



# Great expectations and challenges of artificial intelligence in the screening of diabetic retinopathy

Mingwei Zhao<sup>1,2,3</sup> · Yuzhen Jiang<sup>4,5</sup>

Received: 30 July 2019 / Accepted: 24 September 2019 / Published online: 11 December 2019  
© The Royal College of Ophthalmologists 2019

The global public health burden of diabetes has been growing at a stunning speed due to various risk factors, such as obesity, lack of physical activities, aging, urbanisation and scale of population. The total number of patients with diabetes is estimated to reach 300 million by 2025 and 366 million by 2030 [1, 2]. India, China and the United States will remain the top three nations with the highest numbers of people with diabetes worldwide. A proportion of the affected population is also at risk of diabetic retinopathy (DR). Recent studies revealed that up to over 18% of patients in China [3] and over 21% in India [4] are affected by DR. Effective screening and early intervention are essential for the cost-effective prevention of irreversible blindness from DR [5]. Early diagnosis of DR requires systematic and extensive screening due to the asymptomatic nature of this major cause of preventable blindness. The deficiency in qualified professionals for providing effective screening and treatment of vision threatening DR poses further challenges in both India [6] and China [7]. In the United States, where specialised healthcare has been relatively more accessible, eye care was reported to be available to only <60% of these patients [8].

Promising results from a few research groups worldwide have shown the potential role of deep learning algorithm within quality assured artificial intelligence (AI) assisted systems for DR screening. In a prospective study published earlier this year, Gulshan and colleagues demonstrated the equivalent or better performance of a deep learning algorithm in the diagnosis of moderate or worse DR compared with experienced specialists or trained graders in a cohort of 3049 patients from two hospitals in India [8]. Deep learning algorithm using a convolutional neural network achieved high accuracy in identifying vision threatening DR from over 35,000 images from population-based cohorts of Malays, Caucasian Australians and Indigenous Australians [9]. In another large-scale multiethnic cohort, deep learning showed performance in the diagnosis of DR of various severity comparable with retina specialists or professional ocular image graders [10]. In 2018, IDx-DR developed by researchers from the University of Iowa, became the first FDA-approved AI-based automated DR screening system and achieved satisfactory results in screening referable DR in primary care [11]. In China, AI assisted system completed DR screening of over 28,000 [12] diabetic patients within 1 year from the National Metabolic Management Center that covers 542 hospitals in over 30 provincial administrative regions of China [13].

An ideal AI programme in DR screening should be able to demonstrate both high sensitivity and high specificity to assure an optimal balance between safety and efficiency. Tremendous amount of effort has been devoted to targeting and tackling the barriers to improving the performance of deep learning algorithms, such as the development of methodologies for image standardisation and preprocessing [14], selection of reference standards [6, 15], employment of convolutional neural networks [14], and adoption of suitable database for training and validation. The recently published work by Gulshan and colleagues using the improved Inception-v4 neural network architecture shows encouraging improvement by demonstrating equivalent or better performance in assessing moderate or severe DR

---

✉ Mingwei Zhao  
dr\_zhaomingwei@163.com

<sup>1</sup> Department of Ophthalmology, People's Hospital, Peking University, 11 South Avenue Xi Zhi Men, Beijing 100044, China

<sup>2</sup> Eye diseases and optometry institute, Beijing Key Laboratory of Diagnosis and Therapy of Retinal and Choroid Diseases, 11 South Avenue Xi Zhi Men, Beijing 100044, China

<sup>3</sup> College of Optometry, Peking University Health Science Center, Peking University, 11 South Avenue Xi Zhi Men, Beijing 100044, China

<sup>4</sup> Moorfields Eye Hospital, 162 City Rd, London EC1V 2PD, UK

<sup>5</sup> UCL Institute of Ophthalmology, 43 Bath Street, London EC1V 9EL, UK

compared with manual grading by retinal specialists and trained graders. The impressive model generalisation shown in this study conducted in two different hospitals is important for the application of deep learning algorithms in clinical settings where the consistency in performance can be affected by various factors, such as different cameras used and whether fundus photography was carried out under dilation. As the authors have rightfully pointed out, the high sensitivity and specificity achieved by the automated AI system adopted and improved in this study can potentially serve as part of a well-integrated system combining automated AI system and human graders to increase accessibility in low-resource settings and improve cost-effectiveness in high-resource settings.

Despite these promising study results achieved in the development and assessment of automated systems in DR severity diagnosis especially at the moderate or severe level, challenges remain to adopt automated systems as the ‘gatekeeper’ in extensive screening with wide coverage. For example, challenges remain to improve the capabilities of deep learning algorithm in early DR detection which involves accurately spotting minor changes including fine microaneurysms and small haemorrhages. Furthermore, the validity and accuracy of deep learning algorithms in screening other major causes of blindness such as glaucoma and age-related macular degeneration is also important for extensive application of AI systems in population-based screening programmes. There is also a lack of randomised controlled clinical trials comparing human graders with deep learning incorporated AI systems. However, due to the urgent need for DR screening in large populations in countries like China, we have to balance urgency with accuracy and aim for urgent implementation of AI assisted DR screening in China whilst more validation is being performed.

### Compliance with ethical standards

**Conflict of interest** The authors declare that they have no conflict of interest.

**Publisher’s note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

### References

1. King H, Aubert RE, Herman WH. Global burden of diabetes, 1995–2025: prevalence, numerical estimates, and projections.

- Diabetes Care. 1998;21:1414–31. <https://doi.org/10.2337/diacare.21.9.1414>
2. Wild S, Roglic G, Green A, Sicree R, King H. Global prevalence of diabetes: estimates for the year 2000 and projections for 2030. *Diabetes Care*. 2004;27:1047–53. <https://doi.org/10.2337/diacare.27.5.1047>
3. Song P, Yu J, Chan KY, Theodoratou E, Rudan I. Prevalence, risk factors and burden of diabetic retinopathy in China: a systematic review and meta-analysis. *J Glob Health*. 2018;8:010803 <https://doi.org/10.7189/jogh.08.010803>
4. Gadkari SS, Maskati QB, Nayak BK. Prevalence of diabetic retinopathy in India: The All India Ophthalmological Society Diabetic Retinopathy Eye Screening Study 2014. *Indian J Ophthalmol*. 2016;64:38–44. <https://doi.org/10.4103/0301-4738.178144>
5. Jones CD, Greenwood RH, Misra A, Bachmann MO. Incidence and progression of diabetic retinopathy during 17 years of a population-based screening program in England. *Diabetes Care*. 2012;35:592–6. <https://doi.org/10.2337/dc11-0943>
6. Gulshan V, Rajan RP, 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. <https://doi.org/10.1001/jamaophthalmol.2019.2004>
7. Chen Y, Zhang B, Zhang H. Insights and prospectives of ophthalmologic artificial intelligence (AI) technology. *Chin J Ocul Fundus Dis*. 2019;35:119–23.
8. Abramoff MD, Folk JC, Han DP, Walker JD, Williams DF, Russell SR, et al. Automated analysis of retinal images for detection of referable diabetic retinopathy. *JAMA Ophthalmol*. 2013;131:351–7. <https://doi.org/10.1001/jamaophthalmol.2013.1743>
9. Li Z, Keel S, Liu C, He Y, Meng W, Scheetz J, et al. An automated grading system for detection of vision-threatening referable diabetic retinopathy on the basis of color fundus photographs. *Diabetes Care*. 2018;41:2509–16. <https://doi.org/10.2337/dc18-0147>
10. 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. <https://doi.org/10.1001/jama.2017.18152>
11. van der Heijden AA, Abramoff MD, Verbraak F, van Hecke MV, Liem A, et al. Validation of automated screening for referable diabetic retinopathy with the IDx-DR device in the Hoorn Diabetes Care System. *Acta Ophthalmologica*. 2018;96:63–8. <https://doi.org/10.1111/aos.13613>
12. VoxelCloud (2019) Hundreds of thousands of years at a glimpse: why AI retinal disease screening has become the most practical sector of AI-assisted healthcare? <https://www.chinait.com/ai/36588.html> Accessed 7 July 2019.
13. National Metabolic Management Center (2019) The National Metabolic Management Center: Major Events in May. <https://www.omronmed.com/newsinfo.aspx?pid=112&Cateid=23&Newsid=378> Accessed 6 June 2019.
14. Gargeya R, Leng T. Automated identification of diabetic retinopathy using deep learning. *Ophthalmology*. 2017;124:962–9. <https://doi.org/10.1016/j.ophtha.2017.02.008>
15. 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. <https://doi.org/10.1001/jama.2016.17216>