

Artificial Intelligence for Diagnosis of Inherited Retinal Disease: An Exciting Opportunity and One Step Forward

Tien-En Tan, MBBS (Hons), MMed (Ophth), FRCOphth^{1,2}

Hwei Wuen Chan, MBBS, MMed (Ophth), FAMS, FRCOphth^{3,4}

Mandeep S. Singh, MD, PhD⁵

Tien Yin Wong, MD, PhD^{1,2}

Jose S. Pulido, MD, MBA, MPH⁶

Michel Michaelides, MD (Res), FRCOphth, FACS^{4,7}

Elliott H. Sohn, MD^{8,9}

Daniel Shu Wei Ting, MD (1st Hons), PhD^{1,2}

1. Singapore Eye Research Institute, Singapore, Singapore National Eye Centre, Singapore
2. Duke-National University of Singapore Medical School, Singapore
3. Department of Ophthalmology, National University of Singapore, Singapore
4. UCL Institute of Ophthalmology, University College London, London, UK
5. Wilmer Eye Institute, Johns Hopkins University School of Medicine, Baltimore, Maryland, USA
6. Wills Eye Hospital, Thomas Jefferson University, Philadelphia, Pennsylvania, USA
7. Moorfields Eye Hospital, London, UK
8. University of Iowa Hospitals and Clinics, Iowa City, Iowa, USA
9. Institute for Vision Research, University of Iowa, Iowa City, Iowa, USA

Corresponding Author:

Daniel Shu Wei Ting, MD (1st Hons), PhD

Associate Professor, Duke-NUS Medical School, Singapore

Head, AI and Digital Innovation, Singapore Eye Research Institute

Consultant, Surgical Retina, Singapore National Eye Centre

E-mail: daniel.ting45@gmail.com

Address: 11 Third Hospital Avenue, Singapore 168751

Telephone: (+65) 62277255

Word Count: 1,926 words

INTRODUCTION

Inherited retinal disease (IRD) affects approximately 1 in 3,000 individuals in North America and Europe, and is a significant cause of visual impairment and blindness among children and working-age adults, with huge personal and societal impact.[1,2] Accurate clinical phenotypic and genotypic diagnosis of IRD is challenging, but increasingly important and relevant. Traditionally, genotypic diagnosis has been considered “nice to have”, but not “essential”, with implications usually related to patient prognostication and genetic counselling. However, an accurate genetic diagnosis is now of paramount importance because of rapid advances in potential gene replacement and other therapies for these previously untreatable conditions. In 2017, the first gene therapy for IRD was approved by the United States Food and Drug Administration (FDA) for the treatment of *RPE65*-mediated retinal dystrophy and shortly after by the European Medicines Agency (EMA) as well.[3] Multiple clinical trials are currently underway for other IRDs, including Choroideremia, Stargardt disease, and Retinitis Pigmentosa (RP).[4,5] Besides gene replacement therapy, progress in other areas such as antisense oligonucleotide (AON) therapy and gene editing with clustered regularly interspaced short palindromic repeats (CRISPR) and CRISPR-associated proteins (Cas) also rely on accurate genetic diagnosis.[5,6]

UNMET CLINICAL CHALLENGES

Successful genotypic diagnosis remains elusive for many patients globally, due in part to remaining gaps in knowledge, but also due to limited access to testing, which

remains relatively expensive, along with scarcity and an uneven distribution of institutions with expertise in IRD. In the majority of tertiary centres in the Western world, patients have a high chance of an accurate genetic diagnosis. Recent studies have demonstrated the successful characterization of large cohorts of patients with IRD using systematic clinical phenotyping and genetic testing protocols.[7–10] Typically, historical, clinical, electrophysiologic and multi-modal imaging data are used to assign each patient a clinical phenotypic category and to facilitate the selection of a genetic testing strategy. For example, in 2017, Stone et al reported successful identification of disease-causing genotypes in 76% of 1000 consecutive families with IRD.[10] They used a tiered testing strategy, first relying on focused testing for specific genes, based on the phenotypic category. If needed, they then gradually enlarged the molecular hypothesis in a recursive manner up the classification tree, eventually using whole-exome and whole-genome sequencing, only if required. This tiered testing strategy had greater sensitivity, lower average cost, and a much lower false genotype rate (FGR) than a strategy using whole-exome sequencing for all cases. A phenotype-driven genetic testing strategy may therefore be advantageous, but requires a team of experienced clinicians, well versed in accurate clinical phenotyping, seeing diverse groups of patients with IRD. It also requires major infrastructure such as a certified genetic laboratory and communication between the clinician and testing laboratory. Unfortunately, such expertise and resources are not readily available in most countries except in a select few sub-specialty tertiary referral centres worldwide. Most patients with IRD thus do not have access to an accurate clinical phenotypic and genetic diagnosis – these patients often experience an unacceptably long “diagnostic odyssey”. This major

clinical need presents a unique opportunity for artificial intelligence (AI), and particularly deep learning (DL).

AI AND DL IN MEDICINE

AI and DL techniques have been applied extensively to various technical and medical fields, ranging from medical imaging analysis, to natural language processing and speech.[11] Recently, DL using reinforcement learning has shown promising results in prediction of protein structures, with the potential to greatly increase our understanding of how specific genetic variants cause disease.[12] In ophthalmology, DL techniques have been applied to diagnosis of major ocular diseases such as diabetic retinopathy (DR), age-related macular degeneration and glaucoma from colour fundus photographs, fundus autofluorescence (FAF) images and optical coherence tomography (OCT).[13–17]

AI AND DL FOR INHERITED RETINAL DISEASE

Currently, the application of AI and DL techniques for IRD diagnosis is still at a very nascent stage. Most of these efforts have focused on clinical phenotyping. A few studies have investigated automated classification of IRD versus normal controls or acquired retinal disorders, with AI and DL applied to various modalities such as ultra-widefield (UWF) fundus photographs, FAF or OCT images.[18–21] For example, Masumoto et al utilized 373 UWF pseudocolour and FAF images with DL to diagnose RP cases from controls, while Shah et al used individual OCT B-scans at the fovea (749 scans from 93 individuals) with DL to distinguish Stargardt disease

from controls.[18,21] Both of these studies were limited to binary classifications of single IRD disease cases versus controls. Miere et al developed a DL algorithm that used 389 FAF images to classify patients with 3 different IRDs (RP, Best disease and Stargardt disease) from healthy controls.[20] However, their study was fundamentally limited by the fact that only a minority of this cohort was genetically-confirmed.

In this issue of the *British Journal of Ophthalmology*, Fujinami-Yokokawa et al utilized a Japanese Eye Genetics Consortium dataset of 417 images (fundus photographs and FAF images) from 156 subjects, containing 115 genetically-confirmed cases of *ABCA4*-, *EYS*- and *RP1L1*-associated retinal dystrophies, and 41 normal age-matched controls, to train and validate a DL system for automated classification amongst these 4 categories. They report encouraging overall sensitivity/specificity values for fundus photographs and FAF images of 88.3%/97.4% and 81.8%/95.5%, respectively. They also report area under the receiver operating characteristic curve (AUC) values for fundus photographs and FAF images of 0.708 and 0.703 respectively. The authors conclude that the DL system they have developed could provide accurate, easily accessible diagnosis of these 3 important IRDs in Japan, which may one day help to provide earlier diagnosis, more appropriate referrals, and lower cost of investigation and genetic testing at the general ophthalmologist level. This study represents a step in the right direction for AI in the field of IRD. It tackles the 3 most prevalent genetic causes of IRD in the Japanese population, includes only genetically-confirmed cases, and was developed on a clinically well-characterized IRD cohort. It also uses 2 different and

commonly used imaging modalities in IRD assessment – colour fundus photographs and FAF imaging.

Nevertheless, there are important limitations to this study, which serve to highlight some of the key challenges in applying AI to IRD diagnosis.

First, it is important to note that this is, by nature, an artificially curated dataset.

Typical phenotypes associated with *ABCA4*, *EYS*, and *RP1L1* disease-causing variants are primarily Stargardt disease, RP and occult macular dystrophy respectively, which are markedly different from each other on colour photography and FAF. However, individually, each of these genotypes can result in different, even overlapping, phenotypes. For example, *ABCA4* variants can produce a cone dystrophy/cone-rod dystrophy phenotype which can have similarities with RP especially in advanced stages, while *EYS* variants can also be associated with a cone-rod dystrophy phenotype, and *RP1L1* can also be associated with RP.[22]

Furthermore, these challenges may be compounded by ethnic-specific variants and founder effects for specific IRD genes. There is limited information on the phenotypic heterogeneity within each genetic group in this study. Therefore, it is difficult to be certain if the DL system developed is truly predicting the *causative gene*, or if it is relying largely on notable phenotypic differences to make its classification decisions.

An alternative approach with potential advantages may be to group subjects initially by phenotype (e.g. Stargardt disease) with different disease-causing variants, and use AI-DL within this subset to predict the causative variant. This may also make any DL system developed more clinically applicable – in practice, patients first present with a particular phenotype (e.g. Stargardt disease), and it is the clinician's task to

then solve the genetic basis – a task which one day an AI system may be able to facilitate. Of course, such an analysis would require a much larger sample size, and a very well-characterized cohort of patients, both of which are potential challenges in the field of rare/orphan disease.

Second, this AI system deals with mutations in only 3 genes, in a specific ethnic population. In clinical practice, pathogenic variants in more than 250 genes can cause IRD, with significant ethnic and geographic variation. This AI system in its current form is not yet ready for general deployment on “unknown” IRD patients, that may exhibit significant genotypic and phenotypic heterogeneity.

Third, as with most other DL systems in the early stages of development, this study relies on a selected subset of images of good quality for training and validation. The authors excluded 31.5% of colour fundus photographs, and 60.4% of FAF images in their initial quality assessment. Therefore, the encouraging sensitivity and specificity results reported in their validation need to be interpreted in context. The challenge here, and indeed with other DL systems pushing towards clinical translation, will be to prove that the algorithms developed are robust and generalisable, by testing on external datasets with a variety of image quality and acquisition protocols. Ultimately, DL systems employed as “physician assistive tools”, as opposed to fully autonomous systems (e.g. for DR screening), are less vulnerable to image quality concerns, but this still needs to be addressed in future studies.

Fourth, the AUC values reported in this study of 0.708 and 0.703 are relatively low as compared to DL systems for other ocular conditions. This may be due to

limitations of small sample size or reflect an inherent difficulty in the task at hand.

Future studies with larger sample sizes, and appropriate performance benchmarking against human experts will help to resolve this issue.

FUTURE DIRECTIONS FOR AI APPLICATIONS IN IRD

Evidently, to support further development of AI tools for IRD, there is a need for more well-characterized, sizable datasets. IRDs are individually rare diseases, and there exists also significant geographic variation in relative prevalence patterns of genotypes and phenotypes. Therefore, in order to develop clinically useful and robust AI algorithms, there is a need for greater international collaboration.

Key areas for future development include:

1. Global standardization of nomenclature, classification systems, and data collection for IRD databases by international IRD consortia/consensus groups – this will provide strong ground truth data for AI development
2. Inclusion of additional information from the retinal periphery by greater reliance on UWF colour and FAF images
3. Inclusion of key clinical history data as algorithm input (such as age of onset, family history etc.)
4. Multi-class and multi-modal (including UWF imaging, FAF imaging, OCT, electrophysiology and visual field) AI algorithms for classification of IRD
5. Triaging AI algorithms to stratify the genotype of IRDs to determine suitability for specific gene therapies

6. Predictive AI algorithms on IRD progression
7. Testing of AI algorithms on independent external datasets, with a variety of image quality and acquisition protocols
8. Benchmarking of AI algorithms against expert human performance

CONCLUSIONS

AI systems capable of providing accurate phenotypic or genotypic classification of IRDs based on clinical and multi-modal imaging data could become useful assistive tools for clinicians, particularly in clinics or centres where sub-specialty expertise in IRD is lacking. This would help to fill expertise and resource gaps, and guide general clinicians to order appropriate genetic tests by a targeted testing strategy, which would help to keep testing costs low, minimize the FGR, and potentially, also further improve the rates of genetic diagnosis.

If AI tools can also identify causative sequence variants in certain “difficult” cases and outperform human experts, then analysis of “heatmaps” and reasons for the AI classification decisions may help to identify novel subtle phenotypic abnormalities that point to the underlying genetic diagnosis.

Hopefully, with a concerted effort and greater international collaboration, such effective AI tools for more rapid and accurate diagnosis of IRD can be developed to optimize patient care, with the promise that blindness from IRD can be ultimately prevented or eliminated.

FUNDING STATEMENT

This research received no specific grant from any funding agency in the public, commercial or not-for-profit sectors.

COMPETING INTERESTS STATEMENT

TYW and DSWT are co-inventors, with patents pending, for a deep learning system for diabetic retinopathy, glaucoma, and age-related macular degeneration (SG Non-Provisional Application number 10201706186V), and a computer-implemented method for training an image classifier using weakly annotated training data (SG Provisional Patent Application number 10201901083Y), and are co-founders and shareholders of EyRIS, Singapore. All other authors declare no relevant competing interests.

CONTRIBUTORSHIP STATEMENT

TET, HWC, MSS, TYW, JSP, MM, EHS, DSWT contributed to drafting and revising the manuscript.

REFERENCES

- 1 Moore AT. Genetic Testing for Inherited Retinal Disease. *Ophthalmology* 2017;**124**:1254–5. doi:10.1016/j.ophtha.2017.06.018
- 2 Galvin O, Chi G, Brady L, *et al.* The Impact of Inherited Retinal Diseases in the Republic of Ireland (ROI) and the United Kingdom (UK) from a Cost-of-Illness Perspective. *Clin Ophthalmol* 2020;**14**:707–19. doi:10.2147/OPHTH.S241928
- 3 Russell S, Bennett J, Wellman JA, *et al.* Efficacy and safety of voretigene neparvovec (AAV2-hRPE65v2) in patients with RPE65-mediated inherited retinal dystrophy: a randomised, controlled, open-label, phase 3 trial. *Lancet* 2017;**390**:849–60. doi:10.1016/S0140-6736(17)31868-8
- 4 MacLaren RE, Groppe M, Barnard AR, *et al.* Retinal gene therapy in patients with choroideremia: initial findings from a phase 1/2 clinical trial. *Lancet* 2014;**383**:1129–37. doi:10.1016/S0140-6736(13)62117-0
- 5 Hafler BP. CLINICAL PROGRESS IN INHERITED RETINAL DEGENERATIONS: GENE THERAPY CLINICAL TRIALS AND ADVANCES IN GENETIC SEQUENCING. *Retina (Philadelphia, Pa)* 2017;**37**:417–23. doi:10.1097/IAE.0000000000001341
- 6 Cideciyan AV, Jacobson SG, Drack AV, *et al.* Effect of an intravitreal antisense oligonucleotide on vision in Leber congenital amaurosis due to a photoreceptor cilium defect. *Nat Med* 2019;**25**:225–8. doi:10.1038/s41591-018-0295-0
- 7 Pontikos N, Arno G, Jurkute N, *et al.* Genetic Basis of Inherited Retinal Disease in a Molecularly Characterized Cohort of More Than 3000 Families from the

United Kingdom. *Ophthalmology* Published Online First: 16 April 2020.

doi:10.1016/j.ophtha.2020.04.008

- 8 Weisschuh N, Obermaier CD, Battke F, *et al.* Genetic architecture of inherited retinal degeneration in Germany: A large cohort study from a single diagnostic center over a 9-year period. *Hum Mutat* 2020;**41**:1514–27.
doi:10.1002/humu.24064
- 9 Sharon D, Ben-Yosef T, Goldenberg-Cohen N, *et al.* A nationwide genetic analysis of inherited retinal diseases in Israel as assessed by the Israeli inherited retinal disease consortium (IIRDC). *Hum Mutat* 2020;**41**:140–9.
doi:10.1002/humu.23903
- 10 Stone EM, Andorf JL, Whitmore SS, *et al.* Clinically Focused Molecular Investigation of 1000 Consecutive Families with Inherited Retinal Disease. *Ophthalmology* 2017;**124**:1314–31. doi:10.1016/j.ophtha.2017.04.008
- 11 LeCun Y, Bengio Y, Hinton G. Deep learning. *Nature* 2015;**521**:436–44.
doi:10.1038/nature14539
- 12 Callaway E. ‘It will change everything’: DeepMind’s AI makes gigantic leap in solving protein structures. *Nature* 2020;**588**:203–4. doi:10.1038/d41586-020-03348-4
- 13 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. doi:10.1001/jama.2016.17216

- 14 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. doi:10.1001/jama.2017.18152
- 15 Li Z, He Y, Keel S, *et al.* Efficacy of a Deep Learning System for Detecting Glaucomatous Optic Neuropathy Based on Color Fundus Photographs. *Ophthalmology* 2018;**125**:1199–206. doi:10.1016/j.ophtha.2018.01.023
- 16 Keenan TDL, Chen Q, Peng Y, *et al.* Deep Learning Automated Detection of Reticular Pseudodrusen from Fundus Autofluorescence Images or Color Fundus Photographs in AREDS2. *Ophthalmology* 2020;**127**:1674–87. doi:10.1016/j.ophtha.2020.05.036
- 17 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**:1342–50. doi:10.1038/s41591-018-0107-6
- 18 Shah M, Roomans Ledo A, Rittscher J. Automated classification of normal and Stargardt disease optical coherence tomography images using deep learning. *Acta Ophthalmol* 2020;**98**:e715–21. doi:10.1111/aos.14353
- 19 Miere A, Capuano V, Kessler A, *et al.* Deep learning-based classification of retinal atrophy using fundus autofluorescence imaging. *Comput Biol Med* 2021;**130**:104198. doi:10.1016/j.combiomed.2020.104198
- 20 Miere A, Le Meur T, Bitton K, *et al.* Deep Learning-Based Classification of Inherited Retinal Diseases Using Fundus Autofluorescence. *J Clin Med* 2020;**9**. doi:10.3390/jcm9103303

21 Masumoto H, Tabuchi H, Nakakura S, *et al.* Accuracy of a deep convolutional neural network in detection of retinitis pigmentosa on ultrawide-field images.

PeerJ 2019;**7**:e6900. doi:10.7717/peerj.6900

22 Yang L, Fujinami K, Ueno S, *et al.* Genetic Spectrum of EYS-associated Retinal Disease in a Large Japanese Cohort: Identification of Disease-associated Variants with Relatively High Allele Frequency. *Sci Rep* 2020;**10**:5497.

doi:10.1038/s41598-020-62119-3