A Case for The Use of Artificial Intelligence in Glaucoma Assessment

Joel S. Schuman, MD,1-4 Maria De Los Angeles Ramos Cadena, MD,1 Rebecca McGee, MD,1 Lama A. Al-Aswad, MD, MPH1-5 and Felipe A. Medeiros, MD, PhD6,7 for the Collaborative Community for Ophthalmic Imaging Executive Committee and Glaucoma Workgroup


1 Department of Ophthalmology, NYU Langone Health, NYU Grossman School of Medicine, New York, NY, USA.
2 Departments of Biomedical Engineering and Electrical and Computer Engineering, New York University Tandon School of Engineering, Brooklyn, NY, USA
3 Center for Neural Science, NYU, New York, NY, USA
4 Neuroscience Institute, NYU Langone Health, New York, NY, USA.
5 Department of Population Health, NYU Langone Health, NYU Grossman School of Medicine, New York, NY, USA.
6 Department of Ophthalmology, Duke University School of Medicine, Durham, NC, USA
7 Department of Electrical and Computer Engineering, Pratt School of Engineering, Duke University, Durham, NC, USA

Financial Support: Funding from the National Institutes of Health (Bethesda, MD, USA) R01-EY013178. An unrestricted grant from Research to Prevent Blindness (New York, NY) to the Department of Ophthalmology, NYU Langone Health, NYU Grossman School of Medicine, New York, NY.

Corresponding Author: Joel S. Schuman, MD
222 East 41st Street, Suite 468, New York, NY 10017
Tel: 929-455-5030, Fax: 929-455-5553
joel.schuman@nyu.edu

FDA participates as a member of the Collaborative Community on Ophthalmic Imaging. This article reflects the views of the authors and should not be construed to represent FDA’s views or policies.

Disclosures:

Joel S. Schuman, MD
Aerie Pharmaceuticals, Inc.: Consultant/Advisor, Equity Owner
BrightFocus Foundation: Grant Support
Boehringer Ingelheim: Consultant/Advisor
Carl Zeiss Meditec: Patents/Royalty/Consultant/Advisor
Massachusetts Eye and Ear Infirmary and Massachusetts Institute of Technology: Intellectual Property
National Eye Institute: Grant Support
New York University: Intellectual Property
Ocugenix: Equity Owner, Patents/Royalty
Ocular Therapeutix, Inc.: Consultant/Advisor, Equity Owner
Opticent: Consultant/Advisor, Equity Owner
Perfuse, Inc.: Consultant/Advisor
Regeneron, Inc.: Consultant/Advisor
SLACK Incorporated: Consultant/Advisor
Tufts University: Intellectual property
University of Pittsburgh: Intellectual property

Lama A. Al-Aswad, MD, MPH
Aerie Pharmaceuticals, Inc.: Consultant/Advisor
GlobeChek: Equity Owner
AI Optics: Advisor
New World Medical Inc: Grant Support
Save Vision Foundation: Grant Support
Topcon Medical Systems Inc.: Research support and consultant
Verily: Consultant
Zeiss: Adviser

Michael D. Abramoff, MD, PhD
Digital Diagnostics: ICP
Novago AG, ICP

Bhavna J. Antony, PhD
IBM Research, Employee

Michael Boland, MD, PhD
Carl Zeiss Meditec – consulting

Michael Chiang, MD
NIH: Grant support
NSF: Grant support
Genentech: Grant support
Novartis: Consultant
InTeleretina, LLC: Equity owner

Jeffrey L Goldberg, MD, PhD
Carl Zeiss Meditec – consulting

Naama Hammel, MD
Google Health, Employee

Felipe A. Medeiros, MD, PhD
Aerie Pharmaceuticals (Consultant)
Schuman, et al – AI in Glaucoma

Allergan (Consultant, Financial support)
Annexon (Consultant)
Biogen (Consultant)
Carl ZeissMeditec (Consultant, Financial support)
Galimedix (Consultant)
Google Inc (Financial support)
Heidelberg Engineering (Financial support)
IDx (Consultant)
nGoggle Inc (Patent)
Novartis (Financial support)
Stealth Biotherapeutics (Consultant)
Reichert (Consultant, Financial support)

Louis R Pasquale, MD
Consultant for each:
Eyenovia
Syke Biosciences
Twenty twenty

Harry A. Quigley, MD
Consultant for Sensimed, Intense, IDx, Gore, Kali Care, Equinox; Research Support: Heidelberg, Topcon.

Remo Susanna, MD
Adapt: Patents/Royalty

Jayme Vianna, MD
EadieTech: Consulting

Linda Zangwill, PhD
Grant support: National Eye Institute
Equipment and research support: Carl Zeiss Meditec Inc., Heidelberg Engineering GmbH, Optovue Inc., Topcon Medical Systems Inc.
Patent licensed to: Zeiss Meditec
Abstract

We hypothesize that artificial intelligence applied to relevant clinical testing in glaucoma has the potential to enhance the ability to detect glaucoma. This premise was discussed at the recent Collaborative Community for Ophthalmic Imaging meeting, “The Future of Artificial Intelligence- Enabled Ophthalmic Image Interpretation: Accelerating Innovation and Implementation Pathways,” held virtually September 3-4, 2020.

The Collaborative Community in Ophthalmic Imaging (CCOI) is an independent self-governing consortium of stakeholders with broad international representation from academic institutions, government agencies, and the private sector whose mission is to act as a forum for the purpose of helping speed innovation in healthcare technology. It was one of the first two such organizations officially designated by the FDA in September 2019 in response to their announcement of the collaborative community program as a strategic priority for 2018-2020. Further information on the CCOI can be found online at their website (https://www.cc-oi.org/about).

Artificial intelligence for glaucoma diagnosis would have high utility globally, as access to care is limited in many parts of the world and half of all people with glaucoma are unaware of their illness. The application of artificial intelligence technology to glaucoma diagnosis has the potential to broadly increase access to care worldwide, in essence flattening the Earth by providing expert level evaluation to individuals even in the most remote regions of the planet.
Background

Tools to visualize ocular anatomy and analyze pathology have been paramount in the development of the field of ophthalmology, dating back to the development of the ophthalmoscope more than 170 years ago.

The Ocular Hypertension Treatment Study (OHTS) Confocal Scanning Laser Ophthalmoscopy (CSLO) Ancillary Study demonstrated that baseline measurements of the neuroretinal rim and optic cup, either alone or in combination with age, intraocular pressure, central corneal thickness, and visual field (VF) pattern standard deviation (PSD), could be used to predict the development of primary open-angle glaucoma (POAG).[1] In addition, eyes that progressed to POAG showed a greater rate of change in CSLO ONH parameters when compared to those that did not develop POAG.[2] Similarly, longitudinal studies have demonstrated that the CSLO was better able to identify eyes that demonstrated glaucoma progression compared to traditional measures.[3, 4]

In the 1990s, optical coherence tomography (OCT) was developed as a means for non-invasive, objective, rapid, reproducible, and quantitative sampling of tissue. One initial use of OCT was for evaluation of anterior segment structures.[5] With time, the OCT became an invaluable tool in evaluating numerous ocular posterior and anterior segment structures including the retinal nerve fiber layer (RNFL), optic nerve head (ONH), retina, retinal and optic nerve vasculature, anterior chamber angle, and the aqueous outflow system. OCT measurements, particularly RNFL thickness, macular thickness and rim-based parameters such as Minimum Rim Width (MRW), have high diagnostic sensitivity and specificity to differentiate glaucomatous eyes from healthy eyes.[6-12] At this time the OCT measurements are frequently used in conjunction with measures of optic nerve function and clinical examination to assist
ophthalmologists in the diagnosis of glaucoma. From a regulatory perspective, many of the uses of OCT mentioned in this paragraph represent the practice of medicine and not legally marketed devices or modalities.

In order to expedite and facilitate the development of ophthalmic imaging, the Collaborative Community for Ophthalmic Imaging (CCOI) was formed.[13] The CCOI executive committee then had to decide on the most pressing needs, and created workgroups to address these areas. Glaucoma Imaging was one of the first workgroups formed. What follows is a direct outcome of the CCOI meeting, “The Future of Artificial Intelligence- Enabled Ophthalmic Image Interpretation: Accelerating Innovation and Implementation Pathways,” held virtually September 3-4, 2020.

**Artificial Intelligence**

The term artificial intelligence (AI) refers to the concept of programming computer systems to perform tasks to mimic human cognitive capabilities – such as understanding language, recognizing objects and sounds, learning, and problem solving – by using logic, decision trees, machine learning, or deep learning.[14] Subcategories of AI include machine learning (ML), with deep learning as a particular approach to ML.[15, 16] ML refers to the ability of a machine to learn without needing to be explicitly programmed and can be done with a supervised or unsupervised approach. For example, with supervised learning with classification, labeled examples are used as input into algorithms which then are “trained” to classify those examples into specific categories. A subtype of machine learning, deep learning is a type of AI where a machine has the capability of learning using multilayer neural networks, modeled after the cerebral cortex. Because of the complexity of the networks used, deep learning
can allow for more nuanced decision making but requires more data to train networks. Deep learning appears to have great promise in the field of glaucoma. Medical applications in AI range from detecting disease in input images and signals in electronic health records to analyzing large volumes of data that can be categorized faster than possible with manual review.[17] These advanced analytics techniques typically analyze large and varied datasets that cannot normally be analyzed by humans without specialized software tools, and often discover new patterns in data. While AI can be used for detailed analysis, in some cases it can be difficult to parse out how the algorithms actually arrive at decisions.

**The Potential Role of AI in Glaucoma**

A large proportion of glaucoma remains undiagnosed. It can be a silent disease and often is not diagnosed until the disease is advanced. Fortunately, treatments are available that can prevent or slow disease progression and vision loss once glaucoma has been identified. Glaucoma has relatively low prevalence (~2% over 40 years of age), making it difficult to implement screening in large populations.[18] Currently, there does not exist a screening algorithm with high enough sensitivity to meaningfully identify patients with different stages of disease and determine the next appropriate stage of evaluation. The diagnosis of glaucoma at this time requires individual examination by specialists, adding an additional obstacle in identifying glaucoma in a population. One solution may be to develop and implement more sophisticated automated systems to detect glaucoma. The prevalence of glaucoma is highest in the elderly, which is a rapidly growing segment of the population. At the same time, it is predicted that ophthalmologists will be in ever shorter supply as the number of ophthalmologists trained each year is not increasing. Even if glaucoma is diagnosed, there are sub-populations that remain underserved in terms of medical care; this is an example of issues related to equity in healthcare.
Finally, healthcare costs continue to grow and strain personal, industry, and government budgets.[16] Many of these issues may be addressed, at least in part, by AI, however, as Abramoff, et al. wrote, “Inappropriate bias, increase in health disparities, and thus decreased equity can exist across the entire AI pipeline, including in the choice of intended use of the AI, its design, its validability, its validation and the choice of reference standards, as well as how, and where, it is implemented.”[19]

Patient Perspective

It is also important to take into account the patient’s perspective on technology that is being implemented with data from their eyes.[13] Using fundus imaging in community settings is an easy and relatively quick method to obtain information about their likelihood of having glaucoma. Patients may be reluctant to have their eyes imaged, due to a variety of concerns, including privacy and security. Patients also may not want to take the time to undergo VF testing and may be frustrated by the exam itself. Furthermore, in order for AI to be successful, there must be large databases with images from many patients’ eyes. Screening would require patient consent.

Applications of AI in Ophthalmology

Most published studies of AI applied to OCT imaging have so far focused on diseases impacting the posterior segment of the eye and have relied on interpretation of images.[17] While as of the writing of this manuscript the only currently Food and Drug Administration (FDA) approved AI ophthalmic devices are limited to Diabetic Retinopathy, there are have been many studies published in other areas of ophthalmology. AI has been shown to have a potential role in detecting disease and identifying worsening of disease in a diverse array of retinal pathologies. For example, AI has been applied as a method to distinguish healthy eyes from eyes
with diabetic macular edema (DME) and distinguishing eyes with DME from eyes with age-related macular degeneration.[20-22] Additionally, OCT has been analyzed by AI to differentiate etiologies of disease. One study demonstrated that AI models can identify subretinal fluid that has accumulated from DME.[23] AI applied to OCT technology has also shown utility in detecting RNFL loss in multiple sclerosis and changes in photoreceptors in choroideremia and retinitis pigmentosa.[24] Moreover, AI algorithms have utilized OCT to link structural change to functional changes in glaucoma. In one study, retinal ganglion cell axonal complex optimized approach, which uses the ganglion cell – inner plexiform layer (GCIPL) thickness and RNFL thickness, had high predictive value in determining VF thresholds using OCT image, with an average correlation of 0.74. [25]

**Applications of AI in Glaucoma**

For AI to be effective in glaucoma, it must be capable of assessing the likelihood of the presence or absence of glaucoma and/or detecting the progression of glaucoma.[17] If used to classify patients, AI must be able to distinguish between glaucoma and non-glaucomatous optic neuropathy.[26] To train these types of classification AI algorithms, we need a definition of glaucoma that is objective, validated, and standardized, to serve as a ground truth that can be applied to the examples used to train the system.[27]

AI has a potential role in glaucoma screening to help identify affected patients more efficiently than human experts alone. It would be optimal if eyes with early stage glaucoma or that are at higher risk of developing glaucoma could be flagged by AI,[17] but this is not the initial priority. Instead, AI should be focused on automated identification of moderate glaucoma, when patients begin to experience functional changes or visual field abnormalities and are more likely to experience visual disability over their lifetime. Eventually, the scope of AI in glaucoma
assessment will broaden, but first steps will necessarily be more limited than later applications. This begs the question, however, as to whether society is ready at all for legally marketed AI for glaucoma.

Glaucoma can remain asymptomatic until moderate to severe stages of disease, which can make it even more difficult to diagnose. Studies have demonstrated that the number of undiagnosed cases of glaucoma may range from 60% to 92% in developed countries throughout the world.[28-31] Heijl et al. (2013) found that when screening for glaucoma in Sweden, one-third of subjects with previously undiagnosed glaucoma already had advanced glaucomatous disease in at least one eye.[32] Glaucoma is in theory an ideal disease to use in screening, as it is a chronic disease that progresses slowly and if treated early enough, irreversible visual decline can be prevented. If glaucoma is diagnosed at more advanced stages, it is more difficult to treat and eyes at such stage of disease are more likely to progress to blindness.[33]

In the past, glaucoma screening was not considered to be cost-effective.[34, 35] However, numerous studies have demonstrated that if screening is focused on patients that are high risk for glaucoma, including patients at 68 years of age or older, with family history of glaucoma, or of certain ethnicities, screening becomes more beneficial and cost-effective.[36-41] In addition, the positive predictive value of screening tests increases when the prevalence increases, therefore screening is likely more effective when focused on these high-risk individuals.

On an individual level, identifying glaucoma at an earlier stage may prevent patients from being exposed to the risk and cost of drug escalation and procedures, and may prevent patients from progressing to blindness. Studies have demonstrated that the financial burden of glaucoma increases at later stages of disease.[42, 43] The study by Varma and colleagues (2011)
demonstrated that at patients at early stages of disease incur direct costs through medications and office visits, however at advanced stages, there are additional indirect costs, which are particularly burdensome on health care resources and expensive on a global level.[43]

By flagging eyes with moderate glaucomatous damage, which are more likely to progress, AI will be more cost-effective. Detecting glaucoma at moderate disease rather than early disease would not only still allow for treatment to begin in a timely manner to prevent any additional vision loss and further decline in quality of life, but also would avoid overwhelming the healthcare system. Asaoka and colleagues (2014) showed an area under the receiver operating characteristic curve (AUROC) of 0.79 for pre-perimetric glaucoma using a combination of inputs including total deviation plot, PSD, and MD.[44] This level of performance is not adequate to implement in a screening setting at this time, as it would be costly and increase patient care visits due to a high false positive rate. In contrast, targeting worse-than-mild disease is likely to result in more usable performance by AI systems as the task they are being asked to do is fundamentally “easier” than identifying even the earliest possible stage of disease.

**Parameters for diagnosis: Structure**

Structural changes are followed closely in the diagnosis and progression of glaucoma. The ONH cup-to-disc ratio, RNFL thickness, and GCIPL thickness are fundamental components of the glaucoma work up. A hallmark of glaucoma is cupping of the ONH with progressive neuroretinal rim loss of the optic nerve. However, the ONH cup-to-disc ratio is a more subjective measurement and has high interrater variability. A quantitative structural change that is also monitored in glaucoma consists of OCT imaging of the RNFL thickness and the macular GCIPL thickness.
Important options for AI-assisted diagnosis of glaucoma include fundus photography as well as quantitative OCT image analysis as glaucomatous structural damage has been shown to precede VF loss.[45] Fundus photography is a relatively inexpensive method to screen for eye disease that has already been utilized to assist in detecting retinal pathology.[46] There are also portable fundus cameras that are relatively cheap and can be used to image patient’s eyes in community settings and at primary care offices.

Parameters that are imaged by OCT include the RNFL, ganglion cell / inner-plexiform layer (GCIPL), and ONH. The superior and inferior temporal peripapillary RNFL are commonly associated with glaucoma damage.[16] Using the Diagnostic Innovations in Glaucoma Study, Zangwill and colleagues (2004) were able to show how structural variables can be used in the detection of glaucomatous damage. In one study, it was demonstrated that HRT measurements were able to identify glaucomatous eyes when focusing along the disc margin better than measurements obtained from the peripapillary region.[47] In addition, the Vizzi et al. (2009) used Spectral domain-optical coherence tomographs (SD-OCT) to detect localized RNFL defects in eyes with glaucoma.[48] In another study, Belghith and researchers (2016) demonstrated that even within eyes with advanced glaucomatous disease, structural changes can be detected by analyzing the macular GCIPL, the minimum rim width, and the circumpapillary RNFL.[49] Furthermore, Hammel et al. (2016) utilized structural data, including neuroretinal rim imaged by confocal laser scanning ophthalmoscopy to study the rate and pattern of glaucomatous diseases.[50]

It is also possible that OCT can identify glaucoma using raw image data rather than relying on already defined structural parameters dependent on image segmentation, such as RNFL thickness, ONH measures, ganglion cell complex (GCC) or GCIPL.[86-89]
Parameters for diagnosis: Function

Quantitative measurement of the VF is another parameter that is used for the diagnosis of glaucoma. Glaucomatous neuropathy leads to functional changes that are represented by VF loss. The gold standard for analyzing VF function is Standard Automated Perimetry (SAP).

It is believed that structural changes occur prior to functional changes, and so structural changes may be detected at earlier stages in glaucoma.[51, 52] However, the interpretation of the seminal articles backing this belief have been challenged.[53-55] Further, recent evidence suggests that agreement between OCT (structure) and visual field (function) is better than believed if local regions of deviation maps are compared.[56, 57] In any case, automated perimetry can assist in diagnosis of glaucoma at early stages. VF loss also strongly correlates with quality of life measures. Additionally, monitoring the rate of VF loss may help guide how aggressively to treat glaucoma.

Quantitative OCT prediction of automated full-threshold standard achromatic perimetry can be performed (see below). [51]

Providing an Objective Standard in Glaucoma Diagnosis for Evaluating AI Approaches

AI has been validated using structural data alongside functional assessments with or without clinical examinations.[26] Both approaches offer evidence-based validations for one or more definitions of glaucoma.

Clinical exams plus structural data, i.e. OCT data, and functional analysis through SAP testing requires clinicians to consider past history, clinical exam, OCT, and VF data to form a diagnosis, with diagnosis of each eye determined by one treating expert. Global and sectoral degrees of abnormalities in structure and function would need to be assessed by OCT and VF data, respectively. Subsequently, all analytical data and clinical data are taken into consideration
in order to assist specialists in determining a diagnosis. Given the multitude of data available to aid in the diagnosis of glaucoma, it has been difficult to determine a specific algorithm to diagnose glaucoma. However, a systematic criteria for glaucoma diagnosis is required in order for AI assist in glaucoma diagnosis on a more global scale.

In pursuit of a standard definition of glaucoma, Iyer and colleagues (2020) performed a study that compared clinician diagnosis of definite glaucomatous optic neuropathy to objective data from OCT and visual field among 2500 eyes, with the best sensitivity and specificity identified as OCT RNFL sectoral measurements and VF glaucoma hemifield test (GHT) using specific criteria.[26] The purpose was to determine objective criteria to permit comparisons among glaucoma research studies with standard objective parameters. The criteria included (1) most recent or the preceding pairs of tests with abnormal sectoral OCT with matching abnormal superior/inferior GHT (sensitivity 77%, specificity 98%), (2) abnormal sectoral OCT in the most recent or preceding tests with normal or abnormal and correlating GHT in VF (sensitivity 75%, specificity 98%), (3) corresponding abnormal sectoral OCT and GHT VF abnormalities on the most recent tests (sensitivity 73%, specificity 98%), and (4) abnormal sectoral OCT with matching VF GHT abnormality on the most recent and preceding pair of tests (sensitivity 65%, specificity 99%). While their study demonstrated high specificity for glaucoma detection, the sensitivity ranged from 65% to 77%. There was lower sensitivity for eyes with less visual field loss, and among “possible” glaucomatous neuropathy cases, clinician-objective agreement was lower still. This illustrates that so-called “early” examples of glaucoma represent a heterogeneous group in which clinician agreement as to true status is questionable and longitudinal, repeated examinations would be needed to confirm disease status. A more detailed grading of 1000 eyes as to presence of glaucomatous neuropathy on a 0-100 scale is currently
undergoing analysis by Vianna and Chauhan that will further detail which objective measurements correspond to clinician diagnosis. While AI methods may indeed prove prescient in sorting among such early examples, no single point in time objective method could be identified at this time for comparison.

More recently, Mariottoni and colleagues (2021) described a reference standard that can be used in AI for glaucoma detection.[58] In the study, SD-OCT and SAP tests within the same eye were paired and classified as either having glaucoma or as healthy eyes according to the objective reference standard that the authors created. The standard criteria created by the researchers defined glaucomatous eyes as having loss of global RNFL thickness outside normal limits and SAP GHT outside normal limits, with P-values less than 5%. Localized loss was defined as RNFL thickness outside normal limits in at least one superior sector or inferior sector (temporal and/or nasal) alongside inferior or superior sectors. The study went further to implement this standard in a deep learning algorithm to classify eyes with glaucoma.

A third structure-function approach has been taken by Hood, De Moreas, Tsamis and colleagues who developed an automated and objective method for comparing abnormal regions of deviation maps from OCT (RNFL and GCL) and visual fields (24-2 and 10-2). They demonstrated that this method correctly identified eyes missed by other methods, and these eyes included those with damage near fixation.[27, 56, 57]

Additionally, a recent study by Xiong and researchers (2021) demonstrated that the AI algorithm, FusionNet, utilizing VF data and circular peripapillary OCT scans alone to differentiate patients with glaucomatous optic neuropathy vs. those without glaucomatous optic neuropathy achieved an AUROC of 0.950, outperforming VF data alone, OCT data alone, and two glaucoma specialists.[59] This study was conducted in patients only from China and had a
A higher rate of glaucoma. In addition, the study excluded patients with co-existing retinal or optic nerve diseases. Therefore, further studies are warranted to test the accuracy of FusionNet.

**AI, including Neural Networks for Glaucoma**

*Detection*

A number of studies have looked at the application of AI methods for assisting in glaucoma diagnosis. Davella et al. (2019) reviewed current research on use of structural and functional measures alone and in combination as inputs for AI in glaucoma detection.[60] Many studies analyzing VF data inputs for AI algorithms to detect glaucoma showed promise, but had variability in success, likely related to the level of reliability and reliance on patient performance in VF data[61], and the wide diversity in structural features of the ONH region, including disc size, tilt, and positional entry of axons. Structural data has come in the form of CSLO, fundus color photography, and OCT data. While OCT has shown to be a successful tool in AI for assistance in glaucoma diagnosis, its utility depends on how accurately automated measurements can be made, which can be impacted by blood vessels and tissue reflectance.[62] A combination of both functional and structural data has the potential to differentiate glaucomatous eyes from healthy eyes more efficiently than either alone.

Studies have demonstrated the utility of AI technology with inputs from fundus photographs to assist in the detection of glaucomatous damage.[63-65] In a cross-sectional study, Liu and colleagues (2019) trained a deep learning algorithm with fundus photographs to categorize images as either definite, probable or unlikely glaucomatous optic neuropathy.[65] The model was then validated and evaluated using other images from their database. With primary data sets, the model was found to have an AUROC of 0.996, with a sensitivity of 96.2%
and a specificity of 97.7%. The study also found that the most common cause for false-negative and false-positive outputs were pathologic or high myopia and manual grading.

Additional structural parameters have also been studied for their use in AI for glaucoma detection.[26] Thompson et al. (2019) presented a deep learning program that utilizes minimum optic nerve rim width relative to Bruch’s membrane opening (BMO-MRW).[66] The BMO-MRW has been demonstrated to have a strong correlation to VF loss in glaucoma.[67, 68] In their study, Thompson and colleagues trained the algorithm to predict the BMO-MRW when assessing photos of optic nerves of glaucomatous eyes, eyes of glaucoma suspects, and healthy eyes and showed that the predicted structural data correlated with VF loss (AUROC 0.945).[66] Park et al. (2018) analyzed artificial neural networks (ANN) using macular vessel density, macular GCIPL and the combination of both in ANN, and demonstrated improved performance using the combination (AUROC 0.87).[69] It will be important to confirm the validity of such methods against multiple databases in which the glaucoma diagnosis is standardized.

An important consideration is how the accuracy of AI models using deep learning are strongly impacted by the datasets used to train algorithms. The study by Christopher et al. (2020) evaluated the performance of two deep learning models, the University of California San Diego (UCSD) and the University of Tokyo (UTokyo), on independent datasets from different patient populations.[70] When testing the two AI models, they performed similar when using populations with mostly moderate-to-severe glaucoma. However, adding more mild glaucomatous eyes to train the UTokyo model, which was initially trained with a higher prevalence of severe glaucomatous eyes, increased the algorithms performance. Therefore, it is important to train AI models with populations with stratified severity of glaucoma for them to perform more accurately.[19]
Progression

Published studies have also investigated the use of AI in detecting the progression of glaucoma.[71] Although a variety of methods are used to determine glaucomatous structural and functional changes over time, there is no currently accepted single best definition of progressive change. Subsequently, algorithms need to be validated to detect change over time using longitudinal data on known progressive eyes from various centers. If models are not validated and generalized to different populations, they cannot be utilized in clinical settings. It has been recommended that a minimum sample size of 100 events, but more desirable would be a sample size of 200 events, be used for external validations.[72] Some models that attempted to show external validity did not use large enough sample size or reused the same datasets.[73, 74] More recently, the dynamic structure-function model was shown to predict glaucoma progression over a few visits accurately and is generalizable to other testing indices and populations, however the model is yet to have a clinical functionality.[75]

The first successful AI implemented to detect progression was completed by Lin et al. (2003) by examining VF to detect progression of glaucoma in patients with an established diagnosis of glaucoma (AUROC of 0.92).[76] Other studies have also implemented VF in AI algorithms for detecting glaucoma. A study by Wang and researchers (2019) used a new AI approach to track progression in glaucoma using VF data.[77] The authors found that their method outperformed existing algorithms of detecting glaucoma progression and that their method was also able to detect patterns of progression. Additionally, Dixit et al. (2020) used convolutional long short-term memory network, a network that can pass data between layers and allow the networks to look at data overtime, with VF data to evaluate worsening disease.[71] The study demonstrated that the network was able to detect both global and regional changes in
glaucoma eyes longitudinally, and that the network performed better when VF data was combined with clinical data, compared to VF data alone. In addition, Miri and colleagues (2017) used BMO-MRW to estimate nerve fiber bundles at the ONH and monitor glaucoma progression at the neuroretinal rim.[78]

Yousefi et al. (2020) compared using RNFL data, SAP data and the combination of RNFL with SAP data in machine learning algorithms to analyze which data set provided the most discriminating power to differentiate eyes with glaucoma with progressing disease from eyes with stable disease.[79] The researchers found that using RNFL data provided the most discriminatory power, particularly when analyzing global RNFL data as well as sectoral data from the inferior nasal sector, inferior temporal sector, superior temporal, and temporal sectors together. They did not show an advantage to combining SAP with RNFL data in differentiating eyes that had progressive disease from eyes with stable disease.

**Feature Agnostic, Data Driven Approach**

Another deep learning AI OCT data analysis methodology with regard to glaucoma and glaucoma progression is the use of a feature-agnostic, data driven approach. This method does not require known or assumed disease biomarkers, such as RNFL thickness or GCC or GCIPL thickness. A feature-agnostic approach avoids using secondary analyses that rely upon accurate segmentation of OCT layers and may be more robust when analyzing lower signal strength scans, eyes with advanced disease, and eyes with other retinal pathologies. Studies using this technique have shown promising results.[80-84]

**AI OCT Prediction of Visual Field Results**

Using different AI/Deep Learning approaches, groups have found that OCT is capable of predicting the severity VF damage,[82, 85, 86] including visual thresholds,[87] with good
precision and accuracy. OCT prediction of VFs allows more objective and reproducible data compared to using SAP, that can be unreliable. Moreover, predicting the severity of visual field damage from OCT images opens the opportunity to tailor the frequency of visual field testing to the individual patient, by using the OCT prediction as an indicator of whether the visual field has changed. Additionally, using networks that are trained with OCT data, rather than SAP, is more timely and less subjective as it does not require manual labeling.

**AI OCT and the Optic Disc**

Finally, AI applications with optic disc photos may be used to assess glaucoma status and even to predict OCT RNFL parameters.[88, 89]

**Future Considerations**

For AI to establish a role in glaucoma, collaborations will be necessary to create programs that yield unbiased, reproducible, and accurate results. Training and validation sets must be large, diverse, and clinically verified against outcome or other prognostic standards. Support from many centers globally will be needed with glaucoma patients, glaucoma suspects, and healthy age-matched controls using standardized definitions. Input versus training data will need to be determined to avoid overfitting. These are issues that are central to the Collaborative Community for Ophthalmic Imaging Glaucoma Workgroup, and were discussed in detail at the CCOI public meeting September 4, 2020.

**Conclusions**

There is great potential for the role of AI in glaucoma. Currently, there are a variety of criteria that are used to distinguish a healthy eye from one with glaucoma. Ideally, multiple parameters should be incorporated, including OCT RNFL, VF, macular and disc imaging, to enhance the efficacy of glaucoma diagnosis. As a first step, it would be most important to show
that AI can identify a stage of glaucoma for which clinician agreement and objective criteria are more certain. Then, the search for AI methods that classify the early phases of glaucoma can be sought using eyes in which glaucoma developed over time. One consideration is to not rely on a binary decision of the presence or absence of glaucoma. Instead, using data points to create a likelihood scale may provide more utility in glaucoma screening. It also is important to use crowd-sourced data assessment and multicenter evaluation, as well as multiple evaluators in order to reduce errors and improve accuracy of AI outputs.

The utility of glaucoma screening may be more efficient and cost-effective if focused on populations with a higher prevalence of glaucoma, in which the pre-test probability is higher. It is also vital to ensure the testing has high specificity to avoid overwhelming the healthcare system and inappropriately labeling patients with a potentially blinding eye disease. By choosing lower cut-offs for RNFL thickness, there may be increased post-test probability, and in effect more patients with early-intermediate to intermediate stage glaucoma may be detected, as opposed to very early stage glaucoma. Patients detected at these stages could still allow for earlier detection than would be otherwise possible and such a system may therefore prevent visual field loss. Patients at earlier glaucoma stages that will eventually progress to reach the screening threshold will still be detected before reaching severe stages of glaucoma if screened at different timepoints, and therefore may still be protected from loss of quality vision. Another benefit of screening patients for glaucoma is the potential to detect other macular diseases. Medeiros et al. (2012) found that the average RNFL thickness for moderate glaucoma was 65 microns, which may provide a threshold that can be used to detect patient’s at moderate disease in AI.\[90\] One method that can be used for glaucoma screening would be to utilize OCT data, which is may be less timely and more reliable, in AI to first detect eyes with RNFL thicknesses.
of 65 microns or less, to capture those who likely have moderate to severe glaucoma.

Subsequently, eyes screened as high risk for glaucoma can be evaluated further by using VF and IOP data.

In order for AI to emerge as a widespread technology for glaucoma management, cost savings need to be demonstrated. By detecting glaucoma at an earlier stage, patients may be able to avoid requiring more complicated and costly procedures, thereby reducing costs on both an individual and global scale. In addition, by detecting glaucoma in patients who otherwise would not have presented until they had advanced disease such a system may allow these patients to maintain visual function and also allow patients to continue to be productive. Using Medicare data, it was demonstrated that vision loss secondary to glaucoma is expensive, and more so with greater severity of disease.[91] This study also showed that there is an increased risk for admission into nursing home, depression, accidents, femoral fractures with those that have visual decline from glaucoma compared to those with glaucoma without loss of vision. Therefore, if glaucoma can be detected earlier and patients’ can avoid severe vision loss, there can be cost saving benefits globally.

AI has been shown to have great potential in the detection of glaucoma and glaucoma progression. Clear definitions of these parameters require consensus to enable the widespread development, implementation, and acceptance of AI for glaucoma. Patient autonomy and equity must be considerations in the development and adoption of AI for glaucoma. AI offers the prospect of automating the detection of glaucoma and its progression, as well as improving patient access and reducing costs on an individual and global scale.
References


CCOI Executive Committee

Michael Abramoff, MD, PhD
Department of Ophthalmology and Visual Sciences, Roy J. and Lucille A. Carver College of Medicine, University of Iowa, Iowa City, Iowa, USA

Mark Blumenkranz, MD
Stanford University, Palo Alto, CA, USA

Emily Chew, MD
National Eye Institute, National Institutes of Health, Bethesda, MD, USA

Michael Chiang, MD
National Eye Institute, National Institutes of Health, Bethesda, MD, USA

Malvina Eydelman, MD
US FDA, Washington, DC, USA

David Myung, MD
Stanford University, Palo Alto, CA, USA

Joel S. Schuman, MD
Department of Ophthalmology, NYU Langone Health, NYU Grossman School of Medicine, New York, NY, USA

Carol Shields, MD
Wills Eye Institute, Philadelphia, PA, USA

CCOI Glaucoma Workgroup

Michael Abramoff, MD, PhD
Department of Ophthalmology and Visual Sciences, Roy J. and Lucille A. Carver College of Medicine, University of Iowa, Iowa City, Iowa, USA

Lama Al-Aswad, MD, MPH
Department of Ophthalmology, NYU Langone Health, NYU Grossman School of Medicine, New York, NY, USA

Bhavna J. Antony, PhD
IBM Research, Southbank, Victoria, Australia

Tin Aung, MD, PhD
Singapore Eye Research Institute (SERI)
Singapore National Eye Centre (SNEC)
Duke-NUS Medical School, Singapore

29
Michael Boland, MD, PhD
Massachusetts Eye and Ear and Harvard Medical School, Boston, MA, USA

Tom Brunner, MBA
Glaucoma Research Foundation, San Francisco, CA, USA

Robert T. Chang, MD
Stanford University, Palo Alto, CA, USA

Balwantray Chauhan, PhD
Dalhousie University, Halifax, Nova Scotia, Canada

Michael Chiang, MD
National Eye Institute, National Institutes of Health, Bethesda, MD, USA

D. Hunter Cherwek, MD
Orbis International, USA

David Garway-Heath, MD
Moorfields Eye Hospital, London, England
University College London, London, England

Adrienne Graves, PhD
Glaucoma Research Foundation, San Francisco, CA, USA

Jeffrey L. Goldberg, MD, PhD
Department of Ophthalmology, Stanford University, Palo Alto, CA, USA

Minguang He, MD, PhD
The University of Melbourne, Melbourne, Victoria, Australia

Naama Hammel, MD
Google Health, Palo Alto, CA, USA

Donald Hood, PhD
Columbia University, New York, NY, USA

Hiroshi Ishikawa, MD
Department of Ophthalmology, NYU Langone Health, NYU Grossman School of Medicine,
New York, NY, USA

Chris Leung, MD
LKS Faculty of Medicine, The University of Hong Kong, Hong Kong, China

Felipe Medeiros, MD, PhD
Department of Ophthalmology, Duke Eye Center, Duke University School of Medicine, Durham, NC, USA
Department of Electrical and Computer Engineering, Pratt School of Engineering, Duke University, Durham, NC, USA

Louis Pasquale, MD
Department of Ophthalmology, Icahn School of Medicine at Mount Sinai, NY, NY, USA

Harry A. Quigley, MD
Wilmer Eye Institute, Johns Hopkins University, Baltimore, MD, USA

Calvin W. Roberts, MD
Weill Cornell Medical College, New York, NY, USA
Lighthouse Guild International, New York, NY, USA

Alan L. Robin, MD
Johns Hopkins University, Baltimore, MD, USA
University of Michigan, Ann Arbor, MI, USA

Joel S. Schuman, MD
Department of Ophthalmology, NYU Langone Health, NYU Grossman School of Medicine, New York, NY, USA

Elena Sturman
The Glaucoma Foundation, New York, NY, USA

Remo Susanna, MD
Department of Ophthalmology, University of São Paulo, São Paulo, Brazil

Jayme Vianna, MD
Dalhousie University, Halifax, Nova Scotia, Canada

Linda Zangwill, PhD
Hamilton Glaucoma Center, Shiley Eye Institute, Viterbi Family Department of Ophthalmology, University of California, San Diego, USA