Asia-Pacific Journal of Ophthalmology 13 (2024) 100096

Contents lists available at ScienceDirect

Asia-Pacific Journal of Ophthalmology

journal homepage: www.sciencedirect.com/journal/asia-pacific-journal-of-ophthalmology

Artificial intelligence for retinal diseases

Jennifer I. Lim [a,*], Aleksandra V. Rachitskaya [b], Joelle A. Hallak [a,c], Sina Gholami [d], Minhaj N. Alam [d]

[a] University of Illinois at Chicago, College of Medicine, Department of Ophthalmology and Visual Sciences, Chicago, IL, United States
[b] Department of Ophthalmology at Case Western Reserve University, Cleveland Clinic Lerner College of Medicine, Cleveland Clinic Cole Eye Institute, United States
[c] Department of Ophthalmology and Visual Sciences, College of Medicine, University of Illinois at Chicago, Chicago, IL, United States
[d] University of North Carolina at Charlotte, United States

ARTICLE INFO

Keywords:
Artificial intelligence
Algorithms
Screening
Biomarkers
Deep learning

ABSTRACT

Purpose: To discuss the worldwide applications and potential impact of artificial intelligence (AI) for the diagnosis, management and analysis of treatment outcomes of common retinal diseases.
Methods: We performed an online literature review, using PubMed Central (PMC), of AI applications to evaluate and manage retinal diseases. Search terms included AI for screening, diagnosis, monitoring, management, and treatment outcomes for age-related macular degeneration (AMD), diabetic retinopathy (DR), retinal surgery, retinal vascular disease, retinopathy of prematurity (ROP) and sickle cell retinopathy (SCR). Additional search terms included AI and color fundus photographs, optical coherence tomography (OCT), and OCT angiography (OCTA). We included original research articles and review articles.
Results: Research studies have investigated and shown the utility of AI for screening for diseases such as DR, AMD, ROP, and SCR. Research studies using validated and labeled datasets confirmed AI algorithms could predict disease progression and response to treatment. Studies showed AI facilitated rapid and quantitative interpretation of retinal biomarkers seen on OCT and OCTA imaging. Research articles suggest AI may be useful for planning and performing robotic surgery. Studies suggest AI holds the potential to help lessen the impact of socioeconomic disparities on the outcomes of retinal diseases.
Conclusions: AI applications for retinal diseases can assist the clinician, not only by disease screening and monitoring for disease recurrence but also in quantitative analysis of treatment outcomes and prediction of treatment response. The public health impact on the prevention of blindness from DR, AMD, and other retinal vascular diseases remains to be determined.

Introduction

Artificial intelligence (AI) permeates many aspects of everyday life, extending its reach to include the field of retina. AI algorithms include diagnosis, screening, assessment of disease activity, analysis of response to treatment and prediction of disease progression or treatment outcomes. The first paper demonstrating the clinical utility of deep learning for diagnosis diabetic retinopathy (DR) jolted the field. In this 2016 paper, a Google Deep Mind deep learning algorithm, trained on labeled DR datasets, demonstrated high sensitivity, specificity and accuracy for detection of DR using color fundus photographs.[1] The press speculated whether such an AI system could replace retina specialists. Then, in 2018, the first autonomous DR screening system received Food and Drug Administration (FDA) clearance and became commercially available in the United States.[2] Throughout the world, many AI systems are being developed and hold the potential for improved ocular healthcare and outcomes. Infrastructure, economic considerations, regulatory limitations, security considerations, and ethical considerations have slowed down the pace of implementation of AI applications.[3,4] To date, AI has not replaced retina specialists. Rather, ophthalmologists consider AI as an assistive device for the detection or screening of retinal disease, analysis of treatment results, monitoring for recurrence of disease and prediction of progression or outcomes. This review will delve into each of these aspects. This article is not meant to be an exhaustive review, but rather highlights some retinal disease indications for which approved AI screening systems are in existence (DR), AI analysis algorithms exist or

* Correspondence to: University of Illinois at Chicago, College of Medicine, Department of Ophthalmology and Visual Sciences, 1855 W. Taylor St., Suite 2.50, Mail Code 648, Chicago, IL 60612, United States.
E-mail address: jennylim@uic.edu (J.I. Lim).

https://doi.org/10.1016/j.apjo.2024.100096
Received 28 June 2024; Received in revised form 2 August 2024; Accepted 20 August 2024
Available online 27 August 2024

are in development (AMD, ROP, SCR, surgery), or for which the authors have expertise (DR, AMD, SCR).

Screening for retinal disease

The need to screen for diseases such as DR and age-related macular degeneration (AMD) continues to significantly increase throughout the world, due to both population growth and increased longevity. According to the International Diabetes Federation (IDRF), the prevalence of diabetes mellitus is projected to reach 700 million by 2045.[5] Meanwhile, half of the world's population resides in Asia and the largest groups of DM patients are projected to be in China and India with 116 M and 77 M projected patients. The rise in human longevity combined with the increasing prevalence of DM and increasing population risk factors for DM related to improved socioeconomics is projected to increase the need for eye care substantially.[5–7] Currently, DR is the fifth leading cause of blindness in the world.[6] According to the Global Burden of Disease Study, for diabetic adults 50 years of age and older, the prevalence of blindness and moderate or severe vision impairment from DR was 18.5 % in 2020. The highest DR prevalence was in Africa, North America, the Middle East, and North African continents.

At present, the rates of DR screening remain low, ranging on average from 30 to 50 %. These low screening rates contribute to a delay in the diagnosis and treatment of vision-threatening DR. This missed opportunity to detect retinal disease earlier, when visual acuity is still good, contributes to poorer outcomes. Additionally, although up to half of diabetic patients fail to get an annual DR screening examination, 88 % of diabetic patients visit their primary care provider (PCP) at least once per year (US data). Offering DR screening exams at the PCP office also known as designated point of care screening could increase screening rates. Such a model could mitigate socioeconomic disparities, transportation issues, referral and follow-through issues, and partially overcome the shortage of eye care providers to perform screening in certain areas.[8] In this setting, the ideal screening tool would have a very high sensitivity to detect referable DR and a high enough specificity to be clinically useful. A very high negative predictive value (NPV) is necessary for the system to be unlikely to miss the presence of referable DR. The resulting slightly higher rate of unnecessary referrals that accompanies a lower specificity is an acceptable trade-off in order not to miss potentially blinding DR levels.

The shortage of eye care providers, worldwide, contributes to low screening rates for other retinal conditions. For example, an estimated 25 % of AMD patients remains undiagnosed. In the United States (US), only 12 % of counties have a retinal specialist, and it is estimated that more than 110 million Americans require screening for AMD.[9,10] Worldwide, the shortage of eye care providers hinders public health initiatives to screen for ocular disorders.[9–11] The International Council of Ophthalmology (ICO) survey in 160 countries revealed the estimated global mean ophthalmologist density to be 31.7 per million population, with a range spanning from less than 1 ophthalmologist per million to 182 per million. Furthermore, about 17 % of the global population in 132 countries have access to less than 5 % of the global ophthalmologist population.[9] Two-thirds of the global ophthalmologist population were located in only 13 countries (China, USA, India, Japan, Brazil, Russia, Germany, Italy, Egypt, France, Mexico, Spain and Poland). There is an opportunity for AI screening to provide a partial solution to the demand for ocular providers.

AI algorithms for detection and screening for DR, AMD, ROP, and sickle cell retinopathy exist but vary in the level of validation and applicability to large populations. AI systems have been validated across populations and approved for clinical use in DR screening. The commercially available systems include IDxDR (now Luminetics Core)[12–14], EyeArt[15,16] and AEye systems in the US, European Union (EU), Canada, and many other countries throughout the world. All of these systems have high sensitivity, specificity, accuracy, and userability that led to their approvals. In Singapore[17,18] and South Korea,[19,20] researchers have developed deep learning systems (DLS) that are being used in their public health systems to detect with high accuracy referable eye diseases such as glaucoma, AMD, and DR.

Table 1 outlines studies on autonomous systems approved for DR screening. These systems were evaluated on various datasets, and their performances were reported. In addition to these AI systems, smaller handheld systems and smartphone-based systems are being used and evaluated for DR screening. These devices have shown high rates of sensitivity, specificity, and userability. Some of these are self-contained and do not require the internet, which is practical for use in countries and remote areas where internet service is not readily available or reliable. These handheld systems have been evaluated and are being used in many parts of the world today, including China, Kenya, Zambia, and Thailand.

The first autonomous system approved for DR screening was the IDxDR, now known as Luminetics Core (Digital Diagnostics). The system was tested in primary care clinics in 900 asymptomatic patients who were 22 years of age or older without known DR.[12] Subsequent algorithm enhancements led to increased sensitivity from 87.2 % initially for more than mild DR (mtmDR) to 100 %, and to 100 % sensitivity for vision threatening DR (vtDR). The improvements led to an increase in specificity from 90.7 % for more than mild DR (mtmDR) to 82 % for rDR and 95 % for vtDR.[13] In younger patients, ages 5 to 21 years old, the SEE study showed 85.7 % sensitivity, 79.3 % specificity and 97.5 % imageability in comparison to a retina specialist's grading.[14] This lower accuracy resulted in high rates of false positives because of the shiny ILM in young patients.

The EyeArt system was the second approved autonomous AI DR screening system. The clinical trial compared the AI detection of DR to that of a reading center grading of the fundus photograph. The patients underwent the imaging procedures in primary care, endocrinology, general ophthalmology and retina practices. Only patients 18 years of age or older and without a prior diagnosis of DR were included in the study. Various nonmydriatic cameras were used. The AI system had a sensitivity of 95.5 % and specificity of 86 % with high rates of imageability.[15] In a subgroup analysis, the AI system DR grading was compared to that done by ophthalmologists; the sensitivity for detection of rDR was 96 % for the AI system versus 20.7 % for general ophthalmologists and 59.5 % for retina specialists. Of the falsely negative results, none were vtDR for the Eyeart system or retina specialists, in contrast to 19 % for general ophthalmologists.[16]

The AEYE system was the third system approved in the US, although it was approved for use in the EU prior to the US. The AEYE system includes a desktop as well as a handheld camera version. The desktop sensitivity was 93 % and specificity was 91 %, as compared to 92 %−93 % sensitivity and 89 %−94 % specificity for the handheld model.[26]

In Singapore, the national DR screening program, Singapore Integrated Diabetic Retinopathy Programme (SiDRP), created in 2010, provides primary care clinic-based telescreening for DR, AMD and glaucoma. A DLS system, SELENA, which was created in Singapore, has been in use for DR screening since 2019 as part of SiDRP. SELENA showed excellent diagnostic performance with high sensitivity and specificity for detection of referable and vision threatening DR, glaucoma and AMD.[17,18] Using 494,661 images, the area under the curve (AUC) measurements for the DLS as compared with the ground truth were: 0.931 for AMD, 0.936 for DR and 0.942 for glaucoma. Further external validation showed similar results for detection of rDR in ten additional multiethnic datasets of populations with diabetes, including the community-based Guangdong dataset; the population-based Singapore Malay Eye Study, Singapore Chinese Study, Singapore Indian Study, Beijing Eye Study, and African American Eye Disease Study; and the clinic-based studies from the Royal Victoria Eye and Ear Hospital, Mexican, Chinese University of Hong Kong, and University of Hong Kong.[19]

In SiDRP, nonmydriatic digital retinal images taken by a nurse practitioner are sent from the primary eye care clinic to a centralized

Table 1
Autonomous system approved for DR screening.

Study	Data	Results
IDxDR - Luminetics Core (Digital Diagnostics)	A) 900 asymptomatic patients aged 22 years or older.[13] B) Patients aged 5 to 21 years old.[14]	A) 87.2 % sensitivity, 90.7 % specificity, and 96 % imageability for more than mild DR (mtmDR), 100 % sensitivity for vision-threatening DR (vtDR) and referable DR (rDR), 82 % specificity for rDR and 95 % imageability for vtDR. B) SEE study showed 85.7 % sensitivity, 79.3 % specificity, and 97.5 % imageability.
EyeArt[15]	A) Patients aged 18 years or older without a prior diagnosis of DR. B) Subgroup comparison of AI system compared to the gradings done by ophthalmologists.[16]	A) 95.5 % sensitivity and 86 % specificity, with high rates of imageability. B) 96 % for AI versus 20.7 % for general ophthalmologists and 59.5 % for retina specialists.
AEYE		A) Desktop camera: sensitivity 93 % and specificity 91 %. B) Handheld camera: 92 %−93 % sensitivity and 89 %−94 % specificity.
Singapore Integrated Diabetic Retinopathy Program (SiDRP)[17,18]	494,661 images; detecting DR (using 76,370 images), possible glaucoma (125,189 images), and AMD (72,610 images).	AUC for rDR was 0.936 (95 % CI, 0.925 −0.943), sensitivity was 90.5 % (95 % CI, 87.3 %−93.0 %), and specificity was 91.6 % (95 % CI, 91.0 %−92.2 %). AUC for vtDR was 0.958 (95 % CI, 0.956 −0.961), sensitivity was 100 % (95 % CI, 94.1 %−100.0 %), and specificity was 91.1 % (95 % CI, 90.7 %−91.4 %). Further external validation for rDR showed similar results in ten additional multiethnic datasets of populations with diabetes.
SiDRP[19,20]	39,006 DM patients in Singapore.	Using deterministic sensitivity analyses, the least-expensive model was the semi-automated DLS model with a cost of $62 per person per year as compared to a fully human assessment model costing $77 per person per year, and a fully automated DLS model costing $66 per person per year.
The Google AI-based tool, Automated Retinal Disease Assessment (ARDA)[21]	Trained on high-quality images and performed well during validation studies.	ARDA real world deployment for DR screening in Thailand showed many more ungradable images than in the original study. These were due to low-quality images secondary to dirty lenses, novice camera operators, or bright lighting that interfered with pupil dilatation.
Rwanda Artificial Intelligence for Diabetic Retinopathy Screening (RAIDERS)[22]	DR patients requiring referral were randomized (1:1) to either immediate notification of the need for referral or to delayed communication after a human grader read the image 3 to 5 days later.	The immediate AI communication led to higher adherence to referrals compared with conventional grading by a human (51.5 % versus 39.6 %).
Remidio Non-Mydriatic Fundus on Phone (Remidio Innovative Solutions Pvt Ltd)[23]	Fundus images (posterior pole, nasal, and temporal fields) subjected to automated analysis by the Medios AI (Remidio).	Sensitivity was 100 % and specificity was 88.4 % for rDR. Sensitivity was 85.2 % and specificity was 92.0 % for any DR.
Smartphone-based hand-held device (Eyer, Phelcom Technologies, São Carlos, Brazil)[24]	Two posterior segment images (one centered on the macula and another disc centered (45° field of view) after mydriasis).	DLS sensitivity was 97.8 % and specificity was 61.4 % with an AUC of 0.89. The ungradability rate of images was 17.6 %.
AI system (Pegasus, Visulytix Ltd., UK)[25]	6404 patients, of whom approximately 80 % were diabetic and 65 % were female.	0.894 (95 % CI: 88.0–90.7) area under the receiver operating characteristic (AUROC) for rDR.

reading center, where a non-ophthalmologist assessor evaluates them for the presence of DR. In an economic analysis modelling study of 39,006 DM patients in Singapore, using deterministic sensitivity analyses, the least-expensive model was the semi-automated DLS model with a cost of $62 per person per year as compared to a fully human assessment model costing $77 per person per year and a fully automated DLS model costing $66 per person per year. Implementing such a system could result in $15 million potential annual cost savings in 2050 based on projected DM prevalence.[19,20]

Real-life deployment of AI systems holds lessons. The Google AI-based tool, Automated Retinal Disease Assessment (ARDA), had been trained on high-quality images and performed well during validation studies. When ARDA was deployed for DR screening in Thailand, the algorithm did not perform as well and many more ungradable images were reported. This was attributed to more low-quality images secondary to dirty lenses, novice camera operators, or bright lighting that interfered with pupil dilatation.[21]

AI can impact the adherence with referrals for DR care. In the Rwanda Artificial Intelligence for Diabetic Retinopathy Screening (RAIDERS) study, AI was used to screen and detect DR in patients seen in a primary care clinic. Patients who required DR referral, based on the AI reading, were randomized (1:1) to either immediate notification of the need for referral or to delayed communication after a human grader read the image three to five days later. The immediate AI communication led to a higher adherence to referrals as compared with conventional grading by a human (51.5 % versus 39.6 %).[23]

In addition to these AI systems, there are smaller handheld systems and smartphone-based systems being used and evaluated for DR screening. These devices have shown high rates of sensitivity, specificity and userability.[23–25] Some of these are self-contained and do not require the internet, which are practical for use in countries and remote areas where internet service is not readily available or reliable. These handheld systems that have been evaluated are being used in many parts of the world today. In India, a low-cost Remidio Non-Mydriatic Fundus on Phone (Remidio Innovative Solutions Pvt Ltd) captures fundus images of the posterior pole, nasal, and temporal fields. These images are then subjected to automated analysis by the Medios AI (Remidio), which provides an offline automated analysis of the smartphone image for detection of rDR. A pilot study in India showed that the sensitivity and specificity for rDR were 100.0 % and 88.4 %, respectively, and the sensitivity and specificity for any DR were 85.2 % and 92.0 %, respectively.[23] In Brazil, a smartphone-based hand-held device (Eyer, Phelcom

Technologies, São Carlos, Brazil) was used to capture two posterior segment images (one centered on macula and another disc centered (45° field of view)) after mydriasis induced by 1 % tropicamide eye-drops. The DR classification by the DLS algorithm (PhelcomNet) was compared to a reading center. The DLS sensitivity and specificity were 97.8 % and 61.4 % with an AUC of 0.89. The ungradability rate of images was 17.6 % in this study.[24]

An AI system (Pegasus, Visulytix Ltd.) for detecting rDR and PDR has been tested on images obtained with a 40-degree, nonmydriatic handheld portable fundus camera (Mexican Advanced Imaging Laboratory for Ocular Research (MAILOR) cohort), and on images obtained with a mydriatic 50-degree desktop camera (Indian Diabetic Retinopathy image Dataset (IDRiD) benchmark cohort). The results were compared, and although the handheld images had less detail, the detection of rDR and PDR was good, with better accuracy for PDR than for rDR.[25] For the handheld camera, sensitivity was 81.6 % and specificity was 81.7 % for rDR; sensitivity was 86.6 % and specificity was 87.7 % for PDR. For the desktop camera, sensitivity was 93.4 % and specificity was 94.2 % for rDR, both of which were greater than those using handheld images ($P < 0.001$). For the desktop camera, sensitivity was 83.7 % and specificity was 84.6 % for detecting PDR, which were not significantly different from the handheld camera values. This study shows that handheld cameras combined with AI could be used to screen patients in remote areas where there are shortages of eye care specialists.

The socioeconomic effectiveness and public health impact of AI screening systems are being evaluated. The Care Process for Preventing Vision Loss from Diabetes (CAREVL) used a Markov model to compare the effectiveness of point-of-care autonomous AI DR screening to in-office eye exams. The outcome was the prevention of severe vision loss by five years.[27] The authors noted that the impact on vision depends on several factors, primarily DR prevalence, the AI system, access and adherence to recommended referrals, access and adherence to treatment/management recommendations, and care frictions/imperfections. They estimated that adherence with recommended metabolic recommendations would result in 110 per 100,000 fewer patients progressing to visual loss, and that adherence with ophthalmic treatments would result in 294 per 100,000 fewer patients progressing to visual loss. In their model, the use of AI combined with optimized care was estimated to result in 367 per 100,000 fewer patients progressing to visual loss. This would translate to 110,000 fewer patients with visual loss in the US, where there are 37 million diabetic patients.

A Markov model-based hybrid decision tree compared the costs, effectiveness, and incremental cost-effectiveness ratios (ICER) of AI screening, no screening, and ophthalmologist-based DR screening. The study found that AI screening was more cost-effective than conventional ophthalmologist-based screening. The ICER for AI screening was $180.19, compared to $215.05 for ophthalmologist-based screening. As expected, quality-adjusted life years (QALYs) were greater with AI screening compared to no screening.[28]

Aside from AI screening based upon fundus photographs, there is the potential to use OCT and OCTA images although higher costs of OCT and OCTA machines would limit widespread usage in the world. Nonetheless, researchers have shown AI analysis of OCTA images can result in classification of DR stage. Refinements of AI analysis of OCTA images to utilize artery and vein differential analyses have been developed and have shown improved rates of DR classification.[29–31] Similar studies found that better classification was also achieved with differential artery-vein analysis for SCR.[32] A fully automated AI algorithm is available to perform this analysis on OCTA images.[33,34]

There are fewer AI screening systems for AMD unlike for DR. As discussed earlier in this report, SELENA is a system currently in use to screen for DR, glaucoma, and AMD in Singapore. For the SELENA AMD deep learning model, referable AMD was defined as intermediate AMD (numerous medium-sized drusen, one large drusen, or non-central GA) and/or advanced AMD (central GA or nAMD). The AUC was 0.931 with a sensitivity of 93.2 % and specificity of 88.7 %.[18]

Currently, there are no approved AI AMD screening systems in the US and Europe. In the US, the Collaborative Communities for Ocular Imaging (CCOI) AMD working group is working on the steps needed to develop an AI screening system for widespread use. The CCOI has identified the steps needed for development.[35] Large datasets that have clear labeling, validation of the labeling, attached metadata, and are from a diverse population are needed. The AMD Benchmark Study is in progress.[36]

There are, however, algorithms that can classify AMD based on the AREDS severity score.[37,38] DeepSeeNet, which is publicly available, simulates the human grading process. The algorithm first detects AMD risk factors, such as drusen size and pigmentary abnormalities for each eye, and then calculates a patient-based AMD severity score using the AREDS Simplified Severity Scale.[37] The AUC were high for the detection of large drusen (0.94), pigmentary abnormalities (0.93), and late AMD (0.97). DeepSeeNet outperformed retinal specialists in the detection of large drusen (accuracy: 0.742 vs. 0.696; Cohen's Kappa (κ): 0.601 vs. 0.517) and pigmentary abnormalities (accuracy: 0.890 vs. 0.813; κ: 0.723 vs. 0.535) but showed lower performance in the detection of late AMD (accuracy: 0.967 vs. 0.973; κ: 0.663 vs. 0.754).[37] Such a system may be useful in assisting clinicians for the determination of AMD risk of progression.

AI systems can also identify GA with high accuracy and with an AUC that is non-inferior to retinal specialists.[38,39] AI systems for GA are more useful for early detection of GA or nascent GA. AI systems are inherently faster than human graders. For example, when applied to the analysis of FAF images, the AI system could annotate the GA lesions within 6.06 s, compared to 1.04 min for human graders.

OCT images may also be used for AI-based AMD screening.[40,41] OCT imaging can be used to differentiate AMD from DME and to differentiate AMD requiring treatment from less severe diseases.[41] The CCOI AMD group has identified OCT as a useful test in concert with fundus images. The OCT is useful for the detection of high-risk AMD biomarkers as well as for the detection of late AMD (nAMD and GA).[35]

AI algorithms based on color fundus photos are also being developed for the screening and diagnosis of other retinal diseases, such as sickle cell retinopathy (SCR), retinopathy of prematurity (ROP), and other retinal conditions. The ROP screening system is probably the next system that will receive regulatory approval in the US. The potential global impact of such a system is great because ROP disproportionately affects infants in low- and middle-income countries. The CCOI ROP working group study validated a vascular severity scale to quantify ROP stage and plus disease, and demonstrated that a deep learning system for quantifying ROP stage was possible.[42] Since then, several groups in the US, China, New Zealand, and Australia have developed automated AI systems for diagnosing plus disease in ROP.[43] A multinational validation of an autonomous ROP screening system showed that the AI system performed well in detecting more than mild ROP (mtmROP).

The i-ROP Deep Learning system, developed by the Imaging and Informatics in ROP (i-ROP) consortium, outputs a continuous vascular severity score (VSS) ranging from 1 to 9, based on the learned probability of consensus diagnosis of preplus or plus disease. A multinational validation of an autonomous ROP screening system showed that it performed well in detecting mtmROP and type 1 ROP.[44] The AI system was trained and calibrated using 2530 examinations from 843 infants in the Imaging and Informatics in Retinopathy of Prematurity (i-ROP) study, on two external datasets (6245 examinations from 1545 infants in the Stanford University Network for Diagnosis of ROP [SUNDROP] and 5635 examinations from 2699 infants in the Aravind Eye Care Systems [AECS] telemedicine programs). The sensitivity and specificity were greater than 80 % for detection of mtmROP. Sensitivity was 100 % for detection of type 1 ROP (SUNDROP and AECS exams). Using such an AI system could result in a potential physician workload reduction of 80 %.

There is potential for the VSS to improve the precision and accuracy of diagnosing plus disease.[45,46] Masked graders often disagreed with the original ROP screening examiner on the presence of type 1 ROP. The AI

system VSS could differentiate between type 1 and type 2 ROP disease and represents a solution to the subjectivity of the designation of plus disease and treatment-requiring ROP; the system could aid in the normalization of the assessment of ROP requiring treatment, which is currently very subjective.[46]

Another area in which AI may be useful is SCR. AI may help detect retinal neovascularization, often referred to as a "seafan". An AI-based system would require widefield images, since seafans are usually located in the midperipheral to peripheral retina. Using 1182 ultra-widefield color fundus photographs from 190 patients with sickle cell hemoglobinopathy, a DLS algorithm achieved a sensitivity of 97.4 % and specificity of 97.0 % for detection of seafans.[47] Another AI system has used OCTA images for classifying the level of SCR into early (stage 1 or 2) or advanced (stage 3) stages with good accuracy.[32]

OCTA has also been used in creating AI systems for detection of other retinal diseases. Supervised machine learning, which uses OCTA features to train algorithms to identify SCR using OCTA images, has shown an average accuracy of 95 %. The AI could differentiate between mild sickle-cell retinopathy (stage II) and severe sickle-cell retinopathy (stage III) with an accuracy rate of 97 %.[48] Quantitative analysis of the OCTA images have focused upon vascular parameters, such as blood vessel caliber (BVC), blood vessel tortuosity (BVT), perfusion intensity density (PID), blood vessel density (BVD), vessel area flux (VAF), vessel perimeter index (VPI) for diabetic retinopathy and on blood vessel tortuosity, blood vessel diameter, VPI, foveal avascular zone (FAZ) area, contour irregularity of FAZ and parafoveal avascular density for SCR classification.[49] Using these parameters, machine learning algorithms could result in classification of control vs SCR, SCR mild vs SCR severe with high sensitivity (100 % and 97 % respectively), specificity (100 % and 95 % respectively) and accuracy (100 % and 97 % respectively).[50] Artery-vein differentiation can further improve the classification accuracy.[32]

Use of both OCTA and fundus images as inputs to an AI system for the detection of polypoidal choroidal vasculopathy (PCV) was performed at Peking Union Medical College.[51] Using OCT and fundus images pairs, the AI system had an accuracy of 87.4 %, with a sensitivity of 88.8 % and specificity of 95.6 %, demonstrating perfect agreement with the diagnostic gold standard ($\kappa = 0.828$). This model outperformed the best expert in the diagnosis of PCV.

Assessment of retinal disease activity

Beyond detecting retina conditions, AI can aid in the assessment of treatment response using OCT-algorithms for detection of intraretinal, subretinal, and sub-retinal pigment epithelial (sub-RPE) fluid. OCT has allowed clinicians and researchers to assess biomarkers of disease activity such as intraretinal fluid (IRF), subretinal fluid (SRF), and pigment epithelial detachments (PED) in a variety of retinal pathologies including nAMD, DME, and retinal vein occlusion (RVO).[52] These biomarkers are widely utilized in both clinical trials and clinical practice as it has been established that these anatomic findings correlate highly with functional outcomes.[51]

In clinical trials, assessment of the presence of fluid is commonly utilized as a secondary outcome in establishing the efficacy of a new therapeutic agent or intervention. Additionally, it can be used in defining rescue or retreatment criteria. Depending on the study design, evaluation for the presence of fluid is done by investigators or by the reading center. A study examining treatment decisions based on OCT fluid identification by investigators versus reading center graders in the Comparison of Age-Related Macular Degeneration Treatments Trials (CATT) showed that treating physicians' and reading center's fluid determination agreed in 72.1 % and disagreed in 27.9 % of visits.[53] In fact, disagreement regarding fluid presence can exist even between expert reading center graders.[54] Given how important proper fluid identification is, utilizing artificial intelligence (AI) to objectively identify, localize and quantify fluid is attractive.[55]

In addition, focus is not only placed on central retinal subfield thickness (CSFT) at the beginning and end of treatment but also on fluid fluctuations throughout the treatment course, as these fluctuations can affect visual outcomes. Post hoc analyses of nAMD clinical trials showed that greater variation in retinal thickness during treatment with anti-VEGF was associated with worse visual outcomes and development of fibrosis and macular atrophy.[56] The data being generated by these analyses might lend itself for AI utilization.[57,58]

Furthermore, CSFT is insufficient as a sole biomarker for nAMD activity because it does not include extra-foveal locations of fluid. In addition, CSFT does not differentiate between the locations of fluid, such as IRF, SRF, sub-RPE fluid. A study comparing CSFT and fluid volume measurements has shown that patients with nAMD generally demonstrate the weakest association between CSFT and fluid volume measurements in the central 1 mm, as compared to DME and RVO patients.[59] Instead, the location of the fluid, IRF, SRF or sub-RPE, has more impact on visual acuity. Several studies have shown that the presence of IRF leads to worse visual outcomes.[60,61] In contrast, the relationship between the presence of SRF and its implication for vision is not as clear. Small amounts of SRF are not incompatible with good visual outcomes.[62–64]

Besides the standard biomarkers of IRF and SRF, novel biomarkers are being identified. For instance, hyperreflective foci (HF) have been suggested as biomarkers for DME disease progression and treatment response.[65] Subretinal hyperreflective material (SRHM) is being assessed in nAMD.[66] AI may be a useful tool to quantify these biomarkers at baseline and to follow them over time. AI algorithms can perform this analysis in a fraction of the time required by trained human graders.[67] Moreover, for clinical trials, AI carries the potential benefit of identifying at-risk populations and enriching recruitment.[68]

In clinical practice, the real-world outcomes have consistently been worse than those seen in clinical trials. Although the reasons might be multifactorial, it has been suggested that tolerance of fluid as well as the potential lack of bandwidth in a busy practice to analyze each OCT B-scan closely may contribute to undertreatment.[55] In addition, AI, and in particular, generative adversarial networks (GAN), can also help predict response to treatment and to remove shadows in OCT images.[69]

AI interpretation of the type, location, and amount of OCT retinal fluid might be as good as or superior to human graders. In a study looking at the AREDS2–10 year SD-OCT scans for the presence and absence of IRF and SRF, assessments were done by investigators and Notal OCT Analyzer. The retinal specialists had imperfect accuracy and low sensitivity in detecting retinal fluid compared to the AI-based detection.[70]

In a clinical setting, the use of AI has been shown to demonstrate performance in making a referral recommendation that reaches or exceeds that of experts on a range of sight-threatening retinal diseases.[71] There are several ongoing studies evaluating the role of AI. The RAZORBILL study (NCT04662944) investigates the impact of advanced AI segmentation algorithms on the nAMD disease activity assessment by enriching three-dimensional OCT scans with automated fluid and layer quantification measurements.[72]

A prospective study using fluid monitoring (NCT05093374) is an example of ongoing trials using automated segmentation of retinal fluid volumes.[68] In Europe, the Fluid Monitor (RetInSight) is an algorithm approved for the clinical monitoring of fluid in patients with nAMD. In the U.S., it is being used in investigational studies. The algorithm is linked to the Spectralis HEYEX 2 platform (Heidelberg Engineering) and will soon be linked to the Topcon Triton OCT and Cirrus OCT (Zeiss).[73]

Aside from the utilization of AI-based OCT assessments in clinical trials and clinical practice, AI-powered home OCT is a reality. The Scanly Home OCT device (Notal Vision, Manassas VA) is the first FDA-approved device, which allows patients to perform a home-based macular scan. The scan is then sent to their physicians for evaluation of fluid in between office anti-VEGF treatments. This system could potentially shift the treatment paradigm from a "treat and extend" approach to a

personalized "treat and observe" regimen. Such a system allows for continual monitoring for CNV recurrence and longer inter-visit intervals. Earlier disease detection of recurrent nAMD is theoretically possible.

The potential clinical utility in home OCT for treatment of nAMD is being investigated by the DRCR Retina Network protocol AO. Research is ongoing to determine the optimal threshold level of retinal thickness increase that should trigger a clinical alert to the physician. A potential pitfall of using home OCT is the lack of fundus imaging, which means small amounts of retinal hemorrhage may go undetected. In contrast, imaging in the clinic typically includes a retinal clinical exam, which can detect such issues.

Prior research on the home OCT has shown high rates for both patient userability and retina scans with similar rates of retinal fluid detection as office-based OCT machines. The imaging success rate has been reported to range from 87 % to 93 %. In one study with an imaging success rate of 88 %, higher rates were found for patients with visual acuity greater than 20/320, with a 90 % success rate for VA $>$ 20/320 versus a 50 % success rate for VA $<$ 20/320.[74] In other studies, 86.5 % to 93 % of scans were eligible for fluid quantification.[75] When comparing in-office and home OCT performance, the positive and negative percent agreement rates for the presence of fluid by an expert grader were 98 % and 96 %, respectively. Quantitative agreement between Notal OCT Analyzer (NOA), a deep learning-based algorithm for automated estimation of fluid volumes, and manually graded outputs by an expert grader showed a Pearson correlation of 91.6 %.[74] In another study, fluid assessments performed on the same day using in-office OCT and NOA on the home OCT scans showed that agreement was 96 %.[75] Finally, cases where NOA estimated fluid to be greater than 10-nL were compared to in-office OCT scans; graders identified fluid on the in-office OCT in all cases, effectively demonstrating that there were no false positives.[76]

In summary, AI-based algorithms are able to identify the location, subtype, and fluctuations of essential OCT biomarkers, assist physicians in identifying patient subtypes for clinical trial recruitment, provide quantitative image analysis, and enable personalized clinical care.

Prediction of disease progression

AI deep learning algorithms can detect early AMD biomarkers associated with disease progression, such as pseudodrusen, intraretinal HF and hyporeflective cores.[77] AI algorithms can predict the progression of drusen over time in AMD patients.[78] More significantly, AI algorithms can also predict the likelihood of progression to late stages of AMD, GA or nAMD. For example, there is an AI algorithm that uses a two-step prediction model to accurately predict progression to late AMD within five years. The model, trained and tested on the AREDS2 dataset, utilized DeepSeeNet on color fundus photos to identify AMD features that increased the risk of AMD progression, and then combined patient demographic features to predict progression to late AMD.[79] The C-statistic (representative of the AUC for binary outcomes) for AMD progression at five years from baseline was 0.84. Another study used the Moorfields database on patients with nAMD in one eye, and who also had OCTs of the fellow eye taken every one to twelve months, to create an algorithm that predicted development of nAMD in the fellow eye. This algorithm's prediction of the risk of conversion of the fellow eye to nAMD within six months outperformed that of five out of six retinal specialists.[80] In Europe, the GA Monitor is a commercial software (RetInSight) approved for use in OCT systems for GA monitoring. The GA Monitor quantitates the amount of photoreceptor loss (ellipsoid zone) and RPE loss and can track this over time in GA patients. It identifies patients who will be fast progressors on the basis of a single OCT at presentation.[81] It is likely that other similar systems will gain approval for clinical use throughout the world for AMD and other retinal conditions.

Anti-VEGF treatment requirements and outcome predictions

AI applications have shown great promise not only in predicting progression of AMD but also in assisting in AMD treatment decisions. With the availability of high-resolution OCT data, detailing retinal structures and biomarkers, several AI applications and systems have been developed to assist clinicians in predicting anti-VEGF treatment requirements and clinical outcomes.[82–84]

AI-based systems developed with OCT can involve feature learning and/or deep learning. Feature learning includes predictions and classifications based on pre-determined extraction of biomarkers, such as fluid volume quantification, fibrovascular PED, subretinal hyperreflective material and HF.[68,83] Deep learning includes the ability of neural networks to differentiate between disease states and outcomes. Most AI systems have been developed using standardized clinical datasets. It is important to note that all AI-based systems need to be validated on external patient populations, preferably using real-world data, to ensure generalizability, improve performance, and facilitate translation into clinical practice.

Predicting anti-VEGF treatment requirements

Predicting the required anti-VEGF treatment in patients with neovascular AMD may improve clinical management, ensure that patients are not undertreated, and improve visual outcomes. It can also provide information on the treatment disease burden. Treatment variability and undertreatment have been highlighted by several studies.[85,86] Most importantly, initial vision gained may not be maintained over time in clinical practice.[87] The ability to predict the required anti-VEGF treatment using AI-based approaches may help to identify patients who require a higher frequency of anti-VEGF injections versus those who require a lower frequency of injections. In addition, AI can also be used for treatment predictions over different time intervals. Most of these studies used clinical trial data to develop the AI prediction algorithms.

AI algorithms, using a machine learning approach, have classified patients into low and high treatment frequency.[82] The HARBOR (Harnessing Automatic Real-time Biomedical Observations and Responses) clinical trial dataset, which includes OCT features, visual acuity measurements, and demographic characteristics, was used to predict the anti-VEGF treatment burden over a two-year follow-up period. Patients were classified into low, medium and high injection requirements. Classification of high and low demonstrated an AUC of 0.7 and 0.77, respectively. The most relevant feature for treatment burden prediction was subretinal fluid volume in the central 3 mm; the highest predictive values were those at month 2.[82] Other studies report AUCs of 0.77 and 0.82 for predicting few or many injections, respectively,[88] and for predicting low or high demand in a treat-and-extend setting.[89] While some researchers found OCT fluid, lesion characteristics and treatment trajectory in the first three months of treatment as important features,[88] others were able to predict low demand at the first visit even before the first injection.[89]

Use of real-world data has enabled the development of a fully automated AI algorithm that enables *probabilistic forecasting* (providing uncertainty estimates) of future anti-VEGF treatment frequency.[84] This system highlighted the most relevant imaging biomarkers for these predictions. They provided a measure of predictive uncertainty for each individual prediction (as illustrated in Fig. 1). The researchers extended the previously developed NGBoost algorithm[90] with the addition of a negative binomial distribution as a probability distribution to adequately reflect the needed anti-VEGF injection frequency. Specifically, NGBoost allowed for the prediction of the injection frequency with a mean absolute error of 2.66 injections per year [2.31–3.01]. Feature importance analysis across the machine learning models revealed that the standard deviation of retinal pigment epithelium-drusen complex thickness in the central ETDRS subfield thickness was a top-ranked feature. Another important feature was the

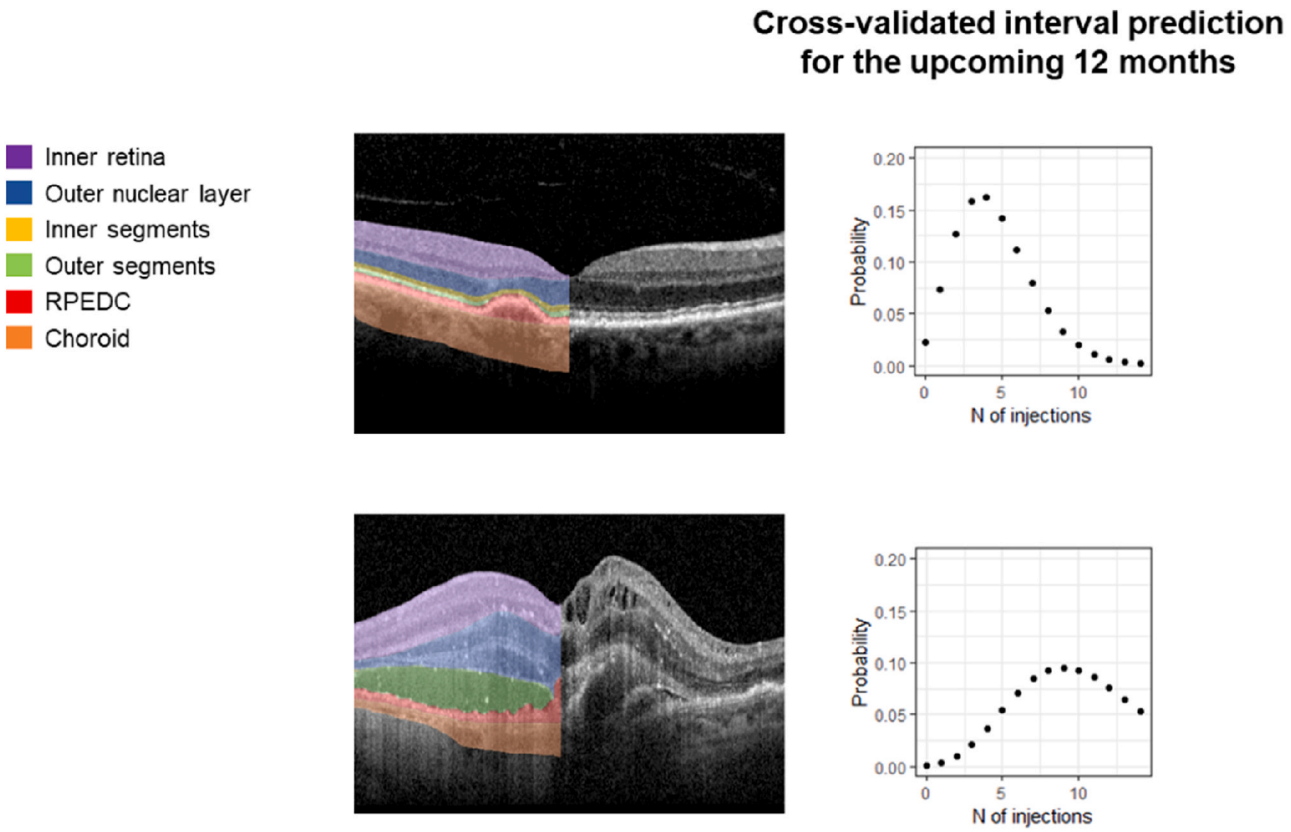

Fig. 1. Figure from: Pfau M, Sahu S, Rupnow RA, Romond K, Millet D, Holz FG, Schmitz-Valckenberg S, Fleckenstein M, Lim JI, de Sisternes L, Leng T, Rubin DL, Hallak JA. Probabilistic Forecasting of Anti-VEGF Treatment Frequency in Neovascular Age-Related Macular Degeneration. Transl Vis Sci Technol. 2021 Jun 1;10(7):30.

standard deviation of the inner segment thickness as an inner nasal ETDRS (Early Treatment Diabetic Retinopathy Study) subfield. The probabilistic prediction of anti-VEGF treatment frequency was compared to other standard machine learning models, showing similar accuracy.

Fig. 1 shows the probabilistic forecasting for two representative patients. For these patients, the forecasted distribution is very much coherent with the clinical imaging characteristics. The figure shows the central SD-OCT B-scan of two patients and the probabilistic forecast for the upcoming twelve months. The upper patient shows a type 1 CNV with no IRF and only subtle SRF (in neighboring B-scans). The predictive model predicts three to four injections per year for this eye (true number of required injections = 2). In contrast, the model predicts seven to eight injections per year for the eye of the lower patient, which is characterized by marked IRF and SRF and a type 2 CNV membrane (true number of required injections = 10). This work highlights the potential of novel algorithms to inform clinical practice, facilitate patient scheduling, and identify patients who may benefit from long-acting treatment modalities.

In addition to utilizing machine and feature extraction learning methods, researchers have utilized a deep learning model to predict the burden of anti-VEGF injections.[91] An end-to-end trainable densely connected neural network (DenseNet) and a recurrent neural network (RNN) were built by sampling 2D-OCT volume images. DenseNET learned retinal spatial features while the RNN integrated information from different time points. In a pro-re-nata (PRN) treatment regimen, the classification task obtained an accuracy of 0.85 in predicting patients with low and high treatment requirements.[91]

Predicting treatment outcomes

Recently, several machine and deep learning methods have been used to predict the response of patients with nAMD to anti-VEGF therapy. Outcomes included predicting visual acuity post anti-VEGF treatment at various time points and OCT features.[92–96]

Standard machine learning techniques using the LASSO feature learning model can predict visual acuity outcomes at three and twelve months after inputting data following three anti-VEGF treatments.[92] The system had a mean absolute error of 5 letters in the 3-month prediction and 8 letters in the 12-month prediction. A 12-month AI tool may help improve adherence to treatment.[92]

Deep learning techniques were utilized in other studies to predict visual outcomes. One study evaluated the predictive ability of OCT imaging biomarkers for cross-sectional and future visual outcomes.[93] AI deep learning algorithms automatically segmented OCT images and predicted visual acuity at distant time points up to twelve months. Most importantly, the study demonstrated that incremental changes in visual acuity after an injection can be predicted.[93]

Several other studies also utilized deep learning to predict visual outcomes, treatment response, OCT images, and to select choice of treatment.[94–98] In one study, a novel convolutional network predicted the 12-month visual outcomes with an AUC of 0.989 and accuracy of 0.936.[94] Another study developed a deep learning architecture named sensitive structure guided network (SSG-Net) to predict short-term anti-VEGF treatment responder/non-responder patients based on OCT images.[97] The model predicted the short-term efficacy of treatment with an accuracy of 84.6 %, AUC of 0.83, sensitivity of 0.692 and specificity of 1.[97]

A generative adversarial network model trained on OCT images was used to predict agent-specific short-term outcomes, specifically to

predict the presence of retinal fluid after treatment.[95] The system had higher sensitivity than human examiners to predict a difference in the efficacy in fluid resolution with anti-VEGF agents.[98] Using real-world data, a machine learning model with computed OCT quantitative features could predict anti-VEGF treatment requirements, visual acuity and morphological outcomes.[96] Best-corrected visual acuity at baseline was the most relevant predictive factor for 1-year visual acuity outcomes. Additionally, the system could predict the development of subretinal fibrosis with an AUC of 0.74.[96]

Lastly, synthetic image generation has also been proposed for AMD. A generative adversarial network generated and evaluated synthetic and individualized post-treatment OCT images that could predict short-term response after anti-VEGF therapy; 92 % of synthetic OCT images were of sufficient quality for clinical interpretation.[98]

The next phase of AI-based systems may include OCTA features, in addition to OCT B scans, to develop multimodal AI-based systems for treatment requirements and outcome prediction. Models may improve predictions by including a combination of imaging, demographic, and clinical information. In summary, AI technology may assist clinicians in customizing the number of injections needed and in selecting the medication most likely to resolve the nAMD features. This paves the pathway towards personalized medical management for nAMD. AI is also being applied to predict outcomes for other retinal diseases and holds promise for increasing the personalization of care for all patients.

Surgical applications

AI applications for retinal surgery include applications related to pre-operative, intraoperative and post-operative aspects. These applications continue to be developed and refined and there are not yet any regulatory agency approved systems.[99] Pre-operative applications include prediction of visual and anatomic outcomes after the surgical intervention.[99–101] For example, a multicenter study using AI deep learning models to predict macular hole status after pars plana vitrectomy (PPVx) surgery with internal limiting membrane (ILM) peeing showed an overall accuracy of 84.7 % with an AUC of 89.32 %.[100]

Another study used a multimodal deep fusion network model (MDFN) to reliably predicted MH closure status (closed or open) one month after PPVx with ILM peeling, based on pre-operative macular OCT images and clinical data (including age, gender, duration of symptoms, minimal diameter of MH, base diameter of MH, height of hole, macular hole index, diameter hole index, hole form factor, and tractional hole index).[101] The AUC of this MH status prediction model was 0.947.

Others have applied AI for prediction of outcomes after rhegmatogenous retinal detachment repair.[102] A multimodal fusion model using ultra-widefield fundus images, macular OCT images, age, gender, and pre-operative BCVA predicted post-operative visual acuity outcomes with an AUC of 0.91 with a mean accuracy of 0.86, sensitivity of 0.94, and specificity of 0.80. Not surprisingly, heatmaps revealed that the macular area for both OCT and ultra-widefield images was the most informative for model predictions.

AI holds potential applications in the performance of surgical procedures.[103–110] Robotic assistance and guidance may enable techniques that have previously been unfeasible due to biological limitations and intrinsic hand tremors.[103–105] Real-time instrument tracking, collision avoidance and surgical education are other areas for which AI may hold significant impact.[104–110] These applications could make surgery safer. Epiretinal membrane peeling, retinal vessel cannulation and subretinal gene therapy are some potential applications.

A robotic surgical system (Preceyes) was first used for human retinal surgery in nAMD patients with subretinal hemorrhage, which was treated with subretinal injections of recombinant tissue plasminogen activator (rt-PA).[110] A surgeon used a remote z-axis control to guide the placement of a thin cannula through the retina and into the subretinal space to deliver the rt-PA. Such intraoperative applications illustrate the use of AI to improve surgical precision. Continued AI research applications are needed not only for operative applications, but also for pre-operative and post-operative applications that will benefit the care of our surgical patients.

Federated learning

Development of AI systems is greatly facilitated by the ability to utilize validated datasets from several sources that include diverse patient populations. Collaboration amongst investigators from different locations may be hindered by institutional concerns about internet security as well as limitations regarding data access. Federated learning provides a collaborative framework in which AI training data is not exchanged. This overcomes limitations on data sharing, such as policy, security, and coalition constraints. Federated learning initially emerged in the field of communications to enhance the training of deep learning (DL) networks using decentralized data.[111–113] Subsequently, it was adopted in the healthcare sector to enable multi-institutional training of models. This approach aims to develop powerful, accurate, safe, robust, and unbiased models while adhering to the Health Insurance Portability and Accountability Act (HIPAA), which mandates the confidential handling of protected health information.[114] A notable development was the federated learning platform created for diagnosing COVID-19 using computed tomography (CT) scans. This platform employed a three-dimensional dense CNN to provide a real-world, globally constructed, and validated clinical tool for CT-based COVID-19 diagnosis, leveraging artificial intelligence.[115]

At its foundation, federated learning involves numerous nodes that collaboratively train a ML or DL model. Each node trains its model locally and shares its parameters using one of two common federated learning communication architectures: centralized or decentralized. In a centralized architecture, a server acts as an orchestrator, collecting model parameters or weights from each node, aggregating them, and then redistributing the updated parameters back to the nodes. In a decentralized architecture, each node directly passes its weights to another, allowing for direct updates to the global parameters by every node.[116]

Recently, several frameworks for developing federated learning algorithms, such as NVIDIA FLARE,[117] Flower,[118] and FedEYE[119] have been developed. There are several challenges in applying federated learning in the medical domain such as model-aggregation policy, participation motivations, hardware or network condition. And more importantly, differences in image acquisition protocols and labeling methodologies across institutions which may lead to the generation of site-specific models that do not fit other sites well and contribute negatively to the global model.[120]

On the other hand, the benefits of federated learning are undeniable. The federated learning framework has the potential to link isolated medical institutions, hospitals, and devices, enabling sharing while ensuring privacy. As the number of wearable devices focused on public health increases, federated learning can leverage medical domain knowledge to personalize the global model for each medical institution and wearable device. Additionally, federated learning is scalable with minimal additional cost, allowing for the training of models using a diverse and augmented set of learning samples.[121]

Several studies have focused on developing model-aggregation policies within a centralized architecture to address the challenges posed by non-independent and identically distributed (non-IID) data, as the quality of the federated learning model degrades if each federated learning node has a unique distribution of data.[122] To create a more generalizable aggregation policy capable of handling heterogeneous data, researchers have proposed the use of reinforcement learning,[123] contrastive learning,[124] and new optimization methods.[125]

Studies on federated learning in ophthalmology support global health collaboration and offer a promising approach to privacy-preserving AI research. Using federated learning, it is possible to train

DL-derived ROP that can identify differences in clinical diagnoses and disease severity across institutions without sharing data. Federated learning could standardize clinical diagnoses and provide objective measurements for image-based diseases.[126] Moreover, a trained federated learning model performs comparably to a centralized model in classification, confirming that federated learning may provide an effective, more feasible solution for interinstitutional learning. Smaller institutions benefit more from collaboration than larger institutions, showing the potential of federated learning for addressing disparities in resource access.[126,127]

In addition, federated learning can enhance domain diversity and generalizability of models using OCT and OCTA images. Federated learning models achieve a AUROC comparable to that of the traditional DL models for microvasculature segmentation and rDR classification.[128,129] Research on the feasibility of utilizing federated learning in identifying AMD demonstrated its practicality and benefits.[120] Different aggregation policies—FedAvg, FedProx,[129] FedMRI,[130] and APFL[131]—alongside deep learning networks such as ResNet and Vision Transformers, proved useful.

More research is needed on implementing federated learning in healthcare. These areas include addressing current barriers to applying federated learning, making it a key strategy for preserving privacy in AI health research,[132] and combining different data types for comprehensive disease diagnosis. There also exists a need to integrate federated learning with blockchain technology to enhance privacy, security, and efficiency.[133]

Conclusion

In summary, AI applications to the retinal field are myriad, ranging from screening and diagnosis to monitoring and predicting treatment response. AI holds the promise of streamlining the assessment of a patient's disease through rapid detection of biomarkers and computation of change over time. AI also improves the granularity of that assessment with quantitative data analysis of FAF, OCT, and OCTA parameters. AI will increase the personalization of treatment, enabling tailored treatment choices and treatment intervals that best address a particular patient's disease state. Not only medical retina patients, but also surgical retina patients, stand to benefit. In order to achieve this universality of benefits for all retina patients, worldwide collaboration remains paramount. It is crucial to include a diverse population—considering race, ethnicity, geography, and socioeconomic factors—in the training of AI models. Striving for widespread collaboration and a culture of inclusivity will help ensure the applicability of AI algorithms for all members of the world.

Funding

Supported by an unrestricted grant from Research to Prevent Blindness and University of Illinois at Chicago, Department of Ophthalmology Core Grant EY01792.

References

1. Gulshan V, Peng L, Coram M, et al. Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. *JAMA*. 2016;316:2402–2410.
2. U.S. Food and Drug Administration. De Novo Classification Request for IDx-DR. Available at: https://www.accessdata.fda.gov/cdrh_docs/reviews/DEN180001.pdf.
3. Tan TF, Thirunavukarasu AJ, Jin L, et al. Artificial intelligence and digital health in global eye health: opportunities and challenges. *Lancet Glob Health*. 2023;11.
4. Pietris J, Lam A, Bacchi S, et al. Health economic implications of artificial intelligence implementation for ophthalmology in Australia: a systematic review. *Asia Pac J Ophthalmol*. 2022;11:554–562.
5. Saeedi P, Petersohn I, Salpea P, et al. Global and regional diabetes prevalence estimates for 2019 and projections for 2030 and 2045: results from the International Diabetes Federation Diabetes Atlas, 9th edition. *Diabetes Res Clin Pract*. 2019;157, 107843.
6. Global Burden of Disease 2019 Blindness and Vision Impairment Collaborators. Vision Loss Expert Group of the Global Burden of Disease Study. Causes of blindness and vision impairment in 2020 and trends over 30 years, and prevalence of avoidable blindness in relation to VISION 2020: the Right to Sight: an analysis for the Global Burden of Disease Study. *Lancet Glob Health*. 2021;9:e144–e160.
7. Teo ZL, Tham YC, Yu M, et al. Global prevalence of diabetic retinopathy and projection of burden through 2045: systematic review and meta-analysis. *Ophthalmology*. 2021;128:1580–1591.
8. Gibson DM. Estimates of the percentage of US adults with diabetes who could be screened for diabetic retinopathy in primary care settings. *JAMA Ophthalmol*. 2019;137:440–444.
9. Resnikoff S, Lansingh VC, Washburn L, et al. Estimated number of ophthalmologists worldwide (International Council of Ophthalmology update): will we meet the needs? *Br J Ophthalmol*. 2020;104:588–592.
10. Bourne RRA, Flaxman SR, Braithwaite T, et al. Magnitude, temporal trends, and projections of the global prevalence of blindness and distance and near vision impairment: a systematic review and meta-analysis. *Lancet Glob Health*. 2017;5:e888–e897.
11. Gunasekeran DV, Wong TY. Artificial intelligence in ophthalmology in 2020: a technology on the cusp for translation and implementation. *Asia Pac J Ophthalmol*. 2020;9:61–66.
12. Abràmoff MD, Lavin PT, Birch M, et al. Pivotal trial of an autonomous AI-based diagnostic system for detection of diabetic retinopathy in primary care offices. *NPJ Digit Med*. 2018;1:39.
13. Shah A, Clarida W, Amelon R, et al. Validation of automated screening for referable diabetic retinopathy with an autonomous diagnostic artificial intelligence system in a Spanish population. *J Diabetes Sci Technol*. 2021;15:655–663.
14. Wolf RM, Liu TYA, Thomas C, et al. The SEE study: safety, efficacy, and equity of implementing autonomous artificial intelligence for diagnosing diabetic retinopathy in youth. *Diabetes Care*. 2021;44:781–787.
15. Ipp E, Liljenquist D, Bode B, et al. Pivotal evaluation of an artificial intelligence system for autonomous detection of referrable and vision-threatening diabetic retinopathy. *JAMA Netw Open*. 2021;4, e2134254.
16. Lim JI, Regillo CD, Sadda SR, et al. Artificial intelligence detection of diabetic retinopathy: subgroup comparison of the EyeArt system with ophthalmologists' dilated examinations. *Ophthalmol Sci*. 2022;3, 100228.
17. Ting DSW, Cheung CYL, Lim G, et al. Development and validation of a deep learning system for diabetic retinopathy and related eye diseases using retinal images from multiethnic populations with diabetes. *JAMA*. 2017;318:2211–2223.
18. Wong DCS, Kiew G, Jeon S, et al. Singapore eye lesions analyzer (SELENA): the deep learning system for retinal diseases. In: Grzybowski A, ed. *Artificial intelligence in ophthalmology*. Cham: Springer; 2021:177–185. https://doi.org/10.1007/978-3-030-78601-4_13.
19. Xie Y, Nguyen QD, Hamzah H, et al. Artificial intelligence for teleophthalmology-based diabetic retinopathy screening in a national program: an economic analysis modelling study. *Lancet Digit Health*. 2020;2:e240–e249.
20. Dismuke C. Comment: progress in examining cost-effectiveness of AI in diabetic retinopathy screening. *Lancet Digit Health*. 2020;2:e212–e213.
21. Widner K, Virmani S, Krause J, et al. Lessons learned from translating AI from development to deployment in healthcare. *Nat Med*. 2023;29:1304–1306.
22. Mathenge W, Whitestone N, Nkurikiye J, et al. Impact of artificial intelligence assessment of diabetic retinopathy on referral service uptake in a low-resource setting: The RAIDERS randomized trial. *Ophthalmol Sci*. 2022;2, 100168.
23. Natarajan S, Jain A, Krishnan R, et al. Diagnostic accuracy of community-based diabetic retinopathy screening with an offline artificial intelligence system on a smartphone. *JAMA Ophthalmol*. 2019;137:1182–1188.
24. Malerbi FK, Andrade RE, Morales PH, et al. Diabetic retinopathy screening using artificial intelligence and handheld smartphone-based retinal camera. *J Diabetes Sci Technol*. 2022;16:716–723.
25. Rogers TW, Gonzalez-Bueno J, Garcia Franco R, et al. Evaluation of an AI system for the detection of diabetic retinopathy from images captured with a handheld portable fundus camera: the MAILOR AI study. *Eye*. 2021;35:632–638.
26. Health AEYE. *AEYE health receives FDA clearance for AI-based autonomous screening for referable diabetic retinopathy*. PR Newswire; 2022. https://www.prnewswire.com/news-releases/aeye-health-receives-fda-clearance-for-ai-based-autonomous-screening-for-referable-diabetic-retinopathy-301678515.html?tc=eml.
27. Channa R, Wolf RM, Abràmoff MD, et al. Effectiveness of artificial intelligence screening in preventing vision loss from diabetes: a policy model. *NPJ Digit Med*. 2023;6:53.
28. Huang XM, Yang BF, Zheng WL, et al. Cost-effectiveness of artificial intelligence screening for diabetic retinopathy in rural China. *BMC Health Serv Res*. 2022;22:260.
29. Alam M, Toslak D, Lim JI, et al. Color fundus image guided artery-vein differentiation in optical coherence tomography angiography. *Investig Ophthalmol Vis Sci*. 2018;59:4953–4962.
30. Son T, Alam M, Liu C, et al. Artery and vein differentiation in retinal optical coherence tomography angiography of macular region. *Ophthalmic Technol XXIX*. 2019;10858.
31. Alam M, Son T, Toslak D, et al. Automated classification and quantitative analysis of arterial and venous vessels in fundus images. *Proc SPIE Int Soc Opt Eng*. 2018;10474:1047426.
32. Alam M, Lim JI, Toslak D, et al. Differential artery-vein analysis improves the performance of OCTA staging of sickle cell retinopathy. *Transl Vis Sci Technol*. 2019;8:3.

33. Abtahi M, Le D, Lim JI, et al. MF-AV-Net: an open-source deep learning network with multimodal fusion options for artery-vein segmentation in OCT angiography. *Biomed Opt Express*. 2022;13:4870–4888.
34. Alam M, Le D, Son T, et al. AV-Net: deep learning for fully automated artery-vein classification in optical coherence tomography angiography. *Biomed Opt Express*. 2020;11:5249–5257.
35. Dow ER, Keenan TDL, Lad EM, et al. From data to deployment: the collaborative community on ophthalmic imaging roadmap for artificial intelligence in age-related macular degeneration. *Ophthalmology*. 2022;129:e43–e59.
36. Domalpally A. Benchmark dataset for age-related macular degeneration. The Collaborative Community on Ophthalmic Innovation (CCOI) 2024 Conference Program. Available at: https://cc-oi.org/page/Program2024.
37. Peng Y, Dharssi S, Chen Q, et al. DeepSeeNet: a deep learning model for automated classification of patient-based age-related macular degeneration severity from color fundus photographs. *Ophthalmology*. 2019;126:565–575.
38. Keenan TD, Dharssi S, Peng Y, et al. A deep learning approach for automated detection of geographic atrophy from color fundus photographs. *Ophthalmology*. 2019;126:1533–1540.
39. Arslan J, Samarasinghe G, Sowmya A, et al. Deep learning applied to automated segmentation of geographic atrophy in fundus autofluorescence images. *Transl Vis Sci Technol*. 2021;10:2.
40. Kermany DS, Goldbaum M, Cai W, et al. Identifying medical diagnoses and treatable diseases by image-based deep learning. *Cell*. 2018;172:1122–1131.
41. Lee CS, Baughman DM, Lee AY. Deep learning is effective for the classification of OCT images of normal versus age-related macular degeneration. *Ophthalmol Retin*. 2017;1:322–327.
42. Campbell JP, Chiang MF, Chen JS, et al. Artificial intelligence for retinopathy of prematurity: validation of a vascular severity scale against international expert diagnosis. *Ophthalmology*. 2022;129:e69–e76.
43. Bai A, Dai S, Hung J, et al. Multicenter validation of deep learning algorithm ROP. AI for the automated diagnosis of plus disease in ROP. *Transl Vis Sci Technol*. 2023;12:13.
44. Coyner AS, Murickan T, Oh MA, et al. Multinational external validation of autonomous retinopathy of prematurity screening. *JAMA Ophthalmol*. 2024;142:327–335.
45. Nguyen TTP, Young BK, Coyner A, et al. Discrepancies in diagnosis of treatment-requiring retinopathy of prematurity. *Ophthalmol Retin*. 2024;8:88–91.
46. Moshfegh DM. Artificial intelligence poised to improve retinopathy of prematurity screening. *Ophthalmol Retin*. 2024;8:1–2.
47. Cai S, Parker F, Urias MG, et al. Deep learning detection of sea fan neovascularization from ultra-widefield color fundus photographs of patients with sickle cell hemoglobinopathy. *JAMA Ophthalmol*. 2021;139:206–213.
48. Alam M, Thapa D, Lim JI, et al. Computer-aided classification of sickle cell retinopathy using quantitative features in optical coherence tomography angiography. *Biomed Opt Express*. 2017;8:4206–4216.
49. Alam M, Thapa D, Lim JI, et al. Quantitative characteristics of sickle cell retinopathy in optical coherence tomography angiography. *Biomed Opt Express*. 2017;8:1741–1753.
50. Alam M, Le D, Lim JI, et al. Supervised machine learning based multi-task artificial intelligence classification of retinopathies. *J Clin Med*. 2019;8:872.
51. Xu Z, Wang W, Yang J, et al. Automated diagnoses of age-related macular degeneration and polypoidal choroidal vasculopathy using bi-modal deep convolutional neural networks. *Br J Ophthalmol*. 2021;105:561–566.
52. Rosenfeld PJ. Optical coherence tomography and the development of antiangiogenic therapies in neovascular age-related macular degeneration. *Investig Ophthalmol Vis Sci*. 2016;57:OCT14–OCT26.
53. Toth CA, Decroos FC, Ying GS, et al. Identification of fluid on optical coherence tomography by treating ophthalmologists versus a reading center in the comparison of age-related macular degeneration treatments trials. *Retina*. 2015;35:1303–1314.
54. DeCroos FC, Toth CA, Stinnett SS, et al. Optical coherence tomography grading reproducibility during the comparison of age-related macular degeneration treatments trials. *Ophthalmology*. 2012;119:2549–2557.
55. Schmidt-Erfurth U, Reiter GS, Riedl S, et al. AI-based monitoring of retinal fluid in disease activity and under therapy. *Prog Retin Eye Res*. 2022;86, 100972.
56. Evans RN, Reeves BC, Maguire MG, et al. Associations of variation in retinal thickness with visual acuity and anatomic outcomes in eyes with neovascular age-related macular degeneration lesions treated with anti–vascular endothelial growth factor agents. *JAMA Ophthalmol*. 2020;138:1043–1051.
57. Schlegl T, Waldstein SM, Bogunovic H, et al. Fully automated detection and quantification of macular fluid in OCT using deep learning. *Ophthalmology*. 2018;125:549–558.
58. Keenan TDL, Chakravarthy U, Loewenstein A, et al. Automated quantitative assessment of retinal fluid volumes as important biomarkers in neovascular age-related macular degeneration. *Am J Ophthalmol*. 2021;224:267–281.
59. Pawloff M, Bogunovic H, Gruber A, et al. Systematic correlation of central subfield thickness with retinal fluid volumes quantified by deep learning in the major exudative macular diseases. *Retina*. 2022;42:831–841.
60. Simader C, Ritter M, Bolz M, et al. Morphologic parameters relevant for visual outcome during anti-angiogenic therapy of neovascular age-related macular degeneration. *Ophthalmology*. 2014;121:1237–1245.
61. Waldstein SM, Wright J, Warburton J, et al. Predictive value of retinal morphology for visual acuity outcomes of different ranibizumab treatment regimens for neovascular AMD. *Ophthalmology*. 2016;123:60–69.
62. Sharma S, Toth CA, Daniel E, et al. Macular morphology and visual acuity in the second year of the comparison of age-related macular degeneration treatments trials. *Ophthalmology*. 2016;123:865–875.
63. Arnold JJ, Markey CM, Kurstjens NP, et al. The role of sub-retinal fluid in determining treatment outcomes in patients with neovascular age-related macular degeneration–a phase IV randomised clinical trial with ranibizumab: the FLUID study. *BMC Ophthalmol*. 2016;16:31.
64. Grechenig C, Reiter GS, Riedl S, et al. Impact of residual subretinal fluid volumes on treatment outcomes in a subretinal fluid–tolerant treat-and-extend regimen. *Retina*. 2021;41:2221–2228.
65. Schreur V, de Breuk A, Venhuizen FG, et al. Retinal hyperreflective foci in type 1 diabetes mellitus. *Retina*. 2020;40:1565–1573.
66. Moraes G, Fu DJ, Wilson M, et al. Quantitative analysis of OCT for neovascular age-related macular degeneration using deep learning. *Ophthalmology*. 2021;128:693–705.
67. Kapoor R, Walters SP, Al-Aswad LA. The current state of artificial intelligence in ophthalmology. *Surv Ophthalmol*. 2019;64:233–240.
68. Mares V, Nehemy MB, Bogunovic H, et al. AI-based support for optical coherence tomography in age-related macular degeneration. *Int J Retin Vitr*. 2024;10:31.
69. Waisberg E, Ong J, Kamran SA, et al. Generative artificial intelligence in ophthalmology. *Surv Ophthalmol*. 2024;S0039-6257(24).
70. Keenan TDL, Clemons TE, Domalpally A, et al. Retinal specialist versus artificial intelligence detection of retinal fluid from OCT: age-related eye disease study 2: 10-year follow-on study. *Ophthalmology*. 2021;128:100–109.
71. De Fauw J, Ledsam JR, Romera-Paredes B, et al. Clinically applicable deep learning for diagnosis and referral in retinal disease. *Nat Med*. 2018;24:1342–1350.
72. Holz FG, Abreu-Gonzalez R, Bandello F, et al. Does real-time artificial intelligence-based visual pathology enhancement of three-dimensional optical coherence tomography scans optimise treatment decision in patients with nAMD? Rationale and design of the RAZORBILL study. *Br J Ophthalmol*. 2023;107:96–101.
73. Pawloff M, Gerendas BS, Deak G, et al. Performance of retinal fluid monitoring in OCT imaging by automated deep learning versus human expert grading in neovascular AMD. *Eye*. 2023;37:3793–3800.
74. Kim JE, Tomkins-Netzer O, Elman MJ, et al. Evaluation of a self-imaging SD-OCT system designed for remote home monitoring. *BMC Ophthalmol*. 2022;22:261.
75. Liu Y, Holekamp NM, Heier JS. Prospective, longitudinal study: daily self-imaging with home OCT for neovascular age-related macular degeneration. *Ophthalmol Retin*. 2022;6:575–585.
76. Blinder KJ, Calhoun C, Maguire MG, et al. Home OCT imaging for newly diagnosed neovascular age-related macular degeneration: a feasibility study. *Ophthalmol Retin*. 2024;8:376–387.
77. Saha S, Nassisi M, Wang M, et al. Automated detection and classification of early AMD biomarkers using deep learning. *Sci Rep*. 2019;9, 10990.
78. Schmidt-Erfurth U, Sadeghipour A, Gerendas BS, et al. Artificial intelligence in retina. *Prog Retin Eye Res*. 2018;67:1–29.
79. Peng Y, Keenan TD, Chen Q, et al. Predicting risk of late age-related macular degeneration using deep learning. *NPJ Digit Med*. 2020;3:111.
80. Yim J, Chopra R, Spitz T, et al. Predicting conversion to wet age-related macular degeneration using deep learning. *Nat Med*. 2020;26:892–899.
81. Mai J, Lachinov D, Reiter GS, et al. Deep learning-based prediction of individual geographic atrophy progression from a single baseline OCT. *Ophthalmol Sci*. 2024;4, 100466.
82. Bogunovic H, Waldstein SM, Schlegl T, et al. Prediction of anti-VEGF treatment requirements in neovascular AMD using a machine learning approach. *Investig Ophthalmol Vis Sci*. 2017;58:3240–3248.
83. Crincoli E, Sacconi R, Querques L, et al. Artificial intelligence in age-related macular degeneration: state of the art and recent updates. *BMC Ophthalmol*. 2024;24:121.
84. Pfau M, Sahu S, Rupnow RA, et al. Probabilistic forecasting of anti-VEGF treatment frequency in neovascular age-related macular degeneration. *Transl Vis Sci Technol*. 2021;10:30.
85. Holz FG, Tadayoni R, Beatty S, et al. Multi-country real-life experience of anti-vascular endothelial growth factor therapy for wet age-related macular degeneration. *Br J Ophthalmol*. 2015;99:220–226.
86. Gillies MC, Campain A, Barthelmes D, et al. Long-term outcomes of treatment of neovascular age-related macular degeneration: data from an observational study. *Ophthalmology*. 2015;122:1837–1845.
87. Khanani AM, Skelly A, Bezlyak V, et al. SIERRA-AMD: a retrospective, real-world evidence study of patients with neovascular age-related macular degeneration in the United States. *Ophthalmol Retin*. 2020;4:122–133.
88. Chandra RS, Ying GS. Evaluation of multiple machine learning models for predicting number of Anti-VEGF injections in the comparison of AMD treatment trials (CATT). *Transl Vis Sci Technol*. 2023;12:18.
89. Gallardo M, Munk MR, Kurmann T, et al. Machine learning can predict anti-VEGF treatment demand in a treat-and-extend regimen for patients with neovascular AMD, DME, and RVO associated macular edema. *Ophthalmol Retin*. 2021;5:604–624.
90. Duan T, Anand A, Ding DY, et al. NGBoost: natural gradient boosting for probabilistic prediction. Proceedings of the 37th international conference on machine learning. *Proc Mach Learn Res*. 2020;119:2690–2700.
91. Romo-Bucheli D, Erfurth US, Bogunović H. End-to-end deep learning model for predicting treatment requirements in neovascular AMD from longitudinal retinal OCT imaging. *IEEE J Biomed Health Inf*. 2020;24:3456–3465.
92. Rohm M, Tresp V, Müller M, et al. Predicting visual acuity by using machine learning in patients treated for neovascular age-related macular degeneration. *Ophthalmology*. 2018;125:1028–1036.

93. Fu DJ, Faes L, Wagner SK, et al. Predicting incremental and future visual change in neovascular age-related macular degeneration using deep learning. *Ophthalmol Retin*. 2021;5:1074–1084.
94. Yeh TC, Luo AC, Deng YS, et al. Prediction of treatment outcome in neovascular age-related macular degeneration using a novel convolutional neural network. *Sci Rep*. 2022;12:5871.
95. Moon S, Lee Y, Hwang J, et al. Prediction of anti-vascular endothelial growth factor agent-specific treatment outcomes in neovascular age-related macular degeneration using a generative adversarial network. *Sci Rep*. 2023;13:5639.
96. Mares V, Schmidt-Erfurth UM, Leingang O, et al. Approved AI-based fluid monitoring to identify morphological and functional treatment outcomes in neovascular age-related macular degeneration in real-world routine. *Br J Ophthalmol*. 2024;108:971–977.
97. Zhao X, Zhang X, Lv B, et al. Optical coherence tomography-based short-term effect prediction of anti-vascular endothelial growth factor treatment in neovascular age-related macular degeneration using sensitive structure guided network. *Graefes Arch Clin Exp Ophthalmol*. 2021;259:3261–3269.
98. Liu Y, Yang J, Zhou Y, et al. Prediction of OCT images of short-term response to anti-VEGF treatment for neovascular age-related macular degeneration using generative adversarial network. *Br J Ophthalmol*. 2020;104:1735–1740.
99. Nuliqiman M, Xu M, Sun Y, et al. Artificial intelligence in ophthalmic surgery: current applications and expectations. *Clin Ophthalmol*. 2023;17:3499–3511.
100. Hu Y, Xiao Y, Quan W, et al. A multi-center study of prediction of macular hole status after vitrectomy and internal limiting membrane peeling by a deep learning model. *Ann Transl Med*. 2021;9:51.
101. Xiao Y, Hu Y, Quan W, et al. Development and validation of a deep learning system to classify aetiology and predict anatomical outcomes of macular hole. *Br J Ophthalmol*. 2023;107:109–115.
102. Guo H, Ou C, Wang G, et al. Prediction of visual outcome after rhegmatogenous retinal detachment surgery using artificial intelligence techniques. *Transl Vis Sci Technol*. 2024;13:17.
103. Mishra K, Leng T. Artificial intelligence and ophthalmic surgery. *Curr Opin Ophthalmol*. 2021;32:425–430.
104. Poh SSJ, Sia JT, Yip MYT, et al. Artificial intelligence, digital imaging, and robotics technologies for surgical vitreoretinal diseases. *Ophthalmol Retin*. 2024;8:633–645.
105. Channa R, Iordachita I, Handa JT. Robotic vitreoretinal surgery. *Retina*. 2017;37:1220–1228.
106. He C, Patel N, Shahbazi M, et al. Toward safe retinal microsurgery: development and evaluation of an RNN-based active interventional control framework. *IEEE Trans Biomed Eng*. 2020;67:966–977.
107. Ramamurthy SR, Dave VP. Robotics in vitreo-retinal surgery. *Semin Ophthalmol*. 2022;37:795–800.
108. Bakshi SK, Lin SR, Ting DSW, et al. The era of artificial intelligence and virtual reality: transforming surgical education in ophthalmology. *Br J Ophthalmol*. 2021;105:1325–1328.
109. Nespolo RG, Yi D, Cole E, et al. Feature tracking and segmentation in real time via deep learning in vitreoretinal surgery: a platform for artificial intelligence-mediated surgical guidance. *Ophthalmol Retin*. 2023;7:236–242.
110. Edwards TL, Xue K, Meenink HCM, et al. First-in-human study of the safety and viability of intraocular robotic surgery. *Nat Biomed Eng*. 2018;2:649–656.
111. McMahan HB, Moore E, Ramage D, et al. Communication-efficient learning of deep networks from decentralized data. In: Proceedings of the 20th international conference on artificial intelligence and statistics (AISTATS). Fort Lauderdale, Florida, United States; 2017.
112. Clark P, Oermann EK, Chen D, et al. Federated AI, current state, and future potential. *Asia Pac J Ophthalmol*. 2023;12:310–314.
113. Teo ZL, Lee AY, Campbell P, et al. Developments in artificial intelligence for ophthalmology: federated learning. *Asia Pac J Ophthalmol*. 2022;11:500–502.
114. Rieke N, Hancox J, Li W, et al. The future of digital health with federated learning. *NPJ Digit Med*. 2020;3:119.
115. Xu Y, Ma L, Yang F, et al. A collaborative online AI engine for CT-based COVID-19 diagnosis. *medRxiv*. 2020.
116. Li Q, Wen Z, Wu Z, et al. A survey on federated learning systems: vision, hype and reality for data privacy and protection. *IEEE Trans Knowl Data Eng*. 2023;35:3347–3366.
117. Roth HR, Cheng Y, Wen Y, et al. Nvidia flare: federated learning from simulation to real-world. *arXiv*. 2022. Available at: https://arxiv.org/abs/2210.13291.
118. Beutel DJ, Topal T, Mathur A, et al. Flower: a friendly federated learning research framework. *arXiv*. 2020. Available at: https://arxiv.org/abs/2007.14390.
119. Yan B, Cao D, Jiang X, et al. FedEYE: a scalable and flexible end-to-end federated learning platform for ophthalmology. *Patterns*. 2024;5, 100928.
120. Ng D, Lan X, Yao MMS, et al. Federated learning: a collaborative effort to achieve better medical imaging models for individual sites that have small labelled datasets. *Quant Imaging Med Surg*. 2021;11:852–857.
121. Xu J, Glicksberg BS, Su C, et al. Federated learning for healthcare informatics. *J Health Inf Res*. 2021;5:1–19.
122. Zhao Y, Li M, Lai L, et al. Federated learning with Non-IID data. *arXiv*. 2018. Available at: https://arxiv.org/abs/1806.00582.
123. Wang H, Kaplan Z, Niu D, et al. Optimizing federated learning on non-IID data with reinforcement learning. In: Proceedings of the IEEE INFOCOM 2020 – IEEE conference on computer communications. Toronto, Canada.
124. Li Q, He B, Song D. Model-contrastive federated learning. *Proc IEEE/CVF Conf Comput Vis Pattern Recognit*. 2021.
125. Karimireddy SP, Kale S, Mohri M, et al. SCAFFOLD: stochastic controlled averaging for federated learning. Proceedings of the 37th international conference on machine learning. *Proc Mach Learn Res*. 2020;119:5132–5143.
126. Hanif A, Lu C, Chang K, et al. Federated learning for multicenter collaboration in ophthalmology: implications for clinical diagnosis and disease epidemiology. *Ophthalmol Retin*. 2022;6:650–656.
127. Lu C, Hanif A, Singh P, et al. Federated learning for multicenter collaboration in ophthalmology: improving classification performance in retinopathy of prematurity. *Ophthalmol Retin*. 2022;6:657–663.
128. Lo J, Yu TT, Ma D, et al. Federated learning for microvasculature segmentation and diabetic retinopathy classification of OCT data. *Ophthalmol Sci*. 2021;1, 100069.
129. Gholami S, Lim JI, Leng T, et al. Federated learning for diagnosis of age-related macular degeneration. *Front Med*. 2023;10, 1259017.
130. Feng CM, Yan Y, Wang S, et al. Specificity-preserving federated learning for MR image reconstruction. *IEEE Trans Med Imaging*. 2023;42:2010–2021.
131. Li T, Sahu AK, Zaheer M, et al. Federated optimization in heterogeneous networks. In: Proceedings of the 3rd MLSys conference. Austin, Texas, United States; 2020.
132. Teo ZL, Lee AY, Campbell P, et al. Developments in artificial intelligence for ophthalmology: federated learning. *Asia Pac J Ophthalmol*. 2022;11:500–502.
133. Nguyen TX, Ran AR, Hu X, et al. Federated learning in ocular imaging: current progress and future direction. *Diagnostics*. 2022;12:2835.