

Artificial intelligence in ophthalmological practice: when ideal meets reality

 Ludwig M Heindl ,¹ Senmao Li,^{1,2} Daniel S W Ting,^{3,4} Pearse A Keane⁵

To cite: Heindl LM, Li S, Ting DSW, *et al.* Artificial intelligence in ophthalmological practice: when ideal meets reality. *BMJ Open Ophthalmology* 2023;**8**:e001129. doi:10.1136/bmjophth-2022-001129



© Author(s) (or their employer(s)) 2023. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ.

¹Department of Ophthalmology, University of Cologne, Köln, Germany

²Department of Ophthalmology, The First Affiliated Hospital of Jinan University, Guangzhou, Guangdong, China

³Singapore National Eye Center, Duke-NUS Medical School, Singapore

⁴Ophthalmology and Visual Sciences Department, Duke-NUS Medical School, Singapore

⁵Medical Retina, Moorfields Eye Hospital NHS Foundation Trust, London, UK

Correspondence to

Dr Ludwig M Heindl; Ludwig.heindl.BJO@gmail.com

‘It was the best of times; it was the worst of times’. Such words never get old in any era. Similarly, in ophthalmology, the same thing is happening. Artificial intelligence (AI) using machine learning and deep learning in ophthalmology have created incredible chemistry,^{1 2} with the US Food and Drug Administration (FDA) approving AI-based diagnostic technology for autonomous diabetic retinopathy screening.³ A human-centred AI study, on diabetic retinopathy screening, conducted in Thailand with the Google team, highlighted many unexpected real-world problems and questions. Further work is required to determine formal cost-benefit analyses, specific workflows for live implementation in diverse healthcare settings and solutions for real-world challenges such as lack of internet connectivity or electronic health records.⁴

Transforming healthcare with AI is a beautiful aspiration, which everyone expects something^{1 2} from. Most patients with stable but chronic eye disease would want to spend less time in crowded waiting rooms waiting for brief and hurried consultations; clinicians wish to focus their limited energy on solving problems; engineers want their code to change the world and device manufacturers want their devices to be the gateway to a virtually connected reality. So, is reality as good as it seems?

The most compelling application of AI in ophthalmology is to aid in diagnosis. Standardised image acquisition in ophthalmology brings a unique advantage to deep learning. From corneal topography to optical coherence tomography angiography, many ophthalmic assistive exams have objective image quality checks on the images at the time of acquisition, which can be used to determine the credibility of the examination results and as high-quality input data for model training. Theoretically, AI-assisted diagnosis applies to most common cases in ophthalmology. Researchers and engineers in ophthalmology have made outstanding contributions to this

effort. If someone follow developments in this field, he/she will know that AI has shown fantastic potential in subspecialties such as diabetic retinopathy, glaucoma, retinal vein occlusion and so on.^{5–8} Many AI-related ophthalmological programmes are going worldwide and many automated eye-disease screening and analysis medical devices have been successfully applied in the clinical practice.⁹ AI can be found in almost all areas of ophthalmology, from the anterior segment of the eye to the fundus.^{10–12} Clinicians have provided many noteworthy labels for machine learning based on their own experience, and we look forward to adding the next label that will make the model even better. As of 2020, there are 94 publicly accessible downloadable ophthalmology databases, and more than half of the image data are retinal fundus photographs, with 18% of these databases not labelled with the relevant diseases collected. Unfortunately, most databases lack basic information (age, sex, ethnicity, etc) and inclusion and exclusion criteria are missing. Barriers to using these data include low visibility, accessibility issues or limited usability due to incomplete metadata, including the lack of critical parameters needed to assess data sources, data quality and diversity of the populations sampled.¹³

Human learning behaviour has not yet been fully explained at the neurological level, and the emerging interdisciplinary field of cognitive neuroscience was born to study this behaviour.¹⁴ Most AI models are disease-centric, and it is challenging to train an AI model to detect normal fundus photos. For training, most AI models require the index case (positive cases) versus control (non-positive cases that usually include normal and other non-index pathologies). On the other hand, the human brain is easier in registering the ‘normal images’ without any pathology. This is unique to the human brain compared with the convolutional neural network we are currently using. Of course different algorithms optimise this feature, for example,



when we use semi-supervised learning, this method requires only a small number of data to be labelled. To unravel the intricate structure behind the data model, the machine must be able to infer patterns between observations for which it has not received explicit tagged information. Semi-supervised learning aims to provide mechanisms for making such connections, which will be essential for achieving this goal.¹⁵ Unfortunately, many semi-supervised learning methods only perform better than their supervised counterparts or base learners in specific cases.^{16 17} But this has received relatively little academic attention.¹⁸ Furthermore, different algorithms fuse to enhance, and many algorithms have been iteratively upgraded to follow the progress of the times and the increase in computing power. First to be noticed are the recent advances in semi-supervised neural networks. Minor variations in the input space should only cause minor variations in the output space. This assumption makes incorporating unsupervised loss terms into the cost function more straightforward than before. This flexibility also accommodates the incorporation of more complex cost terms. Another potential remedy for the lack of robustness of semi-supervised learning methods lies in the application of automated machine learning to the semi-supervised setting. These approaches include meta-learning and neural architecture search as well as automatic algorithm selection and hyperparameter optimisation, which have been prominently and successfully applied to supervised learning, but there has been no application to semi-supervised learning so far.¹⁵ However, there are some issues with the usability. Semi-supervised learning is much less standardised compared with supervised learning. The KEEL software package includes a semi-supervised learning module,¹⁹ and implementations of some transductive graph-based methods exist in scikit-learn.

While the algorithm evolves, it also must match the right application scenario. In the experiments in Bangkok, the environmental light in the clinic obstructed the proper functioning of the diabetic retinopathy diagnostic model.²⁰ Nurses needed to constantly readjust to get an image quality that the machine would recognise, reducing their productivity in an already busy workday. As humans train the models, the models train humans. Could we also consider adding some non-clinical diagnostic factors into the model training? For example, use an AI-enabled automated image optimisation software to improve the luminance, contrast and eye-camera coordination during the image acquisition stage? Such feature points will help increase the model's fit, although they require higher-resolution sensors.

Even if the AI diagnostic performance is deemed to be clinically acceptable in the research and development phase like the above-mentioned example, the real-world AI implementation can possess many challenges, including AI bias due to differences in capturing devices, locations in the specific organs, population/ethnicities; generalisability; data privacy;

ethics and social equities. As we need evidence-based medicine, algorithm training requires constantly expanding the data sample and maintaining a certain update frequency. The FDA uses a framework called Software as Medical Device to review and approve the marketing of AI-based technologies, which evaluates algorithms throughout their life cycle.²¹ Big data and AI may also expose the health risks of some specific populations, leading to an inevitable imbalance in the distribution of health insurance resources and social injustice. Although AI systems can often achieve 'state-of-the-art' performance on 'in silico' testing, these findings are often not replicated in the real world. In this regard, a major focus of clinical AI research is the development of systems which are (1) robust (eg, will work on different machines and in different conditions), (2) reliable (eg, can give some measure of certainty with which they provide an output), (3) safe (eg, can detect rare but potentially serious ophthalmic conditions) and (4) fair (eg, can work equally well in different populations, particularly with regard to age, gender and ethnicity).

Technological optimism and pessimism need to strike a delicate balance about healthcare, a complex relationship between society, ethics and technology that only humans can weigh. Until now, if you ask any clinician whether AI is widely adopted in the clinical practice, his/her answer mostly likely is a 'no'. Medicine is an art, and there are limits to the benefits that a single technological advancement can bring. AI in the clinical optimisation of the patient and doctor experience is its true mission. AI combined with clinical practice can only better step into the business world to achieve a positive cycle of input and output. As clinicians, we all know that 'To cure sometimes. Often relieves. Always comfort'. Taking technology and adding humanistic care may be a condition for AI to move towards medical reality; even if AI is sufficient to solve clinical problems, it cannot understand what is happening in the clinic.

Contributors LMH had the idea for the article. SL performed the literature search and wrote the article. DSWT and PAK helped supervise the project. All authors discussed the results and contributed to the final manuscript.

Funding The authors have not declared a specific grant for this research from any funding agency in the public, commercial or not-for-profit sectors.

Competing interests None declared.

Patient consent for publication Not applicable.

Ethics approval Not applicable.

Provenance and peer review Commissioned; externally peer reviewed.

Open access This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>.

ORCID iD

Ludwig M Heindl <http://orcid.org/0000-0002-4413-6132>

REFERENCES

- 1 Ting DSW, Cheung CY-L, 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;318:2211–23.
- 2 Ting DSW, Peng L, Varadarajan AV, *et al.* Deep learning in ophthalmology: the technical and clinical considerations. *Prog Retin Eye Res* 2019;72:100759.
- 3 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.
- 4 Yuan A, Lee AY. Artificial intelligence deployment in diabetic retinopathy: the last step of the translation continuum. *Lancet Digit Health* 2022;4:e208–9.
- 5 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;316:2402–10.
- 6 Asaoka R, Murata H, Iwase A, *et al.* Detecting preperimetric glaucoma with standard automated perimetry using a deep learning Classifier. *Ophthalmology* 2016;123:1974–80.
- 7 Nagasato D, Tabuchi H, Masumoto H, *et al.* Automated detection of a nonperfusion area caused by retinal vein occlusion in optical coherence tomography angiography images using deep learning. *PLoS One* 2019;14:e0223965.
- 8 Chen JS, Coyner AS, Ostmo S, *et al.* Deep learning for the diagnosis of stage in retinopathy of prematurity: accuracy and generalizability across populations and cameras. *Ophthalmol Retina* 2021;5:1027–35.
- 9 Li F, Pan J, Yang D, *et al.* A multicenter clinical study of the automated fundus screening algorithm. *Trans Vis Sci Tech* 2022;11:22.
- 10 Cui T, Wang Y, Ji S, *et al.* Applying machine learning techniques in Nomogram prediction and analysis for SMILE treatment. *Am J Ophthalmol* 2020;210:71–7.
- 11 Wu X, Huang Y, Liu Z, *et al.* Universal artificial intelligence platform for collaborative management of cataracts. *Br J Ophthalmol* 2019;103:1553–60.
- 12 Zhang K, Liu X, Xu J, *et al.* Deep-learning models for the detection and incidence prediction of chronic kidney disease and type 2 diabetes from retinal fundus images. *Nat Biomed Eng* 2021;5:533–45.
- 13 Khan SM, Liu X, Nath S, *et al.* A global review of publicly available datasets for ophthalmological imaging: barriers to access, usability, and generalisability. *Lancet Digit Health* 2021;3:e51–66.
- 14 Vaadia E. Cognitive Neuroscience. learning how the brain learns. *Nature* 2000;405:523–5.
- 15 van Engelen JE, Hoos HH. A survey on semi-supervised learning. *Mach Learn* 2020;109:373–440.
- 16 Li YF, Zhou ZH. Towards making unlabeled data never hurt. *IEEE Trans Pattern Anal Mach Intell* 2015;37:175–88.
- 17 Singh A, Nowak RD, Zhu X. Unlabeled data: now it helps, now it doesn't. 2008. Available: https://www.cs.cmu.edu/~aarti/pubs/NIPS08_ASingh.pdf
- 18 Zhu X. Semi-supervised learning literature survey. 2008. Available: <http://digital.library.wisc.edu/1793/60444>
- 19 Triguero I, González S, Moyano JM, *et al.* KEEL 3.0: an open source software for multi-stage analysis in data mining. *IJCIS* 2017;10:1238.
- 20 Beede E, Baylor E, Hersch F, *et al.* A human-centered evaluation of a deep learning system deployed in clinics for the detection of diabetic retinopathy. *ACM CHI* 2020:1–12.
- 21 Administration FaD. Artificial intelligence and machine learning in software as a medical device. 2021. Available: <https://www.fda.gov/medical-devices/software-medical-device-samd/artificial-intelligence-and-machine-learning-software-medical-device>