Review Article

Using artificial intelligence for diabetic retinopathy screening: Policy implications

Rajiv Raman, Debarati Dasgupta, Kim Ramasamy[1], Ronnie George[2], Viswanathan Mohan[3], Daniel Ting[4]

Artificial intelligence (AI) has evolved over the last few years; its use in DR screening has been demonstrated in multiple evidences across the globe. However, there are concerns right from the data acquisition, bias in data, difficulty in comparing between different algorithm, challenges in machine learning, its application in different group of population, and human barrier to AI adoption in health care. There are also legal and ethical concerns related to AI. The tension between risks and concerns on one hand versus potential and opportunity on the other have driven a need for authorities to implement policies for AI in DR screening to address these issues. The policy makers should support and facilitate research and development of AI in healthcare, but at the same time, it has to be ensured that the use of AI in healthcare aligns with recognized standards of safety, efficacy, and equity. It is essential to ensure that algorithms, datasets, and decisions are auditable and when applied to medical care (such as screening, diagnosis, or treatment) are clinically validated and explainable. Policy frameworks should require design of AI systems in health care that are informed by real-world workflow and human-centric design. Lastly, it should be ensured that healthcare AI solutions align with all relevant ethical obligations, from design to development to use and to be delivered properly in the real world.

Access this article online

Website: www.ijo.in

DOI: 10.4103/ijo.IJO_1420_21

Quick Response Code:

Key words: Artificial intelligence, diabetic retinopathy, machine learning, policy implications

Diabetes is a global epidemic that results in a heavy health burden to individuals and societies across the world. Diabetic retinopathy (DR) is one of the major causes of avoidable blindness, but the key challenges in heavy populated countries to address DR include a lack of symptoms until the disease has progressed to vision loss.[1,2] Artificial Intelligence (AI) is a branch of computer science in which machines mimic the cognitive function of human mind. Over the past decade, AI has made a breakthrough progress in medical communities especially in ophthalmology.[3,4] Machine learning (ML) and deep learning (DL) algorithms empower computers to diagnose or manage without direct human intervention, by extracting clinically relevant information from medical data. Multiple studies have shown that DL can be leveraged to produce expert-level diagnoses for grading fundus photography images in DR.[5-7] AI systems deployed for DR screening must aim to improve overall population health by reducing the burden of visual impairment; improving DR outcomes and patient satisfaction; lowering overall screening costs; and assisting the ophthalmologist to reduce their workloads.[5] AI systems should also enhance access to eye care, empower patients with diabetes to manage and optimize their health, facilitate and strengthen the relationship and communication that individuals have with their eye care team, and reduce the administrative and cognitive burdens for patients and their eye care team.

Although the potential of AI is profound, there are however unintended, negative consequences. Since AI is designed to accomplish a very specific task on previously curated data from one setting, they have a rather narrow intelligence.[8-10] Most AI models are built on correlations, but due to diverse populations studied, the health data predictions could fail to generalize to different populations or settings and might exacerbate existing inequalities and biases.[11] This perspective will highlight few salient features related to use of AI in DR screening and their policy implications.

AI for DR: Screening vs Clinical Care

In the case of DR, preventive care focuses on screening of retinal images to evaluate the presence or absence of DR or sight-threatening DR. It is part of routine annual diabetes preventive care. Diagnostic care involves treating or investigating sight-threatening DR. It may include treatment for diabetic macular edema or proliferative diabetic retinopathy (PDR), ongoing care of patients with DR, and multimodal imaging tests needed to manage or treat DR.

Though there are emerging roles of AI in multimodal imaging, treatment response prediction, the current utility of AI in DR is limited to preventive care, i.e., in screening.[12] Recently, the American Diabetes Association has outlined the potential use of AI for DR screening.[13] They suggest that AI systems that detect

Sri Bhagwan Mahavir Vitreoretinal Services, Sankara Nethralaya, [2]Glaucoma Services, Sankara Nethralaya, [3]Madras Diabetes Research Foundation and Dr. Mohan's Diabetes Specialities Centre, Chennai, [1]Aravind Eye Care Hospital, Madurai, Tamil Nadu, India, [4]Duke-NUS Medical School Consultant, Vitreo-Retinal Department, Singapore National Eye Center, Singapore

Correspondence to: Dr. Rajiv Raman, Sankara Nethralaya, 18, College Road, Chennai - 600 006, Tamil Nadu, India. E-mail: drrrn@snmail.org

Received: 29-May-2021 Revision: 01-Sep-2021
Accepted: 16-Sep-2021 Published: 29-Oct-2021

This is an open access journal, and articles are distributed under the terms of the Creative Commons Attribution-NonCommercial-ShareAlike 4.0 License, which allows others to remix, tweak, and build upon the work non-commercially, as long as appropriate credit is given and the new creations are licensed under the identical terms.

For reprints contact: WKHLRPMedknow_reprints@wolterskluwer.com

Cite this article as: Raman R, Dasgupta D, Ramasamy K, George R, Mohan V, Ting D. Using artificial intelligence for diabetic retinopathy screening: Policy implications. Indian J Ophthalmol 2021;69:2993-8.

more than mild DR and diabetic macular edema authorized for use by the US Food and Drug Administration (FDA) represent an alternative to traditional screening approaches. AI systems should not be used for patients with known DR, prior DR treatment, or symptoms of vision impairment. Results of eye examinations should be documented and transmitted to the referring health care professional.

AI for DR: Current status

There are several studies using retinal images to test the performance of AI grading systems for detecting DR. AI systems have been found to lower cost, improve diagnostic accuracy, and increase patient access to DR screening.[14]

Studies suggest that iGrading (version 1.1 by Medalytix) and other commercial automated grading systems including Retmarker (version 0.8.2. 2014/02/10 by Critical-Health) and IDx-DR (by IDx) are comparable to that of trained graders.[15] Tufail *et al*.[16] evaluated retinal images from 20,258 patients in routine annual DR screening showed 85% sensitivity by Retmarker and 94% by EyeArt, indicating the potential to replace one or more steps of current DR screening programs.[16] A deep-learning enhanced algorithm (AlexNet/VGGNet) for the automated detection of DR was reported with sensitivity of 96.8%, specificity of 87.0%, and the AUC of 0.980, which is better than the non-DL algorithm system previously reported.[17]

Another major development for DR screening is the DL system (Inception-V3) reported by Gulshan *et al*.;[18] Google AI used a large training dataset of 128,175 and two separate publicly available datasets to test and validate the model. It showed a high sensitivity (≥96%), specificity (≥93%), and area under the receiver operating characteristic curve (AUC) (≥0.99) in the external validation using two public databases.[7] Ting *et al*.[5] developed the VGGNet-based DL system for referable DR, vision-threatening DR, and related eye diseases. AUCs of referable DR ranged from 0.889 to 0.983 with sensitivities varying from 92 to 100% and specificities from 73 to 92% when tested on 10 external datasets with 112,648 images from diverse populations including Chinese, Indian, Malay, Caucasian, Hispanic, and African-American ethnicities.

In April 2018, the US FDA approved an AI algorithm, developed by IDx, used with Topcon Fundus camera (Topcon Medical) for DR identification.[19] More recently in 2020, FDA gave clearance to another algorithm, EyeArt system for DR screening. The technology displayed 96% sensitivity and 88% specificity for the detection of more than minor DR, and 92% sensitivity and 94% specificity for the detection of vision-threatening DR.[20]

Abramoff *et al*.[6] reported the first DL-enhanced algorithm for referable DR and VTDR, which had a sensitivity of 87.2% and a specificity of 90.7% in detection of referable DR (worse than mild DR), and gradeability rate of 96.1%. Kanagasingam *et al*.[21] *described* the performance of a DL algorithm on multiple training dataset including DiaRetDB1, EyePACS, and the Australian Tele-eye care DR database. The outcome showed that the specificity was 92%, and the positive predictive value was 12%. This is a study evaluating an AI-based grading system for DR perform in a real-world setting and showing potential benefits and further deployments in primary care are needed. Li *et al*.[22] developed a DL system to automatically detect the most common sign of DR, retinal hemorrhages, based on 16827 ultra-wide-field fundus (UWF) images.

Multiple studies reported their ability to detect referable DR in automatic DR screening using AI, ranging from 87.2 to 97.5% in sensitivity, and from 0.936 to 0.991 in AUC. Moreover, the probability of missing severe nonproliferative diabetic retinopathy, PDR, or macular edema of current DL algorithms is less than 1%.[23,24] Even experienced ophthalmologists need decades of extensive training to provide high-quality healthcare. Therefore, AI-based DR screening algorithms have the potential to reach or even outperform clinical experts and provide healthcare to large populations, especially in those less-developed areas.[24,25]

In India, telemedicine-based DR screening programs are growing and providing valuable clinical benefit. The Google AI team is working with eye specialists from Aravind Eye Hospital (Madurai) and Sankara Nethralaya (Chennai) on the development of their AI DL system, with the goal of making such technologies available to everyone. Table 1 describes few of the DL-based studies for DR screening.

Aaron *et al*. reported an independent, external, head-to-head automated DR screening algorithm validation study comparing seven AI systems for DR screening and found that the screening performance of state-of-the-art algorithms varied considerably, with substantial differences in overall performance. The sensitivities varied widely (50.98–85.90%) between the algorithms. This study results indicate that there is a need of rigorous testing of all such algorithms on real-world data before clinical implementation.[30]

Challenges in the use of AI for DR screening

Retrospective versus prospective studies versus randomized controlled trials

While existing studies in DR have very large numbers of patients with extensive benchmarking against expert performance, the vast majority of studies have been retrospective, i.e., they use historically labelled data to train and test algorithms. Only through prospective studies we will begin to understand the true utility of AI systems, as performance is likely to be worse when encountering real-world data that differ from that encountered in algorithm training. There are very few randomized controlled trials of AI systems for DR to date. Future studies should aim to use clinical outcomes as trial endpoints to demonstrate longer-term benefit.[31-33]

Metrics do not reflect clinical applicability

The term "AI chasm" has been coined to reflect the fact that accuracy does not necessarily represent clinical efficacy. The AUC of a receiver operating characteristic curve is not necessarily the best metric to represent clinical applicability and is not easily understandable by many clinicians.

Difficulty comparing different algorithms

The comparison of algorithms across studies in an objective manner is challenging as each study's performance is being reported using variable methodologies on different populations.

Challenges related to machine learning science

AI algorithms have the potential of having shortcomings, including inapplicability outside of the training domain and bias. Other factors for consideration include dataset shift, accidentally fitting confounders rather than true signal, and the challenge of generalization to different populations.

Table 1: Deep learning studies for diabetic retinopathy detection

Year	Authors	Dataset Images for Training and Testing/Validation	Reported Outcome	Positive predictive value*	Negative Predictive Value*
2016	Abràmoff *et al*.[6]	Fundus photos 25,000 training, 874 validation	Sensitivity: 96.8% Specificity: 87.0% AUC: 0.980	28%	99.8%
2016	Gulshan *et al*.[7]	Fundus photos 128,175 training Validation: 9963 (EyePACS), 1748 (Messidor)	EyePACS Sensitivity: 97.5% Specificity: 93.4%, AUC: 0.991 Messidor Sensitivity: 96.1% Specificity: 93.9%, AUC: 0.990	44% 45%	99.8% 99.7%
2017	Gargeya *et al*.[26]	Fundus photos 75,137 training 15,000 validation (mixed sources)	Sensitivity: 94% Specificity: 98% AUC: 0.94-0.97	71%	99.6%
2017	Ting *et al*.[5]	Fundus photos 76,370 training Validation: 71,896 images of 14,880 patients	Sensitivity: 90.5% Specificity: 91.6% AUC: 0.936	36%	99.4%
2018	Ramachandran *et al*.[27]	Fundus photos >100,000 training Validation: 485 (Otago), 1200 (Messidor)	Otago Sensitivity: 84.6% Specificity: 79.7% AUC: 0.901 Messidor Sensitivity: 96.0% Specificity: 90.0% AUC: 0.980	18% 34%	98.9% 99.7%
2018	Abràmoff *et al*.[28]	Fundus photos 900 participants for validation	Sensitivity: 87.2% Specificity: 90.7%	33%	99.2%
2019	Bhaskaranand *et al*.[29]	Fundus photos 850,908 images from 101,710 consecutive patient visits	Sensitivity: 91.3% Specificity: 91.1% AUC: 0.965	35%	99.5%
2019	Gulshan *et al*.[18]	Fundus photos 103,634 training 5762 images from 3049 patients at two tertiary sites for validation	Aravind Eye Hospital Sensitivity: 88.9% Specificity: 92.2% AUC: 0.963 Sankara Nethralaya Sensitivity: 92.1% Specificity: 95.2% AUC: 0.980	37% 50%	99.3% 99.5%

*Assuming 5% of the population has sight-threatening diabetic retinopathy

Challenges in generalization to new populations and settings

The majority of AI systems for DR are far from achieving reliable generalizability due to differences in populations to be screened, fundus cameras and skills of photographers. Proper assessment of real-world clinical performance and generalization requires appropriately designed external validation involving testing of an AI system using adequately sized datasets collected from institutions other than those that provided the data for model training.

Algorithmic bias

Algorithmic bias can be divided in three components: model bias (i.e., models selected to best represent the majority), model variance (due to inadequate data from minorities), and outcome noise.

Logistical difficulties in implementing AI systems

Many of the current challenges in translating AI algorithms to clinical practice are related to the fact that most healthcare data are not readily available for ML. DR screening variations including the variations in settings of screening cause logistic difficulties in AI implementation.

Achieving robust regulation and rigorous quality control

An important component to achieving safe and effective deployment of AI algorithms is the development of necessary regulatory frameworks. This poses a challenge given the current pace of innovation and significant risks involved. It is also important to consider the regulatory impact of improvements and upgrades that providers of AI products are likely to develop throughout the life of the product.

Human barriers to AI adoption in healthcare

In order to ensure that this technology can reach and benefit patients, it will be important to maintain a focus on clinical applicability and patient outcomes.

Algorithmic interpretability is at an early stage but rapidly advancing

The effectiveness of AI in DR screening is limited by their inability to "explain" their decision making in an understandable way, though there are some progresses, as some of the algorithms also show the heat maps to explain why the algorithm classified it as DR

Developing a better understanding of interaction between human and algorithm

We have a limited understanding of how humans are affected by algorithms in clinical practice. Whether improved DR screening will reduce visual impairment or will it increase the recall rate or will it result in excessive warnings and alerts is not understood as of now.

Policy principles for AI in DR screening

The following policy principles could be useful for AI for DR screening:

AI Designs

AI designs should be human centric, user friendly, and end-user needs. AI systems should help patients, providers, and other care team members in the process of DR screening. The design, development, and success of AI in healthcare should leverage collaboration and dialogue between caregivers, AI technology developers, and other healthcare

stakeholders in order to have all perspectives reflected in AI solutions. It is important to pay close attention to the preset operating thresholds of the AI algorithm, as this could significantly affect the false positive or false negative rates that could translate into the screening costs.[34-36]

Research

Policy frameworks should support and facilitate research and development of AI in DR screening by prioritizing and providing sufficient funding. Clinical validation and transparency research should be prioritized and address the ethical, social, economic, and legal implications that may result from AI applications in DR screening. Several international AI taskforces have evaluated and proposed standardization of AI reporting guidelines (e.g., CONSORT-AI, SPIRIT-AI, STARD-AI, etc.), ranging from the retrospective to prospective clinical trials.[37]

Quality assurance and oversight

AI in DR screening should align with recognized standards of safety, efficacy, and equity. It should be ensured that AI in DR screening is efficacious, and equitable and algorithms, datasets, and decisions auditable when used in real time are clinically validated. AI developers should consistently utilize rigorous procedures and must be able to document their methods and results. Those developing, offering, or testing healthcare AI systems should be required to provide truthful and easy to understand representations regarding intended use and risks that would be reasonably understood by those intended, as well as expected, to use the AI solution. Adverse events should be timely reported to relevant oversight bodies for appropriate investigation and action.

Access and affordability

Policy frameworks should ensure that AI systems for DR screening are accessible and affordable, and there should be even distribution of resources. Payment and incentive policies must be in place to invest in building infrastructure, preparing personnel and training, as well as developing, validating, and maintaining the AI system.

Ethics

Healthcare AI will only succeed if it is used ethically to protect patients and consumers. Policy frameworks should ensure that healthcare AI solutions align with all relevant ethical obligations, from design to development, to use and encourage the development of new ethical guidelines to address emerging issues as needed. It should be ensured that it is consistent with international conventions on human rights and is equally beneficial to patients across socioeconomic, age, gender and geographic regions. AI for health tools may reveal extremely sensitive and private information about a patient, and it must be ensured that laws protect such information from being used to discriminate against patients.

Collaboration and interoperability

Policy frameworks should enable eased data access and use among policymakers, health AI technology developers and users, and the public.

Modernized privacy and security frameworks

Policy frameworks must be scalable and assure that an individual's health information is properly protected and address the topics of privacy, consent, and modern technological capabilities as a part of the policy development process.

Bias

To address the issue of bias in data, policy framework requires identification, disclosure, and mitigation of bias while encouraging access to databases and promoting inclusion and diversity. The data bias should not cause harm to patients or consumers.

Education

Policy frameworks should support education of patients and consumer for the use and advancement of AI in healthcare and encourage stakeholder engagements to keep frameworks responsive to emerging opportunities and challenges.

National health AI strategy

The cultural, workforce training and education, data access, and technology-related changes will require strong guidance and coordination, and there is a significant role of the government in the regulation, delivery, and payment of healthcare, as well as its role in protecting significant amounts of patient data.

In 2019, the World Health Organization released the evidence-based guidelines for digital health. These guidelines provide nine recommendations on select digital health interventions, which involve the use of a mobile phone or device, provides information on implementation considerations, quality and certainty of extant evidence, factors related to acceptability and feasibility of the intervention, and gaps in the evidence that can inform future research. These guidelines can help provide a roadmap for governments and policymakers in introducing and scaling up digital health interventions to support population health outcomes.[38]

The U.S. FDA issued the "Artificial Intelligence/Machine Learning (AI/ML)-Based Software as a Medical Device (SaMD) Action Plan" from the Center for Devices and Radiological Health's Digital Health Center of Excellence. This considered a total product lifecycle-based regulatory framework for these technologies that would allow for modifications to be made from real-world learning and adaptation while ensuring that the safety and effectiveness of the software as a medical device are maintained.[39]

AI Model development and validation: Policy implications

To develop AI, all stakeholders must understand the needs coming from clinical practice, so that proposed AI systems address the needs of health care delivery. Second, it is necessary that such models be developed and validated through a team effort involving AI experts and health care providers, and they should be careful of the fact that the datasets used to train AI are heterogeneous, complex, and nuanced in ways that are often subtle and institution specific.[40] AI tools are to be monitored for safety and reliability and how they are adapted for different locations and over time. Third, AI systems should be rigorously tested for competency and safety before being deployed at the point of treatment, in the same way as medications, medical devices, and other procedures. Each of the steps of the algorithm development needs proper policies to ensure a clinically useful final product [Fig. 1].

Deploying AI in clinical settings: Policy implications

In evaluating, deciding on, and adopting these tools, health care delivery networks and hospitals will need to address

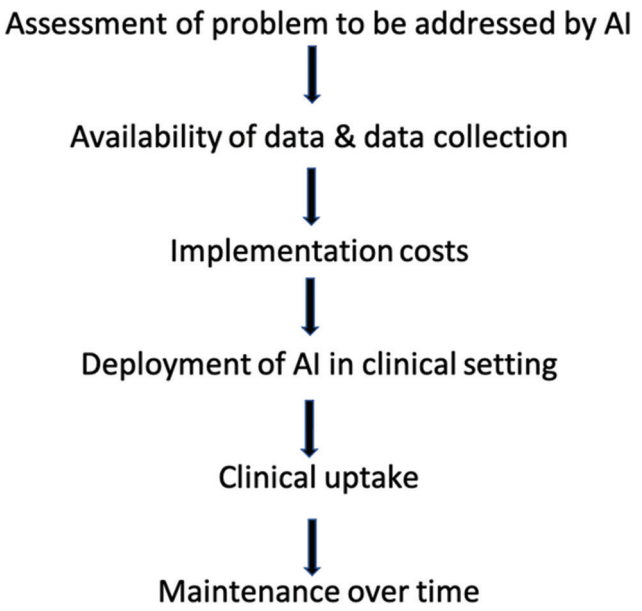

Figure 1: AI algorithm for DR: from development to clinical use

organizational governance, postdevelopment technical problems (i.e., system maintenance), as well as clinical concerns, all of which are critical to effective implementation. For AI applications to be efficient, effective IT governance is needed. To manage AI implementation, health systems must develop or adjust their general IT governance frameworks.[11]

The clinical and administrative leadership of health care systems, with input from all relevant stakeholders, such as patients and the general public must define the near and future needs to measurably improve clinical outcomes. AI is expected to have a positive impact on the healthcare system if these target states are clearly identified. Before deploying AI, health systems should ensure through stakeholder and user engagement especially patients, consumers, and the general public that AI is transparent. It is essential to determine cultural resistance and workflow limitations that may dictate key interpretability and actionability requirements for successful deployment.

Health systems should define standard processes for monitoring and maintaining the performance of AI applications through IT governance and automate those processes, if at all possible, to allow the scalable addition of AI resources for a variety of use cases.

In DR screening, there are two modes, AI is currently being used at the point-of-care for decision making on referrals:

- Fully automatic mode: This is used at sites where screening is done at sites where there is absence of a reliable grader for the acquired retinal images. The referral is based on results of AI. Anyone with referrable DR and ungradable images are referred to an ophthalmologist. All referable images and at least 10% of nonreferable should be preferably graded by an ophthalmologist at a later time.
- Assistive (augmented intelligence) mode: This can be used at sites where there is presence of a reliable grader or an ophthalmologist. The results of AI algorithm are shown to the clinician, and he either accepts the AI grade or modifies the grade based on his clinical judgment.

The same challenges that apply to all health programs apply in deploying AI for DR screening too. While AI has demonstrated far better diagnostic ability and throughput than other techniques, policy makers will have to address other concerns.

Assessment of whether existing resources would be able to deal with the increased referrals from a community-based model. This would include trained personnel, lasers, OPD, and surgical resources in addition to support staff, keeping in mind that the estimated 1200 vitreo-retinal surgeons in India would have to manage 3.5 million referable DR in India alone. Expertise will need to be created to manage nonsurgical DR appropriately. Some attention would also need to be paid to providing care for coexisting systemic conditions such as diabetic nephropathy, which are more prevalent in these individuals.

The number of referrals would depend on the positive predictive value of the tests. It is apparent from Table 1 that even in the best case scenario only 50% of those with referable DR, detected by AI, would actually have referable DR, according to the reference standard. In most cases, this figure is closer to 33%. The other concern in many low resource settings with a significant unoperated cataract population is that 30% of images are ungradable. If these are also to be referred for further evaluation, the health care system must be able to handle the additional load and arrange for appropriate referrals for cataract surgery. Policy makers would need to consider test accuracy and camera quality while making decisions on appropriate technology for their system to minimize avoidable referrals and incorporating it within the larger blindness prevention framework.

Conclusion

AI assisting the DR screening program is a need for countries where there is a demand–supply mismatch, less ophthalmologists, and more patient with diabetes. In last 5 years, there are several algorithms that have shown promise as far as accuracy of DR classification is concerned. However, the health system also needs to be ready to cater with the increased load of cases of sight-threatening DR arising due to increase in screening once AI-based DR screening is deployed. The G20 AI principles provided a framework to guide the design of policy actions. Governments and policy makers should understand the challenges of AI in the health sector; accordingly, steps should be taken to ensure that AI in health is trustworthy.

Financial support and sponsorship

Nil.

Conflicts of interest

There are no conflicts of interest.

References

1. Cho N, Shaw J, Karuranga S, Huang Y, da Rocha Fernandes J, Ohlrogge A, *et al*. IDF Diabetes Atlas: Global estimates of diabetes prevalence for 2017 and projections for 2045. Diabetes Res Clin Pract 2018;138:271-81.
2. Yau JW, Rogers SL, Kawasaki R, Lamoureux EL, Kowalski JW, Bek T, *et al*.; Meta-Analysis for Eye Disease (META-EYE) Study Group. Global prevalence and major risk factors of diabetic retinopathy. Diabetes Care 2012;35:556-64.
3. Jiang F, Jiang Y, Zhi H, Dong Y, Li H, Ma S, *et al*. Artificial intelligence in healthcare: Past, present and future. Stroke Vasc Neurol 2017;2:230-43.

4. Wikipedia. Training, validation, and test sets. 2019. Available from: https://en.wikipedia.org/wiki/Training,_validation,_and_test_sets. [Last accessed on 2019 Nov 13].
5. Ting DSW, Cheung CY, Lim G, Tan GSW, Quang ND, Gan A, *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;318:2211-23.
6. Abràmoff MD, Lou Y, Erginay A, Clarida W, Amelon R, Folk JC, *et al*. Improved automated detection of diabetic retinopathy on a publicly available dataset through integration of deep learning. Invest Ophthalmol Vis Sci 2016;57:5200-6.
7. Gulshan V, Peng L, Coram M, Stumpe MC, Wu D, Narayanaswamy A, *et al*. Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. JAMA 2016;316:2402-10.
8. OECD. 2021. Available from: https://www.oecd.org/els/health-systems/Tackling-Wasteful-Spending-on-Health-Highlights-revised.pdf. [Retrieved 2021 May 12].
9. Yeung K. Recommendation of the Council on Artificial Intelligence (OECD). International Legal Materials 2020; 59:27-34. doi:10.1017/ilm.2020.5
10. OECD iLibrary | Artificial Intelligence in Society. 2021. Available from: https://www.oecd-ilibrary.org/science-and-technology/artificial-intelligence-in-society_eedfee77-en. [Retrieved 2021 May 12].
11. Matheny M, Israni ST, Ahmed M, Whicher D, editors. Artificial Intelligence in Health Care: The Hope, the Hype, the Promise, the Peril. Washington DC: National Academy of Medicine; 2019.
12. Raman R, Ramasamy K, Rajalakshmi R, Sivaprasad S, Natarajan S. Diabetic retinopathy screening guidelines in India: All India Ophthalmological Society diabetic retinopathy task force and Vitreoretinal Society of India consensus statement. Indian J Ophthalmol 2021;69:678-88.
13. Lanzetta P, Sarao V, Scanlon PH, Barratt J, Porta M, Bandello F, *et al*.; Vision Academy. Fundamental principles of an effective diabetic retinopathy screening program. Acta Diabetol 2020;57:785-98.
14. Xie Y, Nguyen Q, Hamzah H, Lim G, Bellemo V, Gunasekeran D, *et al*. Artificial intelligence for teleophthalmology-based diabetic retinopathy screening in a national programme: An economic analysis modelling study. Lancet Digital Health 2020;2:e240-9.
15. Li JO, Liu H, Ting D, Jeon S, Chan R, Kim JE, *et al*. Digital technology, tele-medicine and artificial intelligence in ophthalmology: A global perspective. Prog Retin Eye Res 2021;82:100900.
16. Tufail A, Kapetanakis V, Salas-Vega S, Egan C, Rudisill C, Owen C, *et al*. An observational study to assess if automated diabetic retinopathy image assessment software can replace one or more steps of manual imaging grading and to determine their cost-effectiveness. Health Technol Assess 2016;20:1-72.
17. Li F, Chen H, Liu Z, Zhang X, Wu Z. Fully automated detection of retinal disorders by image-based deep learning. Graefes Arch Clin Exp Ophthalmol 2019;257:495-505.
18. Gulshan V, Rajan R, Widner K, Wu D, Wubbels P, Rhodes T, *et al*. Performance of a deep-learning algorithm vs manual grading for detecting diabetic retinopathy in India. JAMA Ophthalmol 2019;137:987-93.
19. Benjamens S, Dhunnoo P, Meskó B. The state of artificial intelligence-based FDA-approved medical devices and algorithms: An online database. NPJ Digit Med 2020;3:118.
20. Eyenuk Announces FDA Clearance for EyeArt Autonomous AI System for Diabetic Retinopathy Screening-Eyenuk, Inc. ~ Artificial Intelligence Eye Screening. 2021. Available from: https://www.eyenuk.com/us-en/articles/diabetic-retinopathy/eyenuk-announces-eyeart-fda-clearance/. [Retrieved 2021 Sep 12].
21. Kanagasingam Y, Xiao D, Vignarajan J, Preetham A, Tay-Kearney M, Mehrotra A. Evaluation of artificial intelligence–based grading of diabetic retinopathy in primary care. JAMA Netw Open 2018;1:e182665.
22. Li Z, Guo C, Nie D, Lin D, Zhu Y, Chen C, *et al*. Deep learning for detecting retinal detachment and discerning macular status using ultra-widefield fundus images. Commun Biol 2020;3:15.
23. Fenner BJ, Wong RLM, Lam WC, Tan GSW, Cheung GCM. Advances in retinal imaging and applications in diabetic retinopathy screening: A review. Ophthalmol Ther 2018;7:333-46.
24. Bhargava M, Cheung CY, Sabanayagam C, Kawasaki R, Harper CA, Lamoureux EL, *et al*. Accuracy of diabetic retinopathy screening by trained non-physician graders using non-mydriatic fundus camera. Singap Med J 2012;53:715-9.
25. Brzezicki M, Bridger N, Kobetić M, Ostrowski M, Grabowski W, Gill S, *et al*. Artificial intelligence outperforms human students in conducting neurosurgical audits. Clin Neurol Neurosurg 2020;192:105732.
26. Gargeya R, Leng T. Automated identification of diabetic retinopathy using deep learning. Ophthalmology 2017;124:962-9.
27. Ramachandran N, Hong SC, Sime MJ, Wilson GA. Diabetic retinopathy screening using deep neural network. Clin Exp Ophthalmol 2018;46:412-6.
28. Abràmoff MD, Lavin PT, Birch M, Shah N, Folk JC. Pivotal trial of an autonomous AI-based diagnostic system for detection of diabetic retinopathy in primary care offices. NPJ Digital Med 2018;1:39.
29. Bhaskaranand M, Ramachandra C, Bhat S, Cuadros J, Nittala MG, Sadda SR, *et al*. The value of automated diabetic retinopathy screening with the EyeArt system: A study of more than 100,000 consecutive encounters from people with diabetes. Diabetes Technol Ther 2019;21:635-43.
30. Lee A, Yanagihara R, Lee C, Blazes M, Jung H, Chee Y, *et al*. Multicenter, head-to-head, real-world validation study of seven automated artificial intelligence diabetic retinopathy screening systems. Diabetes Care 2021;44:1168-75.
31. Cabitza F, Rasoini R, Gensini GF. Unintended consequences of machine learning in medicine. JAMA 2017;318:517-8.
32. Gianfrancesco MA, Tamang S, Yazdany J, Schmajuk G. Potential biases in machine learning algorithms using electronic health record data. JAMA Intern Med 2018;178:1544-7.
33. Hoff T. Deskilling and adaptation among primary care physicians using two work innovations. Health Care Manage Rev 2011;36:338-48.
34. Gerke S, Minssen T, Cohen G. Ethical and legal challenges of artificial intelligence-driven healthcare. Artif Intell Healthcare 2020:295-336.
35. Ho C, Ali J, Caals K. Ensuring trustworthy use of artificial intelligence and big data analytics in health insurance. Bull World Health Organ 2020;98:263-9.
36. OECD. 2021. Available from: https://www.oecd.org/health/trustworthy-artificial-intelligence-in-health.pdf. [Retrieved 2021 May 12].
37. Ibrahim H, Liu X, Rivera SC, Moher D, Chan AW, Sydes MR, *et al*. Reporting guidelines for clinical trials of artificial intelligence interventions: The SPIRIT-AI and CONSORT-AI guidelines. Trials 2021;22:11.
38. Goodman K, Zandi D, Reis A, Vayena E. Balancing risks and benefits of artificial intelligence in the health sector. Bull World Health Organ 2020;98:230-230A.
39. US Food and Drug Administration releases artificial intelligence/machine learning action plan. J Clin Eng 2021;46:54-5.
40. Gottesman O, Johansson F, Komorowski M, Faisal A, Sontag D, Doshi-Velez F, *et al*. Guidelines for reinforcement learning in healthcare. Nat Med 2019;25:16-8.