A Systematic Review of Global Implementation Practices of Deep Learning in Diabetic Retinopathy Screening

Riddhi Shenoy

Norfolk and Norwich University Hospital, Norwich, Norfolk, UK

Background

• Diabetes is becoming a global epidemic and diabetic retinopathy, a primary cause of blindness worldwide, had an estimated global prevalence of 93 million in 2012[1].

•Early diagnosis and treatment can prevent vision loss from diabetic retinopathy and artificial intelligence (AI) systems has been proposed as one option to increase screening.

•Deep Learning (DL) shows promising diagnostic potential in retrospective studies of Diabetic Retinopathy (DR) screening, with algorithms yielding sensitivities and specificities of over 90% when compared to human grading[1].

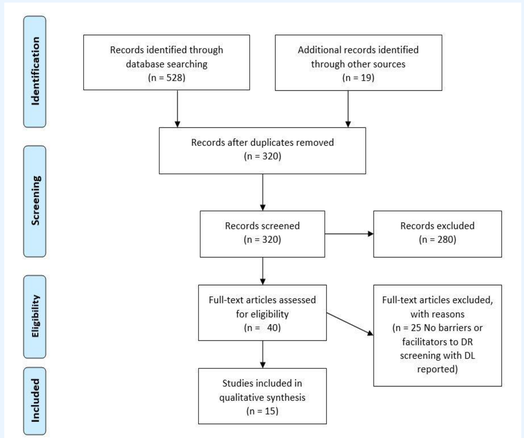
• However, there are significant concerns with implementation particularly with data privacy and quality control. As the practical use of AI is becoming a reality in current diabetic screening practices, there is a need to evaluate current best practices of the implementation of an AI DR screening programme in the literature.

Methods

A preliminary search was conducted to gather key terms to inform creation of the search string. Variations of the terms “artificial intelligence” and “diabetic retinopathy” were combined using Boolean operators where permitted. Additional hand searching of relevant review articles and references of included studies was undertaken. The search string was employed across PubMed, Medline, Embase, and the Health Management Information Consortium databases. Inclusion criteria were determined to be

- full text articles
- published in the last five years
- prospective studies of the use of AI or deep learning in DR screening
- description of barriers or facilitators to implementation.

Figure 1: flow of studies included in the analysis



Key pitfalls of using AI in DR screening

Governance and telecommunications requirement

The large UK study of automated DR screening highlighted that software running issues increased as a larger patient database was created[4]. This was reduced with a cloud-based system but this has significant security implications for confidential patient data storage.

Studies in lower resource settings highlighted that a strong internet connection was required to generate a prediction, but this is impossible with poor connectivity, hence the use of offline systems in lower resource settings[14].

Incorporation into Clinical Workflow

In the study where AI was used to assist physician grading, despite the AI assisted graders self-reporting significantly higher confidence in their diagnoses, the time taken to analyse images was 15-20 seconds longer (average was 50 seconds) with AI assistance compared to unassisted graders and these differences were preserved throughout the course of the study[2]. Another study highlighted staff frustration at the added time required to recapture images deemed ungradable by the AI system [3]. Improvements to the photography conditions such as a cloth hood to increase darkness, turning the light off was difficult to implement as the same physical space often was being used for counselling[3].

Conclusion

- **Best practice for the implementation of AI in DR screening is emerging in the literature.**
- **Further research is required into information governance and data security for the implementation of AI in DR screening, particularly in lower resource settings.**
- **Further consideration is also required into the incorporation of AI into the clinical workflow and ensuring all stakeholders, including patients, technicians and healthcare professionals are involved in decision making surrounding implementation to promote stakeholder buy-in.**

References

1. Grzybowski, A., Brona, P., Lim, G., Ruamviboonsuk, P., Tan, G.S., Abramoff, M. and Ting, D.S., 2020. Artificial intelligence for diabetic retinopathy screening: a review. *Eye*, *34*(3), pp.451-460.
2. Sayres, R., Taly, A., Rahimy, E., Blumer, K., Coz, D., Hammel, N., Krause, J., Narayanaswamy, A., Rastegar, Z., Wu, D. and Xu, S., 2019. Using a deep learning algorithm and integrated gradients explanation to assist grading for diabetic retinopathy. *Ophthalmology*, *126*(4), pp.552-564.
3. Beede, E., Baylor, E., Hersch, F., Iurchenko, A., Wilcox, L., Ruamviboonsuk, P. and Vardoulakis, L.M., 2020, April. A Human-Centered Evaluation of a Deep Learning System Deployed in Clinics for the Detection of Diabetic Retinopathy. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (pp. 1-12).
4. Tufail, A., Kapetanakis, V.V., Salas-Vega, S., Egan, C., Rudisill, C., Owen, C.G., Lee, A., Louw, V., Anderson, J., Liew, G. and Bolter, L., 2016. 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*, *20*(92), pp.1-72.
5. Ribeiro, L., Oliveira, C.M., Neves, C., Ramos, J.D., Ferreira, H. and Cunha-Vaz, J., 2015. Screening for Diabetic Retinopathy in the Central Region of Portugal. Added Value of Automated 'Disease/No Disease'Grading. *Ophthalmologica*, *233*(2), pp.96-103.
6. Natarajan, S., Jain, A., Krishnan, R., Rogye, A. and Sivaprasad, S., 2019. Diagnostic accuracy of community-based diabetic retinopathy screening with an offline artificial intelligence system on a smartphone. *JAMA ophthalmology*, *137*(10), pp.1182-1188.
7. Sosale, B., Sosale, A.R., Murthy, H., Sengupta, S. and Naveenam, M., 2020. Medios–An offline, smartphone-based artificial intelligence algorithm for the diagnosis of diabetic retinopathy. *Indian Journal of Ophthalmology*, *68*(2), p.391.
8. Kanagasingam, Y., Xiao, D., Vignarajan, J., Preetham, A., Tay-Kearney, M.L. and Mehrotra, A., 2018. Evaluation of artificial intelligence–based grading of diabetic retinopathy in primary care. *JAMA network open*, *1*(5), pp.e182665-e182665.
9. Abràmoff, M.D., Lavin, P.T., Birch, M., Shah, N. and Folk, J.C., 2018. Pivotal trial of an autonomous AI-based diagnostic system for detection of diabetic retinopathy in primary care offices. *NPJ digital medicine*, *1*(1), pp.1-8.
10. Bawankar, P., Shanbhag, N., Dhawan, B., Palsule, A., Kumar, D., Chandel, S. and Sood, S., 2017. Sensitivity and specificity of automated analysis of single-field non-mydriatic fundus photographs by Bosch DR Algorithm—Comparison with mydriatic fundus photography (ETDRS) for screening in undiagnosed diabetic retinopathy. *Plos one*, *12*(12), p.e0189854.
11. Rajalakshmi, R., Subashini, R., Anjana, R.M. and Mohan, V., 2018. Automated diabetic retinopathy detection in smartphone-based fundus photography using artificial intelligence. *Eye*, *32*(6), pp.1138-1144.
12. Kim, T.N., Aaberg, M.T., Li, P., Davila, J.R., Bhaskaranand, M., Bhat, S., Ramachandra, C., Solanki, K., Myers, F., Reber, C. and Jalalizadeh, R., 2020. Comparison of automated and expert human grading of diabetic retinopathy using smartphone-based retinal photography. *Eye*, pp.1-9.
13. Keel, S., Lee, P.Y., Scheetz, J., Li, Z., Kotowicz, M.A., Macisaac, R.J. and He, M., 2018. Feasibility and patient acceptability of a novel artificial intelligence-based screening model for diabetic retinopathy at endocrinology outpatient services: a pilot study. *Scientific reports*, *8*(1), pp.1-6.
14. Bellemo, V., Lim, Z.W., Lim, G., Nguyen, Q.D., Xie, Y., Yip, M.Y., Hamzah, H., Ho, J., Lee, X.Q., Hsu, W. and Lee, M.L., 2019. Artificial intelligence using deep learning to screen for referable and vision-threatening diabetic retinopathy in Africa: a clinical validation study. *The Lancet Digital Health*, *1*(1), pp.e35-e44.
15. Thomas, S.A. and Titus, G., 2020. Design of a portable retinal imaging module with automatic abnormality detection. *Biomedical Signal Processing and Control*, *60*, p.101962
16. Van Der Heijden, A.A., Abramoff, M.D., Verbraak, F., van Hecke, M.V., Liem, A. and Nijpels, G., 2018. Validation of automated screening for referable diabetic retinopathy with the IDx-DR device in the Hoorn Diabetes Care System. *Acta ophthalmologica*, *96*(1), pp.63-68.

Key methods for facilitation of DR screening using AI

AI assisting physicians

One prospective study from North America compared the ability of 10 ophthalmologists to grade fundus images with and without AI assistance[2]. AI assisted readings showed significantly increased confidence in gradings, particularly when assisted with an automated grading and a heatmap to show the features on the image used by the AI to decide a particular grading[2].

Quality control methods

A large study in Portugal over 3 years required all AI classified positive fundi with DR to be sent for human assessment[5]. A random subset of AI classified negative fundi were sent to reading centres for human grading as a quality control measure and a percentage of randomly selected images were sent for evaluation by an independent panel to further control quality[5].

A large multi-site study across Thailand ensured all AI diagnoses were reviewed by an ophthalmologist to ensure no missed patients and false negatives prompted nurse follow up[3].

Non-mydriatic fundus photographs

This reduces the need for dilation to streamline workflow and increase patient acceptability[6]. However, insufficient quality images are not gradable by AI systems and several studies have mitigated this by integrating an AI capability to alert the user immediately if the image is of insufficient quality, to allow image recapture in real time[6,7,8,9,10]. A multi-site trial across India mitigated the common issue of image quality and unreadability by using convolutional neural networks and noise mechanisms to generate predictions for even lower quality non-mydriatic images[10]. Indeed, 96.1% of images were fit for evaluation by the algorithm (n=1128), which was beneficial in the lower resource setting with increased physical distance from patients to healthcare[10].

Smartphone use

Several studies reported the benefits of using a smartphone by allowing the operator to use ergonomic touch functions to capture the image and reducing operator mobility to increase image quality [7,11,12].. One study in India integrated the offline automated analysis application into the smartphone-based retinal imaging system, allowing the technician to rapidly generate a diagnosis (DR present/absent) in seconds[6].

Training

A multi-site study undertaken by Google Health in Thailand, included a training programme for non-ophthalmologists to implement the screening[3]. Training included the actual image acquisition, methods to improve image quality, and the input of images into the AI system[3]. A large primary care study in the United States, used a 4-hour training programme to teach primary care workers to take a non-mydriatic image[8].

Clinical Workflow

One study reported patient wait times to be between five minutes to three hours and nurses reported using waiting time for patient education, foot examination or nutritional counselling[3]. There was a large variation in time taken for imaging and AI analysis. One study reported the mean time for the automated screening model was 6.9 +/− 1.4 minutes[13], another reported the nurse took approximately 10 to 15 minutes to obtain images for both eyes, and the AI system provided reading outcomes in less than 3 minutes[8] and one smartphone study, reported analysis in seconds [6]