offer a huge amount of real-world clinical and demographical data, which are essential for the development of AI algorithms. Third, the evolution of AI with the application of deep learning have also contributed to improve its performance. In this review, we will focus on the result of recent studies that applied the use of AI with visual field, retinography, and OCT data to improve the detection of glaucoma progression.
In 2012, Goldbaum et al. tested the ability of a machine learning classifier (MLC) using a Bayesian independent component mixture model to identify progression of glaucomatous visual field defects. The performance of AI was similar to visual field index (VFI), mean deviation (MD), and guided progression analysis (GPA) in glaucoma suspects but had better performance than GPA in glaucoma patients. In 2014, Siamak et al. employed MLC to detect glaucoma progression using longitudinal data from OCT RNFL and SAP. The study compared different MLCs and investigated the performance of the combination of both structural and functional measurements in detecting progression. They included patients with early to moderate glaucoma and found that RNFL parameters alone offered a similar performance compared to other models using RNFL and SAP. In another study, Siamak et al. compared the performance of a Gaussian mixture model and expectation maximization with the commercially available techniques from SAP, such as VFI and GPA, for the detection of glaucoma progression. They also compared the current model with previously described unsupervised learning-based progression detection algorithms such as the Bayesian independent component analysis mixture model. They found that the Gaussian mixture model using expectation maximization performed better than the commercially available SAP progression detection method. In addition, progression detection based on changes in the Gaussian mixture model performed slightly better than other models, while being less computationally complex.

In 2015, Siamak et al. developed an approach for the detection of glaucoma progression using a framework to find a vector that is representative of the progression direction of the sample population. Further analysis of these longitudinal visual fields across the derived vector led to optimal disease progression detection. Compared to the other models mentioned in previous studies, this one had the advantage of requiring only longitudinal data for training, whereas the other models required both cross-sectional and longitudinal data. They found that progression detection using this framework performed slightly better than the previous models and was also more sensitive than SAP VFI and MD.

In 2018, Siamak et al. applied both MLC and statistical analysis to detect glaucomatous progression in series of SAP exams. They included SAP exams from 2,085 eyes of 1,214 subjects (normal and glaucoma suspects) and SAP exams from 133 eyes of 71 glaucoma patients. SAP exams were collected 10 times over 10 weeks. They found that MLC analyses were able to detect progressing eyes earlier than other methods (MD, region-wise, and point-wise analyses) with the advantage of detecting more slowly progressing eyes than other methods.

Lee et al. investigated the performance of different MLCs (extra-trees and random forest models) to predict glaucoma progression in myopic eyes with normal-tension glaucoma (NTG) in a cross-sectional study. They included 155 eyes from 155 myopic NTG patients with axial length higher than 24.00 mm between the ages of 20 and 40 years old. The extra-trees model achieved an area-under-ROC curves (AUC) of 0.881, higher than that of the random forest model (AUC of 0.811, $P = 0.010$).
The extra-trees model also outperformed all the clinical measurements for NTG progression, including average macular ganglion cell-inner plexiform layer (GCIPL) thickness and average RNFL thickness.

3. Recent studies

With the advance of AI algorithms, recent studies are now applying deep learning neural networks to analyze data from glaucoma. Medeiros et al. have developed an innovative approach called machine-to-machine (M2M) using objective measurements from OCT to train the deep learning algorithm rather than using human gradings. They used retinographies to train the algorithm to predict quantitative measurements (RNFL and neuroretinal rim) from OCT. The advantage of these predictions is that, in general, they are more accurate than human gradings, which can suffer from low reproducibility. In 2020, Medeiros et al. performed a retrospective cohort study to evaluate the ability of deep learning using RNFL thickness obtained from retinographies to detect glaucoma progression as measured by RNFL from OCT longitudinally. They included a total of 83,123 pairs of fundus photographs and OCT images collected during 21,232 visits from 8,831 eyes of 5,529 patients with glaucoma or glaucoma suspects. A significant correlation was found between change over time in predicted and observed RNFL thickness. The RNFL predictions showed an AUC of 0.86 to discriminate progressors from nonprogressors. This study was the first to apply deep learning to detect glaucoma progression using fundus photographs.

Siamak et al. developed an analytical pipeline including linear transformation, manifold learning, and unsupervised clustering to improve the detection of glaucoma progression and investigate the patterns of visual field loss. They applied a combination of linear and nonlinear statistical methods to evaluate both local patterns and global defects from SAP exams and created an explainable and clinician-friendly tool with multiple layers of glaucoma knowledge on a simple interpretable 2-dimensional map, which they named as dashboard. After building the dashboard, they were able to identify 32 nonoverlapping clusters. Each cluster on the dashboard corresponded to a particular global functional severity, an extent of visual field loss into different hemifields, and characteristic local patterns of visual field loss. The dashboard developed using AI included a large spectrum of visual field patterns that can aid the clinician to classify the severity of the disease and monitor glaucoma progression.

One of the limitations of OCT and SAP data to monitor glaucoma is that change may occur only after a significant loss of retinal ganglion cells. Therefore, novel methods to discover earlier biomarkers of disease have been reported. Within these approaches, the detection of apoptosing retinal cells (DARC) from Cordero et al. looks promising and is currently in a Phase 2 clinical trial. This technique
uses a molecular marker labelled annexin A5, which has a high affinity for cells undergoing stress and in the early stages of apoptosis. A previous study reported that the number of DARC positively stained cells seen in a retinal fluorescent image could be used to assess glaucoma disease activity.\(^\text{23,24}\) Recently, Normando et al. applied a convolutional neural network (CNN) to aid the detection of DARC spots. The study included images from 40 healthy controls and 20 glaucoma patients. The CNN-aided algorithm was trained and validated using manual counts from control subjects, and then tested on glaucoma eyes. They reported that the model had 97.0% accuracy, 91.1% sensitivity and 97.1% specificity to detect spots when compared to manual grading of controls. In addition to that, it demonstrated 85.7% sensitivity, 91.7% specificity with AUC of 0.89, and a significantly greater DARC count in glaucoma patients who later progressed, based on RNFL thickness measurements from OCT.

Nouri-Mahdavi et al. performed a prospective cohort study with moderate and advanced glaucoma patients to evaluate if the combination of baseline and longitudinal OCT (RNFL and macular ganglion cell-inner plexiform layer [GCIPL]) measurements and demographic data could predict visual field progression using both MLC and elastic net logistic regression analysis (ENR).\(^\text{25}\) They included a total of 104 eyes from 104 patients with ≥ 3 years of follow-up and ≥ 5 visual field examinations. They found that the best MLC predictors included baseline superior hemimacular GCIPL thickness and GCIPL change rates (AUC = 0.81). For ENR, rates of change of superotemporal RNFL sector and GCIPL change rates were the best predictors (AUC = 0.79). Thus, the use of AI could help to predict visual field progression based on baseline and longitudinal structural data from OCT measurements. This finding is especially important in patients with moderate and advanced damage, in which visual field loss could decrease the quality of life more drastically.\(^\text{26}\)

Hood et al. have previously demonstrated that using an individualized region of interest (ROI) approach can outperform instrument in-built defined global peripapillary RNFL thickness measurements to monitor glaucomatous progression.\(^\text{27}\) Bowd et al. have applied an unsupervised deep learning model to detect progression (defined by optic disc stereophotographs) from OCT cube scans using ROI approach and compared the results to RNFL thickness measurements derived from the same cube scans.\(^\text{28}\) They included a total of 44 progressing glaucoma eyes (confirmed by stereophotograph), 189 nonprogressing glaucoma eyes and 109 healthy eyes that were followed for ≥ 3 years with ≥ 4 visits using OCT. The sensitivity for detecting change in progressing eyes was greater for the deep learning models than for global RNFL thickness measurements.

In a retrospective cohort, Shuldiner et al. evaluated the ability of different MLCs to detect fast visual field progressors (MD reduction > 1 dB/year).\(^\text{29}\) They included a total of 175,786 SAP exams (22,925 initial ones) from 14,217 patients who completed 5 reliable visual fields. Among the MLCs, the support vector machine model (AUC 0.72) presented the highest accuracy to predict progression. Interestingly, models
### Table 1. Recent studies on artificial intelligence and glaucoma progression

<table>
<thead>
<tr>
<th>Author (year)</th>
<th>Purpose</th>
<th>Sample size</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Medeiros et al. (2019)</td>
<td>To train a DL algorithm from OCT RNFL data to quantify glaucomatous structural damage on optic disc photographs.</td>
<td>32,820 pairs of optic disc photographs and OCT RNFL scans from 2,312 eyes of 1,198 participants</td>
<td>AUC for discriminating glaucomatous from healthy eyes with the DL predictions was 0.944.</td>
</tr>
<tr>
<td>Thompson et al. (2019)</td>
<td>To train a DL algorithm from OCT BMO-MRW data to quantify glaucomatous structural damage on optic disc photographs.</td>
<td>9,282 pairs of optic disc photographs and OCT optic nerve head scans from 927 eyes of 490 subjects</td>
<td>AUC for discriminating glaucomatous from healthy eyes with the DL predictions was 0.945.</td>
</tr>
<tr>
<td>Medeiros et al. (2020)</td>
<td>To investigate whether predictions of RNFL thickness obtained from a DL model applied to fundus photographs can detect progressive glaucomatous changes over time.</td>
<td>33,466 pairs of fundus photographs and OCT images collected during 7,125 visits from 1,147 eyes of 717 patients</td>
<td>RNFL predictions showed an AUC of 0.86 to discriminate progressors from nonprogressors.</td>
</tr>
<tr>
<td>Yousefi et al. (2020)</td>
<td>To develop an AI dashboard to monitor glaucomatous functional loss.</td>
<td>13,231 VFs from 8,077 subjects were included to develop the AI dashboard. Longitudinal VFs from 287 eyes with glaucoma were used to validate the models</td>
<td>The specificity for detecting ‘likely nonprogression’ was 94% and the sensitivity for detecting ‘likely progression’ was 77%.</td>
</tr>
<tr>
<td>Normando et al. (2020)</td>
<td>To develop an automatic convolutional neural network-aided method of DARC spot detection to enable prediction of glaucoma progression.</td>
<td>40 healthy control and 20 glaucoma patients</td>
<td>85.7% sensitivity and 91.7% specificity with AUC of 0.89, using OCT RNFL measurements as reference standard for progression at 18 months.</td>
</tr>
<tr>
<td>Author (year)</td>
<td>Purpose</td>
<td>Sample size</td>
<td>Results</td>
</tr>
<tr>
<td>--------------</td>
<td>---------</td>
<td>-------------</td>
<td>---------</td>
</tr>
<tr>
<td>Nouri-Mahdavi et al. (2021)</td>
<td>To test the hypothesis that VF progression can be predicted from baseline and longitudinal OCT structural measurements.</td>
<td>104 eyes from 104 patients with ≥ 3 years of follow-up and ≥ 5 VF examinations</td>
<td>Machine learning predictors from baseline superior hemimacular GCIPL thickness and GCIPL change rates had AUC=0.81.</td>
</tr>
<tr>
<td>Bowd et al. (2021)</td>
<td>To compare change over time in eye-specific OCT RNFL regions of interest maps developed using unsupervised DL to RNFL thickness for the detection of glaucomatous progression.</td>
<td>44 progressing and 189 nonprogressing glaucoma eyes (by stereophotograph assessment) and 109 healthy eyes.</td>
<td>Sensitivity for detecting change in progressing eyes with DL regions of interest was 0.90 and specificity for detecting not likely progression in nonprogressing eyes was similar 0.92.</td>
</tr>
<tr>
<td>Shuldiner et al. (2021)</td>
<td>To assess whether ML algorithms can predict eyes that will undergo rapid glaucoma progression based on an initial VF exam</td>
<td>175,786 VFs (22,925 initial VFs) from 14,217 patients who completed ≥ 5 reliable VFs.</td>
<td>The support vector machine model (AUC 0.72) most accurately predicted rapid progression when trained on initial VF data.</td>
</tr>
<tr>
<td>Saeedi et al. (2021)</td>
<td>To develop and test machine learning classifiers for determining visual field progression.</td>
<td>90,713 visual fields from 13,156 eyes were included.</td>
<td>Machine learning classifiers accuracy ranged from 87% to 91% with sensitivity ranging from 0.83 to 0.88 and specificity from 0.92 to 0.96.</td>
</tr>
<tr>
<td>Dixit et al. (2021)</td>
<td>To assess the performance of a convolutional long short-term memory neural network for detecting glaucoma progression based on a longitudinal data set of merged VF and clinical data.</td>
<td>11,242 eyes with ≥ 5 reliable VFs and baseline clinical data (cup-to-disc ratio, central corneal thickness, and intraocular pressure)</td>
<td>The convolutional long short-term memory neural network demonstrated 91% to 93% accuracy. The model that was trained on both VF and clinical data presented an AUC ranging from 0.89-0.93.</td>
</tr>
</tbody>
</table>

DL: deep learning; OCT: optical coherence tomography; RNFL: retinal nerve fiber layer; AUC: area under receiver operating characteristic curve; BMO-MRW: Bruch’s membrane opening-minimum rim width; AI: artificial intelligence; VF: visual field; ML: machine learning; DARC: Detection-of-Apoptosing-Retinal-Cells; GCIPL: ganglion cell-inner plexiform layer
trained on data from the first two visual fields performed no better than models trained on the initial visual field alone. Thus, MLCs presented modest accuracy to predict visual field progression using a single SAP initial exam. However, the authors emphasize that incorporating additional clinical data to the current model could enhance its accuracy.

Saeedi et al. compared 6 different MLCs (logistic regression, random forest, extreme gradient boosting, support vector classifier, CNN, fully connected neural network) to detect visual field progression. In this study, a total of 90,713 visual fields from 13,156 eyes were included. Although they found that MLCs had modest performance, like conventional algorithms (linear regression of MD and VFI, Advanced Glaucoma Intervention Study and Collaborative Initial Glaucoma Treatment Study algorithms, and pointwise linear regression), the MLCs were less subject to class bias, were more balanced, and probably more applicable to a wider range of glaucoma patients with different severities of damage.

Dixit et al. evaluated glaucoma progression using a MLC algorithm trained over a large dataset containing visual field and clinical data. An innovative approach of this study was to identify if a combination of both functional (visual field) and clinical data could improve glaucoma progression detection using a CNN. This retrospective study included 672,123 SAP exams from 213,254 eyes and 350,437 samples of clinical data (cup-to-disc ratio, central corneal thickness, and IOP). They found an accuracy of 91–93% for the CNN when assessing both visual field and clinical data, suggesting that it is possible to improve the ability to detect glaucoma progression when combining clinical data to visual field parameters. A summary of the studies mentioned above can be found in Table 1.

4. Limitations and future perspectives

Monitoring for structural or functional changes is crucial in the management of both glaucoma suspects and those with confirmed disease. Prompt detection of progression can minimize the risks of the patient developing visual impairment, especially in cases with advanced disease or in cases in which central visual field is threatened. The application of several different AI approaches has provided satisfactory accuracy in the detection of progression. Currently, in clinical practice, we combine both structural (fundus photographs and OCT) and functional (SAP) data as well as all clinical (IOP and visual acuity measurements) and demographic data available to define whether the patient is progressing or is at risk for progression. However, in this review we observed that only a few papers investigated the combination of all parameters (structure, function, and clinical data) in algorithms created to detect glaucoma progression. Furthermore, despite the growing amount of research using AI in glaucoma, none of the previously mentioned algorithms are currently available for routine use in clinical practice. One of the barriers to this
adoption is the complexity of the AI algorithms, which is not a familiar language for most of the clinicians.\textsuperscript{32} The incorporation of these new AI developments requires the interaction with manufacturers to create in-built software that can be user-friendly for the clinician.\textsuperscript{33} Currently, we have a huge amount of medical data available in clinics and hospitals that can be shared with increased velocity through internet cloud systems. The use of more representative datasets with multiethnic populations from real-world scenarios could help the development of AI algorithms designed for the detection of glaucoma progression.\textsuperscript{34} However, to do so, collaborative networking for data collection, infrastructure capacity to storage, trained specialists to process and analyze the data and a broader discussion over regulatory laws and cybersecurity are needed.\textsuperscript{32,35}

**Declarations**

**Ethics approval and consent to participate**
Not required.

**Competing interests**
None to declare.

**Funding**
None to declare.

**Acknowledgements**
None to declare.

**References**


3. Medeiros FA, Gracitelli CP, Boer ER, Weinreb RN, Zangwill LM, Rosen PN. Longitudinal changes in quality of life and rates of progressive visual field loss in glaucoma patients. Ophthalmology. Feb 2015;122(2):293-301. [https://doi.org/10.1016/j.ophtha.2014.08.014](https://doi.org/10.1016/j.ophtha.2014.08.014)


