frontiers | Research Topics

Big data and artificial intelligence in ophthalmology

Edited by
Tae-im Kim, Darren Shu Jeng Ting, Yi-Ting Hsieh and Tyler Hyungtaek Rim

Published in
Frontiers in Medicine

ISSN 1664-8714
ISBN 978-2-83251-780-2
DOI 10.3389/978-2-83251-780-2

About Frontiers

Frontiers is more than just an open access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

Frontiers journal series

The Frontiers journal series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the *Frontiers journal series* operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

Dedication to quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews. Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view. By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the *Frontiers journals series*: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area.

Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers editorial office: frontiersin.org/about/contact

Big data and artificial intelligence in ophthalmology

Topic editors

Tae-im Kim — Yonsei University, Republic of Korea
Darren Shu Jeng Ting — University of Nottingham, United Kingdom
Yi-Ting Hsieh — National Taiwan University Hospital, Taiwan
Tyler Hyungtaek Rim — Mediwhale Inc, Republic of Korea

Citation

Kim, T.-i., Ting, D. S. J., Hsieh, Y.-T., Rim, T. H., eds. (2023). *Big data and artificial intelligence in ophthalmology*. Lausanne: Frontiers Media SA.
doi: 10.3389/978-2-83251-780-2

Table of contents

frontiers | Frontiers in Medicine

TYPE Editorial
PUBLISHED 14 February 2023
DOI 10.3389/fmed.2023.1145522

OPEN ACCESS

EDITED AND REVIEWED BY
Jodhbir Mehta,
Singapore National Eye Center, Singapore

*CORRESPONDENCE
Tae-im Kim
✉ TIKIM@yuhs.ac

SPECIALTY SECTION
This article was submitted to Ophthalmology, a section of the journal Frontiers in Medicine

RECEIVED 16 January 2023
ACCEPTED 01 February 2023
PUBLISHED 14 February 2023

CITATION
Thakur S, Rim TH, Ting DSJ, Hsieh Y-T and Kim T-i (2023) Editorial: Big data and artificial intelligence in ophthalmology.
Front. Med. 10:1145522.
doi: 10.3389/fmed.2023.1145522

Editorial: Big data and artificial intelligence in ophthalmology

Sahil Thakur[1], Tyler Hyungtaek Rim[1,2], Darren S. J. Ting[3,4,5], Yi-Ting Hsieh[6,7] and Tae-im Kim[8,9]*

[1]Department of Ocular Epidemiology, Singapore Eye Research Institute, Singapore, Singapore, [2]Mediwhale Inc., Seoul, Republic of Korea, [3]Academic Unit of Ophthalmology, Institute of Inflammation and Ageing, University of Birmingham, Birmingham, United Kingdom, [4]Birmingham and Midland Eye Centre, Birmingham, United Kingdom, [5]Academic Ophthalmology, School of Medicine, University of Nottingham, Nottingham, United Kingdom, [6]Department of Ophthalmology, National Taiwan University Hospital, Taipei, Taiwan, [7]Department of Ophthalmology, College of Medicine, National Taiwan University, Taipei, Taiwan, [8]Department of Ophthalmology, The Institute of Vision Research, Yonsei University College of Medicine, Seoul, Republic of Korea, [9]Department of Ophthalmology, Corneal Dystrophy Research Institute, Yonsei University College of Medicine, Seoul, Republic of Korea

KEYWORDS

big data, ophthalmology, artificial intelligence, deep learning, machine learning

Editorial on the Research Topic
Big data and artificial intelligence in ophthalmology

Big Data and Artificial Intelligence (AI) are rapidly transforming modern healthcare. The combination of these technologies allows for the collection, analysis, and utilization of large amounts of healthcare data in ways that were previously not possible. While there are several challenges to overcome, the potential benefits of using these technologies are significant and include improved patient outcomes, efficient and effective healthcare delivery, and potential of improving access and affordability of healthcare interventions. Ophthalmology is a field that generates a large amount of healthcare data, including images of the retina, cornea, and other eye structures (1). The widespread use of electronic health records has also led to an increase in the amount of patient data available for analysis (2, 3). Thus, ophthalmology is a great medical subspeciality for applications that can utilize big data and AI.

This Research Topic focuses on the utility and the potential of big data and AI in ophthalmology. Authors from a broad spectrum of vision science and ophthalmology associated specialties from several countries, have contributed to this Research Topic. They have highlighted novel uses of large datasets, introduced new perspectives, and have reported AI algorithms with immense translational potential. In this editorial we provide a thematic overview of the exciting and diverse content covered under this Research Topic.

1. Publication trends

Yang et al. have performed a bibliometric analysis of the publication trends in AI in retina from 2012 to 2022 and report interesting findings. Countries like US and China have been leading the research output with maximum number of publications (US:171, China: 149), citations (US:2466, China: 1401), and H-index (US:28, China: 20). However, several issues such as lack of real-world testing of algorithms, meaningful economic impact assessment, use of multimodal imaging data for algorithm development and ethical, regulatory, and legal complexities associated with dataset curation need to be addressed by the researchers. Additionally, adequate representation of different populations in training datasets and algorithm generalizability are other areas of concern that need attention.

2. Anterior segment

In this Research Topic we observe the utility of big data to answer perplexing clinical problems. Kwak et al. used the KNHIS-Senior database ($n = 558{,}147$) to demonstrate that the rate or risk of surgical complications of cataract surgery did not change with tamsulosin use in the Korean elderly population. These findings contradict conventional understanding that intraoperative floppy iris syndrome (IFIS) is frequent in patients taking tamsulosin and can cause significant perioperative or postoperative complications during cataract surgery (4). Though the authors were cautious in interpreting their findings and attributed their results to careful surgeon's effort to respond to perioperative complications and advances in surgical equipment, the big data driven approach of the study shows how conventional observational "wisdom" may sometimes not hold true when tested against benchmarks of real world "evidence." Another study by Ahn et al. (a) demonstrated the value of big data by showing that surgically induced astigmatism (SIA) was higher in the femtosecond assisted cataract surgery + arcuate keratotomy group than the conventional phacoemulsification group (0.886 vs. 0.631 $p < 0.001$). The overcorrection ratios were also higher in the in the femtosecond group (58.9%) vs. the conventional group (48.8%). Though the femtosecond laser was effective when target induced astigmatism (TIA) values are greater than 0.75 D, overcorrection in patients with a lower degree of astigmatism and the angle of error in patients with higher astigmatism may lead to higher postoperative corneal astigmatism. Further research is thus needed to understand the factors that affect astigmatism in femtosecond laser assisted cataract surgery.

Ahn et al. (b) also demonstrated that multi-source ASOCT images can be used for estimating preoperative best corrected visual acuity (BCVA). This AI biomarker can be used as a surrogate for cataract grade as well. The authors also reported that in the subgroup which had an absolute error (AE) ≥ 0.1, subjects had significant vision impairing disease like macular disease/glaucoma or another optic neuropathy. Thus, a more intuitive approach would be to use both anterior and posterior segment imaging to formulate an algorithm that provides an estimate of both pre and postoperative BCVA. Surgeons and patients would then be able to effectively manage expectations and deliver satisfying outcomes.

Qian et al. showed that anterior chamber depth (ACD) can be predicted using smartphone captured images of the anterior segment. The MAE reported by the authors was 0.16 ± 0.13 mm, and R^2 between the predicted and measured ACD was 0.40. The central corneal region was highlighted in the saliency maps indicating that the predicted ACD was correlated with the clinically used site for ACD measurement. Such algorithms that utilize easily available consumer technology for image capture and subsequently can predict important ocular biomarkers have potential for rapid deployment in the real world.

Ahn et al. (c) also demonstrated the possibility of automating hospital workflows by predicting pupil dilation based on medical interview and basic eye examinations data. Using a large well-curated dataset of 56,811 patients over a period of 3 years, the authors demonstrated a sensitivity of 94.2–75.7% and specificity of 96.2–96% for predicting the need of a pupil dilation test only based on basic clinical information. The authors identified that asymptomatic lesions however led to reduced performance of the model, though it is still interesting to see how big data and innovative AI algorithms may improve and automate hospital and clinic workflows.

3. Posterior segment

Lin et al. showed that OCT images could be used to infer VA information using a deep learning algorithm in patients with diabetic macular edema (DME). Traditionally VA is documented using chart based methods that are prone to subjectivity and depend on chart quality and illumination. AI based methods for VA estimation can be thus used for subjects with poor cooperation and provide surrogate functional vision endpoints for monitoring.

Lu et al. demonstrated prediction of axial length classes using choroidal thickness measures from 2D OCT images. The model however requires 6 point choroidal thickness measurement and variables like age, gender, height, and weight for axial length classification. Development of future AI models that can predict axial length without additional demographic information can be explored using larger and multimodal datasets.

Wang et al. demonstrate a deep learning model for DME classification with AUC range of 98.1–95.2%, sensitivity of 96.4–87.4%, and specificity of 90.2–90.1% in three large datasets. The model is novel as it can localize hard exudates along with anatomical landmarks. Diabetic retinopathy (DR) and its associated complications offer several opportunities for AI based algorithm development and potential use due to ease of retinal image capture using portable fundus cameras, well-defined disease labels and high disease prevalence.

In another study using big data from the Korean NHIS-Senior database, Kim et al. showed that in elderly patients with retinal diseases, the vitrectomy group showed the lower mortality from pulmonary causes when compared to those without vitrectomy. The associations were different based on underlying vitreoretinal disease: higher risk of all-cause mortality and vascular causes in patient subgroup with retinal vascular diseases and lower risk of all-cause mortality, vascular causes, and pulmonary causes in those with macular diseases. Such serendipitous results are difficult to explain as in this study a greater proportion of the patients with macular diseases who underwent vitrectomy were current or ex-smokers who regularly consumed alcohol. However, such results inspire discussion and future research into the impact of surgical interventions for ocular disease on patient mortality.

4. APPRAISE survey

Gunasekeran et al. conducted the Acceptance and Perception of Artificial Intelligence Usability in Eye Care (APPRAISE) survey to evaluate the global perspective of ophthalmologists ($n = 1176$) regarding AI, focusing on four major eye conditions, namely DR, age related macular degeneration, glaucoma, and

cataract. This survey highlighted that most respondents (80.9%) believed that the pandemic had played an important role in the willingness to adopt AI tools due to global focus on teleophthalmology and digital solutions for patient care while lockdowns were being implemented. While the goal of AI developers is to provide comprehensive AI solutions for patient care, ophthalmologists are more willing to use AI as clinical assistive tools (88.1%), when compared to clinical decision support tools (78.8%) or diagnostic tools (64.5%). The survey also highlighted the perceived advantages of AI based tools in patient screening (94.5%), improved access (84.7%), affordability (61.9%), quality (69.4%), targeted referrals (87.1%), and reduction of monotonous work (82.7%). Some potential disadvantages of AI like concerns over medical liability for errors (72.5%) and data security/privacy concerns (64.9%) were also mentioned. While AI is often mentioned as a threat to jobs in mainstream media, most survey responders were confident their roles will not be replaced (68.2%). The survey thus provides a comprehensive insight into the current perspective of ophthalmologists regarding AI based tools. The results can be utilized by all stakeholders to facilitate effective communication and formulate targeted interventions to address barriers that hamper development, adoption, and use of AI tools in ophthalmology.

5. Myopia and future trends

Zhang X. et al. report alarming projections of myopia affecting 8.57 million children (7–12 years) and 15.77 million adolescents (13–18 years) by the year 2050 in eastern China. Simple low cost interventions like outdoor activities, frame glasses and eye exercises have high utilization prevalence and can significantly reduce the burden of myopia. AI can also be used as a tool for myopia detection, monitoring and management. In their review Zhang C. et al., highlighted the potential uses of AI in addressing myopia holistically. With the advent of virtual reality (VR) and augmented reality (AR) into consumer realm, there are also opportunities to also develop intelligent digital tools that can aid in behavioral interventions for myopia control. However early detection of myopia and its complications followed by timely therapeutic interventions remain critical for myopia management.

6. Conclusion

This Research Topic has a diverse array of publications that cover a spectrum of topics dealing with the use of big data and applications of AI in ophthalmology. The ideas, algorithms and perspectives discussed and published by the researchers in this topic have potential for considerable impact on shaping the future landscape of ophthalmology. Collaborative efforts built on these foundations can help in translating these innovations across the frontiers of ophthalmology and medicine for effective patient care and welfare.

Author contributions

ST contributed to conception and design of the study and wrote the first draft of the manuscript. All authors contributed to manuscript revision, read, and approved the submitted version.

Acknowledgments

The authors would like to thank the hundreds of colleagues who contributed to this Research Topic. The authors would also like to thank the board and staff of the Frontiers Publishing house for their continuous and unflinching support.

Conflict of interest

TR was employed by Mediwhale Inc.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

References

1. Cheng CY, Soh ZD, Majithia S, Thakur S, Rim TH, Tham YC, et al. Big data in ophthalmology. *Asia Pac J Ophthalmol.* (2020) 9:291–8. doi: 10.1097/APO.0000000000000304

2. Rim TH, Lee G, Kim Y, Tham YC, Lee CJ, Baik SJ, et al. Prediction of systemic biomarkers from retinal photographs: development and validation of deep-learning algorithms. *Lancet Digit Health.* (2020) 2:e526–36. doi: 10.1016/S2589-7500(20)30216-8

3. Rim TH, Lee CJ, Tham YC, Cheung N, Yu M, Lee G, et al. Deep-learning-based cardiovascular risk stratification using coronary artery calcium scores predicted from retinal photographs. *Lancet Digit Health.* (2021) 3:e306–16. doi: 10.1016/S2589-7500(21)00043-1

4. Enright JM, Karacal H, Tsai LM. Floppy iris syndrome and cataract surgery. *Curr Opin Ophthalmol.* (2017) 28:29–34. doi: 10.1097/ICU.0000000000000322

frontiers in Medicine

REVIEW
published: 11 March 2022
doi: 10.3389/fmed.2022.840498

Applications of Artificial Intelligence in Myopia: Current and Future Directions

*Chenchen Zhang, Jing Zhao, Zhe Zhu, Yanxia Li, Ke Li, Yuanping Wang and Yajuan Zheng**

Department of Ophthalmology, The Second Hospital of Jilin University, Changchun, China

With the continuous development of computer technology, big data acquisition and imaging methods, the application of artificial intelligence (AI) in medical fields is expanding. The use of machine learning and deep learning in the diagnosis and treatment of ophthalmic diseases is becoming more widespread. As one of the main causes of visual impairment, myopia has a high global prevalence. Early screening or diagnosis of myopia, combined with other effective therapeutic interventions, is very important to maintain a patient's visual function and quality of life. Through the training of fundus photography, optical coherence tomography, and slit lamp images and through platforms provided by telemedicine, AI shows great application potential in the detection, diagnosis, progression prediction and treatment of myopia. In addition, AI models and wearable devices based on other forms of data also perform well in the behavioral intervention of myopia patients. Admittedly, there are still some challenges in the practical application of AI in myopia, such as the standardization of datasets; acceptance attitudes of users; and ethical, legal and regulatory issues. This paper reviews the clinical application status, potential challenges and future directions of AI in myopia and proposes that the establishment of an AI-integrated telemedicine platform will be a new direction for myopia management in the post-COVID-19 period.

Keywords: artificial intelligence, machine learning, deep learning, telemedicine, myopia

OPEN ACCESS

Edited by:
Tae-Im Kim,
Yonsei University, South Korea

Reviewed by:
Beatrice Gallo,
Epsom and St. Helier University Hospitals NHS Trust, United Kingdom
Chih-Chien Hsu,
Taipei Veterans General Hospital, Taiwan

****Correspondence:***
Yajuan Zheng
zhengyajuan124@126.com

Specialty section:
This article was submitted to Ophthalmology, a section of the journal Frontiers in Medicine

Received: *21 December 2021*
Accepted: *15 February 2022*
Published: *11 March 2022*

Citation:
Zhang C, Zhao J, Zhu Z, Li Y, Li K, Wang Y and Zheng Y (2022) Applications of Artificial Intelligence in Myopia: Current and Future Directions. Front. Med. 9:840498. doi: 10.3389/fmed.2022.840498

INTRODUCTION

With the continuous development of computer technology, big data acquisition and imaging methods, the application of artificial intelligence (AI) in medical fields is expanding. Recently, a large number of AI-related studies have been carried out in many disciplines, such as ophthalmology, radiology, cardiovascularology, and oncology (1–4). Thanks to the development of multimodal imaging, fundus photography and optical coherence tomography (OCT) have provided rich datasets for the development of AI models and have made it possible for AI to flourish in the field of ophthalmology. The study of diseases has expanded from initial diabetic retinopathy (5–8), age-related macular degeneration (9–11), and glaucoma (12–15) to anterior segment diseases, such as refractive error (16–18).

Refractive error, represented by myopia, is becoming a key public health issue. As any degree of myopia will increase the risk of adverse changes in eye tissue, high myopia and pathological myopia (PM) significantly increase the risk of irreversible visual impairment [e.g., glaucoma, retinal detachment, myopic macular degeneration (MMD), and macular choroidal neovascularization] or blindness (19). Early identification of high-risk groups of myopia and regular and repeated follow-up to document the progression of myopia and complications are essential for eye care providers to plan interventions. However, current healthcare systems may not be able to cope with the growing burden. In particular, the COVID-19 pandemic demonstrates the need for remote testing and monitoring. Fortunately, AI technology combined with telemedicine can bridge this gap. To date, studies have integrated AI into all stages of clinical practice of myopia and have achieved positive application effects. This paper introduces the concepts of AI, summarizes the clinical application status, discusses potential challenges and future directions of AI in myopia, and proposes that the establishment of an AI-integrated telemedicine platform will be a new direction of myopia healthcare to provide personalized management throughout the whole process for myopia patients in the post-COVID-19 period.

AI, MACHINE LEARNING, AND DEEP LEARNING

The concept of AI was first proposed by John McCarthy in 1956. Its definition simulates human intelligence through machines (20). Machine learning (ML) is a branch of AI and mainly uses computer system programming to perform tasks or predict results (21). ML has great potential in clinical practice and machine translation (22). Traditional ML algorithms use variables selected by experts as input and usually do not involve large neural networks. They include algorithms such as linear regression, logistic regression, support vector machine, decision tree, and random forest algorithms (23). Deep learning (DL) is a subset of ML. Without special programming, it can automatically extract the rules from known data for the judgment of unknown data; hence, DL can process more complex data (24). DL algorithms usually involve the use of large-scale neural networks, such as artificial neural networks (ANNs), convolutional neural networks (CNNs) and recurrent neural networks (RNNs) (23). Since 2012, the introduction of CNNs has allowed for major breakthroughs in DL in imaging-based applications (e.g., object recognition, image segmentation, and disease classification) (24). VGG, ResNet, Inception and Inception-ResNet are some of the popular CNNs used for classification and are now widely used in medical image recognition (23). Deep CNNs can learn the feature representation from data without human knowledge and have the power to process large training data with high dimensionality. Studies have shown that the accuracy of medical image analysis systems based on DL in disease detection is equal to or even better than that of clinicians or trained personnel (25, 26). Moreover, other studies have proven the potential and feasibility of applying DL algorithms to disease screening and detection (27, 28). The diagnosis of many ophthalmic diseases requires not only symptom evaluation but also imaging information. This feature leads to the widespread use of AI technology represented by DL in clinical ophthalmology (1).

The indexes used to evaluate the quality of an AI model are accuracy, sensitivity and specificity, which are calculated by using four quantitative indexes: true positive, false positive, true negative and false negative (**Table 1**). A receiver operating characteristic curve (ROC) can be drawn with the false positive rate (FPR) as the X-axis and the true positive rate (TPR) as the Y-axis. The area under the curve (AUC) is defined as the area under the ROC curve and generally ranges from 0.5 (for a model with no predictive value) to 1 (for a perfect model) (29) (**Figure 1**).

GLOBAL BURDEN OF MYOPIA

Myopia is one of the most common ophthalmic diseases in the world. It mainly occurs in childhood and early adulthood (30). According to the work of Holden and his coworkers, the global prevalence of myopia is close to 28.3% (2 billion) of the world's population, of which 4.0% (277 million) suffer from high myopia. The "myopia epidemic" is estimated to affect 49.8% (4.758 billion) of the world's population by 2050, with 9.8% (938 million) suffering from high myopia ($\leq$-5.00 D). Of note, Holden et al. standardized to a spherical equivalent of 5.00 D or less for high myopia because it is widely used to identify people at higher risk of pathologic myopia (31). Nature (genetics and heredity) and nurture (environment and lifestyle) are all factors leading to myopia (19). For most people with myopia, the most critical risk factor is likely to be related to modern lifestyles, which include long periods of close-eye activity. The outbreak

TABLE 1 | Common terminologies used to evaluate AI model performance.

		Predicted outcome	
		Disease	**No disease**
Actual outcome	**Disease**	True positive (TP)	False negative (FN)
	No disease	False positive (FP)	True negative (TN)
Remark	Accuracy = (TP+TN)/(TP+FN+FP+TN)		
	Sensitivity = TP/(TP+TN)		
	Specificity = TN/(TN+FP)		
	True positive rate (TPR) = Sensitivity		
	False positive rate (FPR) = 1-Specificity		

Abbreviations: AI, Artificial intelligence; OCT, Optical coherence tomography; ML, Machine learning; DL, Deep learning; ANNs, Artificial neural networks; CNNs, Convolutional neural networks; RNNs, Recurrent neural networks; ROC, Receiver operating characteristic curve; FPR, False positive rate; TPR, True positive rate; AUC, Area under the curve; 5G, 5th generation mobile communication technology; MMD, myopic macular degeneration; PM, Pathologic myopia; LASEK, Laser epithelial keratomileusis; LASIK, Laser *in situ* keratomileusis; SMILE, Small incision lenticular extraction; PIOL, Phakic intraocular lens; IOL, Intraocular lens; AL, Axial length; ACD, Anterior chamber depth; BUII, Barrett Universal II; RBF, Radial basis function.

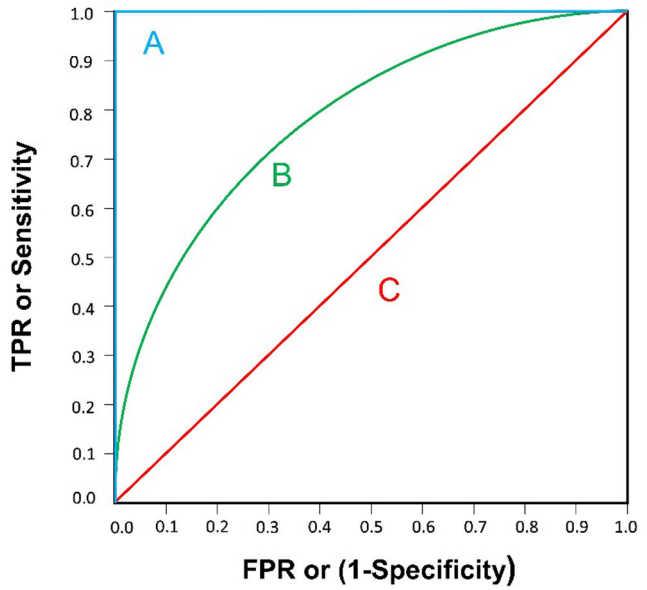

FIGURE 1 | Three examples of ROC curve are illustrated. **(A)** AUC=1: A "perfect" classifier; **(B)** 0.5<AUC<1: A real-world classifier, better than random guess; **(C)** AUC=0.5: Like random guess (e.g., coin tossing), models have no predictive value.

of COVID-19 at the end of 2019 undoubtedly exacerbated the above phenomenon. Research shows that during the COVID-19 pandemic, the reduced time spent outdoors and increased exposure to electronic screens have led to a further increase in the risk of myopia in children (32, 33).

Most cases of myopia are associated with excessive axial growth (19). Retinal damage caused by excessive axial growth is irreversible. Irreversible visual impairments caused by myopia (e.g., glaucoma, retinal detachment, MMD and macular choroidal neovascularization) or blindness not only increase medical costs but also reduce the quality of life of patients, which has caused a global medical and economic burden. Therefore, it is of great significance to comprehensively carry out myopia healthcare services, including the detection, diagnosis, progression prediction and treatment of myopia, as well as the management and prevention of ocular complications and visual impairment in patients with high myopia.

AI IN THE DETECTION AND DIAGNOSIS OF MYOPIA

Refractive Error Assessment

To evaluate refractive error, traditional visual acuity examinations are not only time consuming and laborious but also rely on expensive machines and experienced doctors and technicians. People with expression difficulties (e.g., young children, the elderly, and patients with verbal communication disabilities) have particular difficulties cooperating during an examination (34). In developing countries or impoverished areas, the lack of doctors and medical equipment makes it difficult to accurately evaluate refractive error, and patients are likely to miss the optimal treatment window, resulting in an irreversible loss of vision. Thus, providing timely and high-quality refraction services that are accepted by the general population is extremely needed.

While it is generally difficult for ophthalmologists to evaluate refractive error from a retinal fundus photograph, DL techniques are capable of predicting them fairly accurately. Varadarajan et al. (16) trained a DL algorithm to predict refractive error from retinal fundus photographs. By analyzing attention maps to determine the parts of a photograph most relevant for prediction, they concluded that attention maps consistently highlighted the fovea as a feature that was important for prediction. Tan et al. (35) also reported that by using color fundus photographs, a system consisting of a CNN pretrained with the XGBoost algorithm was able to evaluate refractive error with a high degree of accuracy. Yang et al. (17) trained a DL system to detect myopia automatically from ocular appearance images, and the system obtained an AUC of 0.9270. The research demonstrated the possibility of screening and monitoring refractive status in children with myopia in remote areas.

The Diagnosis of Pathologic Myopia and Complications

PM is accompanied by degenerative changes in the retina, which, if left untreated, can lead to irrecoverable vision loss. It is essential for ophthalmologists to have a sustainable method of monitoring eyes with PM to reduce blinding complications, especially given that many PM patients are young or middle aged. However, the diagnosis of PM, defined as peripapillary atrophy and myopic maculopathy, generally requires a complete examination that includes an assessment of the visual acuity and color fundus photograph acquisition tasks that are labor intensive and skill-dependent (36).

Tan et al. (37) introduced a method to automatically detect PM *via* peripapillary atrophy features by means of variational level sets from fundus photographs. To improve prediction accuracy, Zhang et al. (38) proposed a computer-aided framework based on an ML algorithm for the detection of PM. By analyzing demographic and clinical information, retinal fundus photograph data and genotyping data from 2,258 subjects, this method achieved an AUC of 0.888 and outperformed the detection results obtained from the use of demographic and clinical information (an AUC of 0.607), imaging data (an AUC of 0.852) or genotyping data (an AUC of 0.774) alone, with increases of 46.3%, $p < 0.005$; 4.2%, $p = 0.19$; and 14.7%, $P < 0.005$, respectively. Recently, Hemelings et al. (39) developed a successful approach based on a DL algorithm for the simultaneous detection of PM, with an AUC of 0.9867, and the segmentation of myopia-induced lesions. Other similar studies have also been reported, such as those identifying the different types of lesions of myopic maculopathy automatically from fundus photographs with DL models (40, 41). In addition, OCT macular images were used for the development of CNN models to identify vision-threatening conditions, such as retinoschisis,

macular holes and retinal detachment, in adults with high myopia, and the models obtained good sensitivity and AUC scores (42, 43).

AI IN THE PREDICTION OF MYOPIA PROGRESSION

Considering the potential irreversible disease burden during adulthood, concerns from parents, clinicians and policy makers include the potential progression rate and risk of developing high or even pathological myopia from childhood myopia (44). Thus, predicting myopia progression can provide evidence for transforming clinical practice, health policy-making, and precise individualized interventions regarding the practical control of school-aged myopia.

Lin et al. (45) identified myopia development rules and predicted the onset of myopia and its progression for children and teenagers from clinical measures using a random forest ML model, which had good predictive performance (the AUC ranged from 0.801 to 0.837) for up to 8 years in the future. Yang et al. (46) developed a prediction model to predict myopia in adolescents based on both measurement and behavior data of primary school students, and the model achieved reasonable performance and accuracy. Further research is still required for interpopulation validation to allow these models to be generalized.

AI IN REFRACTIVE SURGERY FOR MYOPIA

The aim of refractive surgery is to correct refractive error in adults with stable myopia and reduce their dependence on corrective aids. Keratorefractive procedures and intraocular procedures are two main forms of refractive surgery. At present, keratorefractive procedures include laser epithelial keratomileusis (LASEK), laser *in situ* keratomileusis (LASIK) and small incision lenticular extraction (SMILE). Intraocular procedures include phakic intraocular lens (PIOL) implantation and cataract surgery (19). To achieve the goal of optimal visual and refractive outcomes and to minimize the risk of postoperative complications, researchers have creatively applied AI to various stages of refractive surgery and achieved ideal results, particularly in the preoperative screening for risk of ectasia following LASIK, guiding the formulation of surgical plans and intraocular lens (IOL) power calculations.

Preoperative Screening

In 1998, Seiler et al. (47) published the first reports of iatrogenic progressive ectasia after LASIK, also known as iatrogenic ectasia. This complication can cause postoperative refraction regression and seriously affect the operation effect. Ectasia occurs due to biomechanical decompensation of the stroma, which may be related to pre-existing biomechanical weakening (e.g., keratoconus, subclinical keratoconus, and forme fruste keratoconus) or a severe impact on the corneal structure (e.g., an attempted treatment for high myopia) (48). Screening before refractive surgeries is extremely important to identify candidates at high risk of iatrogenic ectasia. Xie et al. (49) combined a DL algorithm with corneal tomographic scans to develop the Pentacam InceptionResNetV2 Screening System to screen potential candidates for refractive surgery. They reported a sensitivity of 80% for identifying ectasia suspects, 90% for diagnosing early keratoconus, and an overall diagnostic accuracy of 95% with an AUC of 0.99. To train and develop more accurate AI-based algorithms for identifying candidates at high risk of iatrogenic ectasia, it is necessary to have a longitudinal follow-up and collect more clinical data to train and validate the AI models.

Guiding the Formulation of a Surgical Plan

AI technology can guide a surgeon in selecting the best corneal refractive surgery method to perform on a specific patient. Yoo et al. (50) developed an expert-level multiclass ML model for selecting refractive surgery options for patients. They classified patients into LASEK, LASIK, SMILE and contraindication groups. Using data from 18,480 subjects who intended to undergo refractive surgery, the model was trained to select the optimal refractive surgery type for patients with accuracies of 81 and 78.9% on the internal and external validation datasets, respectively. Cui et al. (51) developed an ML model to recommend a nomogram for SMILE surgery to achieve the optimal postoperative visual outcome. They reported that the efficacy index in the ML group (1.48 ± 1.08) was significantly higher than that in the surgeon group (1.3 ± 0.27) ($t = -2.17$, $P < 0.05$). For high myopia patients who intend to undergo PIOL surgery, which involves the insertion of an additional lens in the anterior segment, it is essential to have correct anterior chamber depth (ACD) measurement (52). ACD measurement is usually obtained with conventional A-scan ultrasound. However, these machines are expensive and cumbersome and may not be available in remote areas. Chen et al. (53) developed a new method for predicting central ACD using a portable smartphone slit lamp device aided by ML. This novel device may provide a new perspective to increase the convenience of ACD measurement.

IOL Power Calculation Related to Myopia

For patients who intend to undergo PIOL implantation or cataract surgery to correct refractive error, accurate IOL power is the key to improving their postoperative visual quality. Ongoing developments in IOL power calculation incorporate new technology and data science to improve the accuracy of IOL selection (54). Compared with the second- and third-generation formulas, fourth-generation formulas, such as the Olsen formula (based on ray tracing) and Barrett Universal II (BUII), show good accuracy and fewer refractive accidents (55). A recent study developed a new XGBoost ML-based calculator for highly myopic eyes, which incorporated the BUII formula results and showed a significant improvement in the percentage of eyes achieving ± 0.25 D of the prediction error compared with the BUII formula alone (18). To date, for high axial myopia, AI-based IOL formulas seem to demonstrate higher levels of accuracy, including the Hill-radial basis function (RBF) calculator and the Kane formula (56–59). The Hill-RBF calculator uses AI

and regression analysis with a very large database of actual postsurgical refractive outcomes to predict IOL power (59). Hill-RBF is based mainly on empirical data; thus, its accuracy is limited by the type of data and eye characteristics from which it is derived (54). To overcome this limitation, Hill-RBF 2.0 expanded the database and improved IOL power prediction for a wider range of eye characteristics, such as high axial myopia, by continuously collecting various eye characteristics and surgical results (57). In September 2020, Hill-RBF 3.0 was released. With the expansion of the Hill-RBF database, the calculator is more likely to obtain a better accuracy in IOL power prediction. The other promising method for IOL calculation is the Kane formula, which incorporates AI with theoretical optics to predict IOL power (54). Studies have shown that the Kane formula has a smaller absolute error than the BUII, Olsen, and Hill-RBF 2.0 formulas (60, 61). In a study of 10,930 eyes in Britain, the Kane formula had the lowest mean absolute prediction error for all ranges of ALs and obtained the smallest absolute error for long eyes (AL>26.0 mm) (60).

AI AND MONITORING DEVICES IN THE BEHAVIORAL INTERVENTION OF MYOPIA

Effective behavioral intervention is as important as early detection to prevent myopia or limit myopia progression. To understand behaviors related to myopic onset and progression, a wearable device named Vivior Monitor (Vivior AG, Zurich, Switzerland) was developed to investigate the visual behavior of children with myopia (6–16 years old) (62). Using ML algorithms, Vivior Monitor identified types of visual activities, such as viewing handheld media, desktop work, and computer work. This research reported that older children spent less time viewing objects at distances, more time using a computer and less time engaging in physical movement. There is no doubt that outdoor activity is the main protective factor against myopia (63, 64). Wearable devices in combination with internet or social network apps aimed at encouraging children to spend more time outdoors are now being developed. The Singapore Eye Research Institute developed a novel wearable fitness tracker (FitSight), which comprises a smartwatch (Sony Smartwatch 3; Sony Corp., Minato, Tokyo, Japan) with a light sensor and an accompanying smartphone app that logs time spent outdoors and sends feedback to parents and children (65). In addition, excessive near-work behavior is one of the most commonly known unhealthy visual behaviors related to myopia, and many studies have shown that it can speed the occurrence and development of myopia (66, 67). Clouclip (Glasson Technology Co. Ltd., Hangzhou, China), a cloud-based sensor device that attaches to the sides of spectacles, can objectively and dynamically monitor the wearer's near-work distance and duration (68, 69). This device can provide a vibration alert when it detects risky near-work-related behaviors, such as particularly short viewing distances or prolonged continuous near-work behavior. Cao et al. (68) collected data from 67 subjects who were assigned to wear Clouclip all day (except for bathing and sleeping) during the experiment; they found that the device can significantly modify near-work behaviors in school-age children and that its effects can last a certain period of time.

AI IN MYOPIA GENETICS

The mechanism of myopia is extremely complex. Nature (genetics and heredity) and nurture (environment and lifestyle) are all factors leading to myopia (19). In recent years, studies on the genetics of myopia have also received considerable attention. By linkage analysis, candidate gene analysis, genome-wide association study (GWAS) and next-generation sequencing (NGS), more than 100 genes and over 20 chromosomal loci have been identified to be associated with myopia or related quantitative traits (70–72). However, the current knowledge about the genetic contributions of the loci and genes to myopia remains limited (73).

To date, studies using big data for genetic analysis and phenotyping correlation have achieved significant progress in various medical fields (74, 75). Genomic readouts, combined with advanced AI, could be a powerful approach for risk prediction in multifactorial diseases such as myopia. At present, both CNNs and RNNs have shown considerable potential in a variety of clinical genomics applications, such as variant calling, genome annotation, and functional impact prediction (76). Given the diversity of myopia with regard to its environmental burden, geographic patterns, and affiliations with different ethnicities and cultural groups worldwide (73), further AI research with larger multiethnic genetic samples from various research institutes will be essential to drive the discovery of new insights into the genetic aspects of myopia and advance AI-genomic applications in managing childhood myopia (77).

NEW MODEL FOR MYOPIA MANAGEMENT: TELEMEDICINE

Telemedicine is a new service model in the medical field that aims to solve the problem of healthcare for people in remote and underdeveloped areas by providing remote medical services (78). The global COVID-19 pandemic is bringing telemedicine to the forefront of ophthalmic medical services (79, 80). With the development of AI technology and the expansion of 5G communication network coverage, AI-integrated telemedicine platforms will gradually become the new normal of post-COVID-19 ophthalmic care. In the clinical application of myopia, AI-integrated telemedicine platforms should mainly focus on the following aspects: reducing the manpower requirements of ophthalmic clinics, reducing the risk of direct physical contact between patients and doctors, and personalizing management throughout the whole process.

Devices based on AI enable non-ophthalmologist health care workers, such as optometrists, nurses and technicians, to perform several tasks, such as assessments of refractive error and measurements of ACDs, individually instead of patients moving through a number of different clinical staff, each performing a specific task. In addition, telemedicine can not only reduce the direct physical contact between patients and doctors but also

prolong the distance of ophthalmic examination. For example, the portable slit lamp examination distance has increased from 18 cm to 55 cm. The examination distance increased from 5 cm for the direct ophthalmoscope to 47 cm for the Glasgow Retinal Imaging Adaptor (Medical Devices Unit, NHS Greater Glasgow & Clyde, UK) (81). These changes can not only satisfy the need for regular and repeated follow-up to monitor and document the refractive status of myopia with high efficiency but also limit exposure risks.

To provide personalized management for myopia patients throughout the whole process, we first need to realize the integration of hospital-community-family health management. Recently, Wu et al. (82) proposed an AI-integrated telemedicine platform to screen and refer patients with cataracts. According to the authors, this telemedicine platform involves self-monitoring at home, primary healthcare and specialized hospital services. Inspired by this platform, we propose a new management model for myopia (**Figure 2**). First, considering that myopia develops primarily during childhood and early adulthood, large-scale refractive error screening of the target population will be carried out regularly with portable devices and technologies based on AI, and the examination data will be stored and documented on telemedicine platforms. Second, AI analysis will be conducted on the collected clinical data, images and potential genomic data to classify the risk of myopia progression in clinically identified individuals and formulate personalized management plans, including visual behavioral interventions for patients with wearable devices (77). Third, home monitoring (using ocular appearance images taken by family members with cell phones and visual acuity tests) can be implemented for patients without myopia-related complications. Home monitoring and community-based primary healthcare institutions (where retinal fundus photographs or OCT scans are captured and used in the telemedicine platform with AI analysis) can be used by myopia patients with non-blinding myopia complications. If the above patients develop pathological myopia or myopia with blinding complications, they can be referred to the specialized hospital *via* a fast tract notification system. Patients initially diagnosed with pathological myopia or blinding complications should be directly transferred to tertiary medical institutions. After treatment, the patient returned home and continued home monitoring. Fourth, for patients requiring surgical treatment, AI-integrated telemedicine can be applied to preoperative screening to determine the risk of ectasia following LASIK and guide the formulation of a surgical plan and IOL power calculation.

CURRENT CHALLENGES AND FUTURE DIRECTIONS

Despite the reported successful clinical applications of AI in myopia, challenges and hurdles are still present. Critical technical and clinical limitations must be surmounted prior to the widespread implementation of AI in myopia.

Standardization of Datasets

Image-based AI technology has made some progress in the application of refractive error assessment, screening, diagnosis and treatment of myopia. However, image-based AI requires large, standardized, labeled data, and ophthalmic open datasets are very small compared to ImageNet's tens of millions of images (13). Obtaining large-scale and high-quality images in a real clinical environment is a great challenge. Technically, more advanced data enhancement methods should be utilized, such as programming simulated lesions to be integrated into normal image data (83) or incorporating real lesions into other locations in normal or abnormal images (84). Recent studies have proposed alternative training methods that can learn from less data. For example, some studies synthesize a large number of random and diverse medical images by generative adversarial networks and report that these images can be used as CNN training datasets in the future (85–87). However, these new methods have not achieved significant success thus far, and their effectiveness needs to be further proven (88). In addition to the amount of data, the quality of images also plays a great part in the performance of AI models (5, 89). Research has reported that poor-quality fundus images that were not removed from the dataset were found to decrease the AUC by 0.1 (90). To surmount this challenge, Wu et al. (23) proposed a quality assessment system for images to select high-quality images. The feasibility of this method needs further study.

Attitude Toward AI

As DL is an end-to-end learning method, that is, inputting original data and outputting results directly without manual coding, DL lacks the ability to explain the detection results and cannot provide an exact judgment basis for the results; this is called the "black box phenomenon." This could reduce the acceptance of test results by ophthalmologists and patients (91). With the development of DL, several approaches are currently available to help improve the interpretability of the results, including occlusion tests (92) and saliency maps (93). However, there is no consensus on which saliency map generation method is most appropriate for ophthalmic imaging data (93). In addition, it is unclear how one should interpret non-traditional features identified by saliency analysis, that is, whether they should be treated as novel biomarkers or erroneous correlations "learned" during training. Processes need to be in place to address such disagreements, such as an independent third party from a multidisciplinary team, as would occur where there is clinical uncertainty (36). Apart from that, education on the implementation and appraisal of AI systems should be included in medical school programs and hospital training to prepare for its adoption when the technology reaches maturation for ophthalmology clinical practice.

Ethical, Legal, and Regulatory Issues

With the increasing use of AI, security and privacy have become issues of concern and involve ethical, legal and regulatory issues (94). For example, an AI algorithm, similar to a human ophthalmologist, is definitely prone to errors. Who is responsible for bearing the legal consequences of an undesirable outcome due to an erroneous judgment made by an AI algorithm? Is it the

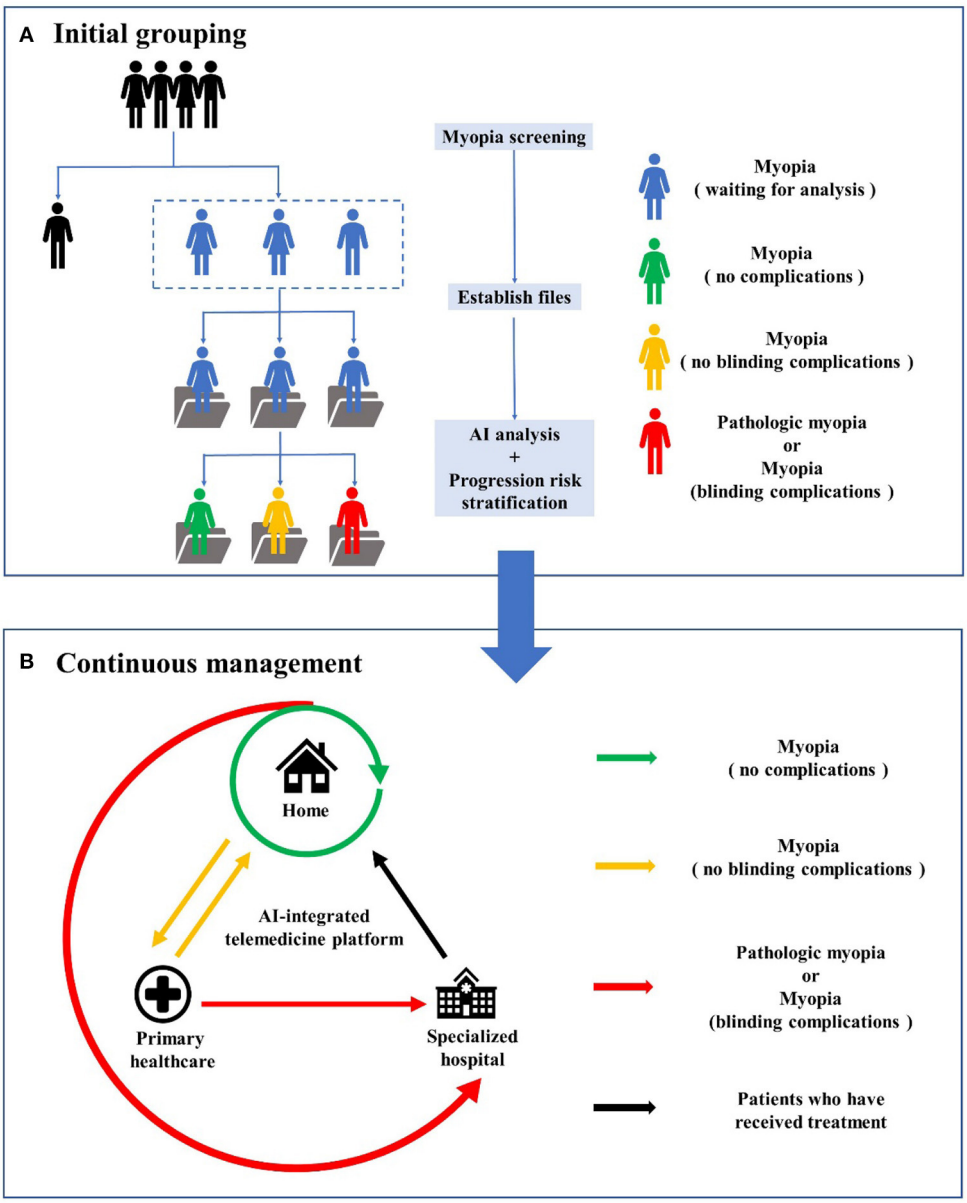

FIGURE 2 | The workflow of AI-integrated telemedicine platform for myopia. **(A)** is the workflow of initial grouping, including myopia screening, files establishing, AI analysis and progression risk stratification for myopia patients. **(B)** is the workflow of continuous management involving self-monitoring at home, primary healthcare and specialized hospital services.

company that develops the algorithm, the individual physician who utilizes the algorithm, or the healthcare organization under which the physician is employed (95)? In addition, protocols and laws aimed at guaranteeing training data and testing data security in AI need to be continually established and improved.

CONCLUSIONS

Given the rapid increases in the prevalence of all levels of myopia in the past three decades and the non-linear rapid COVID-19 disease expansion, there is a need to revolutionize healthcare systems worldwide. Three main areas are the targets for such revolutions: improving efficiency, limiting exposure risk, and providing individualized management for myopic patients. AI is among the most promising solutions to address these issues. Prior to the mass adoption of AI in myopia, AI models need to be further optimized to improve their interpretability, human–machine interactions, generalization abilities, and robustness. It is also necessary to develop relevant clinical standards, integrate large-scale clinical datasets, and develop a standard evaluation framework for AI models in

clinical practice. Moreover, relevant laws and regulations need to be constantly improved to achieve comprehensive supervision of practical applications.

AUTHOR CONTRIBUTIONS

CZ, JZ, and ZZ conceived the idea for the article and contributed to the initial drafting of the manuscript. YL, KL, and YW performed literature search and collected data. YZ was involved in reviewing and editing the manuscript. All authors read and approved the final manuscript.

FUNDING

This work was supported by the Science and Technology Development Plan Project of Jilin Province (No. 20190303186SF).

REFERENCES

1. Ting DSW, Peng L, Varadarajan AV, Keane PA, Burlina PM, Chiang MF, et al. Deep learning in ophthalmology: the technical and clinical considerations. *Prog Retin Eye Res.* (2019) 72:100759. doi: 10.1016/j.preteyeres.2019.04.003
2. Saba L, Biswas M, Kuppili V, Cuadrado Godia E, Suri HS, Edla DR, et al. The present and future of deep learning in radiology. *Eur J Radiol.* (2019) 114:14–24. doi: 10.1016/j.ejrad.2019.02.038
3. Dey D, Slomka PJ, Leeson P, Comaniciu D, Shrestha S, Sengupta PP, et al. Artificial intelligence in cardiovascular imaging: JACC state-of-the-art review. *J Am Coll Cardiol.* (2019) 73:1317–35. doi: 10.1016/j.jacc.2018.12.054
4. Bi WL, Hosny A, Schabath MB, Giger ML, Birkbak NJ, Mehrtash A, et al. Artificial intelligence in cancer imaging: clinical challenges and applications. *CA Cancer J Clin.* (2019) 69:127–57. doi: 10.3322/caac.21552
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. doi: 10.1001/jama.2017.18152
6. 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. doi: 10.1001/jama.2016.17216
7. Gargeya R, Leng T. Automated identification of diabetic retinopathy using deep learning. *Ophthalmology.* (2017) 124:962–9. doi: 10.1016/j.ophtha.2017.02.008
8. 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. doi: 10.1167/iovs.16-19964
9. Burlina PM, Joshi N, Pekala M, Pacheco KD, Freund DE, Bressler NM. Automated grading of age-related macular degeneration from color fundus images using deep convolutional neural networks. *JAMA Ophthalmol.* (2017) 135:1170–6. doi: 10.1001/jamaophthalmol.2017.3782
10. Treder M, Lauermann JL, Eter N. Automated detection of exudative age-related macular degeneration in spectral domain optical coherence tomography using deep learning. *Graefes Arch Clin Exp Ophthalmol.* (2018) 256:259–65. doi: 10.1007/s00417-017-3850-3
11. Schlegl T, Waldstein SM, Bogunovic H, Endstraßer F, Sadeghipour A, Philip AM, et al. Fully automated detection and quantification of macular fluid in OCT using deep learning. *Ophthalmology.* (2018) 125:549–58. doi: 10.1016/j.ophtha.2017.10.031
12. Asaoka R, Murata H, Hirasawa K, Fujino Y, Matsuura M, Miki A, et al. Using deep learning and transfer learning to accurately diagnose early-onset glaucoma from macular optical coherence tomography images. *Am J Ophthalmol.* (2019) 198:136–45. doi: 10.1016/j.ajo.2018.10.007
13. Shibata N, Tanito M, Mitsuhashi K, Fujino Y, Matsuura M, Murata H, et al. Development of a deep residual learning algorithm to screen for glaucoma from fundus photography. *Sci Rep.* (2018) 8:14665. doi: 10.1038/s41598-018-33013-w
14. Hood DC, De Moraes CG. Efficacy of a deep learning system for detecting glaucomatous optic neuropathy based on color fundus photographs. *Ophthalmology.* (2018) 125:1207–8. doi: 10.1016/j.ophtha.2018.04.020
15. Asaoka R, Murata H, Iwase A, Araie M. Detecting preperimetric glaucoma with standard automated perimetry using a deep learning classifier. *Ophthalmology.* (2016) 123:1974–80. doi: 10.1016/j.ophtha.2016.05.029
16. Varadarajan AV, Poplin R, Blumer K, Angermueller C, Ledsam J, Chopra R, et al. Deep learning for predicting refractive error from retinal fundus images. *Invest Ophthalmol Vis Sci.* (2018) 59:2861–8. doi: 10.1167/iovs.18-23887
17. Yang Y, Li R, Lin D, Zhang X, Li W, Wang J, et al. Automatic identification of myopia based on ocular appearance images using deep learning. *Ann Transl Med.* (2020) 8:705. doi: 10.21037/atm.2019.12.39
18. Rampat R, Deshmukh R, Chen X, Ting DSW, Said DG, Dua HS, et al. Artificial intelligence in cornea, refractive surgery, and cataract: basic principles, clinical applications, and future directions. *Asia Pac J Ophthalmol.* (2021) 10:268–81. doi: 10.1097/apo.0000000000000394
19. Baird PN, Saw SM, Lanca C, Guggenheim JA, Smith Iii EL, Zhou X, et al. Myopia. *Nat Rev Dis Primers.* (2020) 6:99. doi: 10.1038/s41572-020-00231-4
20. Mccarthy J, Minsky M, Rochester N, Shannon CEJAM. A proposal for the Dartmouth summer research project on artificial intelligence, August 31, 1955. *AI Magazine.* (2006) 27:12–4. doi: 10.1609/aimag.v27i4.1904
21. Ajani TS, Imoize AL, Atayero AA. An overview of machine learning within embedded and mobile devices-optimizations and applications. *Sensors.* (2021) 21:412. doi: 10.3390/s21134412
22. Bengio Y, Courville A, Vincent P. Representation learning: a review and new perspectives. *IEEE Trans Pattern Anal Mach Intell.* (2013) 35:1798–828. doi: 10.1109/tpami.2013.50
23. Wu X, Liu L, Zhao L, Guo C, Li R, Wang T, et al. Application of artificial intelligence in anterior segment ophthalmic diseases: diversity and standardization. *Ann Transl Med.* (2020) 8:714. doi: 10.21037/atm-20-976
24. LeCun Y, Bengio Y, Hinton G. Deep learning. *Nature.* (2015) 521:436–44. doi: 10.1038/nature14539
25. Hekler A, Utikal JS, Enk AH, Solass W, Schmitt M, Klode J, et al. Deep learning outperformed 11 pathologists in the classification of histopathological melanoma images. *Eur J Cancer.* (2019) 118:91–6. doi: 10.1016/j.ejca.2019.06.012
26. Maron RC, Weichenthal M, Utikal JS, Hekler A, Berking C, Hauschild A, et al. Systematic outperformance of 112 dermatologists in multiclass skin cancer image classification by convolutional neural networks. *Eur J Cancer.* (2019) 119:57–65. doi: 10.1016/j.ejca.2019.06.013
27. Dascalu A, David EO. Skin cancer detection by deep learning and sound analysis algorithms: a prospective clinical study of an elementary dermoscope. *EBioMedicine.* (2019) 43:107–13. doi: 10.1016/j.ebiom.2019.04.055
28. Al-Antari MA, Al-Masni MA, Kim TS. Deep learning computer-aided diagnosis for breast lesion in digital mammogram. *Adv Exp Med Biol.* (2020) 1213:59–72. doi: 10.1007/978-3-030-33128-3_4
29. Deo RC. Machine learning in medicine. *Circulation.* (2015) 132:1920–30. doi: 10.1161/circulationaha.115.001593
30. Morgan IG, French AN, Ashby RS, Guo X, Ding X, He M, et al. The epidemics of myopia: Aetiology and prevention. *Prog Retin Eye Res.* (2018) 62:134–49. doi: 10.1016/j.preteyeres.2017.09.004
31. Holden BA, Fricke TR, Wilson DA, Jong M, Naidoo KS, Sankaridurg P, et al. Global prevalence of myopia and high myopia and temporal trends from 2000 through 2050. *Ophthalmology.* (2016) 123:1036–42. doi: 10.1016/j.ophtha.2016.01.006

32. Wong CW, Tsai A, Jonas JB, Ohno-Matsui K, Chen J, Ang M, et al. Digital screen time during the COVID-19 pandemic: risk for a further myopia boom? *Am J Ophthalmol.* (2021) 223:333–7. doi: 10.1016/j.ajo.2020.07.034
33. Wang J, Li Y, Musch DC, Wei N, Qi X, Ding G, et al. Progression of myopia in school-aged children after COVID-19 home confinement. *JAMA Ophthalmol.* (2021) 139:293–300. doi: 10.1001/jamaophthalmol.2020.6239
34. Amirsolaimani B, Peyman G, Schwiegerling J, Bablumyan A, Peyghambarian N. A new low-cost, compact, auto-phoropter for refractive assessment in developing countries. *Sci Rep.* (2017) 7:13990. doi: 10.1038/s41598-017-14507-5
35. Tan T-E, Ting DS, Liu Y, Li S, Chen C, Nguyen Q, et al. Artificial intelligence using a deep learning system with transfer learning to predict refractive error and myopic macular degeneration from color fundus photographs. *Investig Ophthalmol Vis Sci.* (2019) 60:1478. Available online at: https://iovs.arvojournals.org/article.aspx?articleid=2745932
36. Li JO, Liu H, Ting DSJ, Jeon S, Chan RVP, Kim JE, et al. Digital technology, tele-medicine and artificial intelligence in ophthalmology: a global perspective. *Prog Retin Eye Res.* (2021) 82:100900. doi: 10.1016/j.preteyeres.2020.100900
37. Tan NM, Liu J, Wong DK, Lim JH, Zhang Z, Lu S, et al. Automatic detection of pathological myopia using variational level set. *Annu Int Conf IEEE Eng Med Biol Soc.* (2009) 2009:3609–612. doi: 10.1109/iembs.2009.5333517
38. Zhang Z, Xu Y, Liu J, Wong DW, Kwoh CK, Saw SM, et al. Automatic diagnosis of pathological myopia from heterogeneous biomedical data. *PLoS ONE.* (2013) 8:e65736. doi: 10.1371/journal.pone.0065736
39. Hemelings R, Elen B, Blaschko MB, Jacob J, Stalmans I, De Boever P. Pathological myopia classification with simultaneous lesion segmentation using deep learning. *Comput Methods Programs Biomed.* (2021) 199:105920. doi: 10.1016/j.cmpb.2020.105920
40. Du R, Xie S, Fang Y, Igarashi-Yokoi T, Moriyama M, Ogata S, et al. Deep learning approach for automated detection of myopic maculopathy and pathologic myopia in fundus images. *Ophthalmol Retina.* (2021) 5:1235–44. doi: 10.1016/j.oret.2021.02.006
41. Tan TE, Anees A, Chen C, Li S, Xu X, Li Z, et al. Retinal photograph-based deep learning algorithms for myopia and a blockchain platform to facilitate artificial intelligence medical research: a retrospective multicohort study. *Lancet Digit Health.* (2021) 3:e317–29. doi: 10.1016/s2589-7500(21)00055-8
42. Li Y, Feng W, Zhao X, Liu B, Zhang Y, Chi W, et al. Development and validation of a deep learning system to screen vision-threatening conditions in high myopia using optical coherence tomography images. *Br J Ophthalmol.* (2020). doi: 10.1136/bjophthalmol-2020-317825
43. Sogawa T, Tabuchi H, Nagasato D, Masumoto H, Ikuno Y, Ohsugi H, et al. Accuracy of a deep convolutional neural network in the detection of myopic macular diseases using swept-source optical coherence tomography. *PLoS ONE.* (2020) 15:e0227240. doi: 10.1371/journal.pone.0227240
44. Foo LL, Ang M, Wong CW, Ohno-Matsui K, Saw SM, Wong TY, et al. Is artificial intelligence a solution to the myopia pandemic? *Br J Ophthalmol.* (2021) 105:741–4. doi: 10.1136/bjophthalmol-2021-319129
45. Lin H, Long E, Ding X, Diao H, Chen Z, Liu R, et al. Prediction of myopia development among Chinese school-aged children using refraction data from electronic medical records: a retrospective, multicentre machine learning study. *PLoS Med.* (2018) 15:e1002674. doi: 10.1371/journal.pmed.1002674
46. Yang X, Chen G, Qian Y, Wang Y, Zhai Y, Fan D, et al. Prediction of myopia in adolescents through machine learning methods. *Int J Environ Res Public Health.* (2020) 17:463. doi: 10.3390/ijerph17020463
47. Seiler T, Koufala K, Richter G. Iatrogenic keratectasia after laser *in situ* keratomileusis. *J Refract Surg.* (1998) 14:312–7. doi: 10.3928/1081-597X-19980501-15
48. Ambrósio R Jr. Post-LASIK Ectasia: twenty years of a conundrum. *Semin Ophthalmol.* (2019) 34:66–8. doi: 10.1080/08820538.2019.1569075
49. Xie Y, Zhao L, Yang X, Wu X, Yang Y, Huang X, et al. Screening candidates for refractive surgery with corneal tomographic-based deep learning. *JAMA Ophthalmol.* (2020) 138:519–26. doi: 10.1001/jamaophthalmol.2020.0507
50. Yoo TK, Ryu IH, Choi H, Kim JK, Lee IS, Kim JS, et al. Explainable machine learning approach as a tool to understand factors used to select the refractive surgery technique on the expert level. *Transl Vis Sci Technol.* (2020) 9:8. doi: 10.1167/tvst.9.2.8
51. Cui T, Wang Y, Ji S, Li Y, Hao W, Zou H, et al. Applying machine learning techniques in nomogram prediction and analysis for SMILE treatment. *Am J Ophthalmol.* (2020) 210:71–7. doi: 10.1016/j.ajo.2019.10.015
52. Guerra MG, Silva AMM, Marques SHM, Melo SH, Póvoa JA, Lobo C, et al. Phakic intraocular lens implantation: refractive outcome and safety in patients with anterior chamber depth between 2.8 and 3.0 versus ≥3.0 mm. *Ophthalmic Res.* (2017) 57:239–46. doi: 10.1159/000453528
53. Chen D, Ho Y, Sasa Y, Lee J, Yen CC, Tan C. Machine learning-guided prediction of central anterior chamber depth using slit lamp images from a portable smartphone device. *Biosensors.* (2021) 11:812. doi: 10.3390/bios11060182
54. Xia T, Martinez CE, Tsai LM. Update on intraocular lens formulas and calculations. *Asia Pac J Ophthalmol.* (2020) 9:186–93. doi: 10.1097/apo.0000000000000293
55. Melles RB, Holladay JT, Chang WJ. Accuracy of intraocular lens calculation formulas. *Ophthalmology.* (2018) 125:169–78. doi: 10.1016/j.ophtha.2017.08.027
56. Sramka M, Slovak M, Tuckova J, Stodulka P. Improving clinical refractive results of cataract surgery by machine learning. *PeerJ.* (2019) 7:e7202. doi: 10.7717/peerj.7202
57. Wan KH, Lam TCH, Yu MCY, Chan TCY. Accuracy and precision of intraocular lens calculations using the new Hill-RBF version 2.0 in eyes with high axial myopia. *Am J Ophthalmol.* (2019) 205:66–3. doi: 10.1016/j.ajo.2019.04.019
58. Savini G, Di Maita M, Hoffer KJ, Næser K, Schiano-Lomoriello D, Vagge A, et al. Comparison of 13 formulas for IOL power calculation with measurements from partial coherence interferometry. *Br J Ophthalmol.* (2021) 105:484–9. doi: 10.1136/bjophthalmol-2020-316193
59. Kane JX, Van Heerden A, Atik A, Petsoglou C. Accuracy of 3 new methods for intraocular lens power selection. *J Cataract Refract Surg.* (2017) 43:333–9. doi: 10.1016/j.jcrs.2016.12.021
60. Darcy K, Gunn D, Tavassoli S, Sparrow J, Kane JX. Assessment of the accuracy of new and updated intraocular lens power calculation formulas in 10 930 eyes from the UK National Health Service. *J Cataract Refract Surg.* (2020) 46:2–7. doi: 10.1016/j.jcrs.2019.08.014
61. Connell BJ, Kane JX. Comparison of the Kane formula with existing formulas for intraocular lens power selection. *BMJ Open Ophthalmol.* (2019) 4:e000251. doi: 10.1136/bmjophth-2018-000251
62. Mrochen M, Zakharov P, Tabakcı BN, Tanrıverdi C, Kılıç A, Flitcroft DI. Visual lifestyle of myopic children assessed with AI-powered wearable monitoring. *Investig Ophthalmol Vis Sci.* (2020) 61:82. Available online at: https://iovs.arvojournals.org/article.aspx?articleid=2766581
63. Rose KA, Morgan IG, Ip J, Kifley A, Huynh S, Smith W, et al. Outdoor activity reduces the prevalence of myopia in children. *Ophthalmology.* (2008) 115:1279–85. doi: 10.1016/j.ophtha.2007.12.019
64. Sherwin JC, Reacher MH, Keogh RH, Khawaja AP, Mackey DA, Foster PJ. The association between time spent outdoors and myopia in children and adolescents: a systematic review and meta-analysis. *Ophthalmology.* (2012) 119:2141–51. doi: 10.1016/j.ophtha.2012.04.020
65. Verkicharla PK, Ramamurthy D, Nguyen QD, Zhang X, Pu SH, Malhotra R, et al. Development of the FitSight fitness tracker to increase time outdoors to prevent myopia. *Transl Vis Sci Technol.* (2017) 6:20. doi: 10.1167/tvst.6.3.20
66. Lin Z, Vasudevan B, Mao GY, Ciuffreda KJ, Jhanji V, Li XX, et al. The influence of near work on myopic refractive change in urban students in Beijing: a three-year follow-up report. *Graefes Arch Clin Exp Ophthalmol.* (2016) 254:2247–55. doi: 10.1007/s00417-016-3440-9
67. Sun JT, An M, Yan XB, Li GH, Wang DB. Prevalence and related factors for myopia in school-aged children in Qingdao. *J Ophthalmol.* (2018) 2018:9781987. doi: 10.1155/2018/9781987
68. Cao Y, Lan W, Wen L, Li X, Pan L, Wang X, et al. An effectiveness study of a wearable device (Clouclip) intervention in unhealthy visual behaviors among school-age children: a pilot study. *Medicine.* (2020) 99:e17992. doi: 10.1097/md.0000000000017992
69. Wen L, Cheng Q, Lan W, Cao Y, Li X, Lu Y, et al. An objective comparison of light intensity and near-visual tasks between rural and urban school children in China by a wearable device Clouclip. *Transl Vis Sci Technol.* (2019) 8:15. doi: 10.1167/tvst.8.6.15

70. Verhoeven VJ, Hysi PG, Wojciechowski R, Fan Q, Guggenheim JA, Höhn R, et al. Genome-wide meta-analyses of multiancestry cohorts identify multiple new susceptibility loci for refractive error and myopia. *Nat Genet.* (2013) 45:314–8. doi: 10.1038/ng.2554
71. Fan Q, Barathi VA, Cheng CY, Zhou X, Meguro A, Nakata I, et al. Genetic variants on chromosome 1q41 influence ocular axial length and high myopia. *PLoS Genet.* (2012) 8:e1002753. doi: 10.1371/journal.pgen.1002753
72. Li J, Zhang Q. Insight into the molecular genetics of myopia. *Mol Vis.* (2017) 23:1048–80. Available online at: http://www.molvis.org/molvis/v23/1048/
73. Cai XB, Shen SR, Chen DF, Zhang Q, Jin ZB. An overview of myopia genetics. *Exp Eye Res.* (2019) 188:107778. doi: 10.1016/j.exer.2019.107778
74. Williams AM, Liu Y, Regner KR, Jotterand F, Liu P, Liang M. Artificial intelligence, physiological genomics, and precision medicine. *Physiol Genomics.* (2018) 50:237–43. doi: 10.1152/physiolgenomics.00119.2017
75. Xu J, Yang P, Xue S, Sharma B, Sanchez-Martin M, Wang F, et al. Translating cancer genomics into precision medicine with artificial intelligence: applications, challenges and future perspectives. *Hum Genet.* (2019) 138:109–24. doi: 10.1007/s00439-019-01970-5
76. Dias R, Torkamani A. Artificial intelligence in clinical and genomic diagnostics. *Genome Med.* (2019) 11:70. doi: 10.1186/s13073-019-0689-8
77. Foo LL, Ng WY, Lim GYS, Tan TE, Ang M, Ting DSW. Artificial intelligence in myopia: current and future trends. *Curr Opin Ophthalmol.* (2021) 32:413–24. doi: 10.1097/icu.0000000000000791
78. Waller M, Stotler C. Telemedicine: a Primer. *Curr Allergy Asthma Rep.* (2018) 18:54. doi: 10.1007/s11882-018-0808-4
79. Hollander JE, Carr BG. Virtually perfect? Telemedicine for Covid-19. *N Engl J Med.* (2020) 382:1679–81. doi: 10.1056/NEJMp2003539
80. Ye Y, Wang J, Xie Y, Zhong J, Hu Y, Chen B, et al. Global teleophthalmology with iPhones for real-time slitlamp eye examination. *Eye Contact Lens.* (2014) 40:297–300. doi: 10.1097/icl.0000000000000051
81. Ghazala FR, Hamilton R, Giardini ME, Livingstone IAT. Teleophthalmology techniques increase ophthalmic examination distance. *Eye.* (2021) 35:1780–1. doi: 10.1038/s41433-020-1085-8
82. Wu X, Huang Y, Liu Z, Lai W, Long E, Zhang K, et al. Universal artificial intelligence platform for collaborative management of cataracts. *Br J Ophthalmol.* (2019) 103:1553–60. doi: 10.1136/bjophthalmol-2019-314729
83. Badano A, Graff CG, Badal A, Sharma D, Zeng R, Samuelson FW, et al. Evaluation of digital breast tomosynthesis as replacement of full-field digital mammography using an *in silico* imaging trial. *JAMA Netw Open.* (2018) 1:e185474. doi: 10.1001/jamanetworkopen.2018.5474
84. Cha KH, Petrick N, Pezeshk A, Graff CG, Sharma D, Badal A, et al. Evaluation of data augmentation via synthetic images for improved breast mass detection on mammograms using deep learning. *J Med Imaging.* (2020) 7:012703. doi: 10.1117/1.Jmi.7.1.012703
85. Maspero M, Savenije MHF, Dinkla AM, Seevinck PR, Intven MPW, Jurgenliemk-Schulz IM, et al. Dose evaluation of fast synthetic-CT generation using a generative adversarial network for general pelvis MR-only radiotherapy. *Phys Med Biol.* (2018) 63:185001. doi: 10.1088/1361-6560/aada6d
86. Jin CB, Kim H, Liu M, Jung W, Joo S, Park E, et al. Deep CT to MR synthesis using paired and unpaired data. *Sensors.* (2019) 19:2361. doi: 10.3390/s19102361
87. Yang Q, Yan P, Zhang Y, Yu H, Shi Y, Mou X, et al. Low-Dose CT image denoising using a generative adversarial network with Wasserstein distance and perceptual loss. *IEEE Trans Med Imaging.* (2018) 37:1348–57. doi: 10.1109/tmi.2018.2827462
88. Yi X, Walia E, Babyn P. Generative adversarial network in medical imaging: a review. *Med Image Anal.* (2019) 58:101552. doi: 10.1016/j.media.2019.101552
89. Lee JG, Jun S, Cho YW, Lee H, Kim GB, Seo JB, et al. Deep learning in medical imaging: general overview. *Korean J Radiol.* (2017) 18:570–84. doi: 10.3348/kjr.2017.18.4.570
90. Phan S, Satoh S, Yoda Y, Kashiwagi K, Oshika T. Evaluation of deep convolutional neural networks for glaucoma detection. *Jpn J Ophthalmol.* (2019) 63:276–83. doi: 10.1007/s10384-019-00659-6
91. Stead WW. Clinical implications and challenges of artificial intelligence and deep learning. *JAMA.* (2018) 320:1107–8. doi: 10.1001/jama.2018.11029
92. Zeiler MD, Fergus R. Visualizing and understanding convolutional networks. *arXiv [Preprint].*(2010). arXiv: 1311.2901v3. Available online at: https://arxiv.org/pdf/1311.2901.pdf (accessed November 28, 2013).
93. Liu TYA, Bressler NM. Controversies in artificial intelligence. *Curr Opin Ophthalmol.* (2020) 31:324–8. doi: 10.1097/icu.0000000000000694
94. Shahbaz R, Salducci M. Law and order of modern ophthalmology: Teleophthalmology, smartphones legal and ethics. *Eur J Ophthalmol.* (2021) 31:13–21. doi: 10.1177/1120672120934405
95. Sullivan HR, Schweikart SJ. Are current tort liability doctrines adequate for addressing injury caused by AI? *AMA J Ethics.* (2019) 21:E160. doi: 10.1001/amajethics.2019.160

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers | Frontiers in Medicine

ORIGINAL RESEARCH
published: 04 April 2022
doi: 10.3389/fmed.2022.851644

Diabetic Macular Edema Detection Using End-to-End Deep Fusion Model and Anatomical Landmark Visualization on an Edge Computing Device

Ting-Yuan Wang[1], Yi-Hao Chen[2], Jiann-Torng Chen[2], Jung-Tzu Liu[1], Po-Yi Wu[1†], Sung-Yen Chang[1†], Ya-Wen Lee[1†], Kuo-Chen Su[3] and Ching-Long Chen[2]*

[1] Information and Communications Research Laboratories, Industrial Technology Research Institute, Hsinchu, Taiwan, [2] Department of Ophthalmology, Tri-Service General Hospital, National Defense Medical Center, Taipei, Taiwan, [3] Department of Optometry, Chung Shan Medical University, Taichung, Taiwan

OPEN ACCESS

Edited by:
Darren Shu Jeng Ting,
University of Nottingham,
United Kingdom

Reviewed by:
José Cunha-Vaz,
Association for Innovation and Biomedical Research on Light and Image (AIBILI), Portugal
Gilbert Yong San Lim,
SingHealth, Singapore

***Correspondence:**
Ching-Long Chen
doc30881@mail.ndmctsgh.edu.tw

[†]These authors have contributed equally to this work

Specialty section:
This article was submitted to Ophthalmology, a section of the journal Frontiers in Medicine

Received: 10 January 2022
Accepted: 14 March 2022
Published: 04 April 2022

Citation:
Wang T-Y, Chen Y-H, Chen J-T, Liu J-T, Wu P-Y, Chang S-Y, Lee Y-W, Su K-C and Chen C-L (2022) Diabetic Macular Edema Detection Using End-to-End Deep Fusion Model and Anatomical Landmark Visualization on an Edge Computing Device. Front. Med. 9:851644. doi: 10.3389/fmed.2022.851644

Purpose: Diabetic macular edema (DME) is a common cause of vision impairment and blindness in patients with diabetes. However, vision loss can be prevented by regular eye examinations during primary care. This study aimed to design an artificial intelligence (AI) system to facilitate ophthalmology referrals by physicians.

Methods: We developed an end-to-end deep fusion model for DME classification and hard exudate (HE) detection. Based on the architecture of fusion model, we also applied a dual model which included an independent classifier and object detector to perform these two tasks separately. We used 35,001 annotated fundus images from three hospitals between 2007 and 2018 in Taiwan to create a private dataset. The Private dataset, Messidor-1 and Messidor-2 were used to assess the performance of the fusion model for DME classification and HE detection. A second object detector was trained to identify anatomical landmarks (optic disc and macula). We integrated the fusion model and the anatomical landmark detector, and evaluated their performance on an edge device, a device with limited compute resources.

Results: For DME classification of our private testing dataset, Messidor-1 and Messidor-2, the area under the receiver operating characteristic curve (AUC) for the fusion model had values of 98.1, 95.2, and 95.8%, the sensitivities were 96.4, 88.7, and 87.4%, the specificities were 90.1, 90.2, and 90.2%, and the accuracies were 90.8, 90.0, and 89.9%, respectively. In addition, the AUC was not significantly different for the fusion and dual models for the three datasets ($p = 0.743$, 0.942, and 0.114, respectively). For HE detection, the fusion model achieved a sensitivity of 79.5%, a specificity of 87.7%, and an accuracy of 86.3% using our private testing dataset. The sensitivity of the fusion model was higher than that of the dual model ($p = 0.048$). For optic disc and macula detection, the second object detector achieved accuracies of 98.4% (optic disc) and 99.3% (macula). The fusion model and the anatomical landmark detector can be deployed on a portable edge device.

Conclusion: This portable AI system exhibited excellent performance for the classification of DME, and the visualization of HE and anatomical locations. It facilitates interpretability and can serve as a clinical reference for physicians. Clinically, this system could be applied to diabetic eye screening to improve the interpretation of fundus imaging in patients with DME.

Keywords: diabetic macular edema, hard exudate, optic disc and macula, deep learning, visualization

INTRODUCTION

Diabetes is a prevalent disease that affects ∼476 million people worldwide (1). Diabetic macular edema (DME), characterized by the accumulation of extracellular fluid that leaks from blood vessels in the macula (2), is one of the complications of diabetes mellitus. DME can appear at any stage of diabetic retinopathy (DR) and is the leading cause of severe vision loss in working-age adults with diabetic mellitus (3). The Early Treatment of Diabetic Retinopathy Study (ETDRS) defined the criteria for DME and demonstrated the benefits of laser photocoagulation therapy (4). Currently, with the revolutionary development of intraocular medication, intravitreal injections of anti-vascular endothelial growth factor (anti-VEGF) and steroid agents are the first-line treatment as alternatives to traditional laser photocoagulation as they provide better vision recovery in patients with center-involved macular edema (5–7).

Early diagnosis plays an important role in DME treatment. Moreover, early management such as intensive diabetes control may reduce the risk of progressive retinopathy (8). Early diagnosis and preemptive treatment are facilitated by frequent diabetic eye screening, which reduces the risk of progression to blindness, and the associated socioeconomic burden. To date, owing to developments in the field of ophthalmic imaging, the detection of DME using optical coherence tomography (OCT) imaging is the gold standard in the decision-making process for DME treatment (9). However, limited by various factors, such as the requirements of expensive equipment and highly specialized technicians, OCT imaging is typically readily available in high-income countries. In contrast, retinal photography examination is feasible and affordable in low-income countries and remote areas (10). However, the number of people with diabetes worldwide is increasing yearly and is estimated to reach 571 million by 2025 (1). The rapid growth of diabetic patients is expected to increase the diagnostic burden associated with DME detection. As such, an efficacious and accurate automatic fundus imaging interpretation system is urgently needed.

In the past decade, several studies have focused on DME detection using feature engineering techniques, which extract features by selecting or transforming raw data. Among them, Siddalingaswamy et al. (11) identified DME by detecting hard exudates (HE) and the macula. Subsequently, decisions were made based on the distance between the HE and the macula. Machine learning algorithms have also been applied in several studies for feature extraction in DME classification (12–15). The advantage of feature engineering is that it utilizes a smaller training dataset to achieve satisfactory performance. However, the identification of salient and useful features depends on the experience of clinicians and is thus subjective and limited. In contrast to feature engineering techniques, deep learning, particularly convolutional neural networks (CNNs), is gaining popularity and has achieved significant success in medical imaging applications. This approach can automatically learn feature extraction by using a backbone network mainly comprising convolutional and pooling layers. Several studies have shown that various architectures of CNN can be used to effectively extract features in fundus images for subsequent classification of DR or DME (16–21).

Moreover, given that deep learning models lack interpretability and are viewed as black boxes (22), visualization of the lesion in fundus images is an important issue. Lesion visualization can improve the interpretability of non-ophthalmologist physicians. In addition, visualization is useful to physicians during an initial assessment before a patient is referred to an ophthalmologist for further evaluation, thereby substantially increasing the screening rate and reducing the workload of ophthalmologists. In addition, lesion visualization could help physicians to monitor the status and progression of the disease.

Generally, deep learning models are implemented in cloud computing environments or high-end computers, which provide more computing power and memory space. However, this is usually expensive and requires considerable network resources. These factors limit the application of deep learning models for medical image analysis in remote or resource-limited areas. Thus, an edge device is potentially suitable for the application of deep learning models for medical image analysis in these areas. Previous studies have demonstrated the feasibility of deploying deep learning models for medical image analysis on edge devices (23–25). However, a system with multiple models for disease classification and visualization requires more computing power and memory. Thus, the implementation of such a system on an edge device is challenging.

In this study, we designed an end-to-end deep fusion network model to perform two deep learning tasks, one for the classification of DME and the other for the visualization of HE lesions. We used a private dataset and two open datasets to evaluate the performance of this fusion model. We also added a second object detector model to identify anatomical landmarks (optic disc and macula). These models were deployed on an edge device. The private dataset was used to assess the performance of the models. Overall, this system could be used for diabetic eye screening by non-specialist physicians or in remote or resource-limited areas to improve the early diagnosis of DME. As a

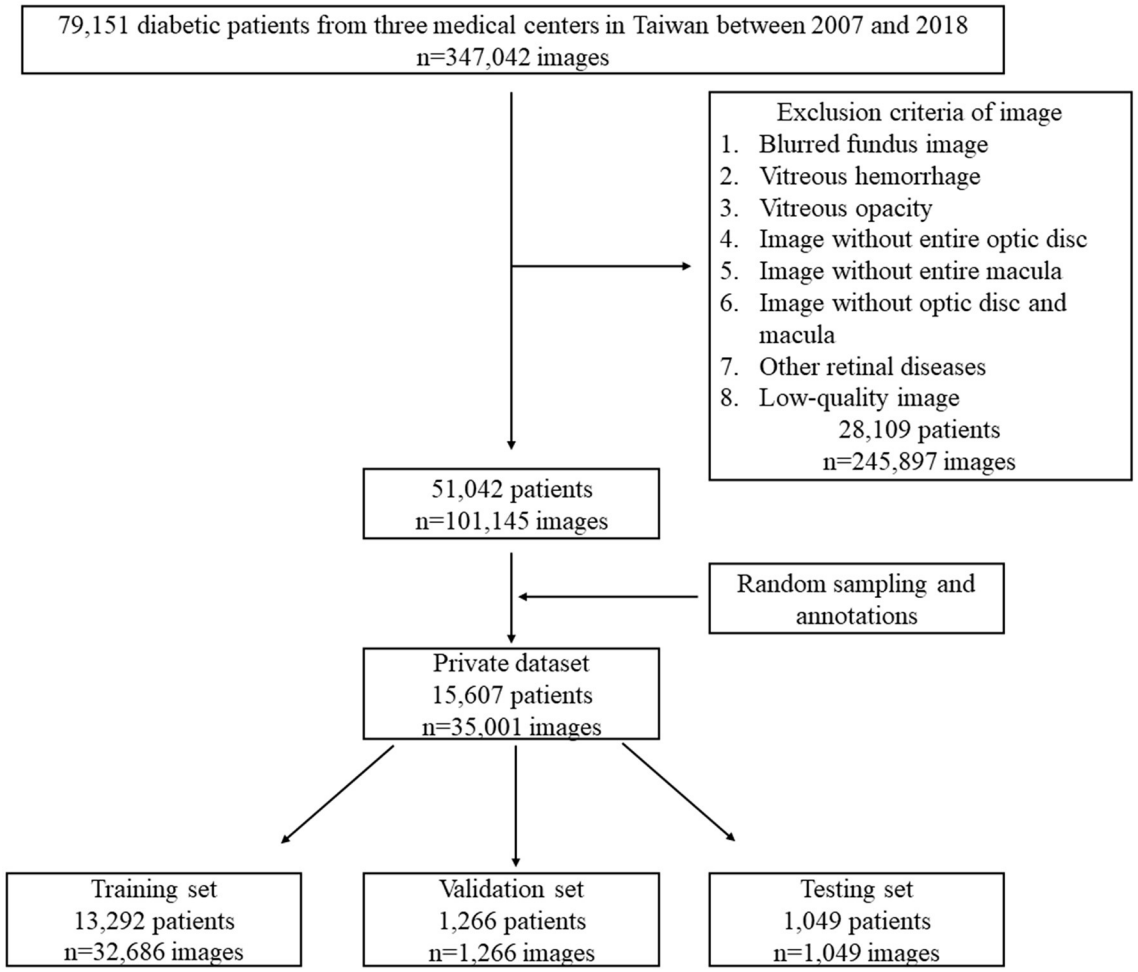

FIGURE 1 | The flowchart of our private dataset.

result, diabetic patients may be referred for early assessment and appropriate treatment, which should lead to better outcomes.

MATERIALS AND METHODS

Private Dataset

We enrolled patients who had a diagnosis of diabetic mellitus according to the ICD-9 codes 250.xx or ICD-10 codes E10-E14 between 2007 and 2018 from three medical centers in Taiwan. Patients younger than 20 years of age and with unknown sex were excluded. The retinal photographs were acquired from ZEISS (VISUCAM 200), Nidek (AFC-330), and Canon (CF-1, CR-DGI, CR2, or CR2-AF) fundus cameras with a 45° field-of-view (FOV) and anonymized owing to the retrospective nature of the study. We collected 347,042 fundus images from 79,151 diabetic patients. For the present study, we included image with optic disc and macula to develop models. Blurred fundus image, vitreous hemorrhage, vitreous opacity, image without entire optic disc, image without entire macula, image without optic disc and macula, other retinal diseases, and low-quality image were excluded, and 101,145 fundus images from 51,042 diabetic patients were left for random sampling and annotations. Finally, 35,001 fundus images from 15,607 patients formed our private dataset for model development (The flowchart shown in **Figure 1**). On our private dataset, the mean age of patients was 57.6 ± 11.8 years and 54.5% were males and 45.5% were females. Eight thousand four hundred and ninety-six patients took only one image and 7,111 patients took more than one image from each eye. The original dimension of the images were 522,728 pixels (724 × 722) to 12,212,224 pixels (4,288 × 2,848). All images were the JPG image format.

Ethical Considerations

The study was reviewed and approved by the institutional review board (IRB) of the three medical centers: Tri-Service General Hospital (IRB: 1-107-05-039), Chung Shan Medical University Hospital (IRB: CSH: CS18087), and China Medical Hospital (IRB: CMUH10FREC3-062). Given that the identities of all patients in three medical centers were encrypted before fundus

TABLE 1 | Dataset profile for the classification task in the private dataset.

Class	Training set		Validation set		Testing set	
	Number of images	Number of patients	Number of images	Number of patients	Number of images	Number of patients
Non-DME	18,921 (57.89%)	10,313	1,140 (90.05%)	1,140	939 (89.51%)	939
DME	13,765 (42.11%)	5,955	126 (9.95%)	126	110 (10.49%)	110
Total	32,686 (100.00%)	16,268	1,266 (100.00%)	1,266	1,049 (100.00%)	1,049

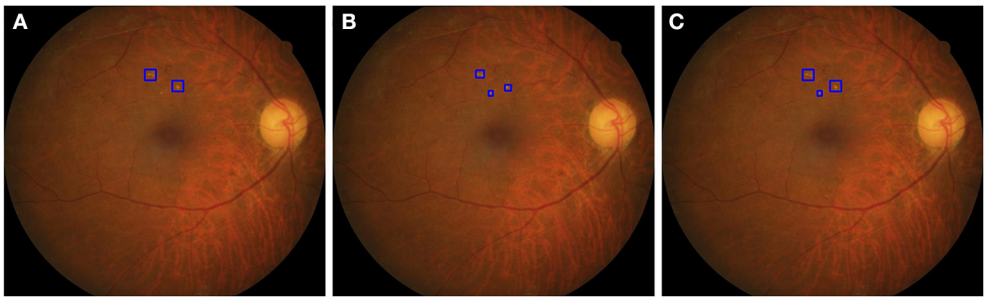

FIGURE 2 | The strategy to obtain a ground truth image. **(A)** The two boxes were annotated by first ophthalmologist. **(B)** The three boxes were annotated by second ophthalmologist. In step 1, the IoU of the top two bounding boxes in **(A,B)** were larger than 0.15, then two larger areas were taken as the GT. In step 2, the IoU of bottom boxes in **(A,B)** were <0.15, the area annotated by the second ophthalmologist was retained as the GT. After step 1 and 2, we obtained a GT image **(C)**.

images were released, the requirement for signed informed consent of the included patients was waived.

Annotations of Private Dataset

Annotating DME Classification for Fundus Image

We recruited 38 ophthalmologists to annotate the fundus images. Each fundus image was annotated by a group of three ophthalmologists. According to the criteria of ETDRS, DME was defined as any HE at or within 1 disc diameter (1DD) of the center of the macula (4). Each ophthalmologist annotated images by using our annotation tool. We used the majority decision of the three ophthalmologists as the ground truth (GT) of the fundus images. Further, the dataset was split into training, validation, and testing sets by patient level to prevent the same patient in different sets (**Figure 1**). Eight thousand four hundred and ninety-six of 15,607 patients took only one image and were randomly sampled to validation set (1,266 patients, 1,266 images) and testing set (1,049 patients, 1,049 images). The rest of these patients and 7,111 of 15,607 patients were reserved as a training set (13,292 patients, 32,686 images). **Table 1** lists the DME and non-DME profiles of these three subsets.

Annotating HE Lesions in Fundus Image

The HE lesions in each fundus image were also annotated by a group of three ophthalmologists (randomly chosen from 38 ophthalmologists) using a bounding box format. However, three resulting annotations may be different from each other in the number, size, and location of the boxes. We adopted the following procedure to obtain a final GT image for training purposes: (Step 1) The bounding boxes for the image labeled by two ophthalmologists were compared. If an HE lesion was annotated and the intersection over union (IoU) $>$ 0.15, then a larger annotated area was taken as the GT; (Step 2) The bounding boxes of an image labeled by two ophthalmologists were compared. If the HE lesion was annotated and the IoU $\leq$ 0.15, then both bounding boxes were retained as the GT. After step 1 and 2, we obtained the first GT image as shown in **Figure 2**. Step 3: First GT image was compared with the image labeled by the third ophthalmologist according to the same method in steps 1 and 2. Then, we obtained the final GT image. In this study, ophthalmologists used bounding boxes to annotate HE lesions in fundus image. The size of the annotated bounding boxes in original images were 9–5,196,672 pixels (9,791.57 $\pm$ 36,966.28 pixels). After resized the image, the size of the annotated bounding boxes in model's input images were 1.50–190,008.85 pixels (1,002.99 $\pm$ 2,719.79 pixels). However, the annotated bounding boxes only indicated whether existed HE lesions and location, not represented the true size of HE lesions. Therefore, the size of the bounding boxes was usually larger than the true size of the HE lesions. In addition, the profiles of HE labels of the three subsets are shown in **Table 2**.

Open Datasets

Two open datasets were used to evaluate the performance and ability of the proposed model to adapt to different datasets.

TABLE 2 | The number of images were annotated HE lesions by ophthalmologists in the private dataset.

	Training set	Validation set	Testing set
Number of images with HE	22,108	583	365
Total	32,686	1,266	1,049

Messidor-1

The Messidor-1 (26) dataset contained 1,200 fundus images from three ophthalmologic departments in France and was annotated with DR and the risk of DME. All images were acquired using a Topcon TRC NW6 non-mydriatic retinal camera with a 45° FOV. Our grading scheme was slightly different from that of Messidor-1, in which DME was graded according to three categories, with 0, 1, 2 representing "no visible HE," "HE presence at least 1DD away from the macula," and "HE presence within 1DD from the macula," respectively. As previously indicated, HE that occurs within 1DD of the center of the macula can serve as a proxy for detecting DME; hence, grades 0 and 1 are equivalent to non-DME and grade 2 is equivalent to DME in our classification scheme.

Messidor-2

The Messidor-2 (26, 27) dataset, as an extension of the Messidor-1 dataset, contained 1,748 (1,744 annotated as gradable) fundus images. In this study, we used 1,744 graded fundus images from the annotated Messidor-2 dataset by Krause et al. (28).

Deep Learning Models

Fusion Model Network

We use EfficientDet-d1 (29) as the object detector because of its great balance between performance and resource usage. Because EfficientDet-d1 employs the feature extraction part of EfficientNet-b1 (30), we can readily use this aspect as the backbone in the fusion model. Lesion detection was implemented using bi-directional feature pyramid network (BiFPN). The classification module consisted of three layers and included a convolutional layer, a global average pooling layer, and a fully connected (FC) layer. The architecture of the fusion model is shown in **Figure 3**.

The fusion model is computationally efficient, as only one convolution layer is needed to extract higher-level features based on the output features obtained from the EfficientDet-d1 backbone. We denote E_{ob} as the loss function of EfficientDet-d1, E_{cl} as the loss function of the classification module, and the loss function for the fusion model is given by Equation (1), where $\omega_{ob} > 0$ and $\omega_{cl} > 0$, which are hyperparameters used to linearly combine the loss functions of the object detector and classifier.

$$\begin{aligned} E_{loss} &= \omega_{ob} \times E_{ob} + \omega_{cl} \times E_{cl} \\ &= \omega_{ob} \times E_{ob} - \omega_{cl}[\alpha(1-p_t)^{\gamma} \log(p_t)] \end{aligned} \quad (1)$$

First, we use the equal weights for ω_{ob} and ω_{cl} in the initial training. Then analyzing the loss value obtained from the object detection model and the classification model. Second, we use the weighting factor (ω_{ob} and ω_{cl}) that is inversely proportional to the loss value of the classifier or object detector to balance the loss, respectively. Finally, we retrain the fusion model using ω_{ob} (= 0.5) and ω_{cl} (= 100) to balance the loss obtained from both models, and avoid overfitting in the classification model or the object detection model. Our results showed that the setting ω_{ob} =0.5 and ω_{cl} =100 achieved a satisfactory balance. The parameters $\alpha \geq 0$ and $\gamma \geq 0$ were also heuristically set to address the large class imbalance encountered during training. In general, α, the weight assigned to the rare class, should be slightly reduced as γ is increased (31). Here we used $\gamma = 2$, $\alpha = 0.25$ as a default setting. The variable p_t is defined in Equation (2), where p is the estimated probability for the binary classification.

$$p_t = \begin{cases} p & \text{if DME is the class label} \\ 1-p & \text{otherwise} \end{cases} \quad (2)$$

Dual Model Approach

For comparison with the fusion model, we implemented a dual model, which consisted of two separate models including an image classifier and an object detector. The two separate models were trained and inferred separately. We used EfficientNet-b1 and EfficientDet-d1 as the image classifier and object detector, respectively, in our dual model for a fair comparison. EfficientNet stacked basic fixed modules and adjusted some hyperparameters such as the number of layers, number of channels, and input image resolution, using a neural architecture search. In addition, EfficientNet achieved state-of-the-art performance on ImageNet without using additional data.

Device and Hyperparameters

All images of private dataset, Messidor-1, and Messidor-2 dataset were preprocessed before feeding our model. Each image was cropped to the fundus image with minimal black region (**Supplementary Figure 1**) and saved in the JPG image format. These cropped images were resized to input image sizes of 640 × 640 pixels. For image augmentation, we randomly flipped the images of private dataset vertically or horizontally. We trained and tested the model on an Intel Xeon E5-2660 v4 computer with 396 GB DRAM and NVIDIA Tesla V100 GPU using PyTorch with an initial learning rate of 0.0001, a dropout rate of 0.2 and a batch size of 16, for both the fusion and dual models. AdamW optimizer was used in the fusion model and dual model (EfficientDet-d1). Adam optimizer was used in the dual model (EfficientNet-b1). Values of weight decay of the fusion model, dual model (EfficientNet-b1), and dual model (EfficientDet-d1) were 0.01, 0.00001, and 0.01, respectively. Based on the setting of dropout and weight decay in the fusion model and dual model (EfficientNet-b1), the loss curves showed without overfitting in training and validation loss (**Supplementary Figure 2**).

To validate the feasibility of deploying our fusion model on an edge device, it was implemented on NVIDIA Jetson Xavier NX with 8GB of memory using PyTorch.

Statistical Analyses

For the evaluation of performance in DME classification, we used metrics of sensitivity, specificity, accuracy, and area under the

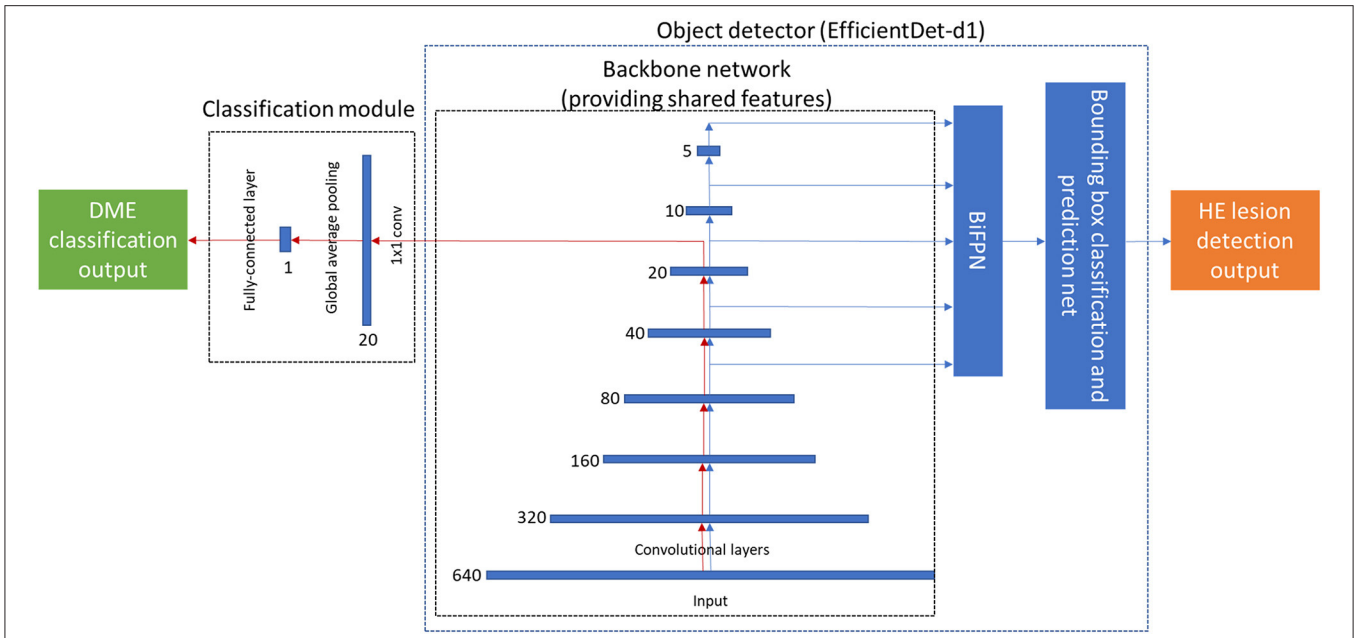

FIGURE 3 | The architecture of the proposed end-to-end deep fusion model. The red arrow denotes the classification path, which forms the same architecture as EfficientNet-b1. The blue arrow denotes the lesion detection path, which has the same architecture as EfficientDet-d1. The number (640, 320, 160, …) near each feature map denotes its resolution.

receiver operating characteristic curve (AUC). All metrics were listed with 95% confidence intervals (CIs). Receiver operating characteristic (ROC) curves were used to illustrate the overall performance using different cutoffs to distinguish between non-DME and DME. A two-proportion z-test was used to compare the two observed proportions obtained from the two models. The DeLong test (32) was used to compare the AUCs. Statistical significance was set at $p < 0.05$. In addition, we evaluated the performance of lesion detection according to Tseng et al. (20).

RESULTS

We trained the fusion model and dual model using the private dataset, and the performance was compared in three aspects: memory usage and execution time, DME classification, and HE detection.

Memory Usage and Execution Time

We investigated the demand for memory and the execution time of the fusion and dual models to process one image from the private testing dataset. We used a command-line utility tool (Nvidia-smi) to evaluate the requirement of memory usage of the fusion model and the dual model to process one fundus image. In addition, the required time of processing one fundus image was calculated by using Python code "time.time()". **Table 3** shows that the fusion model required 1.6 GB of memory, whereas the dual model required 3.6 GB of memory. The mean required time of the fusion and dual model were 2.8 ± 1.5 s and 4.5 ± 1.8 s, respectively. This was averaged over the full testing dataset. These results show that the fusion model reduced the requirement for memory usage and execution time compared to the dual model.

TABLE 3 | The data of memory usage and execution time for the fusion and the dual models to process one image of the private testing dataset.

Resource consumption	Fusion Model	Dual model
Memory (RAM)	1.6 GB	3.6 GB
Time (mean ± standard deviation)	2.8 ± 1.5 s	4.5 ± 1.8 s

DME Classification

The distribution of DME in a private dataset and two open datasets (Messidor-1 and Messidor-2) are shown in **Figure 4**. In **Table 4**, the performance of the fusion and dual models was evaluated using the AUC, sensitivity, specificity, and accuracy. The AUCs of both models were compared using the DeLong test for the three datasets. The result showed that there was no statistically significant difference between the models (p-values of 0.743, 0.942, and 0.114 for the private testing dataset, Messidor-1, and Messidor-2, respectively). Correspondingly, **Figure 5** shows the results of the receiver operating characteristic curves (ROC) of both models for the three datasets. This result demonstrates that the performance of the fusion model is similar to that of the dual model.

HE Lesion Detection

We used fusion and dual models to detect HE lesions on our private testing dataset. We evaluated the performance of these

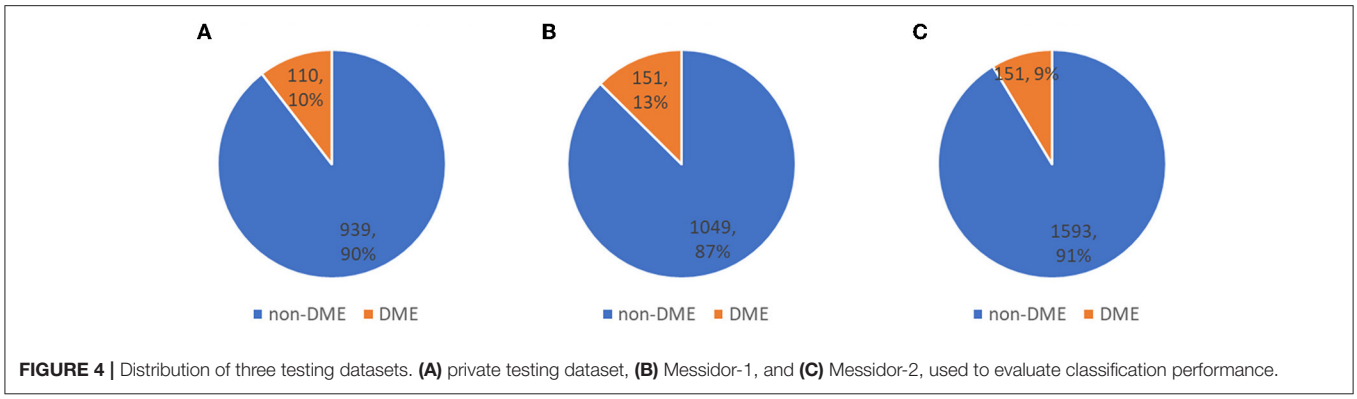

FIGURE 4 | Distribution of three testing datasets. **(A)** private testing dataset, **(B)** Messidor-1, and **(C)** Messidor-2, used to evaluate classification performance.

TABLE 4 | Performance of dual and fusion model for the three datasets.

	Dataset	AUC (%) (95% CI)	Sensitivity (%) (95% CI)	Specificity (%) (95% CI)	Accuracy (%) (95% CI)
Fusion model	Private testing dataset	98.1 (97.3, 98.9)	96.4 (92.9, 99.9)	90.1 (88.2, 92.0)	90.8 (89.1, 92.5)
	Messidor-1	95.2 (93.3, 97.1)	88.7 (83.7, 93.7)	90.2 (88.4, 92.0)	90.0 (88.3, 91.7)
	Messidor-2	95.8 (94.5, 97.1)	87.4 (82.1, 92.7)	90.2 (88.7, 91.7)	89.9 (88.5, 91.3)
Dual model	Private testing dataset	98.0 (97.2, 98.8)	96.4 (92.9, 99.9)	91.8 (90.0, 93.6)	92.3 (90.7, 93.9)
(EfficientNet-b1)	Messidor-1	95.2 (93.2, 97.2)	85.4 (79.8, 91.0)	91.7 (90.0, 93.4)	90.9 (89.3, 92.5)
	Messidor-2	95.1 (93.5, 96.7)	80.8 (74.5, 87.1)	92.7 (91.4, 94.0)	91.7 (90.4, 93.0)

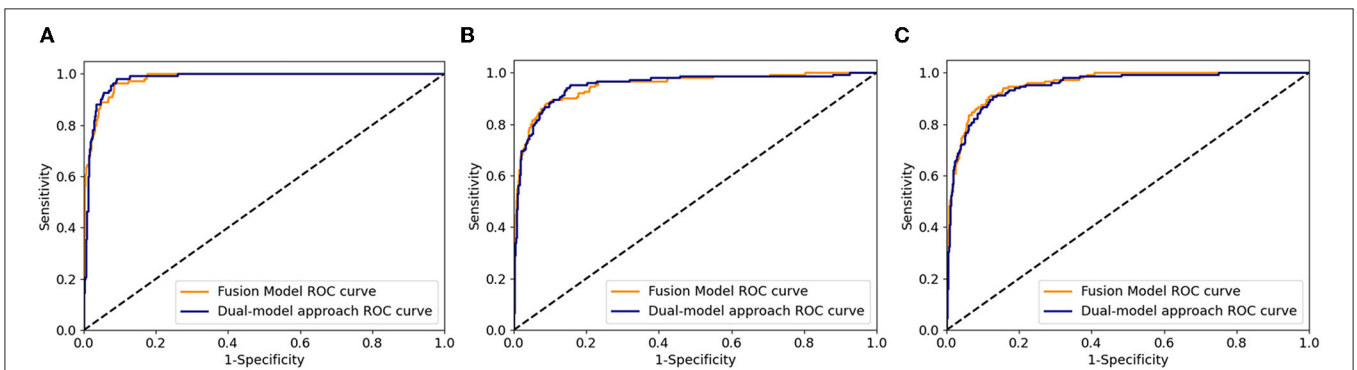

FIGURE 5 | Receiver operating characteristic curves of fusion and dual model for the three datasets. **(A)** private testing dataset, **(B)** Messidor-1, and **(C)** Messidor-2.

models by using true positive, false positive, true negative, and false negative to calculate the accuracy, sensitivity, and specificity. Note that in the HE lesion detection, a true positive image is defined as one of the predicted HE area having an IoU > 0.15 compared to the GT location (as shown in **Figure 6**); a true negative image is defined as both GT and prediction without any lesion detection; a false positive image is defined as GT without any lesion detection but with prediction; and a false negative image is defined as GT with at least one location but no prediction or any prediction location having an IoU ≤ 0.15. In **Table 5**, the results of our private testing dataset revealed that the sensitivity of the fusion model was higher than that of the dual model, and the difference was statistically significant ($p = 0.048$). In addition, the specificity and accuracy of both models were not significantly different ($p = 0.433$ and $p = 0.998$, respectively). This result indicated that the fusion model could detect images with HE lesions more accurately. Furthermore, for lesion visualization, our models could output fundus image with the annotated HE lesion, as shown in **Figure 7**.

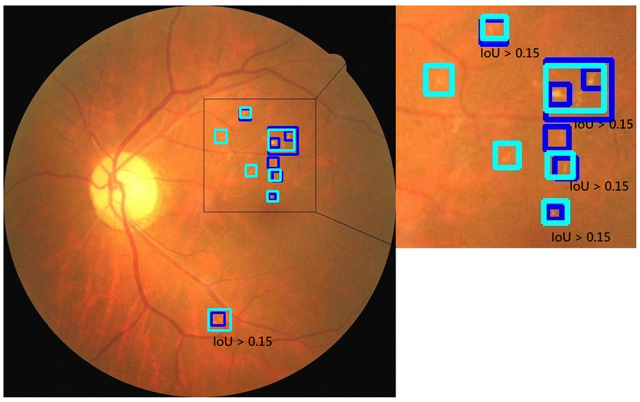

FIGURE 6 | An example image with predicted and GT bounding boxes. The light blue boxes are the prediction result and the dark blue boxes are the GT result. Examples of IoU > 0.15 are marked in image.

TABLE 5 | The performance of HE detection in dual and fusion models.

	Sensitivity (%) (95% CI)	Specificity (%) (95% CI)	Accuracy (%) (95% CI)
Fusion model	79.5 (75.4, 83.6)	87.7 (85.2, 90.2)	86.3 (84.2, 88.4)
Dual model (EfficientDet-d1)	73.0 (68.4, 77.6)	89.2 (86.9, 91.5)	86.4 (84.3, 88.5)

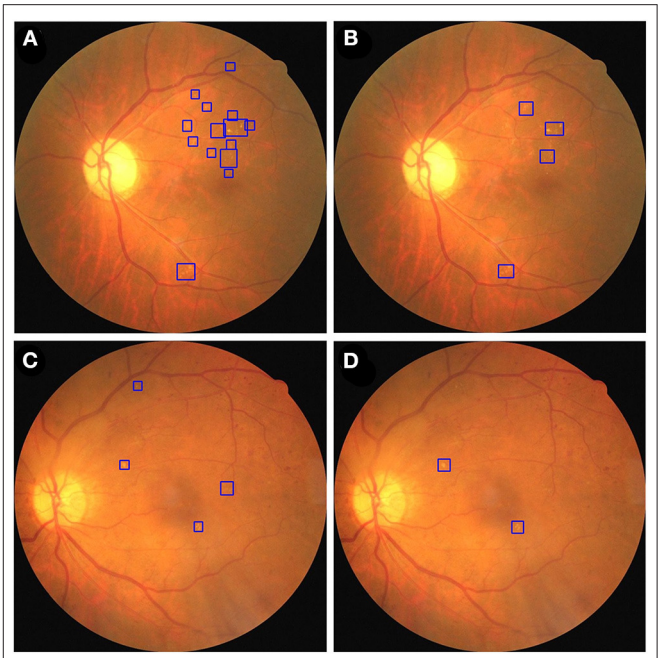

FIGURE 7 | The fusion and dual model annotate HE lesions of two fundus images from the private testing dataset. A fundus image was annotated by fusion model **(A)** and dual model **(B)**. A second fundus image was annotated by fusion model **(C)** and dual model **(D)**. HE lesions are identified using blue bounding boxes.

Optic Disc and Macula Detection

Based on the preceding results, we established a novel end-to-end fusion model that can simultaneously facilitate disease classification and lesion detection. Clinically, anatomical landmarks such as the optic disc and the macula are examined by physicians to determine if there are HE lesions within 1DD from the center of the macula. Thus, we constructed an object detector to detect anatomical landmarks to facilitate advanced visualization. We trained an object detector using YOLOv3 (33) to detect the optic disc and macula. The details of the training process are provided in the **Supplementary Material**. The accuracy of the object detector for the detection of the optic disc and macula was 98.4 and 99.3%, respectively. Furthermore, the object detector could identify the optic disc using a white bounding box and an area within 1DD from the center of the macula using a white circle. These outlined boxes and circles can be integrated into the image results as shown in **Figure 7**. **Figure 8** shows that physicians can instantly ascertain the presence of HE lesions within 1DD from the center of the macula, thereby enabling them to more reliably diagnose DME. Taken together, the results show that lesion visualization can more readily account for the result of DME classification when using the fusion model.

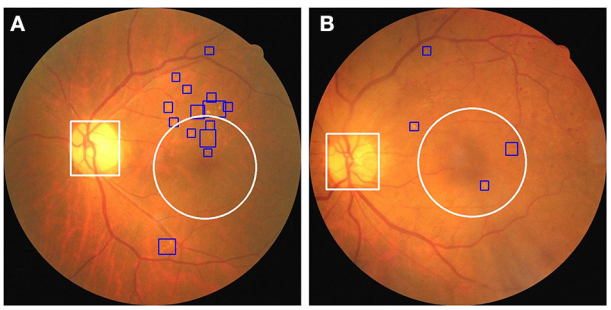

FIGURE 8 | The integration of the visualization of the optic disc and the macula in two fundus images with annotated HE lesions. **(A)** Fundus image from **Figure 7A** annotated with optic disc and macula. **(B)** Fundus image from **Figure 7C** annotated with optic disc and macula. HE lesions are represented as blue bounding boxes. The white circle represents 1DD from the macula center. The white bounding box represents the optic disc.

Implementation on an Edge Device

To verify the feasibility of implementing the entire workflow on an edge device, we tested our fusion model and the anatomical landmark detector on NVIDIA Jetson Xavier NX with 8 GB of memory. The fusion model and the anatomical landmark detector required 7.4 ± 0.02 GB of memory and took 2.53 ± 0.72 s to infer a single fundus image on average. However, the combination of a dual model and an anatomical landmark detector cannot be implemented on edge devices owing to their memory constraints. In addition, we also tested the fusion model on DME classification of the three datasets and HE lesion detection using the NVIDIA Jetson Xavier NX with 8GB of memory. The performance for DME classification and HE lesion

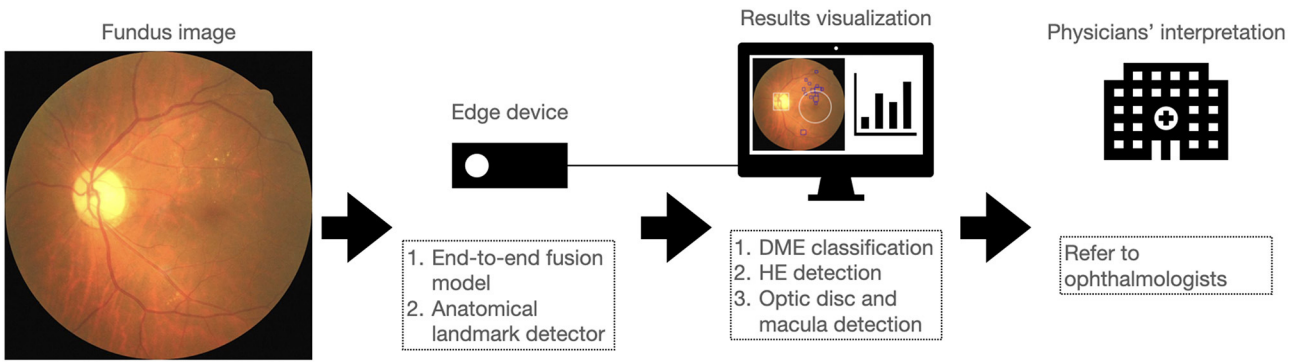

FIGURE 9 | Overview of the proposed approach for implementing a system on an edge device that integrates DME classification, HE detection, and optic disc and macula detection to assist in the interpretation of fundus image by physicians.

detection using the NVIDIA Jetson Xavier NX 8GB of memory was the same as that of the Intel Xeon E5-2660 v4 computer, as shown in **Tables 4**, **5**, respectively.

DISCUSSION

In this study, we proposed a novel end-to-end fusion model to simultaneously facilitate DME classification and HE lesion detection. The performance of the fusion model for DME classification was similar to that of the dual model. The sensitivity of the fusion model for the detection of HE lesions was higher than that of the dual model. We further integrated the detection outputs from the fusion model and the anatomical landmark detector to improve lesion visualization. In addition, we implemented these two models on an edge device to facilitate portability and affordability in remote or resource-limited areas. As shown in **Figure 9**, we report for the first time the integration of the fusion model and a second object detector on an edge device for DME classification, HE detection, and optic disc and macula detection, for lesion visualization and improved interpretability of the AI model. This system allowed physicians not only to obtain the results of DME classification but also to observe the location of HE lesions related to the macula. This might assist physicians in assessing the necessity of referring diabetic patients to ophthalmologists for further examination and treatment.

Recently, several studies have used AI to classify DR with DME or DME only in the Messidor-1 and Messidor-2 datasets (16–19, 34–37). In Messidor-1, Sahlsten et al. (18) proposed an approach based on the ensemble of CNNs with AUC of 95.3%, Sensitivity of 57.5%, Specificity of 99.5%, and Accuracy of 91.6% to detect referable DME. Singh et al. (19) used a hierarchical two-stage ensemble CNN with Sensitivity of 94.7%, Specificity of 97.2%, and Accuracy of 95.5% to grade severity of DME. Ramachandran et al. (34) used a deep neural network software to detect referable DR (moderate DR or DME) achieving AUC of 98.0%, Sensitivity of 96.0%, and Specificity of 90.0%. Li et al. (35) used a cross-disease attention network with AUC of 92.4%, Sensitivity of 70.8%, and Accuracy of 91.2% to jointly grade DR and DME. In Messidor-2, Gulshan et al. (17) used inception-v3 architecture with AUC of 99.0%, Sensitivity of 87.0%, and Specificity of 98.5%. to detect referable DR. Abramoff et al. (16) used the IDx-DR 2.1 device to screen referable DR achieving AUC of 98.0%, Sensitivity of 96.8%, and Specificity of 87.0%. Yaqoob et al. (36) modified ResNet-50 architecture to screen referable DME achieving Accuracy of 96.0%. In 2021, Li et al. (37) used an improved inception-v4 with AUC of 91.7% to detect referable DME. Compared to the performances of above studies in the Messidor-1 and Messidor-2 datasets, the performance of fusion model was AUC of 95.2 and 95.8%, Sensitivity of 88.7 and 87.4%, Specificity of 90.2 and 90.2%, and Accuracy of 90.0% and 89.9% in the Messidor-1 and Messidor-2 datasets. In this study, the classifier of the fusion model was constructed by integrating the EfficientDet-d1 backbone and a classification module. This classifier had the same architecture as EfficientNet-b1. It was determined that the performance of DME classification was similar to that of the original EfficientNet-b1 in the dual model.

In fundus imaging, the determination of the presence and location of HE is useful for physicians in the diagnosis of DME. Several studies have used deep learning to detect HE lesions. Son et al. (38) used a class activation map (CAM) to generate a heatmap to identify the areas that contributed most to the model's decision in classifying DR and other ocular abnormalities. Lam et al. (39) used a sliding window to scan images and a CNN to detect whether HE lesions were present. In addition, Kurilová et al. (40) used the object detector of Faster-RCNN to detect HE lesions in fundus images. In this study, the object detector of our fusion model was modified from EfficientDet-d1, in which the backbone was co-used with the classification module during both the training and inference phases. We found that the performance of EfficientDet-d1 had significantly higher sensitivity for the detection of HEs compared to the original EfficientDet-d1 in the dual model. The higher sensitivity might be because the classification and object detection tasks are complementary in our system.

Typical deep learning models usually lack interpretability, whereas visualization is useful for physicians to assess the result of DME classification by AI. To resolve this problem, we trained another object detector, YOLOv3, to detect anatomical

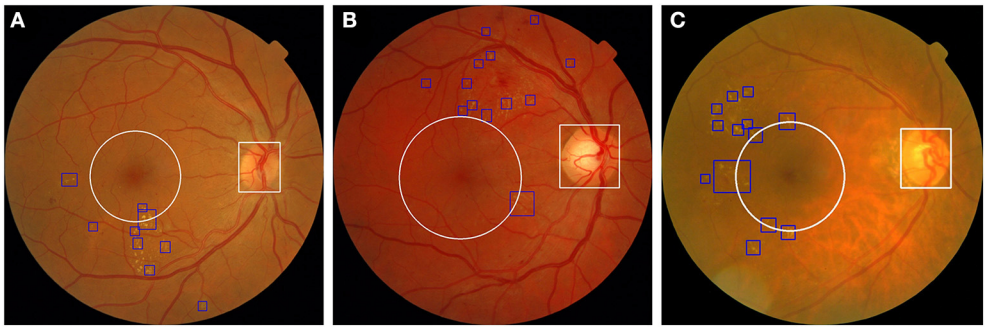

FIGURE 10 | Three fundus images from the open dataset were classified as DME in our system. **(A)** Fundus image from Messidor-1 dataset. **(B,C)** Two fundus images from the Messidor-2 dataset. Our system labeled HE lesions, optic disc, and 1DD from the macula center. The blue bounding box represents the HE. The white circle represents 1DD from the macula center. The white bounding box represents the optic disc.

landmarks (optic disc and macula). Our system integrated the fusion model and another object detector to achieve visualization and increase the interpretability of the AI. We also applied this system to fundus images obtained from open datasets to examine its effect. As shown in **Figure 10**, three fundus images were classified as DME by our system, and it was possible to detect and annotate HE lesions, the optic disc, and 1DD from the macula center. These output fundus images can increase the interpretability of AI results for physicians.

Deep learning models often require large memory usage and computing power. It is difficult to deploy deep learning models on high-end computers in remote areas where resources are limited. Typically, edge devices or cloud computing is utilized to address this issue. However, cloud computing requires network resources. In some remote areas, there was no well-internet service to support cloud computing. Beede et al. (41) discovered that 2 h were required to screen ten diabetic patients using their cloud eye-screening system deployed in Thailand due to sluggish Internet service. Although the edge device is portable and does not require network connections, its small memory size and limited computing power are the primary hindrances. Singh and Kolekar (42) reduced the model size to resolve the storage issue associated with edge devices to classify COVID-19 using computed tomography scans of the chest. In our fusion model, the classifier and object detector co-used the backbone of the object detector. This design reduced the demand for memory usage and the execution time, as shown in **Table 3**. This fusion model is computationally efficient and can be deployed on an edge device with an anatomical landmark detector. In addition, due to traditional fundus camera without appropriate hardware (at least equipped with NVIDIA GeForce GTX 1070 8GB memory), one model to process the data on an edge device could resolve this issue. Therefore, this is the reason why we need to design a deep learning model to process the data in an edge device. Nonetheless, if the computer associated with the fundus camera has appropriate hardware, our model also could integrate into the computer system of camera without needing on an independent edge device.

Our study has several strengths. First, we used a large number of fundus images to train the model. Second, our model yielded satisfactory results for private and open datasets. The model could be implemented on fundus images for different ethnicities. Third, this system facilitates DME classification and the visualization of HE lesions, optic disc, and the macula. Therefore, it is expected that non-ophthalmologist physicians would have more confidence in DME diagnosis determined using AI. Fourth, this system can be deployed on an edge device. This device is portable and affordable. Thus, the proposed system could be applied to diabetic patients in remote or resource-limited areas.

This study has several limitations. First, drusen and the partial features of silicone oil retention are similar to those of HEs. These types of features were not well-trained in our system owing to limited data. This could lead to a false-positive result for DME. Second, we did not integrate the fusion model and anatomical landmark detector into one fusion model. Third, some diseases, such as myelinated fiber layer and optic disc edema, presented blurred boundaries of the optic disc. These diseases could influence the detection of the optic disc and cause inaccurate visualization of 1DD from the macula center.

Based on the obtained results, our future work will involve the application of the proposed system to other object detectors with a backbone that was originally a CNN image classifier, followed by the integration of the fusion model and the anatomical landmark detector into one fusion model on an edge device. Furthermore, we will also train this system to classify the grade of DR and annotate the locations of hard exudates, hemorrhages, soft exudates, microaneurysms, the optic disc, and the macula. This system will grade DR and DME, as well as provide lesion visualization to increase the interpretability of the AI results for physicians.

In conclusion, our system combines a novel end-to-end fusion model with a second object detector to perform DME classification, HE detection, and anatomical localization. It can identify DME and elucidate the relationship between HE and the macula. The entire system can facilitate higher interpretability and serve as a clinical reference for physicians. In addition, it can be implemented on a portable edge device. Clinically, this AI system can be used during the regular examination of DR to improve the interpretation of fundus imaging in patients with DME.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**, further inquiries can be directed to the corresponding author/s.

AUTHOR CONTRIBUTIONS

T-YW and P-YW conceived the fusion model. T-YW, J-TL, and C-LC prepared the manuscript and generated figures and tables, and participated in the experiments involving the fusion model and the DME severity classification model. P-YW conducted the experiments involving the fusion model. S-YC conducted the experiments on the anatomical landmark detector. T-YW, J-TL, C-LC, Y-WL, Y-HC, and J-TC participated in technical discussions. C-LC, Y-HC, and J-TC collected and annotated the data. All authors reviewed the manuscript, contributed to the article and approved the submitted version.

FUNDING

This research was supported by the Industrial Technology Research Institute (Grant Number: K367B82210).

ACKNOWLEDGMENTS

We thank Tri-Service General Hospital, Chung Shan Medical University Hospital, and China Medical University Hospital in Taiwan for providing the fundus image data for this study. Messidor-1 and Messidor-2 are kindly provided by the Messidor program partners (see https://www.adcis.net/en/thirdparty/messidor/) and the LaTIM laboratory.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmed.2022.851644/full#supplementary-material

REFERENCES

1. Lin X, Xu Y, Pan X, Xu J, Ding Y, Sun X, et al. Global, regional, and national burden and trend of diabetes in 195 countries and territories: an analysis from 1990 to 2025. *Sci Rep.* (2020) 10:14790. doi: 10.1038/s41598-020-71908-9
2. Ferris FL III, Patz A. Macular edema. A complication of diabetic retinopathy. *Surv Ophthalmol.* (1984) 28(Suppl.):452–61. doi: 10.1016/0039-6257(84)90227-3
3. Antonetti DA, Klein R, Gardner TW. Diabetic retinopathy. *N Engl J Med.* (2012) 366:1227–39. doi: 10.1056/NEJMra1005073
4. Photocoagulation for Diabetic Macular Edema. Early treatment diabetic retinopathy study report number 1. Early Treatment Diabetic Retinopathy Study research group. *Arch Ophthalmol.* (1985) 103:1796–806. doi: 10.1001/archopht.1985.01050120030015
5. Nguyen QD, Brown DM, Marcus DM, Boyer DS, Patel S, Feiner L, et al. Ranibizumab for diabetic macular edema: results from 2 phase III randomized trials: RISE and RIDE. *Ophthalmology.* (2012) 119:789–801. doi: 10.1016/j.ophtha.2011.12.039
6. Castro-Navarro V, Cervera-Taulet E, Navarro-Palop C, Monferrer-Adsuara C, Hernández-Bel L, Montero-Hernández J. Intravitreal dexamethasone implant Ozurdex® in naïve and refractory patients with different subtypes of diabetic macular edema. *BMC Ophthalmol.* (2019) 19:15. doi: 10.1186/s12886-018-1022-9
7. Korobelnik JF, Do DV, Schmidt-Erfurth U, Boyer DS, Holz FG, Heier JS, et al. Intravitreal aflibercept for diabetic macular edema. *Ophthalmology.* (2014) 121:2247–54. doi: 10.1016/j.ophtha.2014.05.006
8. Lachin JM, Genuth S, Cleary P, Davis MD, Nathan DM. Retinopathy and nephropathy in patients with type 1 diabetes four years after a trial of intensive therapy. *N Engl J Med.* (2000) 342:381–9. doi: 10.1056/NEJM200002103420603
9. Virgili G, Menchini F, Casazza G, Hogg R, Das RR, Wang X, et al. Optical coherence tomography (OCT) for detection of macular oedema in patients with diabetic retinopathy. *Cochrane Database Syst Rev.* (2015) 1:Cd008081. doi: 10.1002/14651858.CD008081.pub3
10. Wong TY, Sun J, Kawasaki R, Ruamviboonsuk P, Gupta N, Lansingh VC, et al. Guidelines on diabetic eye care: the international council of ophthalmology recommendations for screening, follow-up, referral, and treatment based on resource settings. *Ophthalmology.* (2018) 125:1608–22. doi: 10.1016/j.ophtha.2018.04.007
11. Siddalingaswamy PC, Prabhu KG, Jain V. Automatic detection and grading of severity level in exudative maculopathy. *Biomed Eng Appl Basis Commun.* (2011) 23:173–9. doi: 10.4015/S1016237211002608
12. Akram MU, Tariq A, Khan SA, Javed MY. Automated detection of exudates and macula for grading of diabetic macular edema. *Comput Methods Programs Biomed.* (2014) 114:141–52. doi: 10.1016/j.cmpb.2014.01.010
13. Ren F, Cao P, Zhao D, Wan C. Diabetic macular edema grading in retinal images using vector quantization and semi-supervised learning. *Technol Health Care.* (2018) 26:389–97. doi: 10.3233/THC-174704
14. Deepak KS, Sivaswamy J. Automatic assessment of macular edema from color retinal images. *IEEE Trans Med Imaging.* (2012) 31:766–76. doi: 10.1109/TMI.2011.2178856
15. Giancardo L, Meriaudeau F, Karnowski TP, Li Y, Garg S, Tobin KW Jr., et al. Exudate-based diabetic macular edema detection in fundus images using publicly available datasets. *Med Image Anal.* (2012) 16:216–26. doi: 10.1016/j.media.2011.07.004
16. 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. doi: 10.1167/iovs.16-19964
17. 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. doi: 10.1001/jama.2016.17216
18. Sahlsten J, Jaskari J, Kivinen J, Turunen L, Jaanio E, Hietala K, et al. Deep learning fundus image analysis for diabetic retinopathy and macular edema grading. *Sci Rep.* (2019) 9:10750. doi: 10.1038/s41598-019-47181-w
19. Singh RK, Gorantla R. DMENet: diabetic macular edema diagnosis using hierarchical ensemble of CNNs. *PLoS ONE.* (2020) 15:e0220677. doi: 10.1371/journal.pone.0220677
20. Tseng VS, Chen CL, Liang CM, Tai MC, Liu JT, Wu PY, et al. Leveraging multimodal deep learning architecture with retina lesion information to detect diabetic retinopathy. *Transl Vis Sci Technol.* (2020) 9:41. doi: 10.1167/tvst.9.2.41
21. Hsu MY, Chiou JY, Liu JT, Lee CM, Lee YW, Chou CC, et al. Deep learning for automated diabetic retinopathy screening fused with heterogeneous data from EHRs can lead to earlier referral decisions. *Transl Vis Sci Technol.* (2021) 10:18. doi: 10.1167/tvst.10.9.18
22. Alain G, Bengio Y. Understanding intermediate layers using linear classifier probes. *ArXiv [Preprint].* (2017). doi: 10.48550/arXiv.1610.01644

23. Goyal M, Reeves ND, Rajbhandari S, Yap MH. Robust methods for real-time diabetic foot ulcer detection and localization on mobile devices. *IEEE J Biomed Health Inform.* (2019) 23:1730–41. doi: 10.1109/JBHI.2018.2868656
24. Cai W, Zhai B, Liu Y, Liu R, Ning X. Quadratic polynomial guided fuzzy C-means and dual attention mechanism for medical image segmentation. *Displays.* (2021) 70:102106. doi: 10.1016/j.displa.2021.102106
25. Civit-Masot J, Luna-Perejón F, Corral JMR, Domínguez-Morales M, Morgado-Estévez A, Civit A. A study on the use of Edge TPUs for eye fundus image segmentation. *Eng Appl Artif Intell.* (2021) 104:104384. doi: 10.1016/j.engappai.2021.104384
26. Decencière E, Zhang X, Cazuguel G, Lay B, Cochener B, Trone C, et al. Feedback on a publicly distributed image database: the messidor database. *Image Anal Stereol.* (2014) 33:4. doi: 10.5566/ias.1155
27. Abràmoff 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. doi: 10.1001/jamaophthalmol.2013.1743
28. Krause J, Gulshan V, Rahimy E, Karth P, Widner K, Corrado GS, et al. Grader variability and the importance of reference standards for evaluating machine learning models for diabetic retinopathy. *Ophthalmology.* (2018) 125:1264–72. doi: 10.1016/j.ophtha.2018.01.034
29. Tan M, Pang R, Le VQ. EfficientDet: Scalable and Efficient Object Detection. *ArXiv [Preprint].* (2019). doi: 10.48550/arXiv.1911.09070
30. Tan M, Le VQ. EfficientNet: rethinking model scaling for convolutional neural networks. *ArXiv [Preprint].* (2019). doi: 10.48550/arXiv.1905.11946
31. Lin T-Y, Goyal P, Girshick RB, He K, Dollár P. Focal loss for dense object detection. *ArXiv [Preprint].* (2017). doi: 10.48550/arXiv.1708.02002
32. DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. *Biometrics.* (1988) 44:837–45. doi: 10.2307/2531595
33. Redmon J, Farhadi A. YOLOv3: an incremental improvement. *ArXiv [Preprint].* (2018). doi: 10.48550/arXiv.1804.02767
34. Ramachandran N, Hong SC, Sime MJ, Wilson GA. Diabetic retinopathy screening using deep neural network. *Clin Exp Ophthalmol.* (2018) 46:412–6. doi: 10.1111/ceo.13056
35. Li X, Hu X, Yu L, Zhu L, Fu CW, Heng PA. CANet: Cross-disease attention network for joint diabetic retinopathy and diabetic macular edema grading. *IEEE Trans Med Imaging.* (2020) 39:1483–93. doi: 10.1109/tmi.2019.2951844
36. Yaqoob MK, Ali SF, Bilal M, Hanif MS, Al-Saggaf UM. ResNet based deep features and random forest classifier for diabetic retinopathy detection. *Sensors (Basel).* (2021) 21:3883. doi: 10.3390/s21113883
37. Li F, Wang Y, Xu T, Dong L, Yan L, Jiang M, et al. Deep learning-based automated detection for diabetic retinopathy and diabetic macular oedema in retinal fundus photographs. *Eye.* (2021). doi: 10.1038/s41433-021-01552-8
38. Son J, Shin JY, Kim HD, Jung KH, Park KH, Park SJ. Development and validation of deep learning models for screening multiple abnormal findings in retinal fundus images. *Ophthalmology.* (2020) 127:85–94. doi: 10.1016/j.ophtha.2019.05.029
39. Lam C, Yu C, Huang L, Rubin D. Retinal lesion detection with deep learning using image patches. *Invest Ophthalmol Vis Sci.* (2018) 59:590–6. doi: 10.1167/iovs.17-22721
40. Kurilová V, Goga J, Oravec M, Pavlovičová J, Kajan S. Support vector machine and deep-learning object detection for localisation of hard exudates. *Sci Rep.* (2021) 11:16045. doi: 10.1038/s41598-021-95519-0
41. Beede E, Baylor E, Hersch F, Iurchenko A, Wilcox L, Ruamviboonsuk P, et al. 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*. Honolulu, HI: Association for Computing Machinery (2020). p. 1–12. doi: 10.1145/3313831.3376718
42. Singh VK, Kolekar HM. Deep learning empowered COVID-19 diagnosis using chest CT scan images for collaborative edge-cloud computing platform. *Multimed Tools Appl.* (2021) 1–28. doi: 10.1007/s11042-021-11158-7

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers | Frontiers in Medicine

ORIGINAL RESEARCH
published: 25 April 2022
doi: 10.3389/fmed.2022.851536

Vitrectomy and All-Cause and Cause-Specific Mortality in Elderly Patients With Vitreoretinal Diseases: A Nationwide Cohort Study

Yoon Jeon Kim[1†], *Ji Sung Lee*[2,3†], *Yunhan Lee*[1], *Hun Lee*[1*‡], *Jae Yong Kim*[1*‡] *and Hungwon Tchah*[1]

[1] *Department of Ophthalmology, Asan Medical Center, University of Ulsan College of Medicine, Seoul, South Korea,* [2] *Clinical Research Center, Asan Institute for Life Sciences, Asan Medical Center, University of Ulsan College of Medicine, Seoul, South Korea,* [3] *Department of Clinical Epidemiology and Biostatistics, Asan Medical Center, University of Ulsan College of Medicine, Seoul, South Korea*

OPEN ACCESS

Edited by:
Tae-im Kim,
Yonsei University, South Korea

Reviewed by:
Ping Fei,
Shanghai Jiao Tong University, China
Tyler Hyungtaek Rim,
Duke-NUS Medical School, Singapore

***Correspondence:**
Hun Lee
yhun777@hanmail.net
Jae Yong Kim
jykim2311@naver.com

†These authors have contributed equally to this work and share first authorship
‡These authors have contributed equally to this work

Specialty section:
This article was submitted to Ophthalmology, a section of the journal Frontiers in Medicine

Received: *10 January 2022*
Accepted: *16 March 2022*
Published: *25 April 2022*

Citation:
Kim YJ, Lee JS, Lee Y, Lee H, Kim JY and Tchah H (2022) Vitrectomy and All-Cause and Cause-Specific Mortality in Elderly Patients With Vitreoretinal Diseases: A Nationwide Cohort Study. Front. Med. 9:851536. doi: 10.3389/fmed.2022.851536

Purpose: To determine the all-cause and cause-specific mortality in elderly patients with vitreoretinal diseases based on vitrectomy status.

Methods: Elderly patients (aged ≥ 60 years) diagnosed with vitreoretinal diseases between 2003 and 2012 using the Korean National Health Insurance Service-Senior cohort (2002–2015) were included in this nationwide population-based retrospective cohort study. The exposure of interest was vitrectomy, and information on mortality from patient inclusion until December 2015 was obtained. Cox regression modeling was used to assess the association between vitrectomy and mortality. An additional subgroup analysis was performed to investigate the effects of the underlying retinal disease characteristics and comorbidities on mortality.

Results: The study cohort included 152,283 patients (3,313 and 148,970 in the vitrectomy and non-vitrectomy groups, respectively). The adjusted model showed vitrectomy was associated with a decreased risk of pulmonary-cause mortality [hazard ratio (HR), 0.51; $P < 0.001$]; however, no association was observed for all-cause mortality (HR, 0.93; $P = 0.325$). Vitrectomy was associated with increased mortality risk (all-cause: HR, 1.26; $P < 0.001$ and vascular causes: HR, 1.41; $P = 0.003$) among patients with retinal vascular diseases and decreased mortality risk (all-cause: HR, 0.64; $P < 0.001$ and pulmonary causes: HR, 0.35; $P = 0.011$) among patients with macular diseases. There were significant interactions between age and vitrectomy with respect to all-cause mortality among patients with either vitreoretinal disease.

Conclusions: In elderly patients with retinal diseases, the vitrectomy group showed the lower mortality from pulmonary causes with no association for all-cause mortality.

Keywords: vitrectomy, nationwide cohort study, cause specific mortality, all-cause mortality, vitreoretinal disease

INTRODUCTION

In the last few decades, the clinical efficacy and safety of vitrectomy have significantly improved owing to advancements in surgical instruments and equipment (1, 2). In addition, patient discomfort during the postoperative period has also decreased. Accordingly, while vitrectomy was initially only performed for severe cases, such as retinal detachment with large retinal breaks or proliferative vitreous retinopathy, the indications for surgery have been expanding (3). An increasing number of patients worldwide are undergoing surgery, particularly the elderly (4). Therefore, it is worthwhile to analyze whether there is a difference in mortality rates depending on the performance of vitrectomy in the elderly population.

Various previous studies have evaluated the association between ocular surgeries and long-term mortality, particularly between cataract surgery and mortality (5–7). Additionally, some cohort studies have evaluated the survival rates after vitrectomy in patients with proliferative diabetic retinopathy, reporting a 5-year survival rate between 68 and 96% (8, 9) and a 10-year survival rate of 49% (10). However, those studies in relation to vitrectomy included small sample sizes of people from a single group and no comparisons with a non-surgical group. To the best of our knowledge, no previous study has compared mortality rates according to vitrectomy status, particularly in a population-based cohort. Hence, we aimed to evaluate and compare the all-cause and cause-specific mortality of the elderly Korean population based on vitrectomy status. We used data from the Korean National Health Insurance Service Senior cohort (NHIS-Senior) database, which is a nationwide database that covers the entire older-adult Korean population (11). In addition, a subgroup analysis was performed to determine whether the mortality rate differed according to the two main subtypes of vitreoretinal disease (retinal vascular and macular diseases).

METHODS

Study Setting

For this population-based retrospective cohort study, data was obtained from the Korean NHIS-Senior database (2002–2015). As previously described, the Korean NHIS is a national health insurance database that includes all patient data related to healthcare and long-term care services (11). The senior cohort data covers 558,147 individuals randomly sampled from 10% of the approximate 5.5 million Koreans aged $\geq$ 60 years. All participants included in the NHIS-Senior database were followed until 2015 unless they were disqualified for health coverage. The NHIS-Senior database comprises patient data, including age, sex, national health screening, healthcare utilization, disease diagnoses, vitrectomy status, procedures, and prescribed medications, as well as mortality-related information. The patients' healthcare records were not duplicated because all Korean residents receive a unique identification number at birth. The Korean NHIS uses the Korean Standard Classification of Diseases, Seventh Edition (KCD-7) codes and the Korean Electronic Data Interchange (KEDI) codes (12). This database can be used for national healthcare evidence-based analyses that accurately represent the entire elderly population in Korea.

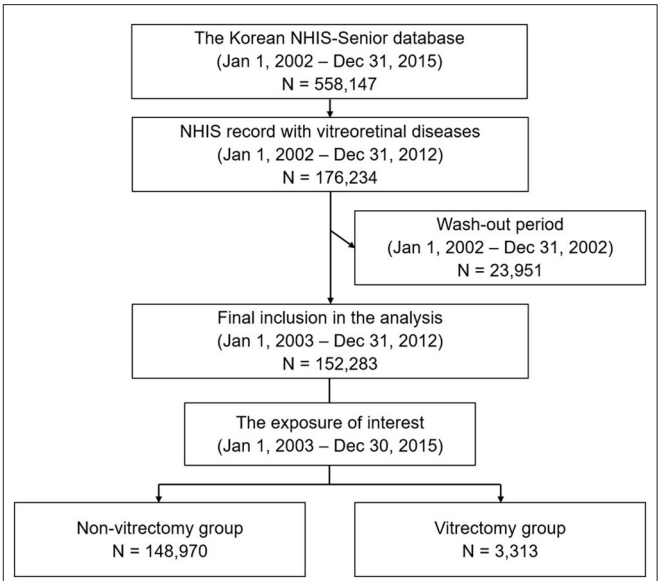

FIGURE 1 | Flow-chart illustrating the study design.

All data in the NHIS-Senior database were de-identified and encrypted to protect the privacy of the participants before use. As the NHIS-Senior database comprises data that are open to the public, the Institutional Review Board (IRB) of the Asan Medical Center waived the requirement for a review of this study (AMC IRB No. 2019-1630).

Study Population

Our target population was Korean older adults who were included in the NHIS-Senior database between January 1, 2002 and December 31, 2015 (n = 558,147) (**Figure 1**). The inclusion criteria were as follows: aged $\geq$ 60 years, with at least one NHIS record between January 1, 2002 and December 31, 2012 (n = 176,234) with a KCD-7 code for a vitreoretinal disease (**Supplementary Table 1**). In addition, we included a wash-out period between January 1, 2002 and December 31, 2002 to reduce the potential impact of surveillance bias (n = 152,283). Patients with the following characteristics were excluded: those with KCD-7 codes for congenital or hereditary vitreoretinal diseases, a ruptured globe, intraocular foreign body, or other retinal/choroidal malignancy.

Exposure

The exposure of interest was a vitrectomy. The eligible subjects were classified into two groups based on whether they underwent a vitrectomy between January 1, 2003 and December 30, 2015 (**Figure 1**). The vitrectomy group comprised all participants with the KEDI code for vitrectomy. Vitrectomy was defined as a total vitrectomy (KEDI code S5121). The non-vitrectomy group was the unexposed group, and comprised all participants with diagnostic codes for vitreoretinal diseases, but with no KEDI code for vitrectomy. The patients in both groups were followed up

starting at the earliest date the diagnostic code for vitreoretinal diseases was assigned.

Outcomes

The primary outcomes of interest were all-cause and cause-specific mortality for the period from patient inclusion to December 31, 2015. Mortality was determined based on an indicator variable in the NHIS-Senior database, which contains information on all-cause and cause-specific mortality. In this study, the causes of death were grouped as cancer, accident-related, vascular, pulmonary, neurologic, infectious, or trauma-related deaths (**Supplementary Table 2**).

We performed a time-dependent analysis to prevent an immortal time bias (13). For the non-vitrectomy group, the time-to-death was calculated as the number of days from the vitreoretinal disease diagnosis to death. For the vitrectomy group, the period between the diagnosis and vitrectomy was considered the follow-up period for vitreoretinal diseases, and the period after surgery was considered the follow-up period for vitrectomy. Therefore, the person-days of follow up before the surgery were classified as the non-surgery group until the surgery day definition was met, and as the surgery group thereafter. Participants who did not have a record of death were followed up until the last known date or December 31, 2015.

Covariates

The collected demographic data included patients' age at the time of vitreoretinal disease diagnosis, sex, area of residence, and income level. The areas of residence were grouped as either metropolitan or provincial areas according to the administrative unit of Korea. Household income was categorized as below or above the twentieth percentile.

Both systemic and ocular comorbidities were included as covariates in this study. The systemic comorbidities were assessed at the time of vitreoretinal disease diagnosis. The Charlson comorbidity index (CCI) score was used as a covariate to represent patients' overall systemic health (14). The CCI is a weighted index of systemic disease burden based on the presence or absence of 17 systemic comorbidities, with a higher CCI score indicating a higher burden of systemic disease. Based on their systemic disease profiles, patients were assigned a CCI score between 0 and 6 that could be used to predict the 1-year mortality risk (6). In addition, data from a self-reported questionnaire (smoking status, alcohol consumption, and regular exercise) was obtained for a subset of patients who underwent a national health screening program, medical check-ups provided every 2 years.

Other ocular comorbidities, including glaucoma (KCD-7 codes H40 and H42) and severe cataract (KCD-7 codes H25.2 and H25.1), were examined. Since objective visual acuity data were not available, patients with diagnostic codes for Morgagnian-type senile cataract and senile nuclear cataract were considered to have severe cataract subtypes. Determining the presence of systemic and ocular comorbidities was based on the availability of KCD-7 codes for these conditions (**Supplementary Table 3**).

Statistical Analysis

The primary objective of this study was to examine the differences in all-cause and cause-specific mortality among patients diagnosed with vitreoretinal diseases between the vitrectomy and non-vitrectomy groups. We used absolute standardized differences (ASD) to compare the baseline characteristics (15). ASD, calculated as differences in the means or proportions divided by a pooled estimate of the standard deviation (SD), is not as sensitive to sample size when compared with the traditional significance testing; hence, it is useful for identifying clinically meaningful differences. An ASD > 0.1 is considered clinically meaningful (16). Cox regression models based on the time-varying covariate vitrectomy status (exposure) were used to determine the covariate-adjusted associations between vitrectomy and time to death from any cause or death attributed to cancer and vascular, pulmonary, neurologic, infectious, or traumatic conditions. We used two models with adjustments for potential confounding factors at baseline. Model 1 was adjusted for age and sex. Model 2 was further adjusted for systemic disease burden as measured using income, area of residence, CCI, and ocular comorbidities. In addition, the adjusted HR of each covariate was used to investigate the effects of age, sex, area of residence, income level, CCI score, glaucoma, and severe cataracts on vitrectomy-related overall mortality. Moreover, subgroup analyses were performed to investigate differences in mortality according to the underlying retinal disease type. The Statistical Analysis System (SAS) program version 9.4 (SAS Institute, Cary, NC, United States) was used for all statistical analyses, and statistical significance was set to a two-sided $p < 0.05$.

RESULTS

Baseline Characteristics

The study cohort included 152,283 patients (mean age, 72.3 $\pm$ 6.2 years), with 3,313 and 148,970 in the vitrectomy and non-vitrectomy groups, respectively. The baseline characteristics of the study cohort are summarized in **Supplementary Table 4**. Most of the patients in both groups were diagnosed with vitreoretinal diseases when aged < 70 years. The older population ($\geq$ 75 years) were not likely to undergo vitrectomy (ASD $= 0.4047$). More patients who underwent vitrectomy lived in metropolitan areas than in provincial areas (ASD $= 0.1085$) and had higher incomes (ASD $= 0.1214$). The remaining demographic factors, including the CCI, were similar between the groups (all ASD < 0.1). A greater proportion of patients had glaucoma in the vitrectomy group (49.4 and 38.5% in the vitrectomy and non-vitrectomy groups, respectively; ASD $=$ 0.2403). The fasting plasma glucose concentration was higher in the vitrectomy group (ASD $= 0.1170$), but other parameters were not different between the groups (all ASD > 0.1).

Mortality Rate

Table 1 shows the mortality rates of the elderly patients with vitreoretinal diseases stratified by vitrectomy status. The crude mortality at any time during the study period was 2.89 deaths per 100 person-years in the vitrectomy group and 3.36 deaths per

TABLE 1 | Mortality rates of elderly patients with vitreoretinal disease stratified by vitrectomy status.

Cause of mortality	Mortality rate, No. of deaths (incidence per 100 person-years; 95% CI)	
	Vitrectomy group	Non-vitrectomy group
All-cause	538 (2.89; 2.66–3.15)	36,627 (3.36; 3.33–3.40)
Cancer	143 (0.77; 0.65–0.91)	9,903 (0.91; 0.89–0.93)
Vascular	143 (0.77; 0.65–0.91)	9,008 (0.83; 0.81–0.84)
Pulmonary	27 (0.15; 0.10–0.21)	3,577 (0.33; 0.32–0.34)
Neurologic	10 (0.05; 0.03–0.10)	1,057 (0.10;0.09–0.10)
Infectious	8 (0.04; 0.02–0.09)	980 (0.09; 0.08–0.10)
Accident or trauma	37 (0.20; 0.14–0.28)	2,366 (0.22; 0.21–0.23)

CI, confidence interval.
Total person-years were 18,589 for the vitrectomy group and 1,090,111 for the non-vitrectomy group.
Poisson regression model with vitreoretinal vitrectomy status as the time-varying covariate.

100 person-years in the non-vitrectomy group. The unadjusted model showed a significant difference in the hazard ratios (HRs) for all-cause and cause-specific mortality between the groups (**Table 2**). Vitrectomy was associated with a lower risk of all-cause mortality [HR, 0.81; 95% confidence interval (CI), 0.74–0.88, $P < 0.001$] and mortality from malignant causes (HR, 0.83; 95% CI, 0.70–0.97, $P = 0.023$), pulmonary causes (HR, 0.40; 95% CI, 0.27–0.58, $P < 0.001$), neurologic causes (HR, 0.46; 95% CI, 0.25–0.86, $P = 0.014$), and infectious causes (HR, 0.44; 95% CI, 0.22–0.88, $P = 0.021$). After adjusting for other variables, the decreased risk of mortality from pulmonary causes present in the vitrectomy group persisted in Model 1 (HR, 0.50; 95% CI, 0.34–0.73, $P < 0.001$) and Model 2 (HR, 0.51; 95% CI, 0.35–0.74, $P < 0.001$). However, the adjusted model showed no association between vitrectomy and all-cause and other cause-specific mortality ($P > 0.05$). Based on this analysis, Kaplan–Meier survival curves were calculated to describe all-cause mortality according to the vitrectomy status (**Figure 2**).

To assess the factors affecting these differences in all-cause mortality between the groups, fully adjusted models were used. A significant interaction between age and vitrectomy-related mortality was found ($P < 0.001$) (**Supplementary Table 5**). The HR in the vitrectomy group was higher for patients < 70 years of age; however, in the vitrectomy group, all-cause mortality decreased in patients aged ≥ 75 years. The lowest mortality in the vitrectomy group was observed for patients aged ≥ 85 years, with a 56% lower HR than that in the non-vitrectomy group (HR, 0.44; 95% CI, 0.25–0.78, $P = 0.005$). The other factors were not associated with mortality rates stratified by vitrectomy status.

Subgroup Analysis

Subgroup analyses were performed based on the etiology of the retinal disease (i.e., retinal vascular or macular diseases) (**Supplementary Tables 6, 7**). For retinal vascular diseases (**Table 3**), the fully adjusted model showed that vitrectomy was associated with a higher risk of mortality from all causes (HR, 1.26; 95% CI, 1.11–1.42, $P < 0.001$) and vascular causes (HR, 1.41; 95% CI, 1.12–1.78, $P = 0.021$). Conversely, for macular diseases (**Table 4**), the fully unadjusted model showed that vitrectomy was associated with a lower risk of mortality from all causes (HR, 0.64; 95% CI, 0.53–0.78, $P < 0.001$), vascular causes (HR, 0.69; 95% CI, 0.48–1.00, $P = 0.052$), and pulmonary causes (HR, 0.35; 95% CI, 0.16–1.79, $P = 0.011$). In both subgroups, considerable interactions between age and vitrectomy-related mortalities were found (**Supplementary Tables 8, 9**).

DISCUSSION

Using the NHIS database, the present study aimed to examine mortality rates according to vitrectomy status in elderly patients diagnosed with vitreoretinal diseases. The overall mortality rate of patients with retinal diseases did not differ according to vitrectomy status; however, the cause-specific mortality was different between groups. Specifically, the risk of mortality owing to pulmonary causes was significantly lower in the vitrectomy group. In addition, the adjusted all-cause and cause-specific mortality in the vitrectomy group were different according to the underlying retinal disease. A strength of this study is that it demonstrates the characteristics and patterns of mortality of patients undergoing vitrectomy and evaluates the associations between vitrectomy and socio-demographic factors.

The association between vitrectomy and mortality was modified by the effects of age in patients with vitreoretinal diseases. Patients who underwent vitrectomy were significantly younger, indicating that younger patients (≤ 70 years) often choose surgical treatment. Meanwhile, patients aged ≥ 85 years showed the lowest mortality risk associated with vitrectomy. In particular, the analysis of the vitreoretinal diseases overall were found to be associated with a decreased risk of pulmonary-related deaths. Though no studies have evaluated the association between vitrectomy and mortality from pulmonary causes, we suggest that patients who undergo vitrectomy have lower rates of mortality from pulmonary causes, possibly since their general state of health is good, meaning they can ambulate independently and receive routine medical care. The other possibility is that vision recovery after surgery is associated with an increase in physical activity, which results in fewer complications and a higher long-term survival rate. Further studies are necessary to better clarify the mechanisms underlying the association between vitrectomy and mortality from pulmonary causes.

Interestingly, the association of mortality and vitrectomy among the patients with vitreoretinal diseases differed from those of mortality and cataract or glaucoma surgery in patients with cataract or glaucoma. While mortality due to pulmonary causes was decreased following vitrectomy, according to our previous studies examining mortality after cataract surgery and glaucoma surgery (17, 18), all-cause mortality and mortality due to vascular and neurologic causes decreased in the cataract surgery group. On the other hand, all-cause mortality and, in particular, mortality from neurologic causes, increased in the glaucoma surgery group. These differences might be caused by the differences in the characteristics of the ophthalmic diseases and the indications for surgical decisions.

TABLE 2 | Hazard ratios for all-cause and cause-specific mortality in elderly patients with vitreoretinal disease stratified by vitrectomy status.

Cause of mortality	Unadjusted Cox model [a]		Adjusted Cox model 1 [a,b]		Adjusted Cox model 2 [a,c]	
(No. of participants)	Hazard ratio (95% CI)	*P*-value	Hazard ratio (95% CI)	*P*-value	Hazard ratio (95% CI)[b,c]	*P*-value
All-cause	0.81 (0.74–0.88)	<0.001	0.94 (0.86–1.02)	0.142	0.96 (0.88–1.04)	0.325
Cancer	0.83 (0.70–0.97)	0.023	0.87 (0.74–1.03)	0.103	0.88 (0.75–1.04)	0.143
Vascular	0.86 (0.73–1.02)	0.083	1.04 (0.88–1.23)	0.623	1.07 (0.91–1.26)	0.422
Pulmonary	0.40 (0.27–0.58)	<0.001	0.50 (0.34–0.73)	<0.001	0.51 (0.35–0.74)	<0.001
Neurologic	0.46 (0.25–0.86)	0.014	0.58 (0.31–1.08)	0.088	0.58 (0.31–1.08)	0.088
Infectious	0.44 (0.22–0.88)	0.021	0.53 (0.26–1.06)	0.075	0.53 (0.26–1.06)	0.074
Accident or trauma	0.89 (0.65–1.24)	0.496	0.97 (0.70–1.34)	0.849	0.98 (0.71–1.36)	0.909

CI, confidence interval.

[a] *Cox model with vitreoretinal vitrectomy status as the time-varying covariate.*

[b] *Adjusted for age and sex.*

[c] *Adjusted for age, sex, income, area of residence, Charlson comorbidity index (0, 1, 2, 3, 4, ≥ 5), glaucoma, and cataract severity.*

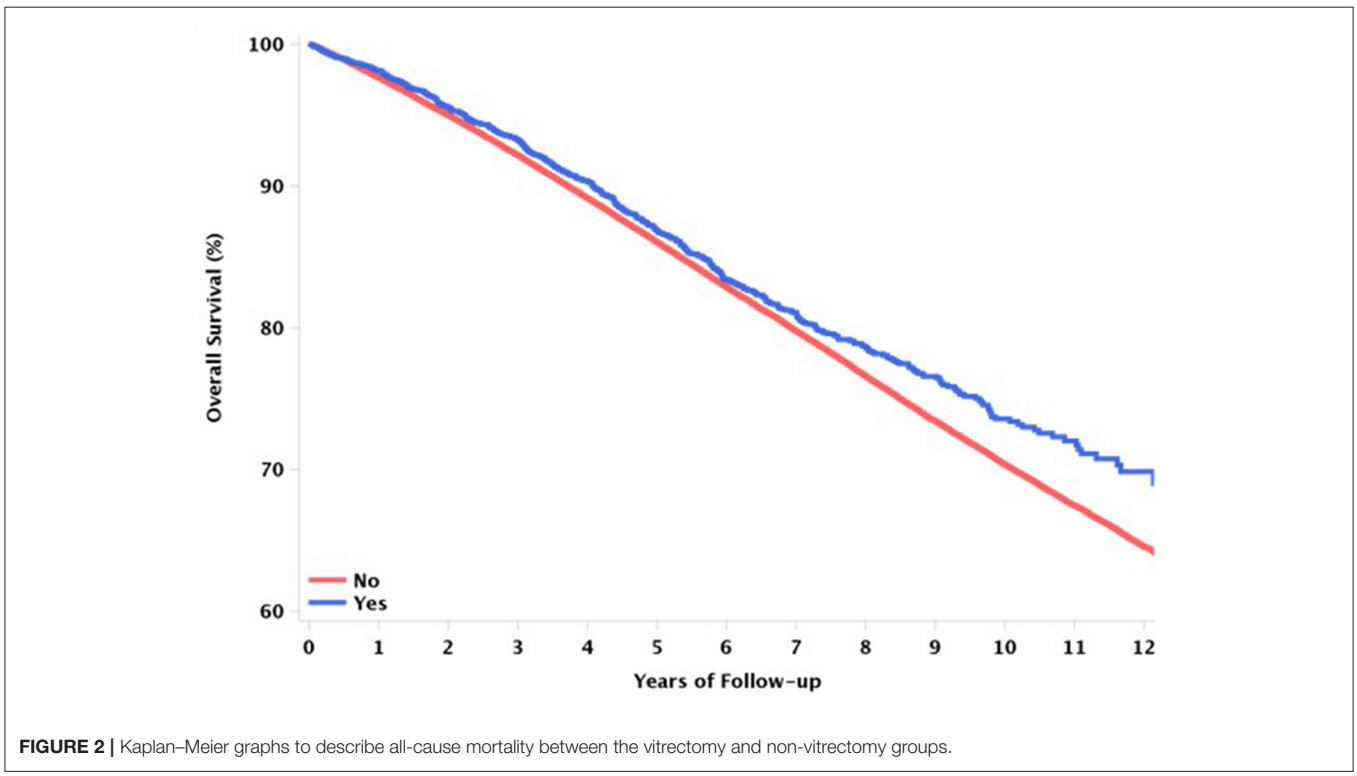

FIGURE 2 | Kaplan–Meier graphs to describe all-cause mortality between the vitrectomy and non-vitrectomy groups.

To analyze the differences in mortality in the vitrectomy group according to the underlying retinal disease etiology, patients were categorized into two subgroups, which accounted for more than half of all vitreoretinal diseases. When patients with retinal vascular disease and macular disease were examined separately, different associations were observed. As expected, our results showed that patients in the older adult population with retinal vascular diseases who underwent vitrectomy had a higher CCI score and systemic vascular risk and, consequently, had a higher risk of mortality from all causes and vascular causes owing to the cardiovascular risk after adjusting for demographic characteristics and systemic and ocular comorbidities. Consistent with our results, a recently published study suggested that the severity of diabetic retinopathy may provide valuable insights into patients' risk of mortality from all causes and vascular causes (19). This tendency was particularly evident in the younger patient group (aged 60–69), and it can be speculated that older adults often choose surgery when vascular risk factors are relatively well controlled.

Furthermore, we demonstrated decreased mortality in the vitrectomy group from all causes and pulmonary causes for patients with macular diseases after adjusting for demographic characteristics and systemic and ocular comorbidities. In addition, a decline in the mortality rate following vitrectomy was observed, especially in the elderly. This is a rather contradictory result, considering that a greater proportion of the patients with

TABLE 3 | Hazard ratios for all-cause and cause-specific mortality in elderly patients with retinal vascular diseases stratified by vitrectomy status.

Cause of mortality	Unadjusted Cox model [a]		Adjusted Cox model 1 [a,b]		Adjusted Cox model 2 [a,c]	
(No. of participants)	Hazard ratio (95% CI)[a]	*P*-value	Hazard ratio (95% CI)	*P*-value	Hazard ratio (95% CI)	*P*-value
All-cause	1.04 (0.92–1.16)	0.561	1.19 (1.06–1.34)	0.004	1.26 (1.11–1.42)	<0.001
Cancer	0.90 (0.69–1.16)	0.408	0.94 (0.72–1.22)	0.624	0.99 (0.76–1.29)	0.960
Vascular	1.12 (0.89–1.41)	0.317	1.35 (1.07–1.69)	0.010	1.41 (1.12–1.78)	0.003
Pulmonary	0.60 (0.36–1.00)	0.052	0.75 (0.45–1.26)	0.278	0.78 (0.47–1.31)	0.356
Neurologic	0.37 (0.12–1.17)	0.090	0.48 (0.15–1.49)	0.204	0.48 (0.15–1.51)	0.209
Infectious	0.54 (0.20–1.46)	0.227	0.65 (0.24–1.75)	0.398	0.68 (0.25–1.82)	0.439
Accident or trauma	0.79 (0.45–1.39)	0.413	0.84 (0.48–1.48)	0.547	0.88 (0.50–1.55)	0.653

CI, confidence interval.
[a]Cox model with vitreoretinal vitrectomy status as the time-varying covariate.
[b]Adjusted for age and sex.
[c]Adjusted for age, sex, income, area of residence, Charlson comorbidity index (0, 1, 2, 3, 4, ≥ 5), glaucoma, and cataract severity.

TABLE 4 | Hazard ratios for all-cause and cause-specific mortality in elderly patients with macular diseases stratified by vitrectomy status.

Cause of mortality	Unadjusted Cox model [a]		Adjusted Cox model 1 [a,b]		Adjusted Cox model 2 [a,c]	
(No. of participants)	Hazard ratio (95% CI)	*P*-value	Hazard ratio (95% CI)	*P*-value	Hazard ratio (95% CI)	*P*-value
All-cause	0.46 (0.38–0.56)	<0.001	0.65 (0.54–0.79)	<0.001	0.64 (0.53–0.78)	<0.001
Cancer	0.62 (0.44–0.87)	<0.001	0.74 (0.53–1.04)	0.083	0.73 (0.52–1.02)	0.069
Vascular	0.48 (0.33–0.69)	<0.001	0.70 (0.48–1.01)	0.059	0.69 (0.48–1.00)	0.052
Pulmonary	0.23 (0.10–0.51)	<0.001	0.36 (0.16–0.81)	0.014	0.35 (0.16–0.79)	0.011
Neurologic	0.25 (0.06–1.00)	0.050	0.38 (0.09–1.52)	0.172	0.36 (0.09–1.46)	0.153
Infectious	0.46 (0.15–1.42)	0.174	0.64 (0.20–1.99)	0.439	0.62 (0.20–1.92)	0.403
Accident or trauma	0.94 (0.54–1.64)	0.838	1.20 (0.69–2.08)	0.525	1.21 (0.70–2.11)	0.492

CI, confidence interval.
[a]Cox model with vitreoretinal vitrectomy status as the time-varying covariate.
[b]Model 1: adjusted for age and sex.
[c]Model 2: adjusted for age, sex, income, area of residence, Charlson comorbidity index (0, 1, 2, 3, 4, ≥ 5), glaucoma, and cataract severity.

macular diseases who underwent vitrectomy were current or ex-smokers who regularly consumed alcohol. The main aim of surgical treatment for patients with macular disease is to facilitate better vision in most cases. Older patients who underwent vitrectomy to restore their vision may have been in good general health and thus had lower mortality rates from and vascular causes. Taken together, vitreoretinal surgeons should focus more on improving visual restoration with macular diseases.

This study has some limitations. The main limitation was the observational nature of the study. Though an association between vitrectomy and mortality was observed, it was difficult to determine the extent to which the diagnosis of retinal disease or the performance of vitrectomy contributed to the outcomes. Second, various factors, including surgeons' and patients' preferences and disease severity, must be considered when interpreting our results, however, these are unavoidable limitations to using claims data. In addition, the NHIS-Senior database does not provide information on visual acuity measurements, preoperative retinal status, or the postoperative clinical course. Therefore, we could not determine if the difference between the long-term survival rates in the vitrectomy group was directly associated with the ocular changes following vitrectomy. Third, several variables could be confirmed only for a subset of patients who underwent the national health screening program and approximately 50% of these data were missing. Fourth, we did not consider the number of surgeries in our analysis. Patients with severe diseases are more likely to need more surgeries, which may be related to mortality rates. However, since we included only the time point of the first surgery in the analysis without considering the number of surgeries and the overall course, these aspects should be considered when interpreting the study results. Lastly, we focused only on residents of South Korea; therefore, the observed findings cannot be generalized to other ethnic groups. Despite these limitations, this study is valuable. This was the first study to report a significant association between vitrectomy and long-term survival based on all-cause and cause-specific mortality in older adults using a nationwide population-based database. This study also included a large sample; the data were obtained from the NHIS-Senior database, and selection bias was relatively low because Korea has a single public insurance system that covers the entire population (20, 21).

This nationwide cohort study showed that the vitrectomy group had the lower mortality from pulmonary causes with no association for all-cause mortality. In addition, associations varied according to two types of vitreoretinal disease: higher risk

of all-cause mortality and vascular causes in the patients with retinal vascular diseases and lower risk of all-cause mortality, vascular causes, and pulmonary causes in those with macular diseases. Though causality requires further study and analysis, the effect of vitrectomy on the mortality rates of patients can be inferred from the current observations. Further, the expected direction of changes of the future could be predicted based on the different mortality rate after vitrectomy in macular and retinal vascular disease.

DATA AVAILABILITY STATEMENT

Korean claims data used for this study is available only when accessed on Cloud, therefore the database set cannot be exported from the system. Further inquiries can be directed to the corresponding authors.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Asan Medical Center IRB. Written informed consent for participation was not required for this study in accordance with the national legislation and the institutional requirements.

AUTHOR CONTRIBUTIONS

YJK, HL, and JYK contributed to conception and design of the study. JSL performed the statistical analysis. YJK wrote the first draft of the manuscript. HL, JYK, and HWT wrote sections of the manuscript. All authors contributed to manuscript revision, read, and approved the submitted version.

FUNDING

This work was supported by the Korea Medical Device Development Fund grant funded by the Korea government (the Ministry of Science and ICT, the Ministry of Trade, Industry and Energy, the Ministry of Health and Welfare, the Ministry of Food and Drug Safety) (Project Number: 9991006821, KMDF_PR_20200901_0148), Korean Fund for Regenerative Medicine funded by Ministry of Science and ICT, and Ministry of Health and Welfare (21C0723L1-11, Republic of Korea), a grant from the Asan Institute for Life Sciences, Asan Medical Center, Seoul, Korea (2021IL0034, 2021IP0060), and a grant (NRF-2018R1D1A1B07043010) from the National Research Foundation of Korea.

ACKNOWLEDGMENTS

We would like to thank Editage (www.editage.co.kr) for their English language editing assistance.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmed.2022.851536/full#supplementary-material

REFERENCES

1. Kim MJ, Park KH, Hwang JM Yu HG, Yu YS, Chung H. The safety and efficacy of transconjunctival sutureless 23-gauge vitrectomy. *Korean J Ophthalmol.* (2007) 21:201–7. doi: 10.3341/kjo.2007.21.4.201
2. Yang SJ, Yoon SY, Kim JG, Yoon YH. Transconjunctival sutureless vitrectomy for the treatment of vitreoretinal complications in patients with diabetes mellitus. *Ophthalmic Surg Lasers Imaging.* (2009) 40:461–6. doi: 10.3928/15428877-20090901-04
3. Wubben TJ, Talwar N, Blachley TS, Gardner TW, Johnson MW, Lee PP, et al. Rates of Vitrectomy among Enrollees in a United States Managed Care Network, 2001–2012. *Ophthalmology.* (2016) 123:590–8. doi: 10.1016/j.ophtha.2015.11.001
4. Kim JY, Rim TH, Kim SS. Trends of Pars Plana Vitrectomy Rates in South Korea: A Nationwide Cohort Study. *Korean J Ophthalmol.* (2017) 31:446–51. doi: 10.3341/kjo.2016.0070
5. Fong CS, Mitchell P, Rochtchina E, Teber ET, Hong T, Wang JJ. Correction of visual impairment by cataract surgery and improved survival in older persons: the Blue Mountains Eye Study cohort. *Ophthalmology.* (2013) 120:1720–7. doi: 10.1016/j.ophtha.2013.02.009
6. Tseng VL Yu F, Lum F, Coleman AL. Cataract Surgery and Mortality in the United States Medicare Population. *Ophthalmology.* (2016) 123:1019–26. doi: 10.1016/j.ophtha.2015.12.033
7. Tseng VL, Chlebowski RT Yu F, Cauley JA Li W, Thomas F, Virnig BA, Coleman AL. Association of cataract surgery with mortality in older women: findings from the women's health initiative. *JAMA Ophthalmol.* (2018) 136:3–10. doi: 10.1001/jamaophthalmol.2017.4512
8. Liu E, Estevez J, Kaidonis G, Hassall M, Phillips R, Raymond G, et al. Long-term survival rates of patients undergoing vitrectomy for diabetic retinopathy in an Australian population: a population-based audit. *Clin Exp Ophthalmol.* (2019) 47:598–604. doi: 10.1111/ceo.13466
9. Banerjee PJ, Moya R, Bunce C, Charteris DG, Yorston D, Wickham L. Long-term survival rates of patients undergoing vitrectomy for proliferative diabetic retinopathy. *Ophthalmic Epidemiol.* (2016) 23:94–8. doi: 10.3109/09286586.2015.1089578
10. Shukla SY, Hariprasad AS, Hariprasad SM. Long-term mortality in diabetic patients with tractional retinal detachments. *Ophthalmol Retina.* (2017) 1:8–11. doi: 10.1016/j.oret.2016.09.002
11. Kim YI, Kim YY, Yoon JL, Won CW, Ha S, Cho KD, et al. Cohort profile: national health insurance service-senior (NHIS-senior) cohort in Korea. *BMJ Open.* (2019) 9:e024344. doi: 10.1136/bmjopen-2018-024344
12. Ryu SY, Kim J, Hong JH, Chung EJ. Incidence and characteristics of cataract surgery in South Korea from 2011 to 2015: a nationwide population-based study. *Clin Exp Ophthalmol.* (2020) 48:319–27. doi: 10.1111/ceo.13705
13. Levesque LE, Hanley JA, Kezouh A, Suissa S. Problem of immortal time bias in cohort studies: example using statins for preventing progression of diabetes. *BMJ.* (2010) 340:b5087. doi: 10.1136/bmj.b5087
14. Charlson ME, Pompei P, Ales KL, MacKenzie CR, A. new method of classifying prognostic comorbidity in longitudinal studies: development and validation. *J Chronic Dis.* (1987) 40:373–83. doi: 10.1016/0021-9681(87)90171-8
15. Austin PC. Balance diagnostics for comparing the distribution of baseline covariates between treatment groups in propensity-score matched samples. *Stat Med.* (2009) 28:3083–107. doi: 10.1002/sim.3697
16. Mamdani M, Sykora K, Li P, Normand SL, Streiner DL, Austin PC, et al. Reader's guide to critical appraisal of cohort studies: 2. Assessing potential for confounding. *BMJ.* (2005) 330:960–2. doi: 10.1136/bmj.330.7497.960
17. Kim JY, Chung HS, Lee JS, Lee H, Tchah H. Relationship between cataract surgery and mortality in elderly patients with cataract: nationwide population-based cohort study in South Korea. *J Pers Med.* (2021) 11:1128. doi: 10.3390/jpm11111128

18. Lee SY, Lee H, Lee JS, Han SA, Kim YJ, Kim JY, et al. Association between glaucoma surgery and all-cause and cause-specific mortality among elderly patients with glaucoma: a nationwide population-based cohort study. *Sci Rep.* (2021) 11:17055. doi: 10.1038/s41598-021-96063-7
19. Modjtahedi BS, Wu J, Luong TQ, Gandhi NK, Fong DS, Chen W. Severity of diabetic retinopathy and the risk of future cerebrovascular disease, cardiovascular disease, all-cause mortality. *Ophthalmology.* (2020) 128:1169–79. doi: 10.1016/j.ophtha.2020.12.019
20. Seong SC, Kim YY, Park SK, Khang YH, Kim HC, Park JH, et al. Cohort profile: the National Health Insurance Service-National Health Screening Cohort (NHIS-HEALS) in Korea. *BMJ Open.* (2017) 7:e016640. doi: 10.1136/bmjopen-2017-016640
21. Lee J, Lee JS, Park SH, Shin SA, Kim K. Cohort Profile: The National Health Insurance Service-National Sample Cohort (NHIS-NSC), South Korea. *Int J Epidemiol.* (2017) 46:e15. doi: 10.1093/ije/dyv319

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers | Frontiers in Medicine

ORIGINAL RESEARCH
published: 17 May 2022
doi: 10.3389/fmed.2022.871382

Artificial Intelligence for the Estimation of Visual Acuity Using Multi-Source Anterior Segment Optical Coherence Tomographic Images in Senile Cataract

Hyunmin Ahn[1], *Ikhyun Jun*[1,2], *Kyoung Yul Seo*[1], *Eung Kweon Kim*[2,3] *and Tae-im Kim*[1,2]*

[1] Department of Ophthalmology, Institute of Vision Research, Yonsei University College of Medicine, Seoul, South Korea, [2] Corneal Dystrophy Research Institute, Yonsei University College of Medicine, Seoul, South Korea, [3] Saevit Eye Hospital, Goyang, South Korea

OPEN ACCESS

Edited by:
Michele Lanza,
University of Campania Luigi Vanvitelli, Italy

Reviewed by:
Rajiv Raman,
Sankara Nethralaya, India
Mukharram M. Bikbov,
Ufa Eye Research Institute, Russia
Emmanuel Bui Quoc,
Assistance Publique Hopitaux de Paris, France
Shiri Soudry,
Rambam Health Care Campus, Israel
Georgios A. Kounis,
GNEMS, Greece
Yingting Zhu,
Sun Yat-sen University, China

***Correspondence:**
Tae-im Kim
tikim@yuhs.ac

Specialty section:
This article was submitted to Ophthalmology, a section of the journal Frontiers in Medicine

Received: 08 February 2022
Accepted: 04 April 2022
Published: 17 May 2022

Citation:
Ahn H, Jun I, Seo KY, Kim EK and Kim T-i (2022) Artificial Intelligence for the Estimation of Visual Acuity Using Multi-Source Anterior Segment Optical Coherence Tomographic Images in Senile Cataract. Front. Med. 9:871382. doi: 10.3389/fmed.2022.871382

Purpose: To investigate an artificial intelligence (AI) model performance using multi-source anterior segment optical coherence tomographic (OCT) images in estimating the preoperative best-corrected visual acuity (BCVA) in patients with senile cataract.

Design: Retrospective, cross-instrument validation study.

Subjects: A total of 2,332 anterior segment images obtained using swept-source OCT, optical biometry for intraocular lens calculation, and a femtosecond laser platform in patients with senile cataract and postoperative BCVA ≥ 0.0 logMAR were included in the training/validation dataset. A total of 1,002 images obtained using optical biometry and another femtosecond laser platform in patients who underwent cataract surgery in 2021 were used for the test dataset.

Methods: AI modeling was based on an ensemble model of Inception-v4 and ResNet. The BCVA training/validation dataset was used for model training. The model performance was evaluated using the test dataset. Analysis of absolute error (AE) was performed by comparing the difference between true preoperative BCVA and estimated preoperative BCVA, as ≥0.1 logMAR ($AE_{\geq 0.1}$) or <0.1 logMAR ($AE_{<0.1}$). $AE_{\geq 0.1}$ was classified into underestimation and overestimation groups based on the logMAR scale.

Outcome Measurements: Mean absolute error (MAE), root mean square error (RMSE), mean percentage error (MPE), and correlation coefficient between true preoperative BCVA and estimated preoperative BCVA.

Results: The test dataset MAE, RMSE, and MPE were 0.050 ± 0.130 logMAR, 0.140 ± 0.134 logMAR, and 1.3 ± 13.9%, respectively. The correlation coefficient was 0.969 ($p < 0.001$). The percentage of cases with $AE_{\geq 0.1}$ was 8.4%. The incidence of postoperative BCVA > 0.1 was 21.4% in the $AE_{\geq 0.1}$ group, of which 88.9% were in the underestimation group. The incidence of vision-impairing disease in the underestimation group was 95.7%. Preoperative corneal astigmatism and lens thickness were higher, and nucleus cataract was more severe ($p < 0.001$, 0.007, and 0.024, respectively) in $AE_{\geq 0.1}$

than that in $AE_{<0.1}$. The longer the axial length and the more severe the cortical/posterior subcapsular opacity, the better the estimated BCVA than the true BCVA.

Conclusions: The AI model achieved high-level visual acuity estimation in patients with senile cataract. This quantification method encompassed both visual acuity and cataract severity of OCT image, which are the main indications for cataract surgery, showing the potential to objectively evaluate cataract severity.

Keywords: artificial intelligence, cataract, convolutional neural network, optical coherence tomography, visual acuity

INTRODUCTION

Cataract is the leading cause of blindness, with ∼12.6 million cases of cataract worldwide (1). The visual impairment caused by cataract can be treated with advanced cataract surgery, which can ensure a progressively better quality of vision and fewer complications than in the past (2–5). The most important indications for cataract surgery are preoperative visual acuity and cataract grading, and advances in surgical technology have expanded the scope of the surgery to even include less severe cataracts (5).

Although the cataract grading system shows a good correlation with surgical difficulty, indicating that the surgery becomes more challenging as the cataract grading increases (6, 7), it shows limitations in reflecting the patient's visual symptoms, especially in cases with nuclear cataract and cortical opacity (8). Cataract grading depends on the subjective competence of the investigator (9). However, visual acuity reflects the patient's symptoms and influences surgical difficulty; therefore, the surgery becomes more challenging also as the visual acuity decreases (10, 11). Moreover, preoperative visual acuity can serve as an important predictor of postoperative vision in various diseases (12–14).

Artificial intelligence (AI) is being increasingly used in medicine, and ophthalmology is one of the most active fields for its clinical application (15). Recent studies have attempted to use AI for cataract grading with various methods, including slit-lamp photography, fundus photography, and optical coherence tomography (OCT) (16–22), and the results suggest that AI-based cataract grading shows acceptable performance with 70–90% accuracy. However, an AI-based approach for evaluation of visual acuity in patients with cataracts is still lacking. An approach linking objective image data with the subjective symptoms represented by visual acuity is particularly relevant, since the resultant method would encompass both visual acuity and cataract grade, which are the main indications for cataract surgery. Therefore, we attempted to implement an AI model that can evaluate cataract severity based on visual acuity by using multi-source OCT data and to assess the applicability of this AI model in actual clinical practice.

Abbreviations: AE, absolute error; AI, artificial intelligence; BCVA, best-corrected visual acuity; CNN, convolutional neural network; LOCS, Lens Opacities Classification System; MAE, mean absolute error; MPE, mean percentage error; OCT, optical coherence tomography; RMSE, root mean square error.

METHODS

The study was conducted at the Department of Ophthalmology, Severance Hospital, Yonsei University College of Medicine in accordance with the ethical standards of the Declaration of Helsinki, and institutional review board approval was obtained for the study protocol (4-2021-1697). The institutional review boards waived the need for informed consent because of the retrospective and de-identified nature of the study.

Participants and Dataset

All medical records of patients who underwent cataract surgery between January 2019 and December 2021 were reviewed. The demographic and clinical information of the patients, including age, sex, and clinical history, was collected. We defined distinct inclusion criteria for the training/validation and test datasets. For the training/validation dataset, we collected 2,332 anterior segment OCT images of 2,332 eyes in 2,332 patients with senile cataract alone whose 1-month postoperative best-corrected visual acuity (BCVA) was 0.0 logMAR or better; the images were obtained between January 2019 and December 2020 by using swept-source OCT (ANTERION® swept-source OCT; Heidelberg Engineering, Heidelberg, Germany), optical biometry for intraocular lens calculation (IOLMaster® 700; Carl Zeiss Meditec AG, Jena, Germany), and a femtosecond laser platform (LenSx® laser system; Alcon Laboratories, Inc., Fort Worth, TX, USA). Through this process, we aimed to develop a pure cataract analyzer. In the test dataset, to evaluate the application of the analyzer in actual clinical practice, we collected 1,002 anterior segment OCT images of 1,002 eyes of 1,002 patients who underwent cataract surgery between January 2021 and December 2021; the images were obtained using optical biometry (IOLMaster® 700) and another femtosecond laser platform (CATALYS™ Precision Laser System; Johnson & Johnson Inc., New Brunswick, NJ, USA). When multi-axial images were obtained from one device, a vertical image was selected. When multi-source anterior segment OCT images of a patient were available, a study image was selected randomly using Python version 3.8. When data for both eyes of a patient were available, a study eye was selected randomly. In the training/validation and test datasets, the patients with no anterior segment OCT images of the crystalline lens, not obtained due to corneal opacity or other reasons, were excluded. All OCT images in the training/validation and test datasets were labeled with the preoperative BCVA (**Figure 1**).

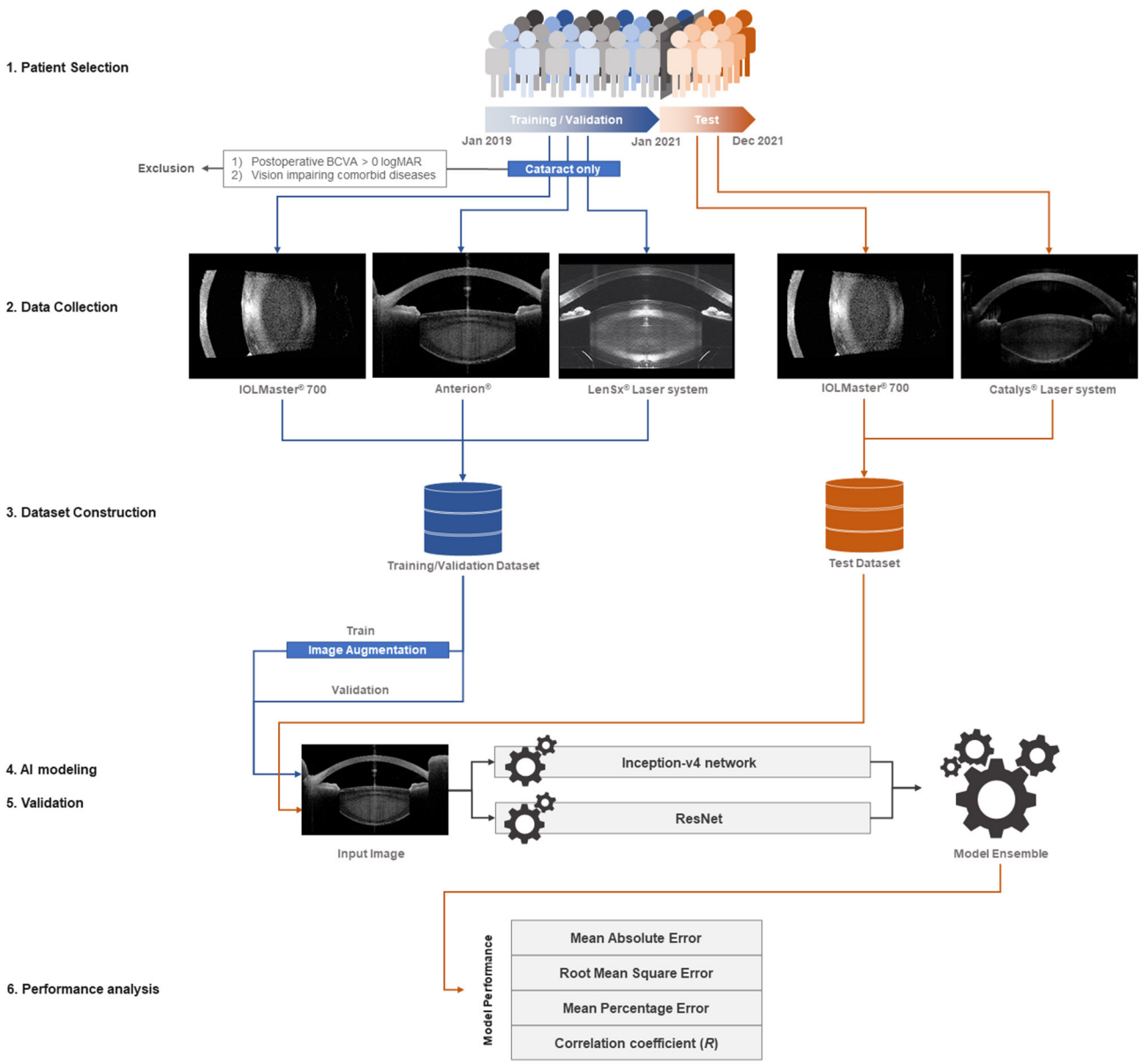

FIGURE 1 | Study flow diagram. For the training/validation dataset, we collected 2,332 anterior segment OCT images of 2,332 eyes in as many patients with isolated senile cataract, whose 1-month postoperative best-corrected visual acuity (BCVA) was 0.0 logMAR or better. The images were obtained between January 2019 and December 2020 using swept-source OCT (ANTERION®), optical biometry for intraocular lens calculation (IOLMaster® 700), and a femtosecond laser platform (LenSx® Laser System). With this process, we aimed to develop a pure cataract analyzer. For the test dataset, we collected 1,002 anterior segment OCT images of 1,002 eyes in as many patients who underwent cataract surgery between January 2021 and December 2021 to evaluate the application of the analyzer in actual clinical practice. The images were obtained using optical biometry (IOLMaster® 700) and a different femtosecond laser platform (CATALYS™ Precision Laser System). The multi-source OCT images varied in size, contour, and direction; an image augmentation method (ImageDataGenerator from the Keras library) was used for the training dataset. The AI modeling was based on an ensemble of the Inception-v4 convolutional neural network (CNN) and ResNet via the stacking technique. The mean absolute error, root mean square error, mean percentage error, and correlation coefficient between the actual and predicted preoperative BCVA were calculated to evaluate the model performance.

Artificial Intelligence Modeling

In this study, to improve the model performance and to analyze the technical error, we manipulated the AI modeling through reduction of input image size, use of image augmentation in the training dataset, and modification of the model architecture. The image size was reduced from 100 to 1% in 1% increments while maintaining the same height-to-width ratio. The multi-source OCT images varied in size, contour, and direction; the image augmentation method using ImageDataGenerator from Keras library was used for the training dataset. Finally, AI modeling

was based on an ensemble of the Inception-v4 convolutional neural network (CNN) and ResNet via the stacking technique (23, 24). In the high-resolution and large-scale images, such as the OCT images in this study, a very deep CNN architecture can be expected to perform well until a certain level (25). However, exploding calculation and gradient vanishing are problems associated with very deep CNNs. To overcome these problems, architectures, such as the Inception network and ResNet, can be considered (23, 24). Ensemble modeling of the Inception network and ResNet was implemented using an aggregating method with a weighted average, and the hyperparameters of the model were modified to ensure that the model performed flexibly according to the proportion of the input image shape.

Clinical Assessments

Clinical assessments were used to evaluate AI performance and perform error analysis in the medical approach. All patients underwent detailed preoperative examinations, including slit-lamp biomicroscopy, non-contact tonometry, ophthalmoscopy, and manifest refraction for BCVA. The mean corneal power and corneal astigmatism were measured using autokeratometry (Topcon KR-800A; Topcon Corporation, Tokyo, Japan). Intraocular lens calculation was performed using the IOLMaster® 700, and axial length and lens thickness were measured. An A-scan ultrasound biometry was used for intraocular lens calculation when the IOLMaster® 700 was unavailable. Cataract grading was performed using the Lens Opacities Classification System (LOCS) III by an expert surgeon (T.K.) who evaluated the opacity of the cortex, nucleus, and posterior subcapsular portion of the crystalline lens (26). Postoperative examinations, BCVA, slit-lamp examination, and pupil-dilation were conducted 1 month after the cataract surgery. Vision-impairing disease was defined as a clearly diagnosed disease in the detailed pre- and postoperative examinations that satisfied all of the following criteria: (1) postoperative BCVA > 0.1 logMAR, (2) persistent disease (that leaves an irreversible visual sequelae), (3) existing before cataract surgery, and (4) not a complication of cataract surgery.

Model Performance

To evaluate the model performance, the mean absolute error (MAE), root mean square error (RMSE, which is influenced by large errors), and mean percentage error (MPE, which shows in percentage how much the forecasts of a model differ from the actual values) between the actual preoperative BCVA and predicted preoperative BCVA were calculated as follows:

$$MAE = \frac{\Sigma \left| y - \hat{y} \right|}{n}$$

$$RMSE = \sqrt{\frac{\Sigma \left(y - \hat{y} \right)^2}{n}}$$

$$MPE = \frac{\sum \left(y - \hat{y} \right)}{n} \times 100\ (\%)$$

y: true preoperative BCVA
$\hat{y}$: estimated preoperative BCVA
n: number of images in the test dataset

TABLE 1 | Demographics and characteristics of the training/validation dataset and test dataset.

	Training/validation dataset (n = 2,332)	**Test dataset (n = 1,002)**
OCT image source	ANTERION® (n = 580) IOL-master® 700 (n = 1166) LenSx® laser system (n = 586)	IOL-master® 700 (n = 621) CATALYS™ laser system (n = 381)
Underlying disease	Senile cataract only	Senile cataract and/or other vision-impairing condition
Age (years, mean ± SD)	67.5 ± 9.1	69.1 ± 7.9
Sex (% of female)	64.5	63.9
Preoperative BCVA (logMAR, mean ± SD)	0.22 ± 0.15	0.20 ± 0.51
Postoperative BCVA (logMAR, mean ± SD)	0.00 ± 0.00	0.08 ± 0.12
Axial length (mm, mean ± SD)*	24.3 ± 1.2 (n = 2,186)	24.4 ± 1.3 (n = 948)
Lens thickness (mm, mean ± SD)*	4.3 ± 0.6 (n = 2,186)	4.3 ± 0.7 (n = 948)
LOCS III cataract grade		
Cortical (mean ± SD)	2.9 ± 1.3	3.0 ± 1.0
Nucleus (mean ± SD)	3.0 ± 0.6	3.0 ± 0.5
Posterior subcapsule (mean ± SD)	0.9 ± 1.5	1.0 ± 1.3

BCVA, best corrected visual acuity; LOCS III, Lens Opacities Classification System III; logMAR, logarithm of the minimum angle of resolution; OCT, optical coherence tomography; SD, standard deviation.
**Considered the values of optical biometry.*

A correlation analysis was conducted between the true and estimated preoperative BCVAs with Pearson correlation coefficient.

Error Analysis

The participants were initially classified into two groups based on the absolute error (AE) between true preoperative BCVA and estimated preoperative BCVA, one with an AE of at least 0.1 logMAR ($AE_{\geq 0.1}$), and the other with AE under 0.1 logMAR ($AE_{<0.1}$). Next, the $AE_{\geq 0.1}$ group was further divided into the underestimation group, in which the estimated preoperative BCVA was lower than the true preoperative BCVA, and the overestimation group, in which the estimated preoperative BCVA was higher than the true preoperative BCVA. Ordinal interference was considered if the clinical and statistical sequences were consistent.

TABLE 2 | Performance of the artificial intelligence model for prediction of preoperative best corrected visual acuity in patients with senile cataract.

Performance parameter	Value
MAE (logMAR, mean ± SD)	0.050 ± 0.130
RMSE (logMAR, mean ± SD)	0.140 ± 0.134
MPE (%, mean ± SD)	1.3 ± 13.9
Correlation coefficient (*R*)	0.969 ($p < 0.001$)

logMAR, logarithm of the minimum angle of resolution; MAE, mean absolute error; MPE, mean percentage error; RMSE, root mean square error; SD, standard deviation.

Statistical Analysis

Comparative analyses of clinical assessments were conducted between the groups and between the subgroups with independent *t*-tests for continuous variables and Fisher's exact tests for categorical variables. The values from the A-scan ultrasound biometry were excluded in the comparisons of axial length and lens thickness. Statistical significance was set at $p < 0.05$.

RESULTS

Table 1 shows the overall demographics and characteristics of the training/validation and test datasets. Although the OCT image source and underlying disease differed between the datasets, no remarkable differences were observed in age, sex, preoperative BCVA, axial length, lens thickness, and LOCS III cataract grade. The postoperative BCVA differed between the datasets ($p < 0.001$).

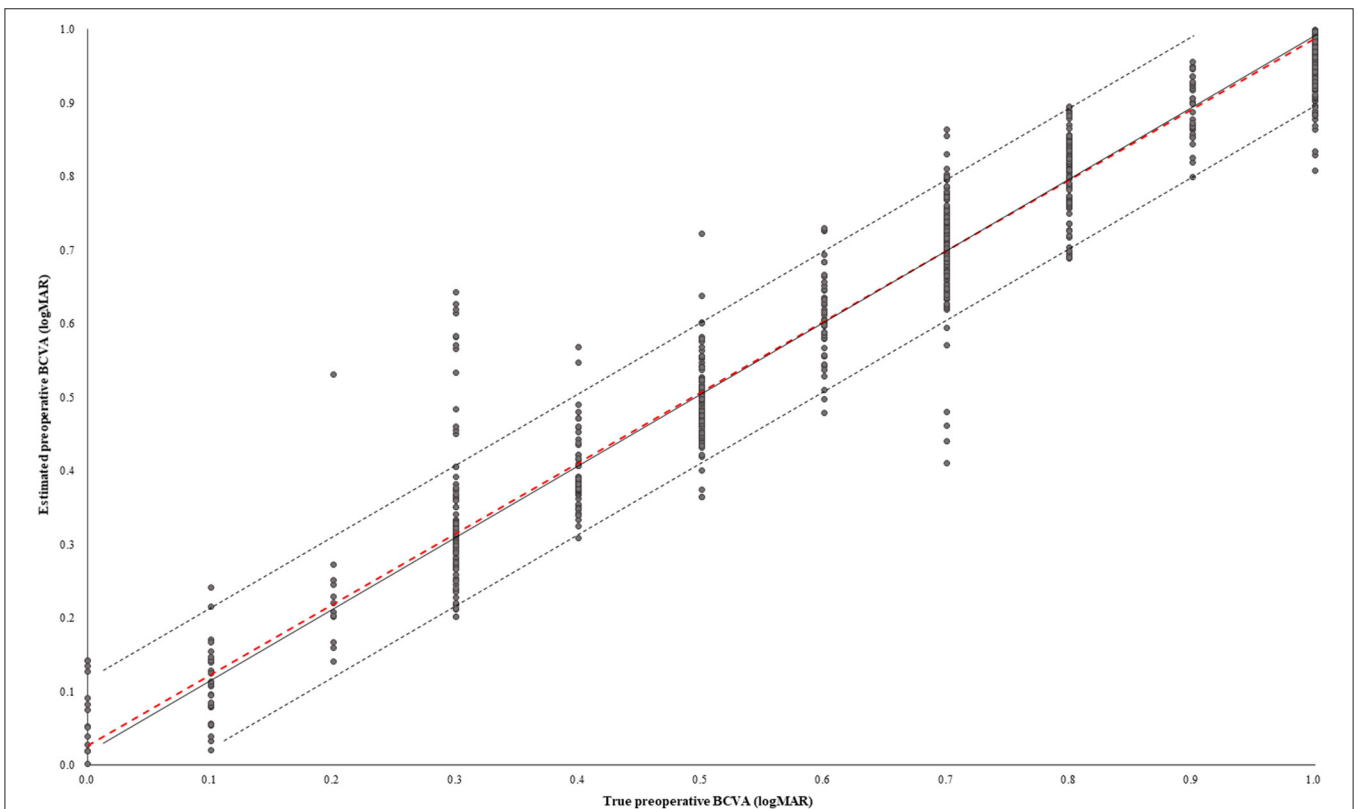

FIGURE 2 | Scatter plot between true preoperative BCVA and estimated preoperative BCVA. The solid black line shows that the true and estimated preoperative BCVAs were equal. The black dotted line shows the absolute error (AE) of the true and estimated preoperative BCVAs of 0.1 logMAR. The red dotted line describes the Pearson correlation coefficient. In the test dataset, 91.6% of cases had AE under 0.1 logMAR.

In the test dataset, the MAE, RMSE, and MPE of the model performance were 0.050 ± 0.130 logMAR, 0.140 ± 0.134 logMAR, and 1.3 ± 13.9%, respectively (**Table 2**). The correlation coefficient (R) between the true preoperative BCVA and estimated preoperative BCVA was 0.969 ($p < 0.001$) (**Figure 2**). The percentage of cases in the AE ≥ 0.1 group was 8.4%, and 1.9% had AE ≥ 0.2.

The preoperative BCVA values were significantly different between the $AE_{\geq 0.1}$ and $AE_{<0.1}$ groups (0.30 ± 0.08 vs. 0.12 ± 0.10, respectively; $p < 0.001$), and corneal astigmatism, lens thickness, and nucleus cataract in the LOCS III grading were significantly different ($p < 0.001$, <0.001, and 0.024, respectively) (**Table 3**). Postoperative BCVA in the $AE_{\geq 0.1}$ group was worse than that in the $AE_{<0.1}$ group ($p < 0.001$). In the $AE_{\geq 0.1}$ group, the incidence of postoperative BCVA > 0.1 was 21.4%, which was higher than that in the $AE_{<0.1}$ group, and the proportion of vision-impairing disease was also higher than in the $AE_{<0.1}$ group. The percentage of cases with vision-impairing diseases was significantly higher in the $AE_{\geq 0.1}$ group than in the $AE_{<0.1}$ group (54.8 vs. 12.9%; $p < 0.001$).

The $AE_{\geq 0.1}$ group was further divided into underestimation and overestimation groups (**Table 4**). Pre- and postoperative BCVAs were significantly worse in the underestimation group (0.39 ± 0.10 *vs.* 0.15 ± 0.09 in overestimation, and 0.10 ± 0.15 vs. 0.01 ± 0.05; $p < 0.001$ and $p < 0.001$, respectively). The proportion of cases with postoperative BCVA > 0.1 was higher in the underestimation group ($p < 0.001$). The proportion of cases with vision-impairing disease was 95.7% in the underestimation group. Axial length was longer, and cortical opacity and posterior subcapsular opacity were more severe in the underestimation group ($p < 0.001$, 0.045, and 0.002, respectively).

Error analysis for model performance was conducted by performing comparative analyses between the $AE_{\geq 0.1}$ and $AE_{<0.1}$ groups and between the underestimation and overestimation subgroups of the $AE_{\geq 0.1}$ group (**Table 5**). Preoperative corneal astigmatism, lens thickness, and nuclear opacity in the LOCS III grading were higher in the $AE_{\geq 0.1}$ group. Increased BCVA and vision-impairing disease were more significant in the underestimation subgroup. Axial length was longer, and the cortical/posterior subcapsular opacity was more severe in the overestimation subgroup than in the underestimation subgroup.

In the AI modeling process, reducing the original image decreased the model performance, and the neural network could not estimate the BCVA when the area was reduced by ∼≥90% (32% of width and height) (**Figure 3A**). Image augmentation increased the model performance in terms of the MAE (**Figure 3B**). The ensemble model had the lowest MAE, and Inception-v4 and ResNet were almost similar in terms of MAE (**Figure 3C**).

DISCUSSION

With the test dataset of this study, over 90% of cases could be estimated in their BCVA under 0.1 logMAR of AE. Most of the underestimation errors were caused by vision-impairing disease, and about half of the decreased model performance could be

TABLE 3 | Comparison of the patients showing absolute error of BCVA of 0.1 and over ($AE_{\geq 0.1}$) with those showing absolute error under 0.1 ($AE_{<0.1}$) in the test dataset.

	$AE_{\geq 0.1}$ ($n = 84$)	$AE_{<0.1}$ ($n = 918$)	*p*-value
OCT image source (% of CATALYS^{TM} laser system)‡	35.7	35.9	0.966*
Mean age (years, mean ± *SD*)§	70.5 ± 8.2	69.0 ± 7.9	0.180
Sex (% of female)‡	64.3	63.8	0.934
Preoperative BCVA (logMAR, mean ± SD)§	0.30 ± 0.08	0.12 ± 0.10	<0.001*
Preoperative corneal power (D, mean ±*SD*)§	43.9 ± 1.6	43.9 ± 1.2	1.000
Preoperative corneal astigmatism (D, mean ± SD)§	1.2 ± 1.1	0.8 ± 0.5	<0.001*
Axial length (mm, mean ±*SD*)†§	24.6 ± 0.9 ($n = 75$)	24.4 ± 1.4 ($n = 873$)	0.225
Lens thickness (mm, mean ± SD)†§	4.6 ± 0.2 ($n = 75$)	4.3 ± 0.7 ($n = 873$)	<0.001*
LOCS III cataract grade			
Cortical (mean ± SD)§	2.9 ± 1.1	3.0 ± 1.0	0.436
Nucleus (mean ± SD)§	3.1 ± 0.3	3.0 ± 0.3	0.024*
Posterior subcapsule (mean ± SD)§	0.9 ± 1.6	1.0 ± 1.3	0.491
Postoperative BCVA (logMAR, mean ± SD)§	0.07 ± 0.05	0.01 ± 0.09	<0.001*
Postoperative BCVA > 0.1 (%)‡	21.4	6.8	<0.001*
Vision-impairing disease (%)‡	54.8	12.9	<0.001*
Corneal disease, central (%)‡	4.8	0.0	<0.001*
Macular disease (%)‡	26.2	7.2	<0.001*
Glaucoma and other optic neuropathy (%)‡	28.6	6.3	<0.001*

**p-value < 0.05.*
†*Considered the values of optical biometry only.*
‡*Analyzed by Fisher's exact test.*
§*Analyzed by independent t-test.*
BCVA, best corrected visual acuity; D, diopter; LOCS III, Lens Opacities Classification System III; logMAR, logarithm of the minimum angle of resolution; SD, standard deviation.

TABLE 4 | Comparison of the overestimation and underestimation groups among patients showing an absolute error of 0.1 and over ($AE_{\geq 0.1}$) in BCVA.

	Underestimation ($n = 46$)	Overestimation($n = 38$)	*p*-value
OCT image source (% of CATALYS^{TM} laser system)‡	37.0	34.0	0.794
Mean age (years, mean ± SD)§	69.6 ± 7.5	71.8 ± 8.2	0.203
Sex (% of females)‡	65.2	63.2	1.000
Preoperative BCVA (logMAR, mean ± SD)§	0.39 ± 0.10	0.15 ± 0.09	<0.001*
Preoperative corneal power (D, mean ± SD)§	43.8 ± 1.4	44.2 ± 1.7	0.240
Preoperative corneal astigmatism (D, mean ± SD)§	1.2 ± 1.1	1.1 ± 1.0	0.677
Axial length (mm, mean ± SD)†	25.1 ± 1.0 ($n = 37$)	24.2 ± 0.8	<0.001*
Lens thickness (mm, mean ± SD)†	4.5 ± 0.2 ($n = 37$)	4.6 ± 0.2	<0.034*
LOCS III cataract grade			
Cortical (mean ± SD)§	3.3 ± 1.0	2.6 ± 1.0	>0.045*
Nucleus (mean ± SD)§	3.0 ± 0.5	3.1 ± 0.2	0.775
Posterior subcapsule (mean ± SD)§	1.5 ± 1.6	0.3 ± 0.6	>0.002*
Postoperative BCVA (logMAR, mean ± SD)§	0.10 ± 0.15	0.01 ± 0.05	<0.001*
Postoperative BCVA >0.1 [% of each group, (% of $AE_{\geq 0.1}$)]‡	34.8 (88.9)	5.3 (11.1)	<0.001*
Vision-impairing disease [% of each group, (% of $AE_{\geq 0.1}$)]‡	95.7 (95.7)	5.3 (4.3)	<0.001*

p-value < 0.05.
†*Considered the values of optical biometry only.*
‡*Analyzed by Fisher's exact test.*
§ *Analyzed by independent t-test.*
BCVA, best corrected visual acuity; D, diopter; LOCS III, Lens Opacities Classification System III; SD, standard deviation.

TABLE 5 | Error analysis of model performance in clinical practice.

Characteristics	$AE_{\geq 0.1]}$ vs. $AE_{<0.1}$	Estimation group		
		Under	$AE_{<0.1}$ (fair)	Over
OCT image source	−	−	−	−
Age	−	−	−	−
Sex	−	−	−	−
Preoperative BCVA	+	+	−	−
Preoperative corneal astigmatism	+	+	−	+
Postoperative BCVA	+	+	−	−
Postoperative BCVA over 0.1	+	+	−	−
Vision-impairing disease	+	+	−	−
Axial length†	−	+^{longer}	+^{intermediate}	+^{shorter}
Lens thickness‡	+	+	−	+
LOCS III grade				
Cortex†	−	+^{higher}	+^{intermediate}	+^{lower}
Nucleus	+	+	−	+
Posterior subcapsule†	−	+^{higher}	+^{intermediate}	+^{lower}

†*Ordinal interference.*
‡*Statistically different between the underestimation and overestimation groups but failed to list in order of lens thickness.*

explained by a clinical approach. The technical issues, image resolution, diversity of image forms in the training dataset, and model architecture also affected the model performance.

Objective clinical assessment for estimation of the subjective visual symptoms represented by visual acuity, which is considered the ultimate goal of ophthalmologic interventions, can be utilized in various ways from clinical to experimental; however, this is very difficult for clinicians due to confounding factors (27). AI is expected to help human evaluators perform difficult tasks (28). This study suggested that cataract severity can be quantified as visual acuity via AI using OCT images.

In this study, cross-instrument validation was performed using different combinations of anterior segment OCT image sources for the training/validation dataset and the test dataset. To enhance the model performance, the OCT images used in the training/validation dataset were obtained from three devices with different detection methods, image resolutions, and image

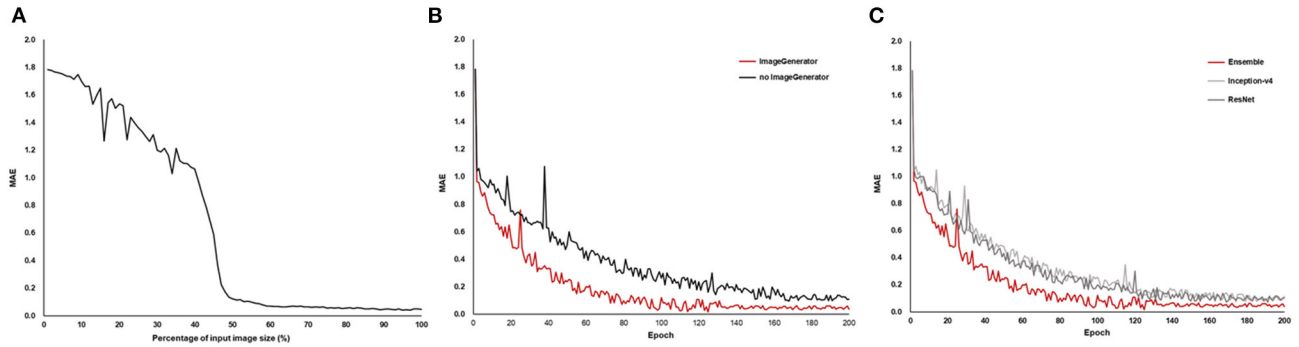

FIGURE 3 | Mean absolute error (MAE) of the model performance in the AI training process. As the size of the OCT image reduced, MAE increased **(A)**, and image augmentation using the image generator showed a lower MAE **(B)**. The ensemble model of Inception-v4 and ResNet showed the lowest MAE compared to the individual architecture **(C)**.

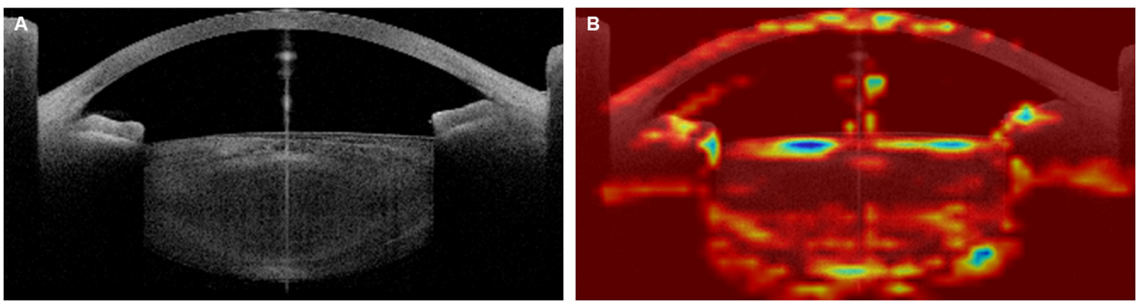

FIGURE 4 | Example of heatmap analysis of the Artificial Intelligence (AI) attention area. Original OCT image **(A)** and Heatmap image **(B)**. The preoperative best-corrected visual acuity (BCVA) of this 67-year-old patient, who showed a senile cataract with grade 3 cortical opacity and grade 2 nuclear opacity in the Lens Opacities Classification System III, was 0.50 logMAR, and the AI-estimated BCVA was 0.52 logMAR.

directions. The test dataset was constructed to evaluate the practical application of the AI model in real-world scenarios by including all patients within a certain period and using commonly available OCT images in the process of cataract surgery and its preparation.

Model performance and error analysis were conducted from both clinical and engineering perspectives. Although medical issues have rarely been mentioned in previous AI research, for clinical application of medical AI, error analysis from the clinical perspective is indispensable (29, 30). The results suggest that the predisposing disease and the conditions, such as high values of corneal astigmatism, axial length, lens thickness, and severity of each cataract subtype, which are well known to cause visual impairment, were the main factors underlying errors in the medical approach (8, 31–34). For engineering issues, this study suggested that low-resolution images led to the degradation of the model performance (see **Figure 3A**), and highlighted the importance of high-resolution images in analyzing precise medical observations (35). Multi-source images were used, and the transformation for direction and size through image augmentation in the training dataset was shown to yield better model performance (36). Although the AI model using a unified material may show a different performance from that using multi-sources materials, we focused on the versatility of the AI model by using OCT images acquired with multiple instruments.

The primary goal of cataract surgery is to restore the best possible vision as well as remove the natural crystalline lens (37). Thus, postoperative visual acuity is the target indicator of cataract surgery. Wei et al. reported the use of AI with fundus OCT for predicting postoperative visual acuity in patients with high myopia (38), and the result showed the lowest MAE of 0.16 logMAR and RMSE of 0.24 logMAR. However, this performance level was still insufficient for clinical application, and the findings implied that visual acuity was not determined by a single defined etiology. Clinicians frequently encounter patients with visual impairment without a specific pathologic lesion. This phenomenon may be caused by medical issues and unknown predisposing factors, including developmental problems, such as amblyopia, an extraocular disease, such as brain lesion, or a temporary problem, such as dry eye disease. Thus, approaches based on a single etiological factor showed limited ability to predict the postoperative BCVA. Since multiple factors, including the presence of both cataract and comorbid disease, can influence the accuracy of prediction of postoperative BCVA with AI based on the preoperative BCVA, our study could serve as the basis for future studies. Future studies should aim to predict postoperative

visual acuity using a combination of preoperative visual acuity, anterior segment OCT, and other examinations.

Similar to previous medical AI research (16, 39–41), this study did not completely overcome the limitations in interpreting the neural network. In this study, the attention areas of the AI model were analyzed using heatmap analysis (**Figure 4**), and we tried to analyze the relationships between the attention areas and the error analysis in the medical approach. Although model training was conducted using patients with only a clinical diagnosis of cataract, the AI also recognized the cornea that was laid on the visual axis and the angle in the anterior chamber. Thus, some factors affecting the corneal shape (e.g., corneal refractive surgery), anterior chamber angle (e.g., glaucoma), and lens thickness may be related.

The strengths of this study were not limited to the excellent performance of the AI model. Our attempt to prove the importance of clinical approach in model performance was successful and our findings suggest that the addition of various clinical information in AI modeling is crucial for improving model performance.

In conclusion, the AI developed using OCT images from multiple sources showed excellent performance in estimating visual acuity in patients with senile cataracts. This quantification method encompasses both visual acuity and cataract severity of the OCT images, which are the main indications for cataract surgery, and has the potential to allow objective evaluation of cataract severity. This AI model can be used when it is difficult to express or measure the subjective visual acuity due to various causes, such as an inability to communicate. Additionally, we would like to emphasize that this was a preliminary study to expand the prediction of visual acuity after cataract surgery in patients with other diseases, possibly accompanied by visual impairment.

DATA AVAILABILITY STATEMENT

The datasets presented in this article are not readily available because of privacy and ethical concerns, neither the data nor the source of the data can be made available. Requests to access the datasets should be directed to HA, overhyun31@gmail.com.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Severance Hospital Clinical Research Ethics Committee. Written informed consent for participation was not required for this study in accordance with the national legislation and the institutional requirements.

AUTHOR CONTRIBUTIONS

HA and T-iK: conceptualization. HA: methodology, software, formal analysis, investigation, writing—original draft preparation, and visualization. IJ, KS, EK, and T-iK: validation. IJ, KS, and T-iK: resources. T-iK: writing—review and editing and supervision. All authors contributed to the article and approved the submitted version.

REFERENCES

1. Flaxman SR, Bourne RR, Resnikoff S, Ackland P, Braithwaite T, Cicinelli MV, et al. Global causes of blindness and distance vision impairment 1990–2020: a systematic review and meta-analysis. *Lancet Global Health.* (2017) 5:e1221–34. doi: 10.1016/S2214-109X(17)30393-5
2. Lundström M, Dickman M, Henry Y, Manning S, Rosen P, Tassignon MJ, et al. Femtosecond laser–assisted cataract surgeries reported to the European Registry of Quality Outcomes for Cataract and Refractive Surgery: Baseline characteristics, surgical procedure, and outcomes. *J Cataract Refract Surg.* (2017) 43:1549–56. doi: 10.1016/j.jcrs.2017.09.029
3. Lundström M, Behndig A, Kugelberg M, Montan P, Stenevi U, Thorburn W. Decreasing rate of capsule complications in cataract surgery: eight-year study of incidence, risk factors, and data validity by the Swedish National Cataract Register. *J Cataract Refract Surg.* (2011) 37:1762–7. doi: 10.1016/j.jcrs.2011.05.022
4. Lundström M, Barry P, Henry Y, Rosen P, Stenevi U. Visual outcome of cataract surgery; study from the European Registry of Quality Outcomes for Cataract and Refractive Surgery. *J Cataract Refract Surg.* (2013) 39:673–9. doi: 10.1016/j.jcrs.2012.11.026
5. Lundström M, Dickman M, Henry Y, Manning S, Rosen P, Tassignon MJ, et al. Changing practice patterns in European cataract surgery as reflected in the European Registry of Quality Outcomes for Cataract and Refractive Surgery 2008 to 2017. *J Cataract Refract Surg.* (2021) 47:373–8. doi: 10.1097/j.jcrs.0000000000000457
6. Galan A, Tavolato M, Babighian S. Grading the surgical difficulty of cataract phacoemulsification. *Ophthal Surg Lasers Imaging Retina.* (2009) 40:361–5. doi: 10.3928/15428877-20096030-02
7. Lapid-Gortzak R. Gauging the difficulty of phacoemulsification: new grading systems. *Expert Rev Ophthalmol.* (2009) 4:455–6. doi: 10.1586/eop.09.41
8. Stifter E, Sacu S, Benesch T, Weghaupt H. Impairment of visual acuity and reading performance and the relationship with cataract type and density. *Invest Ophthalmol Vis Sci.* (2005) 46:2071–5. doi: 10.1167/iovs.04-0890
9. Karbassi M, Khu PM, Singer DM, Chylack Jr L. Evaluation of lens opacities classification system III applied at the slitlamp. *Optomet Vis Sci.* (1993) 70:923–8. doi: 10.1097/00006324-199311000-00009
10. McGwin Jr G, Scilley K, Brown J, Owsley C. Impact of cataract surgery on self-reported visual difficulties: comparison with a no-surgery reference group. *J Cataract Refract Surg.* (2003) 29:941–8. doi: 10.1016/S0886-3350(02)01846-1
11. Dooley IJ, O'Brien PD. Subjective difficulty of each stage of phacoemulsification cataract surgery performed by basic surgical trainees. *J Cataract Refract Surg.* (2006) 32:604–8. doi: 10.1016/j.jcrs.2006.01.045
12. Chu CJ, Johnston RL, Buscombe C, Sallam AB, Mohamed Q, Yang YC, et al. Risk factors and incidence of macular edema after cataract surgery: a database study of 81984 eyes. *Ophthalmology.* (2016) 123:316–23. doi: 10.1016/j.ophtha.2015.10.001
13. Yoeruek E, Deuter C, Gieselmann S, Saygili O, Spitzer MS, Tatar O, et al. Long-term visual acuity and its predictors after cataract surgery in patients with uveitis. *Eur J Ophthalmol.* (2010) 20:694–701. doi: 10.1177/112067211002000409
14. Ostri C, Lund-Andersen H, Sander B, La Cour M. Phacoemulsification cataract surgery in a large cohort of diabetes patients: visual acuity outcomes and prognostic factors. *J Cataract Refract Surg.* (2011) 37:2006–12. doi: 10.1016/j.jcrs.2011.05.030
15. Li J-PO, Liu H, Ting DS, Jeon S, Chan RP, Kim JE, et al. Digital technology, tele-medicine and artificial intelligence in ophthalmology: a global perspective. *Prog Retin Eye Res.* (2021) 82:100900. doi: 10.1016/j.preteyeres.2020.100900
16. Prevedello LM, Halabi SS, Shih G, Wu CC, Kohli MD, Chokshi FH, et al. Challenges related to artificial intelligence research in medical imaging and the

importance of image analysis competitions. *Radiol Artif Intelligence.* (2019) 1:e180031. doi: 10.1148/ryai.2019180031
17. Xu X, Zhang L, Li J, Guan Y, Zhang L. A hybrid global-local representation CNN model for automatic cataract grading. *IEEE J Biomed Health Inform.* (2019) 24:556–67. doi: 10.1109/JBHI.2019.2914690
18. Zhang H, Niu K, Xiong Y, Yang W, He Z, Song H. Automatic cataract grading methods based on deep learning. *Comput Methods Programs Biomed.* (2019) 182:104978. doi: 10.1016/j.cmpb.2019.07.006
19. Xiong L, Li H, Xu L. An approach to evaluate blurriness in retinal images with vitreous opacity for cataract diagnosis. *J Healthc Eng.* (2017) 2017:1–16. doi: 10.1155/2017/5645498
20. Yang J-J, Li J, Shen R, Zeng Y, He J, Bi J, et al. Exploiting ensemble learning for automatic cataract detection and grading. *Comput Methods Programs Biomed.* (2016) 124:45–57. doi: 10.1016/j.cmpb.2015.10.007
21. Guo L, Yang J-J, Peng L, Li J, Liang Q. A computer-aided healthcare system for cataract classification and grading based on fundus image analysis. *Comput Industry.* (2015) 69:72–80. doi: 10.1016/j.compind.2014.09.005
22. Zhang X, Xiao Z, Higashita R, Chen W, Yuan J, Fang J, et al. A novel deep learning method for nuclear cataract classification based on anterior segment optical coherence tomography images. In *IEEE.* (2020). p. 662–8. doi: 10.1109/SMC42975.2020.9283218
23. Szegedy C, Ioffe S, Vanhoucke V, Alemi AA. Inception-v4, inception-resnet and the impact of residual connections on learning. In: *Thirty-First AAAI Conference on Artificial Intelligence* (2017*).*
24. He K, Zhang X, Ren S, Sun J. *Deep Residual Learning for Image Recognition.* (2016). p. 770–8.
25. Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:14091556*. (2014). Available online at: https://arxiv.org/abs/1409.1556
26. Chylack LT, Wolfe JK, Singer DM, Leske MC, Bullimore MA, Bailey IL, et al. The lens opacities classification system III. *Arch Ophthalmol.* (1993) 111:831–6. doi: 10.1001/archopht.1993.01090060119035
27. Mangione CM, Orav EJ, Lawrence MG, Phillips RS, Seddon JM, Goldman L. Prediction of visual function after cataract surgery: a prospectively validated model. *Arch Ophthalmol.* (1995) 113:1305–11. doi: 10.1001/archopht.1995.01100100093037
28. Ahuja AS. The impact of artificial intelligence in medicine on the future role of the physician. *PeerJ.* (2019) 7:e7702. doi: 10.7717/peerj.7702
29. Ting DS, Lee AY, Wong TY. An ophthalmologist's guide to deciphering studies in artificial intelligence. *Ophthalmology.* (2019) 126:1475–9. doi: 10.1016/j.ophtha.2019.09.014
30. Szolovits P. *Artificial Intelligence in Medicine*. New York, NY: Routledge (2019).
31. Wolffsohn JS, Bhogal G, Shah S. Effect of uncorrected astigmatism on vision. *J Cataract Refract Surg.* (2011) 37:454–60. doi: 10.1016/j.jcrs.2010.09.022
32. Hashimoto S, Yasuda M, Fujiwara K, Ueda E, Hata J, Hirakawa Y, et al. Association between axial length and myopic maculopathy: the Hisayama Study. *Ophthalmol Retina.* (2019) 3:867–73. doi: 10.1016/j.oret.2019.04.023
33. Fujiwara A, Shiragami C, Manabe S, Izumibata S, Murata A, Morizane Y, et al. Normal values of retinal sensitivity determined by macular integrity assessment. *Invest Ophthalmol Vis Sci.* (2014) 55:5875–75. Available online at: https://iovs.arvojournals.org/article.aspx?articleid=2271519
34. Westall CA, Dhaliwal HS, Panton CM, Sigesmund D, Levin AV, Nischal KK, et al. Values of electroretinogram responses according to axial length. *Documenta Ophthalmol.* (2001) 102:115–30. doi: 10.1023/A:1017535207481
35. Kannojia SP, Jaiswal G. Effects of varying resolution on performance of CNN based image classification: an experimental study. *Int J Comput Sci Eng.* (2018) 6:451–6. doi: 10.26438/ijcse/v6i9.451456
36. Shorten C, Khoshgoftaar TM. A survey on image data augmentation for deep learning. *J Big Data.* (2019) 6:1–48. doi: 10.1186/s40537-019-0197-0
37. Brad H. Feldman SH. *Cataract*. American Academy of Ophthalmology (2021). Available online at: https://eyewiki.aao.org/Cataract (accessed January 23, 2022).
38. Wei L, He W, Wang J, Zhang K, Du Y, Qi J, et al. An optical coherence tomography-based deep learning algorithm for visual acuity prediction of highly myopic eyes after cataract surgery. *Front Cell Dev Biol.* (2021) 9:1195. doi: 10.3389/fcell.2021.652848
39. Kallianos K, Mongan J, Antani S, Henry T, Taylor A, Abuya J, et al. How far have we come? Artificial intelligence for chest radiograph interpretation. *Clin Radiol.* (2019) 74:338–45. doi: 10.1016/j.crad.2018.12.015
40. Alsharqi M, Woodward W, Mumith J, Markham D, Upton R, Leeson P. Artificial intelligence and echocardiography. *Echo Res Pract.* (2018) 5:R115–25. doi: 10.1530/ERP-18-0056
41. Mendelson EB. Artificial intelligence in breast imaging: potentials and limitations. *Am J Roentgenol.* (2019) 212:293–9. doi: 10.2214/AJR.18.20532

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers | Frontiers in Medicine

ORIGINAL RESEARCH
published: 19 May 2022
doi: 10.3389/fmed.2022.882131

Relationship Between Tamsulosin Use and Surgical Complications of Cataract Surgery in Elderly Patients: Population-Based Cohort Study

Jiehoon Kwak, Jung Yeob Han, Su Young Moon, Sanghyu Nam, Jae Yong Kim, Hungwon Tchah and Hun Lee**

Department of Ophthalmology, Asan Medical Center, University of Ulsan College of Medicine, Seoul, South Korea

OPEN ACCESS

Edited by:
Darren Shu Jeng Ting, University of Nottingham, United Kingdom

Reviewed by:
Aristeidis Konstantinidis, University Hospital of Alexandroupolis, Greece
Yi-Ting Hsieh, National Taiwan University Hospital, Taiwan

****Correspondence:***
Jae Yong Kim
jykim2311@amc.seoul.kr
Hun Lee
yhun777@gmail.com

Specialty section:
This article was submitted to Ophthalmology, a section of the journal Frontiers in Medicine

Received: *23 February 2022*
Accepted: *25 April 2022*
Published: *19 May 2022*

Citation:
Kwak J, Han JY, Moon SY, Nam S, Kim JY, Tchah H and Lee H (2022) Relationship Between Tamsulosin Use and Surgical Complications of Cataract Surgery in Elderly Patients: Population-Based Cohort Study. Front. Med. 9:882131. doi: 10.3389/fmed.2022.882131

Purpose: Although several previous studies have investigated the relationship between tamsulosin use and surgical complications of cataract surgery, no population-based cohort study has been conducted for the Asian population. We aimed to investigate the relationship between tamsulosin use and surgical complications of cataract surgery in the Korean elderly population.

Methods: This nationwide population-based retrospective cohort study included elderly patients (≥60 years) who had undergone cataract surgery in the period from 2003 to 2015. Baseline characteristics were age, sex, income, residence, and systemic, and ocular comorbidities (glaucoma, myopia, eye trauma, diabetes mellitus with ophthalmic manifestations, severe cataract, age-related macular degeneration). The exposure of interest was tamsulosin use within 1 year before cataract surgery. Logistic regression model was used to evaluate the relationship of tamsulosin use with surgical complications of cataract surgery.

Results: The rate of surgical complications of cataract surgery was 0.88% (375/42,539) in the non-tamsulosin group and 0.83% (71/8,510) in the tamsulosin group. The groups showed no significant difference in the risk of surgical complications of cataract surgery in the unadjusted model [odds ratio (OR) = 0.946; 95% confidence interval (CI):0.733–1.220; $P = 0.669$]. Additionally, tamsulosin use was not significantly associated with surgical complications of cataract surgery in the fully adjusted model accounting for age, income, residence, and systemic and ocular comorbidities (OR = 0.997; 95% CI: 0.749–1.325; $P = 0.981$).

Conclusions: The rate or risk of surgical complications of cataract surgery does not change with tamsulosin use. We suggest that better surgical techniques and surgeons' cognizance of the patient's tamsulosin use could improve surgical outcomes, without increasing surgical complications.

Keywords: cataract surgery, tamsulosin, surgical complication, KNHIS-Senior cohort, cataract (senile)

INTRODUCTION

Tamsulosin is a subtype-selective alpha (1A and 1D) adrenoceptor antagonist that induces relaxation of smooth muscles in the prostate and bladder (1). It is commonly used to treat symptomatic benign prostate hyperplasia (BPH) and kidney stones and was approved for use in the United States in 1997 and in South Korea in 2006. It has been prescribed globally to treat acute urinary retention caused by BPH, and systemic side effects, e.g., hypotension, have been relatively uncommon (1).

Intraoperative floppy iris syndrome (IFIS), which comprises intraoperative progressive miosis, iris prolapse, and iris billowing, is frequent in patients taking tamsulosin (2–4). The prevalence of IFIS is 2%, and most cases are related to tamsulosin use (5, 6). In addition, patients who are administered with tamsulosin preoperatively tend to develop miotic pupils, iris prolapse at the incision margin, and hypotonic iris during cataract surgery (2, 7, 8). Such anatomical or functional changes induced by tamsulosin increase the difficulty of cataract surgery, which could lead to perioperative or postoperative complications (2, 8, 9). Nonetheless, owing to the surgeon's effort to respond to the risk posed by tamsulosin use and advances in surgical equipment, surgical outcomes of cataract surgery improved in patients taking tamsulosin (9–12).

Previous studies evaluated surgical complications of cataract surgery associated with tamsulosin use (2, 6). A recent population-based study on yearly cataract surgical complication rates of patients taking tamsulosin demonstrated that the risk of cataract surgical complications decreased with time with or without tamsulosin use (13). However, it is unclear whether those findings are generalizable to the Korean elderly population, and no population-based cohort study has been conducted for the Asian population. Therefore, the aim of the present study was to investigate the relationship between tamsulosin use and surgical complications of cataract surgery in the Korean elderly population using the Korean National Health Insurance Service-Senior cohort (NHIS-Senior) database.

METHODS

Study Design and Data Source

This was a population-based retrospective cohort study conducted using the KNHIS-Senior database. The health insurance system in South Korea is a nationwide universal single-payer system managed by the KNHIS. The KNHIS-Senior database is provided by the KNHIS. It includes the data of 558,147 individuals randomly sampled from 10% of the approximate 5.5 million South Korean people aged ≥60 years, including information on age, sex, general health examinations, hospital and pharmacy visits, disease diagnoses and status, procedures, and prescribed medications (14, 15). All participants included in the NHIS-Senior database were followed-up until 2015 unless they were disqualified for health coverage reasons, such as death or emigration.

The KNHIS uses Korean Electronic Data Interchange (KEDI) and Korean Standard Classification of Diseases (KCD) codes, a system similar to the International Classification of Diseases (16). As the NHIS-Senior database comprises publicly accessible data, the Institutional Review Board of Asan Medical Center (University of Ulsan College of Medicine) instead of approved the waiver of reviewing this study (2020-1194). This study was conducted according to the ethical principles outlined in the Declaration of Helsinki. The requirement for obtaining informed consent was waived because anonymized and de-identified data were used for analyses.

Study Population

We selected the target population among those who were included in the NHIS-Senior database from 1 January 2002 to 31 December 2015 (n = 558,147). Initially, we applied the wash-out period of between 1 January and 31 December 2002 to reduce the potential risk of surveillance bias. The inclusion criterion was the presence of at least one NHIS record from 1 January 2003 to 15 December 2015 with the following conditions (n = 54,236): a KEDI code for cataract surgery and men aged ≥60 years in this period. Eligible subjects were classified into tamsulosin and non-tamsulosin groups according to tamsulosin use within 1 year before cataract surgery. Patients with the following characteristics were excluded: age < 60 years; procedures combined with vitrectomy or glaucoma surgery; and prior ocular procedures, including intraocular surgery or intravitreal injections within 1 year before cataract surgery; or retinal laser procedures within 5 years before cataract surgery. Patients who underwent simultaneous bilateral cataract surgery were excluded to avoid confounding.

The exposure of interest was tamsulosin use within 1 year before cataract surgery, except for tamsulosin medication on the same day. The tamsulosin group (n = 8,510) comprised participants with a KEDI code for cataract surgery and tamsulosin use within 1 year before cataract surgery. For each patient, cataract surgery was defined as the simultaneous claim of extracapsular or intracapsular extraction (KEDI code: S5111) or phacoemulsification (KEDI code: S5119) and primary intraocular lens implantation (KEDI code: S5117) on the same day (KEDI codes: S5111 + S5117 and S5119 + S5117). The non-tamsulosin group (n = 42,539) comprised participants with a KEDI code for cataract surgery but without tamsulosin use within 1 year before cataract surgery. Additionally, to investigate relationship between the use of alpha antagonist and surgical complications of cataract surgery in the Korean elderly population, we performed the comparison analysis between alpha antagonist (terazosin, alfuzosin, doxazosin, silodosin, and tamsulosin) group and non-alpha antagonist group.

Surgical Complication Events as the Outcome Measure

Surgical complications of cataract surgery, which is associated with tamsulosin use, included posterior capsule rupture (PCR), dropped lens fragments, retinal detachment, and suspected endophthalmitis (13). The use of anterior vitrectomy for intraoperative PCR was documented if the KEDI code S5122 was reported between the cataract surgery day and 2 weeks after cataract surgery. The use of total vitrectomy for intraoperative PCR, dropped lens fragments, or suspected endophthalmitis was

documented if the KEDI code S5121 was reported between 1 day and 2 weeks after cataract surgery. Similarly, the use of retinal detachment operation for retinal detachment was documented if the KEDI code S5130 was reported between 1 day and 2 weeks after cataract surgery. Even though KEDI codes S5122 and S5122 have been also indicated in other vitreoretinal diseases, we made an assumption that vitrectomy within 2 weeks after cataract surgery is highly related with secondary surgery which is associated with surgical complication from cataract surgery. Only the first event was included in cases of multiple complications.

Covariates

Demographics included age at the time of cataract surgery, sex, residence, and income level; the residential area was divided into metropolitan (Seoul and large cities) and provincial regions (small cities and rural areas) according to the administrative unit of Korea. Household income was categorized as below or above 20% of the income. Both systemic and ocular comorbidities were included as covariates and assessed at the time of cataract surgery. Systemic comorbidities included diabetes (KCD codes: E10–E14), hypertension (KCD codes: I10–I15), BPH (KCD code: N40), and vascular disease (KCD codes: I20–I25, I61, I63–I66, I67.2, I67.8, I69, I70, I73, and I74). Ocular comorbidities included glaucoma (KCD codes: H40 and H42), myopia (KCD codes: H52.1 and H44), eye trauma (KCD code: S05), diabetes mellitus (DM) with ophthalmic manifestations (KCD codes: E10.3, E11.3, E12.3, E13.3, and E14.3), severe cataract (KCD codes: H25.2 and H25.1), and age-related macular degeneration (KCD codes:

TABLE 1 | Baseline characteristics of subjects who underwent cataract surgery according to tamsulosin use in the Korean elderly population.

Variable	Non-tamsulosin group (n = 42,539)	Tamsulosin group (n = 8,510)	P-value	ASD[a]
Age (years)				0.3289
<70	8,348 (19.62)	760 (8.93)	<0.001	
70–80	26,242 (61.69)	5,550 (65.22)	<0.001	
80–90	7,541 (17.73)	2,103 (24.71)	<0.001	
≥90	408 (0.96)	97 (1.14)	<0.001	
Mean ± SD	74.44 ± 5.78	76.19 ± 5.34	<0.001	0.3160
Residence			0.047	0.0235
Metropolitan	16,798 (39.48)	3,458 (40.63)		
Provincial	25,741 (60.51)	5,052 (59.36)		
Income			0.140	0.0175
Below 20 percentiles	8,568 (20.14)	1,774 (20.84)		
Above 20 percentiles	33,971 (79.86)	6,736 (79.15)		
Diabetes	23,002 (54.07)	5,759 (67.67)	<0.001	0.2814
Hypertension	29,911 (70.31)	7,017 (82.46)	<0.001	0.2888
BPH	16,989 (39.94)	8,410 (98.82)	<0.001	1.6607
Vascular disease	24,931 (58.61)	6,536 (76.8)	<0.001	0.3967
Glaucoma	280 (0.66)	93 (1.09)	<0.001	0.0467
Myopia	6,470 (15.21)	1,513 (17.78)	<0.001	0.0693
Eye trauma	1,537 (3.61)	375 (4.41)	<0.001	0.0404
DM with ophthalmic manifestations	1,299 (3.05)	350 (4.11)	<0.001	0.0570
Severe cataract	14,513 (34.12)	3,186 (37.44)	<0.001	0.0693
Age-related macular degeneration	1,541 (3.62)	511 (6.0)	<0.001	0.1115
Event of complication	375 (0.88)	71 (0.83)	0.669	0.0051

SD, standard deviation; BPH, Benign prostatic hyperplasia; DM, diabetes mellitus.
Data are expressed as the mean ± SD, or n (%).
[a]*ASD of > 0.1 is considered meaningful imbalances.*

TABLE 2 | Odds Ratio of complication event of cataract surgery in the Korean elderly population with cataract surgery according to tamsulosin use.

Non-tamsulosin vs. tamsulosin	Odds ratio	95% CI	P-value
Crude (no adjustment)	0.946	0.733–1.220	0.669
Adjusted for age	0.950	0.735–1.228	0.694
Adjusted for age, income, residence, systemic, and ocular comorbidities[a]	0.997	0.749–1.325	0.981

CI, confidence interval.
[a]*Diabetes, hypertension, benign prostatic hyperplasia, vascular disease, glaucoma, myopia, eye trauma, DM with ophthalmic manifestations, severe cataract, and age-related macular degeneration.*

H35.30, H35.31, and H35.39). The presence of severe cataract was recognized as an indicator of poor vision because visual acuity data were not available (17, 18). Patients with diagnostic codes for brunescent cataract and morgagnian cataract were considered to have severe cataract (17, 18).

Statistical Analysis

A logistic regression model was used to evaluate the relationship between tamsulosin use and surgical complications of cataract surgery. We used two models of adjustment to account for potential confounding factors. Model 1 was adjusted for age (<70, 70–80, 80–90, and ≥90 years). Model 2 was further adjusted for income, residence, and systemic and ocular comorbidities. All statistical analyses were conducted using SAS software, version 9.4 (SAS Institute, Cary, NC, United States). Statistical significance was considered at a two-sided p-value < 0.05. The absolute standardized difference (ASD) was used to compare baseline characteristics. An ASD > 0.1 was considered to be clinically meaningful (19).

RESULTS

Table 1 summarizes baseline characteristics. The study cohort included 51,049 patients, 42,539 of whom were in the non-tamsulosin group whereas 8,510 in the tamsulosin group. The largest proportion of patients in both groups was 70–80 years old at the time of cataract surgery (61.69 and 65.22%). Compared to patients in the non-tamsulosin group, those in the tamsulosin group were slightly older (ASD = 0.3289) and had a significantly higher proportion of systemic diseases, such as diabetes, hypertension, and vascular disease (ASD = 0.2814, 0.2888, and 0.3967, respectively). In terms of ocular comorbidity, the tamsulosin group had a higher proportion of age-related macular degeneration (ASD = 0.1115). The development of surgical complications of cataract surgery did not differ significantly between tamsulosin and non-tamsulosin groups [71 (0.83%) cases in the tamsulosin group and 375 (0.88%) cases in the non-tamsulosin group; ASD = 0.0051]. In addition, the rate of surgical complications of cataract surgery did not differ between alpha antagonist (terazosin, alfuzosin, doxazosin, silodosin, and tamsulosin) group and non-alpha antagonist group (ASD = 0.0174; **Supplementary Table 1**).

Table 2 shows the odds ratio (OR) of surgical complications in the Korean elderly population with cataract surgery according to tamsulosin use. There was no significant difference in OR of surgical complication events between tamsulosin and non-tamsulosin groups in the unadjusted model [OR = 0.946; 95% confidence interval (CI): 0.733–1.220; $P = 0.669$]. Even after adjusting for age, OR of surgical complication events did not differ significantly between tamsulosin and non-tamsulosin groups (OR = 0.950; 95% CI: 0.735–1.228; $P = 0.694$). Additionally, tamsulosin use was not significantly associated with surgical complications of cataract surgery in the fully adjusted model accounting for age, income, residence, systemic and ocular comorbidities (OR = 0.997; 95% CI: 0.749–1.325; $P = 0.981$).

TABLE 3 | Effects of Calendar Year and Covariates on complication event of cataract surgery in the tamsulosin group.

Tamsulosin group (n = 8,510)	Odds ratio	95% CI	P-value
Calendar year (per additional year)	0.985	0.907–1.07	0.719
Patient–level effects (vs. age < 70)			
Age 70–80 yrs	1.556	0.538–4.502	0.415
Age 80–90 yrs	1.653	0.535–5.105	0.382
Age ≥ 90 yrs	3.805	0.661–21.891	0.135
Diabetes	1.270	0.739–2.185	0.387
Hypertension	0.869	0.459–1.645	0.667
Vascular disease	1.100	0.598–2.025	0.760
Glaucoma	<0.001		0.986
Myopia	0.419	0.180–0.971	0.043
Eye trauma	0.303	0.042–2.198	0.238
DM with ophthalmic manifestations	0.646	0.156–2.673	0.546
Severe cataract	1.115	0.689–1.803	0.659
Age–related macular degeneration	1.839	0.811–4.168	0.144

CI, confidence interval; DM, diabetes mellitus.

TABLE 4 | Effects of Calendar Year and Covariates on complication event of cataract surgery in the non-tamsulosin group.

Non-tamsulosin group (n = 42,539)	Odds ratio	95% CI	P-value
Calendar year (per additional year)	1.017	0.982–1.052	0.342
Patient-level effects (vs. age < 70)			
Age 70–80 yrs	0.783	0.592–1.036	0.087
Age 80–90 yrs	0.938	0.667–1.318	0.712
Age ≥ 90 yrs	1.186	0.474–2.969	0.715
Diabetes	0.941	0.755–1.172	0.586
Hypertension	0.862	0.68–1.094	0.222
Vascular disease	0.967	0.766–1.221	0.778
Glaucoma	0.842	0.208–3.402	0.809
Myopia	0.786	0.577–1.069	0.124
Eye trauma	1.849	1.215–2.814	0.004
DM with ophthalmic manifestations	0.750	0.368–1.526	0.427
Severe cataract	1.335	1.083–1.645	0.007
Age-related macular degeneration	1.022	0.589–1.771	0.939

CI, confidence interval; DM, diabetes mellitus.

Furthermore, the use of alpha antagonist was not associated with surgical complications of cataract surgery in the fully adjusted model accounting for age, income, residence, systemic and ocular comorbidities (OR = 0.813; 95% CI: 0.624–1.059; $P = 0.130$; **Supplementary Table 2**).

In the tamsulosin group, myopia was associated with decreased surgical complications of cataract surgery (OR = 0.419; 95% CI: 0.180–0.971; $P = 0.043$; **Table 3**). Both eye trauma (OR = 1.849; 95% CI: 1.215–2.814; $P = 0.004$) and severe cataract (OR = 1.335; 95% CI: 1.083–1.645; $P = 0.007$) were associated with increased surgical complications of cataract surgery in the non-tamsulosin group (**Table 4**).

DISCUSSION

This nationwide population-based cohort study demonstrated that surgical complications of cataract surgery were not significantly affected by tamsulosin use, although there was a trend of increasing rate of surgical complications of cataract surgery after adjusting for demographics and systemic and ocular comorbidities. In patients with tamsulosin use within 1 year preceding cataract surgery, no specific risk factor was associated with the surgical complication event during cataract surgery.

IFIS and other surgical complications have been reported since the worldwide use of tamsulosin began for acute urinary retention in the old age group, which is the most common age group undergoing cataract surgery. Previous studies on surgical complications of cataract surgery in patients taking tamsulosin demonstrated that IFIS occurred in 2% of cataract surgeries, and adjunctive measures for pupil dilation were ineffective compared to non-tamsulosin users (5). Cataract surgical complications, including retinal detachment, loss of lens fragment, and endophthalmitis, were significantly more prevalent (OR = 2.33; CI: 1.22–4.43) in tamsulosin users (2). A previous study demonstrated that doxazosin (an alpha blocker for BPH) was related to higher risks of PCR and vitreous loss (OR = 1.51; CI: 1.09–2.07; adjusted model) (20). However, unlike those previous results, a recent large population study from Canada and our large cohort study showed that tamsulosin use was not associated with increased cataract surgical complications (13).

In a recently published population-based study, Campbell et al. showed that the risk of surgical complications of cataract surgery, such as PCR, dropped lens fragments, retinal detachment, and suspected endophthalmitis, significantly decreased with time from 2003 to 2013 in patients with tamsulosin use within 1 year preceding cataract surgery (OR = 0.95/year; 95% CI: 0.91–0.99/year; $P = 0.010$) (13). The risk also decreased in patients without tamsulosin use within 1 year preceding cataract surgery. However, those results did not reflect direct comparison results between patients with and without tamsulosin use. In our study, there was no significant difference in the risk of surgical complications of cataract surgery between tamsulosin and non-tamsulosin groups in the unadjusted model (OR = 0.946; 95% CI: 0.733–1.220; $P = 0.669$). In addition, tamsulosin use was not significantly associated with surgical complications of cataract surgery in the fully adjusted model accounting for age, income, residence, and systemic and ocular comorbidities (OR = 0.997; 95% CI: 0.749–1.325; $P = 0.981$). Nevertheless, surgeons should devote efforts to avoid adverse surgical events, including PCR and dropped lens fragments. Considering that IFIS increases surgical difficulties, which might lead to PCR and vitreous prolapse, we included partial anterior vitrectomy on the day of cataract surgery to manage such surgical complications, which can be considered to be our novelty.

After the introduction of IFIS caused by tamsulosin and cognizance of significant risks posed by tamsulosin use, several surgical techniques were introduced worldwide (10, 11, 21). Although our study lacks the usage of adjunctive measures, such as drugs or device during cataract surgery, we suggest that comparable results between tamsulosin and non-tamsulosin groups can be attributed to efforts to respond to the risk posed by tamsulosin use using sophisticated surgical instruments or an intraoperative epinephrine injection into the anterior chamber. Particularly, viscoadaptive ophthalmic viscosurgical devices, fluidic parameter optimization, mechanical pupil expansion devices, and intensive pharmacologic pupil dilation can be applied to increase the efficiency and safety of cataract surgery in patients with tamsulosin use (22–24). Recently introduced femtosecond laser-assisted cataract surgery (FLACS) can be helpful for safe cataract surgery in eyes with a small pupil due to tamsulosin use (25, 26). Conrad-Hengerer et al. demonstrated that surgically dilating small pupils before femtosecond laser using intracameral epinephrine, viscomydriasis, and pupil expander can assist safe anterior capsulotomy and nuclear fragmentation (25). Although our cohort data did not include FLACS as a variable in cataract surgery, such a cutting-edge technique can improve surgical outcomes. Moreover, surgeons' recognition of the perioperative risk during cataract surgery in patients with tamsulosin use could decrease the complication rate (9).

Our study included various variables related to cataract surgery: age, socioeconomic state, systemic disease (DM, hypertension, BPH, and vascular disease), and ocular comorbidities (glaucoma, myopia, eye trauma, DM with ophthalmic manifestations, severe cataract, and age-related macular degeneration). The presence of ocular trauma, severity of cataract, presence of myopia, and DM increase surgical complications of cataract surgery. Among them, ocular trauma is associated with anatomical deformities of the iris, zonules, and lens, which can subsequently lead to surgical difficulties during cataract surgery and increased surgical complications (20, 27). Similarly, severe cataract could make phacoemulsification difficult and be more vulnerable to surgical complications (20). Lacking of information in cataract grading with slit lamp examination which is based on the LOCS grading system is the limitation of our study. Thus, we hypothesized that including the severe cataract as covariates can be meaningful after defining the severe cataract using the KCD codes (H25.1 and H25.2) in order to investigate the effect of severity of cataract on the relationship between tamsulosin use and surgical complications of cataract surgery. However, in the tamsulosin group, no significant risk factor was associated with surgical complications of cataract surgery. Protective association was noted between surgical complications and myopia in the tamsulosin group. In contrast, in the non-tamsulosin group, both eye trauma and severe cataract were associated with surgical complications of cataract surgery. We assumed that surgeons' recognition of possible intraoperative complications related to tamsulosin use can lead to use careful surgical maneuver and supportive techniques, which eventually decrease surgical complications in patients with ocular trauma and severe cataract.

Our study had some limitations. First, this study was mainly limited by its observational nature. Second, as this study was based on data from a medical insurance claims database, identification of patients with cataract surgery and diagnostic accuracy of systemic and ocular comorbidities

might be inaccurate compared to information obtained from medical charts. Moreover, the NHIS-Senior database cannot provide information on cataract grading, objective visual acuity, axial length, presence of pseudoexfoliation syndrome, or postoperative inflammation grade. In addition, there was a lack of availability of certain covariates including surgeons' experience, metabolic profiles, body mass index, alcohol intake, smoking status, and physical activity, thereby proposing the need for further studies including various covariates. Third, among various types of alpha antagonist, including tamsulosin, terazosin, silodosin, doxazosin, etc., only the use of tamsulosin was included in the current study because the tamsulosin is the most commonly prescribed medication among them for treating BPH in South Korea (**Supplementary Table 1**). When interpreting the results, clinicians should be in cautious since it may lead to a bias and overall results cannot be generalized to patients taking medication other than tamsulosin. Nevertheless, the rate of surgical complications of cataract surgery did not differ between alpha antagonist group and non-alpha antagonist group (**Supplementary Tables 1**–**4**). Fourth, the primary outcome of surgical complications was defined as secondary vitrectomy surgery within 14 postoperative days. The NHIS-Senior database lacks information on cataract surgery-related minor anatomical and/or functional complications, such as iris prolapse, iris atrophy, and pupil abnormality. Therefore, our study had a limitation of overlooking cataract surgery-related minor anatomical and/or functional complications, which are not indications for the secondary operation. Therefore, overall surgical complication rates (0.87%) might have been underestimated. Delayed onset complications, such as delayed endophthalmitis and intraocular lens dislocation, were possibly excluded. Additionally, glaucoma filtering surgery due to increased intraocular pressure or intraocular lens sulcus insertion due to posterior capsular rupture might be possibly excluded. We counted the first adverse event after cataract surgery, so additional surgical procedures for complications might have been missed. Finally, we focused only on South Korean residents. Therefore, observed findings cannot be generalized to other ethnic groups.

Within these limitations, this is the first report on the relationship between oral tamsulosin use and surgical complications of cataract surgery in the elderly South Korean patients using a nationwide, general population-based database. Moreover, this study used a large sample size of the NHIS-Senior database. Selection bias was relatively low because the entire Korean population was enrolled in the same insurance system.

In summary, despite concerns regarding perioperative and postoperative complications in cataract surgery related to tamsulosin use, our study demonstrated no statistically significant difference in surgical complication events of cataract surgery between tamsulosin and non-tamsulosin groups. Better techniques to manage a difficult cataract surgery and surgeons' cognizance of tamsulosin use could improve surgical outcomes.

DATA AVAILABILITY STATEMENT

The datasets presented in this article are not readily available. The data that support the findings of this study are available from NHIS, but restrictions apply to the availability of these data, which were used under license for the current study, and so are not publicly available. Data are however available from the authors upon reasonable request and with permission of NHIS. Requests to access the datasets should be directed to yhun777@gmail.com.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Institutional Review Board of Asan Medical Center and University of Ulsan College of Medicine approved the waiver of reviewing this study (2020-1194). Written informed consent for participation was not required for this study in accordance with the national legislation and the institutional requirements.

AUTHOR CONTRIBUTIONS

Conceptualization, formal analysis, investigation, writing—original draft preparation, and writing—review and editing: JKw, JH, SM, SN, JKi, HT, and HL. Methodology and data curation: JKw, JKi, HT, and HL. Supervision: JKi and HL. Project ad-ministration: JKi, HT, and HL. Funding acquisition: HL. All authors have read and agreed to the published version of the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

This research was supported by the Korea Medical Device Development Fund grant funded by the Korea government (the Ministry of Science and ICT, the Ministry of Trade, Industry and Energy, the Ministry of Health & Welfare, the Ministry of Food and Drug Safety) (Project Number: 9991006821, KMDF_PR_20200901_0148), by Korean Fund for Regenerative Medicine funded by Ministry of Science and ICT, and Ministry of Health and Welfare (21C0723L1-11, Republic of Korea), and by a grant from the Asan Institute for Life Sciences, Asan Medical Center, Seoul, Korea (2021IP0061-2, 2022IP0019-1). The funding agencies had no role in the design or conduct of this study; the collection, management, analysis, or interpretation of the data; preparation, review, or approval of the manuscript; or in the decision to submit the manuscript for publication.

ACKNOWLEDGMENTS

We would like to thank Editage (www.editage.co.kr) for their English language editing assistance.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmed.2022.882131/full#supplementary-material

REFERENCES

1. Dunn CJ, Matheson A, Faulds DM. Tamsulosin: a review of its pharmacology and therapeutic efficacy in the management of lower urinary tract symptoms. *Drugs Aging.* (2002) 19:135–61. doi: 10.2165/00002512-200219020-00004
2. Bell CM, Hatch WV, Fischer HD, Cernat G, Paterson JM, Gruneir A, et al. Association between tamsulosin and serious ophthalmic adverse events in older men following cataract surgery. *JAMA*. (2009) 301:1991–6. doi: 10.1001/jama.2009.683
3. Cheung CM, Awan MA, Sandramouli S. Prevalence and clinical findings of tamsulosin-associated intraoperative floppy-iris syndrome. *J Cataract Refract Surg.* (2006) 32:1336–9. doi: 10.1016/j.jcrs.2006.03.034
4. Christou CD, Tsinopoulos I, Ziakas N, Tzamalis A. Intraoperative floppy iris syndrome: updated perspectives. *Clin Ophthalmol.* (2020) 14:463–71. doi: 10.2147/OPTH.S221094
5. Chang DF, Campbell JR. Intraoperative floppy iris syndrome associated with tamsulosin. *J Cataract Refract Surg.* (2005) 31:664–73. doi: 10.1016/j.jcrs.2005.02.027
6. Chang DF, Braga-Mele R, Mamalis N, Masket S, Miller KM, Nichamin LD, et al. ASCRS white paper: clinical review of intraoperative floppy-iris syndrome. *J Cataract Refract Surg.* (2008) 34:2153–62. doi: 10.1016/j.jcrs.2008.08.031
7. Shtein RM, Hussain MT, Cooney TM, Elner VM, Hood CT. Effect of tamsulosin on iris vasculature and morphology. *J Cataract Refract Surg.* (2014) 40:793–8. doi: 10.1016/j.jcrs.2013.10.031
8. Haridas A, Syrimi M, Al-Ahmar B, Hingorani M. Intraoperative floppy iris syndrome (IFIS) in patients receiving tamsulosin or doxazosin-a UK-based comparison of incidence and complication rates. *Graefes Arch Clin Exp Ophthalmol.* (2013) 251:1541–5. doi: 10.1007/s00417-013-2260-4
9. Chang DF, Osher RH, Wang L, Koch DD. Prospective multicenter evaluation of cataract surgery in patients taking tamsulosin (Flomax). *Ophthalmology.* (2007) 114:957–64. doi: 10.1016/j.ophtha.2007.01.011
10. Gurbaxani A, Packard R. Intracameral phenylephrine to prevent floppy iris syndrome during cataract surgery in patients on tamsulosin. *Eye.* (2007) 21:331–2. doi: 10.1038/sj.eye.6702172
11. Nguyen DQ, Sebastian RT, Kyle G. Surgeon's experiences of the intraoperative floppy iris syndrome in the United Kingdom. *Eye.* (2007) 21:443–4. doi: 10.1038/sj.eye.6702616
12. Flach AJ. Intraoperative floppy iris syndrome: pathophysiology, prevention, and treatment. *Trans Am Ophthalmol Soc.* (2009) 107:234–9.
13. Campbell RJ, El-Defrawy SR, Gill SS, Whitehead M, Campbell ELP, Hooper PL, et al. Evolution in the risk of cataract surgical complications among patients exposed to tamsulosin: a population-based study. *Ophthalmology.* (2019) 126:490–6. doi: 10.1016/j.ophtha.2018.11.028
14. Seong SC, Kim YY, Park SK, Khang YH, Kim HC, Park JH, et al. Cohort profile: the National Health Insurance Service-National Health Screening Cohort (NHIS-HEALS) in Korea. *BMJ Open.* (2017) 7:e016640. doi: 10.1136/bmjopen-2017-016640
15. Lee J, Lee JS, Park SH, Shin SA, Kim K. Cohort profile: the National Health Insurance Service-National Sample Cohort (NHIS-NSC), South Korea. *Int J Epidemiol.* (2017) 46:e15. doi: 10.1093/ije/dyv319
16. Ryu SY, Kim J, Hong JH, Chung EJ. Incidence and characteristics of cataract surgery in South Korea from 2011 to 2015: a nationwide population-based study. *Clin Exp Ophthalmol.* (2020) 48:319–27. doi: 10.1111/ceo.13705
17. Tseng VL, Chlebowski RT Yu F, Cauley JA Li W, Thomas F, et al. Association of cataract surgery with mortality in older women: findings from the women's health initiative. *JAMA Ophthalmol.* (2018) 136:3–10. doi: 10.1001/jamaophthalmol.2017.4512
18. Tseng VL Yu F, Lum F, Coleman AL. Cataract surgery and mortality in the United States medicare population. *Ophthalmology.* (2016) 123:1019–26. doi: 10.1016/j.ophtha.2015.12.033
19. Mamdani M, Sykora K, Li P, Normand SL, Streiner DL, Austin PC, et al. Reader's guide to critical appraisal of cohort studies: 2. Assessing potential for confounding. *BMJ.* (2005) 330:960–2. doi: 10.1136/bmj.330.7497.960
20. Narendran N, Jaycock P, Johnston RL, Taylor H, Adams M, Tole DM, et al. The cataract national dataset electronic multicentre audit of 55,567 operations: risk stratification for posterior capsule rupture and vitreous loss. *Eye.* (2009) 23:31–7. doi: 10.1038/sj.eye.6703049
21. Shugar JK. Use of epinephrine for IFIS prophylaxis. *J Cataract Refract Surg.* (2006) 32:1074–5. doi: 10.1016/j.jcrs.2006.01.110
22. Nderitu P, Ursell P. Iris hooks versus a pupil expansion ring: operating times, complications, and visual acuity outcomes in small pupil cases. *J Cataract Refract Surg.* (2019) 45:167–73. doi: 10.1016/j.jcrs.2018.08.038
23. Balal S, Jbari AS, Nitiapapand R, Cook E, Akhtar W, Din N, et al. Management and outcomes of the small pupil in cataract surgery: iris hooks, malyugin ring or phenylephrine? *Eye (Lond).* (2021) 35:2714–8. doi: 10.1038/s41433-020-01277-0
24. Hashemi H, Seyedian MA, Mohammadpour M. Small pupil and cataract surgery. *Curr Opin Ophthalmol.* (2015) 26:3–9. doi: 10.1097/ICU.0000000000000116
25. Conrad-Hengerer I, Hengerer FH, Schultz T, Dick HB. Femtosecond laser-assisted cataract surgery in eyes with a small pupil. *J Cataract Refract Surg.* (2013) 39:1314–20. doi: 10.1016/j.jcrs.2013.05.034
26. Jun JH, Bang SP, Yoo YS. Joo CK. Efficacy of 0015% intracameral epinephrine for significant miosis induced by photodisruption during femtosecond laser-assisted cataract surgery. *Medicine.* (2018) 97:e11693. doi: 10.1097/MD.0000000000011693
27. Tuft SJ, Minassian D, Sullivan P. Risk factors for retinal detachment after cataract surgery: a case-control study. *Ophthalmology.* (2006) 113:650–6. doi: 10.1016/j.ophtha.2006.01.001

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers | Frontiers in Medicine

ORIGINAL RESEARCH
published: 23 June 2022
doi: 10.3389/fmed.2022.912214

Smartphone-Acquired Anterior Segment Images for Deep Learning Prediction of Anterior Chamber Depth: A Proof-of-Concept Study

*Chaoxu Qian[1,2], Yixing Jiang[3], Zhi Da Soh[1,4], Ganesan Sakthi Selvam[3], Shuyuan Xiao[2], Yih-Chung Tham[1,4,5], Xinxing Xu[3], Yong Liu[3], Jun Li[6], Hua Zhong[2] and Ching-Yu Cheng[1,4,5]**

[1] Singapore Eye Research Institute, Singapore National Eye Centre, Singapore, Singapore, [2] Department of Ophthalmology, The First Affiliated Hospital of Kunming Medical University, Kunming, China, [3] Institute of High Performance Computing, Agency for Science, Technology and Research (A*Star), Singapore, Singapore, [4] Department of Ophthalmology, Yong Loo Lin School of Medicine, National University of Singapore, Singapore, Singapore, [5] Ophthalmology and Visual Sciences Academic Clinical Program (Eye ACP), Duke-NUS Medical School, Singapore, Singapore, [6] Department of Ophthalmology, The Second People's Hospital of Yunnan Province, Kunming, China

OPEN ACCESS

Edited by:
Tyler Hyungtaek Rim,
Duke-NUS Medical School, Singapore

Reviewed by:
Tae Keun Yoo,
Korea Air Force Academy,
South Korea
Shuning Li,
Capital Medical University, China

***Correspondence:**
Ching-Yu Cheng
chingyu.cheng@duke-nus.edu.sg

Specialty section:
This article was submitted to Ophthalmology, a section of the journal Frontiers in Medicine

Received: 04 April 2022
Accepted: 09 May 2022
Published: 23 June 2022

Citation:
Qian C, Jiang Y, Soh ZD, Sakthi Selvam G, Xiao S, Tham Y-C, Xu X, Liu Y, Li J, Zhong H and Cheng C-Y (2022) Smartphone-Acquired Anterior Segment Images for Deep Learning Prediction of Anterior Chamber Depth: A Proof-of-Concept Study. Front. Med. 9:912214. doi: 10.3389/fmed.2022.912214

Purpose: To develop a deep learning (DL) algorithm for predicting anterior chamber depth (ACD) from smartphone-acquired anterior segment photographs.

Methods: For algorithm development, we included 4,157 eyes from 2,084 Chinese primary school students (aged 11–15 years) from Mojiang Myopia Progression Study (MMPS). All participants had with ACD measurement measured with Lenstar (LS 900) and anterior segment photographs acquired from a smartphone (iPhone Xs), which was mounted on slit lamp and under diffuses lighting. The anterior segment photographs were randomly selected by person into training (80%, no. of eyes = 3,326) and testing (20%, no. of eyes = 831) dataset. We excluded participants with intraocular surgery history or pronounced corneal haze. A convolutional neural network was developed to predict ACD based on these anterior segment photographs. To determine the accuracy of our algorithm, we measured the mean absolute error (MAE) and coefficient of determination (R^2) were evaluated. Bland Altman plot was used to illustrate the agreement between DL-predicted and measured ACD values.

Results: In the test set of 831 eyes, the mean measured ACD was 3.06 ± 0.25 mm, and the mean DL-predicted ACD was 3.10 ± 0.20 mm. The MAE was 0.16 ± 0.13 mm, and R^2 was 0.40 between the predicted and measured ACD. The overall mean difference was −0.04 ± 0.20 mm, with 95% limits of agreement ranging between −0.43 and 0.34 mm. The generated saliency maps showed that the algorithm mainly utilized central corneal region (i.e., the site where ACD is clinically measured typically) in making its prediction, providing further plausibility to the algorithm's prediction.

Conclusions: We developed a DL algorithm to estimate ACD based on smartphone-acquired anterior segment photographs. Upon further validation, our algorithm may be further refined for use as a ACD screening tool in rural localities where means of assessing ocular biometry is not readily available. This is particularly important in China where the risk of primary angle closure disease is high and often undetected.

Keywords: primary angle-closure glaucoma, glaucoma, anterior chamber depth, smartphone, deep learning

INTRODUCTION

Primary angle-closure glaucoma (PACG) is a significant cause of vision loss in Asia. It was estimated that the number of people aged 40–80 years with PACG worldwide was 23.36 million in 2020, of which Asia accounted for 76.8% of cases (1). Bilateral blindness affected 5.3 million people with PACG in 2020, the majority of whom were from Asian regions (2). Thus, screening for people with high risks of PACG is important to provide timely interventions, particularly in Asian countries (3).

Anterior chamber depth (ACD), the distance from corneal endothelium to the anterior crystalline lens capsule, is an important biometric dimension to assess the risk of angle closure development. A population-based study reported that ACD was a significant risk factor for angle closure amongst Mongolia and Chinese (4). Another population-based longitudinal study in China demonstrated that shallow ACD was independently associated with angle closure development over a 6-year period (5). Anterior chamber depth alone may provide a simple and effective way to distinguish eyes with angle closure from those with open angles (6), and has been suggested as a quick screening tool for detecting primary angle closure disease (PACD) (7, 8).

Currently, the methods used for ACD measurement include A-Scan ultrasound, slit-lamp biomicroscopy, non-contact partial coherence interferometry [e.g., IOLMaster (Carl Zeiss AG, Oberkochen, Germany), Lenstar (Lenstar LS 900®, Haag-Streit AG, Switzerland), Pentacam (Oculus System, Wetzlar, Germany)], and anterior segment optical coherence tomography (AS-OCT) (6, 9). However, the need for technical expertise, along with the cost and lack of portability, limit their usage in community screening (8, 9). The advent of artificial intelligence has made tremendous breakthroughs in ophthalmic imaging and shown great capabilities in disease diagnosis and screening (10). In recent times, Chen et al. developed a machine learning algorithm to predict ACD from images captured by a smartphone mounted with a portable slit lamp ($n = 66$) (11). In brief, the portable slit lamp was placed in front of the eye parallel to the cornea. The slit beam focused on the mid-peripheral iris surface, not too center nor too peripheral. Multiple images were captured in ∼1 mm steps from nasal to temporal. Although their algorithm-predicted ACD showed moderate correlation with the measured ACD measurements, the need for manual maneuvering across the cornea with a 1 mm slit was subjective and time-consuming.

The availability of portable smartphones with cameras has become a tool for ophthalmologists in clinics (11–13). Using smartphones to take anterior segment photographs provide good reproducibility (12), and could provide clinicians with a simple and quick way to obtain anterior segment photographs for evaluation in rural or less-resourced areas.

In the present study, we aimed to develop and validate a DL algorithm for quantitative prediction of ACD from anterior segment photographs that were captured by a smartphone. This approach may provide clinicians with a mean to obtain ACD measurements in settings where biometers and advanced imaging tools are not readily available.

METHODS

Study Population

The Mojiang Myopia Progression Study (MMPS) is a longitudinal school-based study that evaluates the onset and progression of myopia in school-aged children in rural China. Details of the methodology have been described previously (14–17). In brief, this study was conducted in Mojiang, a small country in Yunnan Province in the Southwestern part of China. A total of 2,432 elementary students (response rate 90.2%) and 2,346 middle school students (response rate 93.5%) were enrolled in the MMPS. The baseline examinations were conducted in 2016 and the MMPS participants were followed annually. The data used for the present study were from 2,195 elementary students participated in the 5-year follow up visit in 2020 (response rate 99.1%).

All study procedures were performed in accordance with the tenets of the Declaration of Helsinki. Ethics approval was obtained from the institutional review board of Kunming Medical University. Written informed consent was obtained from at least one parent or legal guardian of each participant.

Anterior Chamber Depth and Ocular Biometry Measurements

Anterior chamber depth (ACD), from corneal endothelium to lens surface, was obtained using the Lenstar LS 900 (Lenstar LS 900®, Haag-Streit AG, Switzerland), a non-invasive, non-contact optical low-coherence reflectometry biometer. Other ocular biometry measurements including central corneal thickness (CCT), lens thickness (LT), axial length (AL), keratometry readings of flattest and steepest meridian (K1 and K2) were also recorded simultaneously. Refractive error was measured before and after cycloplegia using an autorefractor (RM-8000, Topcon Co., Tokyo, Japan). **Supplementary Figure 1** shows the diagram of the human eye and the details of ocular biometry measurements.

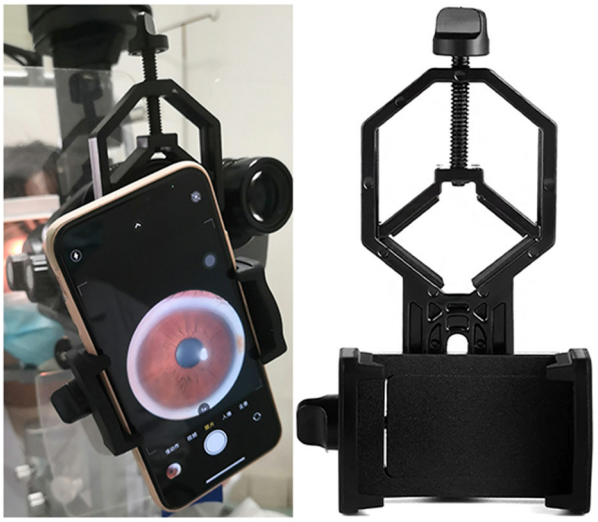

FIGURE 1 | Smartphone mounted on slit lamp in use. Anterior segment photographs were captured on study eyes using a smartphone (iPhone Xs, Apple Inc, CA, USA) attached to a slit lamp. The smartphone was fixed on the eyepiece with an adapter (Celestron 81035, Celestron Acquisition LLC, CA, USA), making the camara lens in line with the eyepiece. We used the default mode of iPhone camara with a minimal magnification (1 X) to take photographs. A Bluetooth trigger for a one-tap image capture was fixed on the joystick making the procedure of taking photographs quickly and stably. Diffuse illumination of slit-lamp was used at 45-degree angle, with magnification set at 16X.

Anterior Segment Photographs Acquisition

Anterior segment photographs were captured on study eyes before cycloplegia using a smartphone (iPhone Xs, Apple Inc, CA, USA) attached to a slit lamp (**Figure 1**). The smartphone was fixed on the eyepiece with an adapter (Celestron 81035, Celestron Acquisition LLC, CA, USA), making the camara lens in line with the eyepiece. In this study, we captured the anterior segment photographs with the light source from the slit lamp always to the left of the pupil. We used the default mode of iPhone camara with a minimal magnification (1 X) to take photographs. A Bluetooth trigger for a one-tap image capture was fixed on the joystick making the procedure of taking photographs quickly and stably. Diffuse illumination of slit-lamp was used at 45-degree angle, with magnification set at 16 X.

Inclusion and Exclusion Criteria

The MMPS participants who had both anterior segment photographs and ACD measurements were included in this study. Participants who had pronounced opacities of the central cornea, and/or history of intraocular surgery were excluded.

Development of the Deep Learning Algorithm

Neural Network Architecture

Residual Network 34 (ResNet-34) architecture was adopted in this project (18). Several modifications were introduced to ResNet-34 to finetune the model for ACD prediction. Firstly, the fully connected layer was replaced by a linear layer with an output channel of one for the regression task. No activation function was added after the linear layer. Then, the first convolutional layer was changed to one which takes in 4-channel images. Finally, the adopted ResNet-34 ended with one fully connected layer.

Data Preprocessing and Augmentation

Preprocessing of images was done to clean image data for model input (19). It decreases model training time and increases the model's inference speed. This process will not significantly affect the model's performance. OpenCV was used for image pre-processing in the present study. The first step for image pre-processing was cropping images to regions of interest (ROI). The original color photographs were first converted into grayscale ones and binarized using simple thresholding. Then the bounding rectangle of foreground was identified and used as ROI for the original color photographs. The images were resized to (200,200,3) after cropping, and the brightness was increased by 20%. Histogram equalization was then used to balance the RGB values of an image to enhance the contrast of images, followed by a change of color space from 3-channel to 4-channel. The last step was image normalization which scales the pixel values to zero means and unit variances. Consequently, the final input to the neural network is of size (200, 200, 4).

Image augmentation is a process to create new training examples out of the existing training data (20). This helps to adjust the current training data to generalize to other situations which allows the model to learn from a wider array of situations. To mitigate overfitting, data augmentation was used during training stage. Specifically, random rotation from −35 to 35 degrees, randomly horizontal flip with a probability of 0.5 and vertical flip with a probability of 0.1 were used.

Training Details and Evaluation Metrics

The dataset was randomly split into a training set and a test set with a ratio of 4:1. The batch size used is 16. Random shuffling was used for the training set. Pytorch (21), an open-source software library for DL, was used in the training and evaluation of the models. The model was trained on TITAN XP powered GPU server. Transfer learning was adopted, the ResNet-34 was loaded with a pretrained model which was trained on the ImageNet dataset which consists of 1,000 classes of objects. The modifications discussed in the architecture part were applied after loading the pre-trained weights. Adam optimizer with a learning rate of 4e-4 was used to train the model for 200 epochs (22). Mean absolute error (MAE) was used as the loss function.

Heat Map Generation

In order to further interpreting how the DL algorithm worked, we generated heat maps using Gradient-weighted Class Activation Mapping (Grad-CAM) algorithm (23, 24). Highlighting the important regions in hotter color, heat maps help visualization of the regions that the algorithm uses for its prediction. After normalizing the heat maps for individual images to [0, 1], we obtained the averaged heat maps across all images for an aggregated visualization.

TABLE 1 | Demographic and clinical characteristics of the eyes in this study.

	Training samples	Testing samples	Total
Number of individuals	1,667	417	2,084
Numbers of eyes	3,326	831	4,157
Age (years)	11.6 ± 0.53	11.7 ± 0.67	11.6 ± 0.56
Gender, % Female	46%	46%	46%
Anterior chamber depth, mm	3.05 ± 0.25	3.06 ± 0.25	3.06 ± 0.26
Central corneal thickness, mm	536.41 ± 30.91	536.61 ± 33.09	536.44 ± 31.34
Lens thickness, mm	3.45 ± 0.19	3.47 ± 0.19	3.45 ± 0.19
Axial length, mm	23.48 ± 0.93	23.57 ± 1.01	23.49 ± 0.94
Keratometry readings of flattest meridian	42.79 ± 1.41	42.70 ± 1.46	42.77 ± 1.42
Keratometry readings of steepest meridian	43.88 ±1.56	43.77 ± 1.59	43.86 ± 1.57

Data presented as mean ± SD.

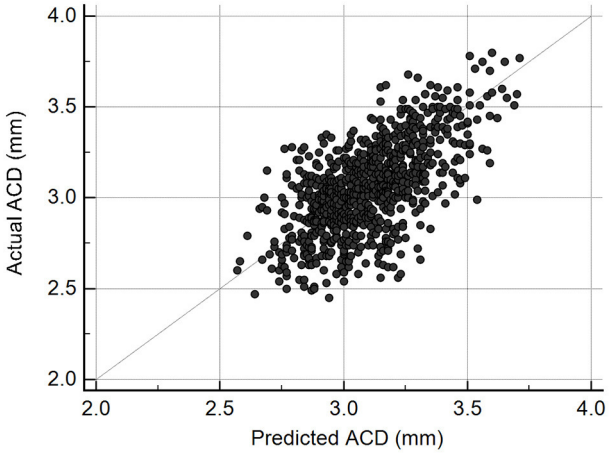

FIGURE 2 | Scatterplot illustrating the relationship between deep learning-predicted and actual anterior chamber depth (ACD) measurements from Lenstar ($n = 831$, $r = 0.63$, $P < 0.001$).

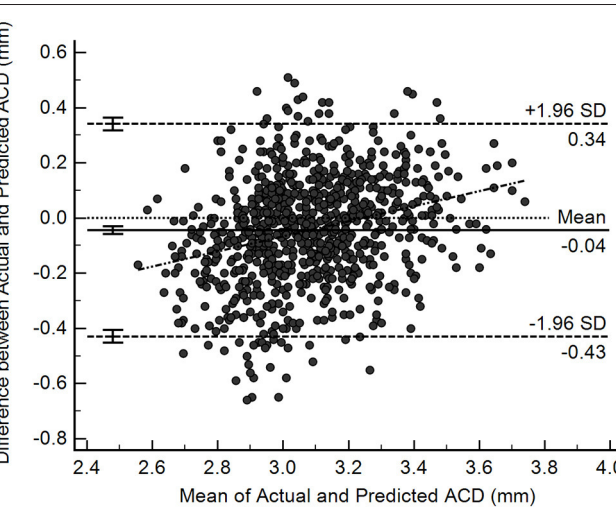

FIGURE 3 | Bland-Altman plots illustrating agreement between deep learning-predicted and actual anterior chamber depth (ACD) measurements from Lenstar ($n = 831$).

Statistical Analysis

The Pearson's correlation coefficient (r) was used to evaluate the correlation between predicted and measured ACD values. The MAE and coefficient of determination (R^2) were used to evaluate the accuracy of prediction from the algorithm. Bland-Altman plot was used to illustrate the agreement between predicted and measured ACD values.

RESULTS

Of the 4,390 eyes of the MMPS 2,195 participants, we excluded 233 eyes (118 without ACD values, 115 eyes without anterior segment photographs or with poor image quality), and 4,157 eyes from 2,084 participants with both ACD values and anterior segment photographs were used to build our DL algorithm. The anterior segment photographs from these eyes were randomly distributed into a training set (3,326 photographs) and test set (831 photographs) based on a 4:1 ratio at individual level. The demographic and clinical characteristics of the eyes are presented in **Table 1**. The mean actual ACD in the training and test set were 3.05 ± 0.25 mm and 3.06 ± 0.25 mm, respectively.

The scatter plot presented in **Figure 2** shows there was a good correlation ($r = 0.63$, $P < 0.001$) between ACD predictions from the DL algorithm and actual Lenstar measurements in the test set of 831 eyes. The mean difference was −0.04 ± 0.20 mm, and MAE was 0.16 ± 0.13 mm. If we set measurements less than 2.80 mm as shallow ACD (25, 26), the MAE of eyes with shallow ACD was 0.26 ± 0.16 mm ($n = 134$), and the MAE of eyes with ACD ≥ 2.80 mm was 0.14 ± 0.11 mm ($n = 697$).

Figure 3 shows the Bland-Altman plot evaluation of the agreement between predicted and measured ACD in the test samples ($n = 831$). The overall mean difference was −0.04 ± 0.20 mm, with 95% limits of agreement ranging between −0.43 and 0.34 mm. Nevertheless, there was a mild but statistically significant proportional bias ($r = 0.27$, $P < 0.001$), suggesting that at smaller range of ACD the predictions tend to give higher values than measured ACD, while at larger range of ACD, the predictions trend to give lower values than measured ACD.

Figure 4 shows examples of smartphone-obtained anterior segment photograph, the corresponding heatmap of the present neural network and the averaged heatmap crossed all images ($n = 831$). The averaged heatmap shows that the algorithm utilized regions of the central cornea in making its prediction.

DISCUSSION

In this study, we developed a novel DL algorithm to quantitatively predict ACD through smartphone-acquired anterior segment photographs. The predicted ACD showed

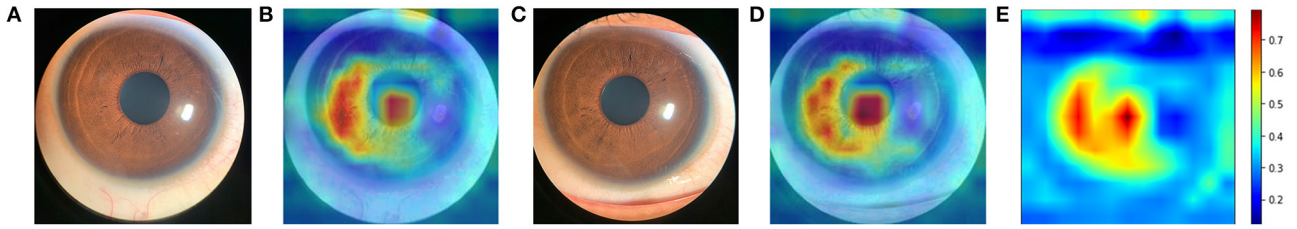

FIGURE 4 | Averaged heatmap shows the regions of the anterior segment photograph that were most important for the deep learning algorithm predictions in test set. Hotter colors (reds) indicate higher activity while cooler colors (blues) represent lower activity. **(A)** an example of original anterior segment photograph (right eye) obtained by iPhone Xs; **(B)** the heatmap of the corresponding photograph; **(C)** an example of original anterior segment photograph (left eye); **(D)** the heatmap of the corresponding photograph; **(E)** averaged heat map crossed all images ($n = 831$).

good agreement with the measured ACD values. To our knowledge, this may be the first investigation to demonstrate that a DL algorithm can potentially predict the ACD through smartphone-acquired anterior segment photographs.

Our novel DL algorithm successfully predicted ACD through smartphone captured anterior segment photographs. The MAE of the predictions in test set was only 0.16 ± 0.13 mm (RMSE = 0.20 mm). The MAE of eyes with shallow ACD was bigger than the MAE of eyes with ACD ≥ 2.80 mm. That may be because of the number of eyes with ACD < 2.80 mm in training set is only 498, much less than the number of eyes with ACD ≥ 2.80 mm ($n = 2{,}828$). The average difference of measured and predicated ACD was −0.04 ± 0.20 mm ($P = 0.000$). However, this difference was significant statistically but not clinically as the difference was small. We captured two photographs for 50 eyes for assessing repeatability and reproducibility. For group one, the MAE of predicted ACD was 0.14 ± 0.09 mm, with 95% limits of agreement ranging between −0.36 and 0.16 mm, repeatability coefficient was 0.33 mm. For group two, the MAE was 0.14 ± 0.10 mm, with 95% limits of agreement ranging between −0.34 and 0.07 mm, repeatability coefficient was 0.33 mm. The MAE and repeatability coefficient were similar when ACD were predicted using two different photographs. **Supplementary Figure 2** showed the Bland-Altman plot of the predicted ACD from group one and group two. The mean difference was −0.04 ± 0.09 mm, with 95% limits of agreement ranging between −0.20 and 0.13 mm. For the 50 eyes photographed twice, the distribution of predicted ACD was showed in **Supplementary Figure 3.**

A previous study that utilized machine learning to predict ACD from slit lamp images captured with a smartphone also reported a RMSE of 0.20 mm (11). However, in that study, the images used for prediction required manual maneuvering of a narrow slit (0.1 mm) which was subjective and time-consuming. In contrast, our study involved the development of a deep-learning algorithm that was trained on a much larger dataset and without manual maneuvering. Furthermore, our anterior segment photographs were captured under diffuse illumination, which suggested a two-dimensional image without slit illumination can be used to predict a third dimensional parameter, the ACD.

The overall mean difference between measured and predicted ACD in test set was −0.04 ± 0.20 mm, with 95% limits of agreement of −0.43 to 0.34 mm. Study focused on the repeatability of Lenstar showed that for ACD measurement, mean standard deviation between three consecutives measurements was 0.029, coefficient of variation was 1.06% and intraclass correlation coefficient was 0.991 (27). A previous study evaluated the agreement of ACD (ACD measurement were all from corneal epithelium to the anterior crystalline lens) measured by different instruments, including partial coherence laser interferometry (IOLMaster), scanning peripheral anterior chamber analyzer (SPAC) and anterior segment OCT (AS-OCT) (28). The 95% limits-of-agreement was: AS-OCT vs SPAC, −0.44 to 0.51 mm; AS-OCT vs. IOLMaster: −0.37 to 0.25 mm; SPAC vs. IOLMaster: −0.57 to 0.50 mm (28). Another study found that the 95% limits of agreement of ACD between Lenstar and IOL Master in eyes with cataract was −0.12 to 0.38 mm, in eyes with clear lens was −0.33 to 0.63 mm (29). The extent of agreements reported by the authors was similar to ours. Therefore, the mean difference between measured and predicted ACD is unlikely to be clinically significant. Although there was a proportional bias of our results, similar trends were observed between different methods for ACD measurement (28).

The generated saliency maps showed that the algorithm mainly utilized central corneal region in making its prediction, which was similar to another DL algorithm that predicted shallow ACD (binary classification) from Scheimpflug images (30). The hottest region was congruent with the actual measurement site of ACD which is centered on the cornea, along the visual axis from the corneal endothelium to the anterior crystalline lens capsule. Iris also played a role in making predictions. We speculate that iris was an important panel for the algorithm, like clinicians evaluate the anterior chamber in real world. The upper and right side of the iris were less used by the algorithm, that was because of the eyelid and reflex of the light make these parts less important. Randomly selected heatmaps with MAE ≤ 0.2 mm are presented in **Supplementary Figure 4**. We also investigated those images with poor predictions. The poor predictions were mainly attributed to dilated pupils. Randomly selected heatmaps with MAE > 0.2 mm are presented in **Supplementary Figure 5**. In the present study, we only excluded those participants with pronounced opacities of the central cornea, and/or with intraocular surgery history. Images with small eyelids, obscured by eyelashes, and dilated pupils were all included, to make the dataset closer to the real-word dataset, and to make the algorithm

more generalizable. The MAE of the predictions with dilated pupil ($n = 47$) and un-dilated pupil ($n = 784$) was 0.22 ± 0.15 mm and 0.15 ± 0.12 mm, respectively.

Anterior chamber depth has been demonstrated to be a screening tool for angle closure glaucoma (6–8). Devereux et al. reported that, using a screening cutoff of < 2.22 mm, ACD got a sensitivity of 85% and specificity of 84% for detecting occludable angles (8). A recent study presented a higher sensitivity of 90.2% and specificity of 85.2% using the same cutoff value for distinguishing PACD from normal eyes (6). Angle closure glaucoma is an important public health problem in Asians due to its higher rate of visual morbidity. Most patients with PACG are asymptomatic, up to 64.7% of PACG cases are undetected in Asia (31). China accounts for 48% of angle closure glaucoma worldwide (2), and 90% of the cases with primary angle closure in rural China are undiagnosed (32). Gonioscopy is the current gold standard of anterior chamber angle examination. However, gonioscopy is time consuming and requires technical expertise, which limits its feasibility in large-scale population-based screening (33). ASOCT and ultrasound biomicroscopy (UBM) can help to assess the anterior chamber angle, but they are bulky, expensive and need experienced technicians. The flashlight test and van Herick's test are simple to operate. However, these two methods were reported to be of limited use as screening tests for detecting occludable angles (34).

Smartphones are increasingly used in clinical settings to provide high quality images (35, 36). Coupled with DL algorithms, smartphones may be used for detecting ocular diseases. For example, smartphone based anterior segment photographs and retinal images for cataract grading, glaucoma and diabetic retinopathy detection have been reported (12, 36, 37). There are plenty of advantages for smartphones used in clinics. Since smartphones are widely available, they provide a low-cost and universally accessible method to capture high resolution ocular images. Smartphones usually have a large data storage capacity and do not require extra computers for image storage or processing. In addition, the images captured by smartphones can be easily transmitted wirelessly for consultation in real time. These advantages make smartphone a useful tool in clinics and can bring great benefits for tele-consultation or screenings in remote areas. A previous study successfully developed a machine learning system using anterior segment images captured by digital camera under visible wavelength to diagnose anterior segment eye abnormalities (38). It is conceivable that eye images captured under nature light by smartphone without extra equipment could provide many useful information for ophthalmologist with the help of artificial intelligence. As such it may be used by a wide potential audience and locations, especially in rural area and developing countries.

There are several strengths in the present proof-of-concept study. First, this may be the first study to use DL to quantitatively predict ACD through smartphone-acquired anterior segment photographs. The generated saliency maps showed that the algorithm mainly utilized central corneal region in making its prediction, which was congruent with the actual measurement site of ACD. Secondly, by using merely a smartphone we obtained high quality of anterior segment photographs. Simple instrument makes more cost effective and sustainable. These images were captured under diffuse illumination without slit beam, which makes the procedure much easier and reproduceable.

There are also some limitations in our study. Participants were all from a school-based cohort study aged 11–15 years old, and there were no PACD patient included. Hence, further training of the algorithm involving eyes of older participants, and PACD eyes are needed. Nevertheless, the present study is a proof-of-concept study, which demonstrated that smartphone-acquired anterior segment images can potentially be used to estimate ACD *via* DL.

CONCLUSION

In conclusion, we developed a novel method to estimate ACD using DL algorithm based on smartphone-acquired anterior segment photographs. Further refinement and training involving older participants PACD eyes are still needed, followed up further external validations. This is particularly important in China where the risk of PACG is high and often undetected, leading to increased risk of vision impairment.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available upon further inquiries to the corresponding author.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Kunming Medical University. Written informed consent to participate in this study was provided by the participants' legal guardian/next of kin.

AUTHOR CONTRIBUTIONS

All authors listed have made a substantial, direct, and intellectual contribution to the work and approved it for publication.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmed.2022.912214/full#supplementary-material

Supplementary Figure 1 | Diagram of the human eye. CCT, central corneal thickness; ACD, anterior chamber depth (back of cornea to front of lens); LT, lens thickness; AL, axial length.

Supplementary Figure 2 | Bland-Altman plots illustrating agreement between the predicted ACD from one eye using two anterior segment photographs ($n = 50$).

Supplementary Figure 3 | Overlapping histogram of the measurements. For the 50 eyes photographed twice, the distribution of predicted anterior chamber depth overlapped well.

Supplementary Figure 4 | Randomly selected heatmaps with MAE ≤ 0.2 mm. The regions of the central corneal were most important for the deep learning algorithm making predictions. RE, Right eye; LE, Left eye; GT, Ground truth; Pred, prediction (mm).

Supplementary Figure 5 | Randomly selected heatmaps with MAE > 0.2 mm. The hottest regions were not the central corneal. Most of them had mid-dilated pupils. RE, Right eye; LE, Left eye; GT, Ground truth; Pred, prediction (mm).

REFERENCES

1. Tham YC Li X, Wong TY, Quigley HA, Aung T, Cheng CY. Global prevalence of glaucoma and projections of glaucoma burden through 2040: a systematic review and meta-analysis. *Ophthalmology.* (2014) 121:2081–90. doi: 10.1016/j.ophtha.2014.05.013
2. Quigley HA, Broman AT. The number of people with glaucoma worldwide in 2010 and 2020. *Br J Ophthalmol.* (2006) 90:262–7. doi: 10.1136/bjo.2005.081224
3. Zhang N, Wang J, Chen B, Li Y, Jiang B. Prevalence of primary angle closure glaucoma in the last 20 years: a meta-analysis and systematic review. *Front Med.* (2021) 7:624179. doi: 10.3389/fmed.2020.624179
4. Aung T, Nolan WP, Machin D, Seah SK, Baasanhu J, Khaw PT, et al. Anterior chamber depth and the risk of primary angle closure in 2 east asian populations. *Arch Ophthalmol.* (2005) 123:527–32. doi: 10.1001/archopht.123.4.527
5. Zhang Y, Zhang Q, Thomas R, Li SZ, Wang NL. Development of angle closure and associated risk factors: the Handan eye study. *Acta Ophthalmol.* (2022) 100:e253–61. doi: 10.1111/aos.14887
6. Ma P, Wu Y, Oatts J, Patlidanon J, Yu Y, Ying GS, et al. Evaluation of the diagnostic performance of swept-source anterior segment optical coherence tomography in primary angle closure disease. *Am J Ophthalmol.* (2022) 233:68–77. doi: 10.1016/j.ajo.2021.06.033
7. Nolan WP, Baasanhu J, Undraa A, Uranchimeg D, Ganzorig S, Johnson GJ. Screening for primary angle closure in Mongolia: a randomised controlled trial to determine whether screening and prophylactic treatment will reduce the incidence of primary angle closure glaucoma in an East Asian population. *Br J Ophthalmol.* (2003) 87:271–4. doi: 10.1136/bjo.87.3.271
8. Devereux JG, Foster PJ, Baasanhu J, Uranchimeg D, Lee PS, Erdenbeleig T, et al. Anterior chamber depth measurement as a screening tool for primary angle-closure glaucoma in an East Asian population. *Arch Ophthalmol.* (2000) 118:257–63. doi: 10.1001/archopht.118.2.257
9. Konstantopoulos A, Hossain P, Anderson DF. Recent advances in ophthalmic anterior segment imaging: a new era for ophthalmic diagnosis? *Br J Ophthalmol.* (2007) 91:551–7. doi: 10.1136/bjo.2006.103408
10. Ting DSJ, Foo VH, Yang LWY, Sia JT, Ang M, Lin H, et al. Artificial intelligence for anterior segment diseases: emerging applications in ophthalmology. *Br J Ophthalmol.* (2021) 105:158–68. doi: 10.1136/bjophthalmol-2019-315651
11. Chen D, Ho Y, Sasa Y, Lee J, Yen CC, Tan C. Machine learning-guided prediction of central anterior chamber depth using slit lamp images from a portable smartphone device. *Biosensors.* (2021) 11:182. doi: 10.3390/bios11060182
12. Chen DZ, Tan CW. Smartphone imaging in ophthalmology: a comparison with traditional methods on the reproducibility and usability for anterior segment imaging. *Ann Acad Med Singap.* (2016) 45:6–11.
13. Mohammadpour M, Mohammadpour L, Hassanzad M. Smartphone assisted slit lamp free anterior segment imaging: a novel technique in teleophthalmology. *Cont Lens Anterior Eye.* (2016) 39:80–1. doi: 10.1016/j.clae.2015.09.005
14. Pan CW, Wu RK, Liu H, Li J, Zhong H. Types of lamp for homework and myopia among Chinese school-aged children. *Ophthalmic Epidemiol.* (2018) 25:250–6. doi: 10.1080/09286586.2017.1420204
15. Pan CW, Qiu QX, Qian DJ, Hu DN Li J, Saw SM, et al. Iris colour in relation to myopia among Chinese school-aged children. *Ophthalmic Physiol Opt.* (2018) 38:48–55. doi: 10.1111/opo.12427
16. Qian DJ, Zhong H, Li J, Liu H, Pan CW. Spectacles utilization and its impact on health-related quality of life among rural Chinese adolescents. *Eye.* (2018) 32:1879–85. doi: 10.1038/s41433-018-0197-x
17. Pan CW, Wu RK, Wang P, Li J, Zhong H. Reduced vision, refractive errors and health-related quality of life among adolescents in rural China. *Clin Exp Optom.* (2018) 101:758–63. doi: 10.1111/cxo.12680
18. He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition.* Las Vegas, NV (2016). p. 770–8. doi: 10.1109/CVPR.2016.90
19. Moradmand H, Aghamiri SMR, Ghaderi R. Impact of image preprocessing methods on reproducibility of radiomic features in multimodal magnetic resonance imaging in glioblastoma. *J Appl Clin Med Phys.* (2020) 21:179–90. doi: 10.1002/acm2.12795
20. Shorten C, Khoshgoftaar TM. A survey on image data augmentation for deep learning. *Journal of big data.* (2019) 6:1–48. doi: 10.1186/s40537-019-0197-0
21. Paszke A, Gross S, Massa F, Lerer A, Bradbury J, Chanan G, et al. Pytorch: an imperative style, high-performance deep learning library. *Adv Neural Inf Process Syst.* (2019) 32. Available online at: https://proceedings.neurips.cc/paper/2019; https://proceedings.neurips.cc/paper/2019/hash/bdbca288fee7f92f2bfa9f7012727740-Abstract.html
22. Kingma DP, Ba J. Adam: a method for stochastic optimization. *arXiv* preprint. (2014). Available online at: https://arxiv.org/abs/1412.6980
23. Selvaraju RR, Cogswell M, Das A, Vedantam R, Parikh D, Batra D. Grad-cam: visual explanations from deep networks via gradient-based localization. In: *Proceedings of the IEEE international conference on computer vision.* Venice (2017). p. 618–26. doi: 10.1109/ICCV.2017.74
24. Selvaraju RR, Das A, Vedantam R, Cogswell M, Parikh D, Batra D. Grad-CAM: why did you say that?. *arXiv preprint.* (2016). Available online at: https://arxiv.org/abs/1611.07450
25. Xu G, Wu G, Du Z, Zhu S, Guo Y, Yu H, et al. Distribution of white-to-white corneal diameter and anterior chamber depth in Chinese myopic patients. *Front Med.* (2021) 8:732719. doi: 10.3389/fmed.2021.732719
26. Niu L, Miao H, Han T, Ding L, Wang X, Zhou X. Visual outcomes of Visian ICL implantation for high myopia in patients with shallow anterior chamber depth. *BMC Ophthalmol.* (2019) 19:121. doi: 10.1186/s12886-019-1132-z
27. Wang Q, Ji X, Lu D, Zhu Y, Whelchel A, Wang J, et al. Comparison of A-Scan ultrasonography and the Lenstar optical biometer in Guinea pig eyes. *Exp Eye Res.* (2021) 207:108578. doi: 10.1016/j.exer.2021.108578
28. Lavanya R, Teo L, Friedman DS, Aung HT, Baskaran M, Gao H, et al. Comparison of anterior chamber depth measurements using the IOLMaster, scanning peripheral anterior chamber depth analyser, and anterior segment optical coherence tomography. *Br J Ophthalmol.* (2007) 91:1023–6. doi: 10.1136/bjo.2006.113761
29. Hoffer KJ, Shammas HJ, Savini G. Comparison of 2 laser instruments for measuring axial length. *J Cataract Refract Surg.* (2010) 36:644–8. doi: 10.1016/j.jcrs.2009.11.007
30. Qian Z, Xie X, Yang J, Ye H, Wang Z, Chen J, et al. Detection of shallow anterior chamber depth from two-dimensional anterior segment photographs using deep learning. *BMC Ophthalmol.* (2021) 21:341. doi: 10.1186/s12886-021-02104-0
31. Soh Z, Yu M, Betzler BK, Majithia S, Thakur S, Tham YC, et al. The global extent of undetected glaucoma in adults: a systematic review and meta-analysis. *Ophthalmology.* (2021) 128:1393–404. doi: 10.1016/j.ophtha.2021.04.009
32. Liang Y, Friedman DS, Zhou Q, Yang XH, Sun LP, Guo L, et al. Prevalence and characteristics of primary angle-closure diseases in a rural adult Chinese population: the handan eye study. *Invest Ophthalmol Vis Sci.* (2011) 52:8672–9. doi: 10.1167/iovs.11-7480
33. Desmond T, Tran V, Maharaj M, Carnt N, White A. Diagnostic accuracy of AS-OCT vs gonioscopy for detecting angle closure: a systematic review and meta-analysis. *Graefes Arch Clin Exp Ophthalmol.* (2022) 260:1–23. doi: 10.1007/s00417-021-05271-4
34. Thomas R, George T, Braganza A, Muliyil J. The flashlight test and van Herick's test are poor predictors for occludable angles. *Aust N Z J Ophthalmol.* (1996) 24:251–6. doi: 10.1111/j.1442-9071.1996.tb01588.x
35. Madanagopalan VG, Raman R. Commentary: artificial intelligence and smartphone fundus photography-are we at the cusp of revolutionary changes in retinal disease detection? *Indian J Ophthalmol.* (2020) 68:396–7. doi: 10.4103/ijo.IJO_2175_19
36. Nakahara K, Asaoka R, Tanito M, Shibata N, Mitsuhashi K, Fujino Y, et al. Deep learning-assisted (automatic) diagnosis of glaucoma using a smartphone. *Br J Ophthalmol.* (2022) 106:587–92. doi: 10.1136/bjophthalmol-2020-318107
37. Karakaya M, Hacisoftaoglu RE. Comparison of smartphone-based retinal imaging systems for diabetic retinopathy detection using deep learning. *BMC Bioinformatics.* (2020) 21:259. doi: 10.1186/s12859-020-03587-2

38. Mahesh Kumar SV and Gunasundari R. Computer-aided diagnosis of anterior segment eye abnormalities using visible wavelength image analysis based machine learning. *J Med Syst*. (2018) 42:128. doi: 10.1007/s10916-018-0980-z

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers | Frontiers in Medicine

ORIGINAL RESEARCH
published: 28 June 2022
doi: 10.3389/fmed.2022.850284

Predicting Axial Length From Choroidal Thickness on Optical Coherence Tomography Images With Machine Learning Based Algorithms

*Hao-Chun Lu[1,2], Hsin-Yi Chen[3,4], Chien-Jung Huang[3], Pao-Hsien Chu[5], Lung-Sheng Wu[5] and Chia-Ying Tsai[3,4]**

[1] Graduate Institute of Business and Management, Chang Gung University, Taoyuan, Taiwan, [2] Division of Cardiology, Department of Internal Medicine, Chang Gung Memorial Hospital, Taoyuan, Taiwan, [3] Department of Ophthalmology, Fu Jen Catholic University Hospital, New Taipei City, Taiwan, [4] School of Medicine, College of Medicine, Fu Jen Catholic University, New Taipei City, Taiwan, [5] Division of Cardiology, Department of Internal Medicine, Chang Gung Memorial Hospital, Chang Gung University College of Medicine, Taipei, Taiwan

Purpose: We formulated and tested ensemble learning models to classify axial length (AXL) from choroidal thickness (CT) as indicated on fovea-centered, 2D single optical coherence tomography (OCT) images.

Design: Retrospective cross-sectional study.

Participants: We analyzed 710 OCT images from 355 eyes of 188 patients. Each eye had 2 OCT images.

Methods: The CT was estimated from 3 points of each image. We used five machine-learning base algorithms to construct the classifiers. This study trained and validated the models to classify the AXLs eyes based on binary (AXL $<$ or $>$ 26 mm) and multiclass (AXL $<$ 22 mm, between 22 and 26 mm, and $>$ 26 mm) classifications.

Results: No features were redundant or duplicated after an analysis using Pearson's correlation coefficient, LASSO-Pattern search algorithm, and variance inflation factors. Among the positions, CT at the nasal side had the highest correlation with AXL followed by the central area. In binary classification, our classifiers obtained high accuracy, as indicated by accuracy, recall, positive predictive value (PPV), negative predictive value (NPV), F1 score, and area under ROC curve (AUC) values of 94.37, 100, 90.91, 100, 86.67, and 95.61%, respectively. In multiclass classification, our classifiers were also highly accurate, as indicated by accuracy, weighted recall, weighted PPV, weighted NPV, weighted F1 score, and macro AUC of 88.73, 88.73, 91.21, 85.83, 87.42, and 93.42%, respectively.

Conclusions: Our binary and multiclass classifiers classify AXL well from CT, as indicated on OCT images. We demonstrated the effectiveness of the proposed classifiers and provided an assistance tool for physicians.

Keywords: high myopia, choroidal thickness, axial length, machine learning, ensemble learning, optical coherence tomography (OCT)

OPEN ACCESS

Edited by:
Yi-Ting Hsieh,
National Taiwan University Hospital, Taiwan

Reviewed by:
Minhaj Nur Alam,
Stanford University, United States
Ailin Song,
Duke University, United States

***Correspondence:**
Chia-Ying Tsai
chiaying131@gmail.com

Specialty section:
This article was submitted to Ophthalmology, a section of the journal Frontiers in Medicine

Received: *07 January 2022*
Accepted: *25 May 2022*
Published: *28 June 2022*

Citation:
Lu H-C, Chen H-Y, Huang C-J, Chu P-H, Wu L-S and Tsai C-Y (2022) Predicting Axial Length From Choroidal Thickness on Optical Coherence Tomography Images With Machine Learning Based Algorithms. Front. Med. 9:850284. doi: 10.3389/fmed.2022.850284

INTRODUCTION

Myopia is a common disease among Asian people, and its incidence has increased worldwide. Holden et al. (1) estimated that the global prevalence of myopia would reach 49.8% in 2050 along with 9.8% for high myopia, and the myopia rate in East Asia would increase from 51.6 to 65.3%, the highest in the world, in the next 3 decades. Among Taiwanese schoolchildren evaluated between 1983 and 2017, the myopia rate quintupled from 5.37 to 25.41% for 7-year-olds and more than doubled (from 30.66 to 76.67%) for 12-year-olds (2).

Eyes with a spherical equivalence (SE) of less than -6.00 D were defined as having high myopia, and high myopia is correlated with axial length longer than 26.0 mm (3, 4). High myopia is associated with increased risks of cataract, glaucoma, retinal detachment, and maculopathy (5). These ocular complications of high myopia become more common with advanced age and may eventually lead to blindness (5, 6). Morgan et al. (6) suggested that the elongated AXL is the underlying mechanism of myopia development and progression. Choroid, located at the exterior of the retina and which provides blood supply to the outer portion of the retina, has been reported to be thinner in myopic than emmetropic eyes and is related to AXL elongation (7). Choroid thinning not only correlates with myopia progression but is also related to other complications, such as staphyloma and chorioretinal atrophy in high myopia (7–14). In addition to longer AXL, CT is also lower in older adults and in women (15).

Artificial intelligence (AI) is being used in medicine. In ophthalmology, color fundus images are commonly used for machine training in disease diagnosis, such as for diabetic retinopathy (DR), (16, 17) age-related macular disease (AMD), (18, 19) and glaucoma (20). Asaoka et al. (21) classified open-angle glaucoma and healthy eyes using deep learning algorithm trained on color fundus images from 159 patients (including 51 with glaucoma). Hemelings et al. identified pathologic myopia from color fundus images by means of Convolutional Neural Network (CNN) (22). Optical coherence tomography (OCT) has become one of the most effective imaging modalities in the diagnosis of various retinal conditions by providing high-resolution, cross-sectional images of the entire retina and choroid (23). The long wavelength (870 nm) used for scanning in spectral-domain OCT (SD-OCT) enables better penetration and ensures high-resolution retina and choroid images. Machine learning and deep learning have been successfully applied in OCT images for biomarker identification in AMD (24). Since myopia is a rising problem in ophthalmology, OCT images have been used for AI prediction in myopic eyes recently (25–27).

In this study, we focused on the relationship between CT and AXL. Since OCT is a common exam in clinics for patients with retinal diseases, glaucoma, and cataract surgery, it is meaningful to access more information from the existed exam images. With SD-OCT images from eyes with different refraction status and AXL, we investigated the utility of machine learning algorithms for predicting AXL and proposed a multiclass classifier of AXL by means of the CTs (28). In this study, five machine learning base algorithms [3 layers backpropagation neural network (BPN), support vector machine (SVM), random forest (RF), adaptive boosting (AdaBoost), extreme gradient boosting (XGBoost)] are used to construct classifiers for binary and multiclass classifications. The proposed classifiers can quickly and accurately predict the axial length by means of the choroid thickness (CT) and help us to understand the contribution of choroidal change in the etiology of myopia.

MATERIALS AND METHODS

This retrospective cohort study adhered to the tenets of the Declaration of Helsinki. This study was approved by the Institutional Review Board of Fu Jen Catholic University Hospital (FJUH).

DATA SETS

Participants

Patients with OCT image findings taken from and who underwent AXL evaluation in CY Tsai's and CJ Huang's clinics in the ophthalmology department at FJUH at any period from Sep. 2017 to Dec. 2019 were included in this study. We collected comprehensive information for participants' sex, age, body height, body weight, and best-corrected visual acuity (29, 30). Patients with incomplete data or retinopathies, such as diabetic retinopathy, age-related macular degeneration, and history of previous photodynamic therapy, were excluded from the study.

Optical Coherence Tomography Machine and Scanning Settings

Spectralis SD-OCT equipment (Heidelberg Engineering, Heidelberg, Germany) was used to evaluate CT in both eyes; OCT was performed in the daytime. Cross-sectional and longitudinal scanning was performed in each eye (**Figure 1**). The SD-OCT uses a super luminescence diode with an average wavelength of 870 nm as a light source, an 8-um axial resolution, and a 10-μm transverse resolution in tissue. The position of fovea was defined as the anatomical depression of macula. The CT was measured at 6 points: central fovea, 3 mm nasal, and 3 mm temporal to the fovea at cross-sectional image and central fovea, 3 mm superior, and 3 mm inferior to the fovea at longitudinal image (**Figure 2**). Each image was measured by 2 investigators independently and rechecked by a third investigator.

AXL Measurement

The AXL of the eyes was evaluated by a non-contact technique by using a Lenstar LS 900 platform (HAAG-Streit, Mason, OH, United States).

Features

This data set had 11 features (**Table 1**), which were (1) participants' gender, age, height, and weight; (2) 3 cross-sectional CTs; (3) 3 longitudinal CTs; and (4) AXL. The pairwise scatter plots of all features with binary and multiclass classifications are shown in **Figures 3**, **4**, respectively. **Figures 3**, **4** clearly show that the relationships of all pairs of two features are almost non-linear.

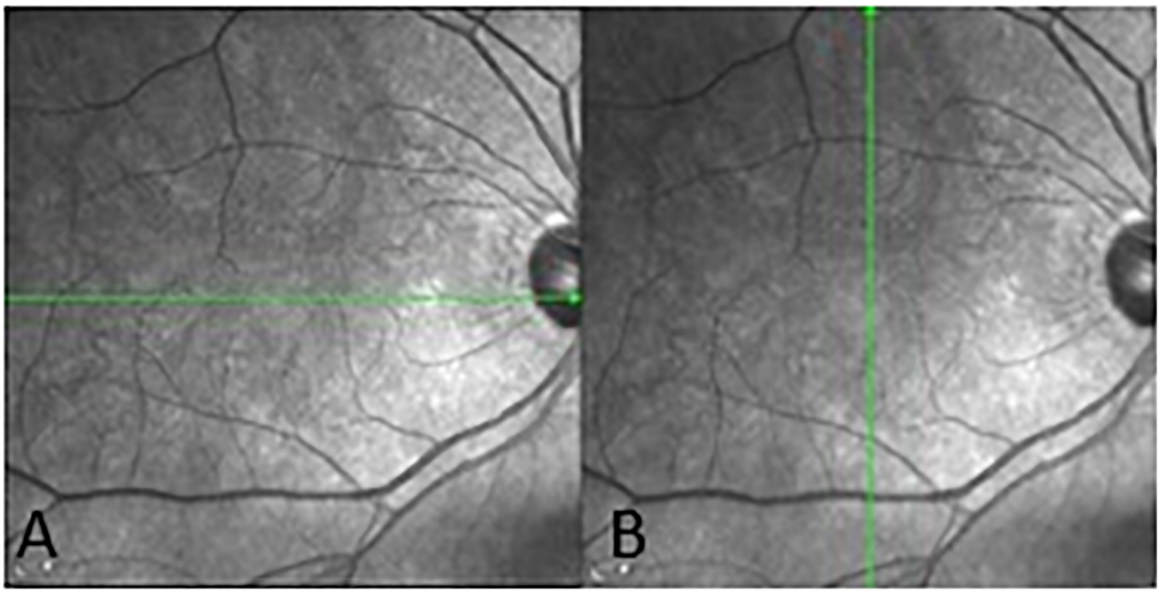

FIGURE 1 | Cross-sectional **(A)** and longitudinal **(B)** choroidal images from SD-OCT.

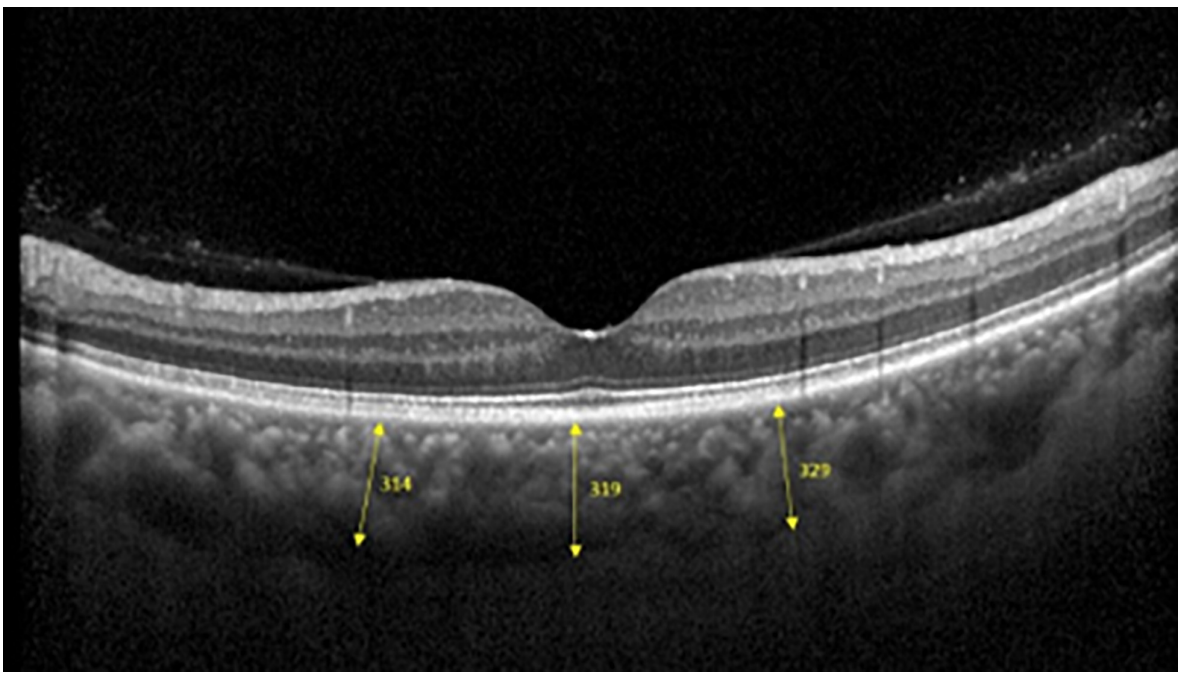

FIGURE 2 | Three positions at which choroid thicknesses was indicated in OCT images.

Binary and Multiclass Classification of Axial Length

This study trained and validated the classifiers, which predicted the class of AXL of each eye using binary and multiclass classifications. In binary classification, we classified eyes into AXL $<$ 26 mm and AXL$>$ 26 mm; in multiclass classification, we classified eyes into AXL $<$ 22 mm, 22 mm$>$ AXL $<$ 26 mm, and AXL $>$ 26 mm (31–34).

Development of Classifiers by Machine Learning Algorithms

Algorithm Selection

In this study, we analyzed 710 OCT images from 355 eyes of 188 patients. However, 710 images are quite low for a CNN algorithm. An appropriate number of samples depends on the specific problem, and it should be tested for each case

TABLE 1 | Features in this study.

No.	Feature name	Description	Data type
1.	Gender	0 for male and 1 for female.	Nominal
2.	Age	The age of subject.	Continuous
3.	Height	The height of subject (cm).	Continuous
4.	Weight	The weight of subject (kg).	Continuous
5.	Choroid-LU	Up thicknesses of longitudinal choroid.	Continuous
6.	Choroid-LM	Middle thicknesses of longitudinal choroid.	Continuous
7.	Choroid-LD	Down thicknesses of longitudinal choroid.	Continuous
8.	Choroid-CT	Temporal thicknesses of cross sections choroid.	Continuous
9.	Choroid-CM	Middle thicknesses of cross sections choroid.	Continuous
10.	Choroid-CN	Nasal thicknesses of cross sections choroid.	Continuous
11.	AXL	Axial length of eyes	Continuous

individually. But a rough rule of thumb is to train a CNN algorithm with a data set larger than 5,000 samples for effective generalization of the problem. Our previous study used data augmentation to increase this study's image samples and utilized a CNN algorithm to construct the image classifier through OCT images. However, the image classifier obtains a low accuracy. For obtaining satisfactory results, this research does not use simple algorithms to construct the linear classifier and selects state-of-the-art or strong algorithms to construct the non-linear classifier. Therefore, the selected algorithms are BNN, SVM, RF, AdaBoost, and XGBoost. RF, AdaBoost, and XGBoost are also the ensemble learning.

Essentially, ensemble learning algorithms feature the combination of several weak classifiers to form a strong one with bagging or boosting approaches. The bagging approach trains many individual models in a parallel way, and each model is trained by a random subset of the data. Boosting approach trains a bunch of individual models in a sequential manner, and each individual model learns from mistakes made by the previous model. The Ensemble learning algorithms obtain less bias, less variance, and better results than traditional machine learning in general. Friedman et al. (35) indicated that boosting approach results in dramatic performance improvements and no additional requirements for the dataset and classifiers.

The RF, AdaBoost, and XGBoost are based on the bagging, boosting, and hybrid bagging and boosting approaches. AdaBoost, one of the first boosting algorithms adapted to solve practical problems, uses multiple iterations to create a strong learner by iteratively adding weak learners. Gradient boosting, a generalization of AdaBoost, is one of the most powerful techniques for building predictive models. The main objective of gradient boosting is to minimize the loss function by adding weak learners using a gradient descent algorithm. XGBoost is an extension of gradient-boosted decision trees and has the following advantages: regularized learning, gradient tree boosting, and shrinkage with column subsampling. Since the used ensemble learning algorithms in this study always have the hyperparameter—*n_estimators* (the number of estimators), the *n_estimators* means the number of the individual model will be performed. Therefore, ensemble learning algorithms always spend much more time than BNN and SVM calculation time.

Classifiers Construction Process

The processes of this study are exhibited in **Figure 5**: process 1 (preprocess) and process 2 (primary processes for each algorithm). Before we constructed the classifiers, the data set was preprocessed by using process 1. This study utilized 5 algorithms (BNN, SVM, RF, AdaBoost, and XGBoost) to predict myopia by means of the CTs. We constructed 2 classifiers for binary and multiclass classifications for each algorithm. Without loss of generality, all models constructed by each algorithm were executed by process 2. Finally, this study obtained the appropriate features, suggesting resample methods, and the appropriate values of hyperparameters for each algorithm with the target classifications. The details of gray steps exhibited in **Figure 5** are described in subsections *feature standardization*, *data splitting*, *feature selection*, *hyperparameter Optimization*, and *Oversampling of Imbalanced Data*.

Feature Standardization

To reduce the training phase's processing time, we standardized numerical features by removing their means and scaling to unit variance through the formula as follows.

Feature with normalization = (feature – feature's mean)/feature's standard deviation.

Data Splitting

The 355 tuples in this imbalanced data were collected from 188 patients with 355 eyes. Each tuple represented a completely eyeball's six choroidal thicknesses based on cross-sectional and longitudinal scanning images. **Table 4** indicates that the proportion of AXL $<$ 26 mm to AXL $>$ 26 mm was 0.7944:0.2056 in binary classification. In multiclass classification, the proportion of AXL $<$ 22 mm, between 22 and 26 mm, and $>$ 26 mm was 0.0394:0.7549:0.2056. We split the imbalanced data into training and test sets based on a uniform random distribution, and the percentage ratio of training and test sets followed 80 and 20% with patient level (no patient across both training and test sets), where each set shared a similar proportion of all categories. The training set was used in *feature selection*, *hyperparameter optimization*, and o*versampling*. Finally, the test set was used to evaluate all metrics of each set {algorithm, hyperparameters search method, oversampling method} in Evaluation step of process 2 in **Figure 2** and the comparisons of AXL class prediction between humans and classifiers.

Feature Selection

This study used 3 methods: Pearson's correlation coefficient (Pearson), variance inflation factor (VIF), and least absolute shrinkage and selection operator (Lasso) to evaluate and select the appropriate features based on the training set. In the training set, one tuple only contained one eye's features. Therefore, the feature selection can purely evaluate the relationships of the features within one eye. The three feature selection methods were performed sequentially independently, and any feature detected as redundant or useless by any method will be removed in this step.

Pearson's r indicated the linear relationship between a given feature and class label. The VIF measured how substantially

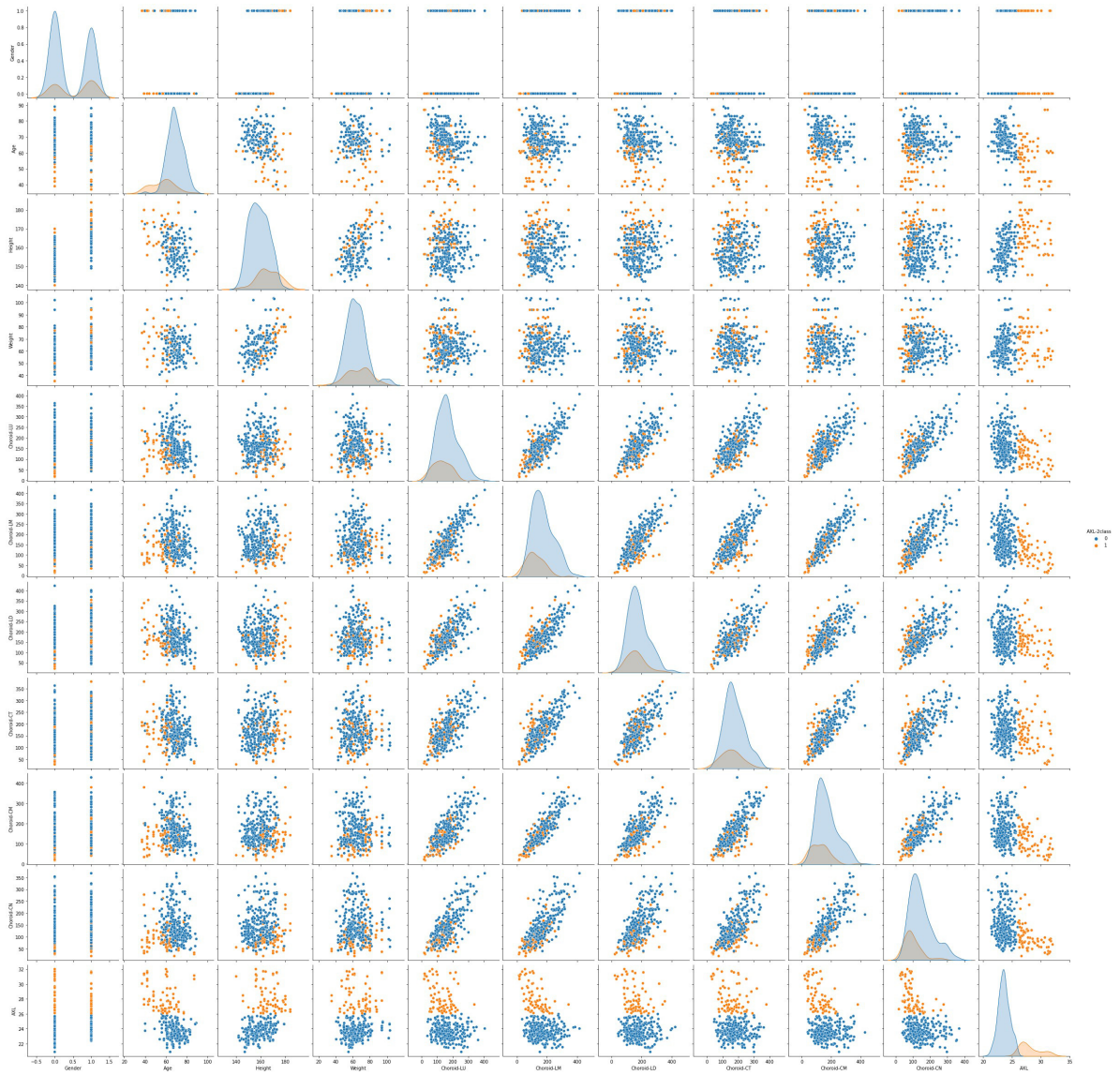

FIGURE 3 | Pairwise scatter plots of all features with binary classification.

the variance of an independent feature was influenced by other independent features. If the VIF of the target feature was > 10, we eliminated the target feature. Lasso performs covariate selection by forcing the sum of the absolute value of the regression coefficients to be less than a fixed value, which forced certain coefficients to be 0. The order of importance of input features made the fitted model more interpretable. LASSO utilized the L1 penalty to select the most feature at once based on a given lambda value. Compared with LASSO, Elastic net used a penalty mixed L1 and L2 norms, and Elastic net is hard to obtain the clear order of importance input features.

Finally, 100 different lambda values (1.00, 0.99,. . ., 0.01) with descending order were used in LASSO, and we only kept the lambda values that LASSO selection an additional new significant feature in S **Table 5**.

Hyperparameter Optimization

Each algorithm had its hyperparameters that need to be tuned because the appropriate hyperparameters were very different for the algorithm applied in a different dataset. That is, different hyperparameters will most influence the performances of an algorithm. The search scopes for consideration of hyperparameters for all algorithms in this study (BNN, SVM, RF, AdaBoost, XGBoost) were listed in **Supplementary Appendix-I**. Since the combinations of hyperparameters for each algorithm were numerous, this study used three search methods [grid search, random search, and Hyperopt search proposed by Bergstra et al. (36)] to search the appropriate hyperparameters. Each search method searched and evaluated the possible hyperparameters among the scopes in **Supplementary Appendix-I**, and the search method obtained

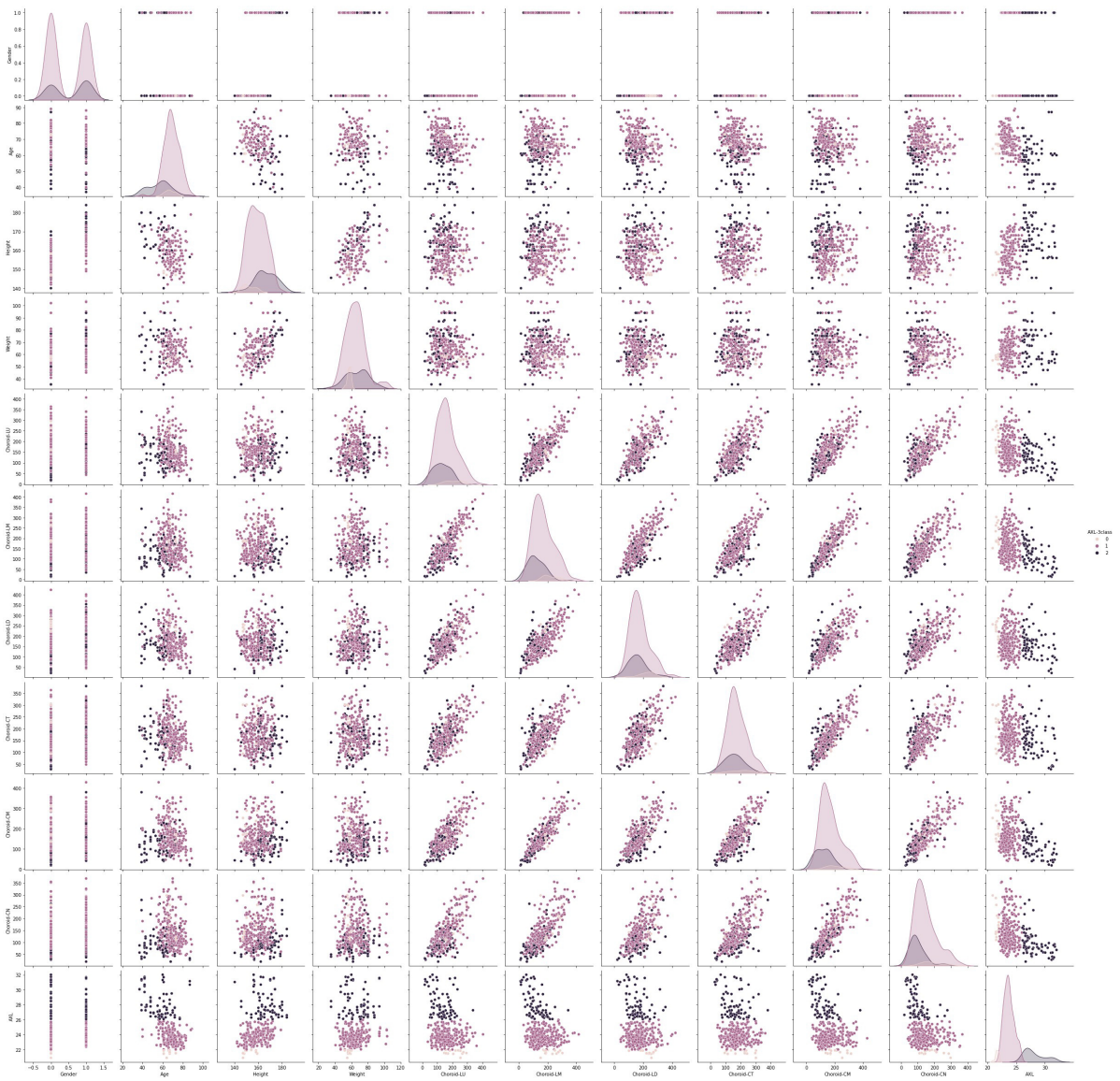

FIGURE 4 | Pairwise scatter plots of all features with multiclass classification.

the best hyperparameters of the algorithm with the target training set by the main matric—F1-score. Finally, each algorithm will pick the best hyperparameters among the three search methods.

Grid search searched the performances with all combinations of hyperparameters with the specific fixed values in their scopes. Random and Hyperopt searches selected the hyperparameters' with the possible real numbers from the specific intervals; therefore, random search and Hyperopt search selected more floating values for the hyperparameters that were not listed in the grid search. The numbers of grid search and random search were the same. The speed of Hyperopt search was much slower than grid search and random search because Hyperopt analyzes and improves the values of hyperparameters after each iteration; therefore, the search number of Hyperopt was 1/2 that of grid search. In **Supplementary Appendix-I**, the search numbers of {grid search, random search, and Hyperopt} for BNN, SVM, RF, AdaBoost, and XGBoost are {162, 162, 81}, {20,100, 20,100, 10,050}, {15,000, 15,000, 7,500}, {24,000, 24,000, 12,000}, and {25,920, 25,920, 12,960}, respectively. Therefore, there were 212,955 searches for binary classification, and the total number of searches in this study is 425,910 for both binary and multiclass classifications.

For each search in the three search methods, stratified fivefold cross-validation (CV) was used to evaluate the performance of the current hyperparameters. Although the fivefold CV will take a 5-times validation time than the holdout method, the trained model will not be easy to overfit for a specific validation set and reduce the bias and variance of the performance estimate.

TABLE 2 | Size of original data sets and oversampled data sets.

Oversampling Set		None			ROS/SMOTE/ADASYN		
		Training(80%)	Test(20%)	Total	Training	Test	Total
Binary	AXL < 26 mm	226	56	282	226	56	282
	AXL > 26 mm	58	15	73	226	15	241
	Sum	284	71	355	452	71	523
Multiclass	AXL < 22 mm	11	3	14	215	3	218
	22 mm>AXL < 26 mm	215	53	268	215	53	268
	AXL> 26 mm	58	15	73	215	15	230
	Sum	284	71	355	645	71	716

Oversampling for Imbalanced Data

For each split of fivefold CV, we oversampled the training fold to avoid the imbalanced issue. The use of this technique increased the number of samples of the smaller-sized categories for the sample sizes to be consistent among all categories. The oversampled samples of smaller-sized categorized of the training fold appeared only in the training fold. This study used three oversampling techniques as follows.

Random oversampling (ROS): Randomly sample the tuples in the categories of smaller sample sizes.

Synthetic minority oversampling technique (SMOTE): For the categories of smaller sample sizes, find a sample x and its k-nearest neighbor samples $x_j (j = 1, ..., k)$. Select one individual x_j' from x_j and create a new sample based on the linear combination of x_i and x_j' .

Adaptive synthetic sampling (ADASYN): ADASYN is a technique based on the SMOTE algorithm for generating synthetic data. The difference between ADASYN and SMOTE is that ADASYN implements a methodology that detects those samples of the minority class found in spaces dominated by the majority class to generate samples in the lower density areas of the minority class. ADASYN focuses on those samples of the minority class that are difficult to classify because they are in a low-density area.

Finally, the numbers of training and validation folds with the three oversampling techniques and original data set are presented in **Table 2**.

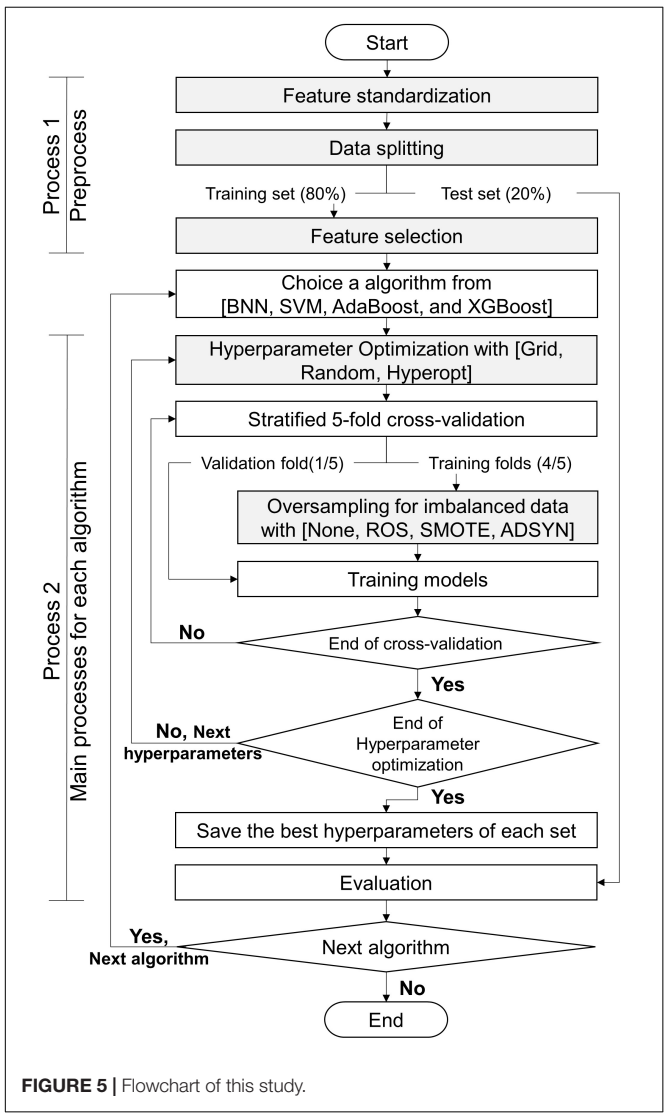

FIGURE 5 | Flowchart of this study.

Algorithm Evaluation and Statistical Analysis

Evaluating Metrics

Because the data set in this study was imbalanced and had many classes, accuracy alone was not sufficient to indicate the classifiers' performance. Therefore, accuracy, recall (sensitivity), PPV (precision), NPV, F1 score, Specificity, and AUC were used for binary classification. For multiclass classification, this study used accuracy, weighted recall (sensitivity), weighted PPV, weighted NPV, weighted F1 score, weighted Specificity, and macro AUC where macro AUC is the macro average of multiple one-vs-rest AUCs. Five algorithms (BPN, SVM, AdaBoost, XGBoost, RF), three hyperparameter optimizations (Grid Search, Random Search, Hyperopt), and four oversampling techniques (None, ROS, SMOTE, and ADASYN) were used to construct classifiers with training set based on the best hyperparameters obtained in Hyperparameter Optimization for binary and multiclass classifications. The numbers of training

and test sets with the three oversampling techniques and original data set are presented in **Table 2**. In total, 120 experiment results with test set are reported for binary and multiclass classification. The details of the experiment results of binary and multiclass classifications are listed in **Supplementary Appendix-II,III** respectively.

AXL Prediction

For evaluation of AXL prediction from human eyes, 2 ophthalmologists and two medical students were asked to predict AXL of the eyes by the 10 features (age, sex, height, weight, choroidal thickness from the 6 points). They were blind to the AXL during prediction, and the results were checked and compared with the results from AI algorithm with *t*-test by another person in our group.

RESULTS

Demographics

In this study, 710 OCT images of 355 eyes (172 left and 183 right) of 188 patients were collected. All patients had complete sex, age, body height, and body weight data in their medical records. In total, 87 (46.28%) men and 101 (53.72%) women participated. The average age was 66.49 ± 9.73 years, average body height was 159.8 ± 8.72 cm, and average body weight was 64.8 ± 12.1 kg (**Table 3**). The average AXL was 24.55 ± 2.26 mm at the right eye and 24.61 ± 2.21 mm at the left eye. The average CT at the central fovea, 3 mm nasal, and 3 mm temporal to the fovea on cross-sectional image was 161.46 ± 75.17 mm, 137.88 ± 69.29 mm, and 170.01 ± 65.56 mm at the right eye, respectively; 162.48 ± 72.16 mm, 139.58 ± 67.30 mm, 175.86 ± 66.50 mm at the left eye, respectively. The average CTs at the central fovea, 3 mm superior, and 3 mm inferior to the fovea on longitudinal image were 161.56 ± 74.9, 172.67 ± 71.56, and 158.16 ± 66.01 mm at the right eye, respectively; 162.64 ± 71.03, 161.24 ± 69.22, and 175.32 ± 62.97 mm at the left eye, respectively. In our binary classification, there were 282 (79.44%) eyes with AXL $<$ 26 mm and 73 (20.56%) eyes with AXL $\geq$ 26 mm; in the multiclass classification, there were 14 (3.94%) eyes with AXL $<$ 22 mm, 268 (75.49%) between 22 and 26 mm, and 73 (20.56%) eyes with AXL $\geq$ 26 mm (**Table 4**).

Results of Pearson, VIF, and LASSO

The Pearson results of all eyes revealed that the 10 features were substantially correlated with the AXL (all $P < 0.05$). Age had the strongest correlation, followed by height, Choroid-CN, Choroid-LM, Choroid-CM, Choroid-LU, Choroid-LD, Choroid-CT, gender, and weight. Height and weight had positive correlations with AXL, and men had longer AXL, but the other features had negative correlation with AXL (**Supplementary Tables 1**, **2**). Regarding the CT between the left and right eye, Choroid-CM and Choroid-CN had symmetric properties between the left and right eye because their *r* coefficients were higher than others. However, the others were asymmetric (**Supplementary Table 3**). The results of VIF revealed that the 10 features exhibited no multicollinearity because no feature's VIF was $>$ 10 (**Supplementary Table 4**).

The sequence of the features' coefficient becoming non-zero under Lasso with decreasing α is listed as follows: age, Choroid-CN, Choroid-LM, Height, Choroid-CM, Choroid-LU, weight, Choroid-CT, Choroid-LD, and gender. Height and male gender are positively correlated with AXL. Among the 6 CTs, Choroid-CT and Choroid-LD features were positively correlated to AXL, compared with other features. However, based on the Lasso with $\alpha = 0.01$, the coefficients of gender, choroid-LD, and weight were relatively small (**Supplementary Table 5**). After the Pearson, VIF, and Lasso analyses were conducted, all of the 10 features were found to be non-redundant. Therefore, all 10 of the features were used for constructing the classifiers.

Results of Proposed Classifiers

Classifier 3 obtained the best PPV and Specificity of 90.91 and 98.25%. Classifier 4 obtains the best PPV and Specificity of 92.21 and 93.37%. **Tables 6**, **7** list some classifiers with the best metrics for binary and multiclass classifications, respectively. All metrics in **Tables 6**, **7** are calculated from an independent test set described in the Data Splitting. Classifier 1 (SVM with random search and ROS oversampling), Classifier 2 (AdaBoost with random search and ADASYN oversampling), and Classifier 3 (AdaBoost with Hyperopt search and ROS oversampling) have different best metrics for binary classification. Classifier 1 obtains the best recall, NPV, and AUC of 100, 100, and 95.61%, respectively. Classifier 2 obtains the best accuracy and F1-score of 94.37 and 86.67%. Classifier 3 obtained the best PPV and Specificity of 90.91 and 98.25%.

For multiclass classification, Classifier 4 (SVM with grid random and SMOTE oversampling), Classifier 5 (AdaBoost with grid search and ROS oversampling), Classifier 6 (XGBoost with grid search and SMOTE oversampling), and Classifier 7 (XGBoost with random search and ROS oversampling) have different best metrics. Classifier 4 obtains the best PPV and Specificity of 92.21 and 93.37%. Classifier 5 obtains the best accuracy, weighted recall, and weighted F1 score of 88.73, 88.73, and 87.43%, respectively. Classifier 6 and Classifier 7 obtain the best macro AUC (93.51%) and weighted NPC, respectively.

Among all metrics, F1-score is the main metric in this study because F1-score seeks the balance of Recall and PPV for the imbalanced dataset. In clinical application, it can help doctors utilize the ensemble learning with the balance of positive prediction and effective medical resource use. However, assessing a model with the best F1-score and poor other metrics is inappropriate. It is still very important to comprehensively consider all metrics. Based on **Tables 6**, **7**, the proposed Classifiers 2 and 5 are excellent models for detecting myopia with binary and multiclass classifications. It is possible to classify AXL $>$ or $<$ 26 mm by CTs with the proposed Classifier 2 because Classifier 2 has no low performances of all metrics. AXL $<$ 22 mm, between 22 and 26 mm, $\geq$ 26 mm can be classified based on CTs with the proposed Classifier 5 because Classifier 5 has good performances for all metrics.

Based on **Table 6**, Classifier 1 (SVM) has the best recall, NPV, and AUC but a poor PPV. Classifier 2 (AdaBoost) obtains the

TABLE 3 | Characteristics of participants.

Feature	Number (%)	Feature	Number (%)
Gender		Height (cm, mean = 159.8, SD = 8.72)	
Male	87 (46.3%)	<150	24 (12.8%)
Female	101 (53.7%)	150–159.9	67 (35.6%)
Age (mean = 66.5, SD = 9.73)		160–169.9	65 (34.6%)
<40	3 (1.6%)	170–179.9	29 (15.4%)
40–49	10 (5.3%)	>179.9	3 (1.6%)
50–59	19 (10.1%)	Weight (kg, mean = 64.8, SD = 12.1)	
60–69	85 (45.2%)	<50	14 (7.4%)
70–79	58 (30.9%)	50–59.9	54 (28.7%)
>79	13 (6.9%)	60–69.9	55 (29.3%)
		70–79.9	45 (23.9%)
		80–89.9	12 (6.4%)
		>89.9	8 (4.3%)

best accuracy and F1-score, and it also has the second-best recall, PPV, NPV, F1-score, and specificity. Additional, the gaps of recall, NPV, and AUC between Classifier 1 and Classifier 2 are small. Because Classifier 2 is more stable than Classifiers 1 and 3 and has no low performances of all metrics, Classifier 2 is recommended to classify AXL > or < 26 mm by CTs.

Based on **Table 7**, Classifier 4 (SVM) has the best the best PPV (weighted) and specificity (weighted) but very poor accuracy, recall (weighted), and NPV (weighted). Classifier 6 (XGBoost) and Classifier 7 (XGBoost) respectively, have the best AUC (macro) and NPV (weighted) but medium remaining metrics. Classifier 5 (AdaBoost) obtains the best Accuracy, recall, and F1-score, and it also has the second-best PPV (weighted), NPV (weighted), specificity (weighted), and AUC (macro). Because Classifier 5 is excellent and stable than other three Classifiers (4, 6, and 7), Classifier 5 can be used to classify AXL < 22 mm, AXL between 22 and 26 mm, and AXL ≥ 26 mm by CTs.

The appropriate values of hyperparameters of classifiers in **Tables 5**, **6** are obtained by Hyperparameter Optimization, and the details of values of hyperparameters are listed in **Supplementary Appendix-IV**.

The AXL Prediction

To compare the results of AXL prediction based on the 10 features between the proposed classifiers and ophthalmologists, we recruited 2 ophthalmologists and 2 medical students to predict AXL in binary and multiclass classification based on the same 10 features. In the results, the accuracy was 48.61–69.44%, PPV (weighted) was 61.50–76.08%, Recall (weighted) was 48.61–69.44%, F1 score (weighted) was 54.29–71.92%, NPV (weighted) was 38.77–53.71%, and AUC (macro) was 49.07–63.03%. The results were considerably less accurate than those from our developed classifiers in **Tables 5**, **6**.

The Comparisons of AXL Class Prediction Between Humans and Classifiers

To compare the AXL class prediction between the proposed classifiers and ophthalmologists, we recruited two ophthalmologists (OPHs) and two medical students to compare binary and multiclass classification based on the same test set with ten features (without AXL feature). Since the test set size is over 30, this study used the test of proportion to verify the performance of results. The null and alternative hypotheses are $H_0 : p_{\text{Human}} \geq p_{\text{AI}}$ and $H_1 : p_{\text{Human}} < p_{\text{AI}}$, where p_{Human} and p_{AI} are the metrics of human performances and proposed classifiers, respectively. The comparison and test results are listed in **Tables 7**, **8**.

In **Tables 7**, **8**, the human performances' results of accuracy, recall, PPV, NPV, F1-score, specificity, and AUC, respectively, are 47.89–80.28%, 6.67–67.61%, 5.56–72.85%, 38.91–90.38%, 6.06–69.04%, 36.98–85.71%, and 38.15–75.30%. Compared with the same metrics of proposed classifiers 2 and 5 in **Tables 6**, **7**, all tests of proportion rejected H0. It demonstrated that the proposed classifiers outperform the human performances.

DISCUSSION

In this study, we proposed that Classifiers 1–6 can predict AXL by means of patients' age, sex, height, weight, and CT measured from OCT images. Studies have reported that CT is negatively correlated with AXL, and people with high myopia tend to have a thinner choroid (7–11). However, few studies have assessed the prediction of AXL by means of CT. In the proposed classifiers, the binary prediction has accuracy, recall, PPV > 90%, and NPV > 85%; multiclass prediction has accuracy, recall, PPV, and NPV > 80%, which is substantially better than prediction by ophthalmologists in this study. The 10 selected features were correlated with AXL, and the correlation was confirmed by the Pearson, VIF, and Lasso analyses. In the Pearson and Lasso analyses, age had the highest negative correlation with AXL. This observation may result from the difference in

TABLE 4 | Class label criteria in terms of AXL.

Binary classification		
Class	Rule	Number (%)
0	AXL < 26 mm	282 (79.44%)
1	AXL ≥ 26 mm	73 (20.56%)
Class	Multiclass ClassificationRule	Number (%)
0	AXL < 22 mm	14 (3.94%)
1	22 mm ≤ AXL < 26 mm	268 (75.49%)
2	AXL ≥ 26 mm	73 (20.56%)

prevalence of myopia among various age groups. Studies have reported that from 1983 to 2017, the prevalence of myopia in the same age group substantially increased in Taiwan, (2) which leads to more incidence of myopic eyes in younger patients in our cohort. Another possible reason for this correlation is that whereas older patients could have various conditions that require ophthalmic clinic follow-up, younger patients seldom have severe eye disease that requires clinic visits and image studies, except for those patients with myopia or high myopia who were at risk of retinal complications. Among the six choroid locations, CT at the nasal side in cross-section (Choroid-CN) was the thinnest. This result is compatible with those of El-Shalzly et.al. (9) and Gupta et al. (13) which have demonstrated that CT at the nasal side was thinner in patients with myopia and emmetropia. Furthermore, CT at the nasal side also has the highest negative correlation with AXL both in Pearson and Lasso analyses. Although the exact mechanism requires further investigation, this result demonstrated that CT at the nasal side is essential for AXL prediction and possibly essential in myopia development.

Regarding myopia prediction, Varadarajan et al. (37) developed a model to predict SE from color fundus images. Shi et al. used CNN to predict myopia with absolute mean error of 1.115 D in SE from a color fundus image (38). We chose AXL classifications as our prediction. Our patients' average age was ≈

TABLE 5 | Superior performance in binary classification.

Classifier	Algorithm	Hyper. Opt.	Over sampling	Accuracy	Recall	PPV	NPV	F1-score	Specificity	AUC
1	SVM	Random	ROS	92.96%	100%	73.68%	100%	84.85%	91.22%	95.61%
2	AdaBoost	Random	ADASYN	94.37%	92.86%	81.25%	98.18%	86.67%	94.73%	93.80%
3	AdaBoost	Hyperopt	ROS	92.30%	71.43%	90.91%	93.33%	80.00%	98.25%	84.84%

TABLE 6 | Superior performances in multiclass classification.

Classifier	Algorithm	Hyper. Opt.	Over sampling	Accuracy	Recall (weighted)	PPV (weighted)	NPV (weighted)	F1-score (weighted)	Specificity (weighted)	AUC (macro)
4	SVM	Random	SMOTE	78.87%	78,87%	92.21%	62.56%	83.17%	93.37%	88.71%
5	AdaBoost	Grid	ROS	88.73%	88.73%	86.16%	82.28%	87.43%	74.75%	93.06%
6	XGBoost	Grid	SMOTE	85.92%	85.92%	83.06%	78.89%	84.32%	65.07%	93.42%
7	XGBoost	Random	ROS	87.32%	87.32%	84.96%	85.83%	85.78%	68.27%	84.64%

TABLE 7 | The comparison and test results in binary classification.

	Item	Accuracy	Recall	PPV	NPV	F1-score	Specificity	AUC
Student 1	Metric	80.28%	66.67%	52.63%	90.38%	58.82%	83.93%	75.30%
	p-value	0.006**	0.000***	0.000***	0.023*	0.000***	0.019*	0.001**
Student 2	Metric	56.34%	6.67%	5.56%	73.58%	6.06%	69.64%	38.15%
	p-value	0.000***	0.000***	0.000***	0.000***	0.000***	0.000***	0.000 ***
OPH 1	Metric	77.46%	60.00%	47.37%	88.46%	52.94%	82.14%	71.07%
	p-value	0.002**	0.000***	0.000***	0.010*	0.000***	0.010*	0.000***
OPH 2	Metric	80.28%	60.00%	52.94%	88.89%	56.25%	85.71%	72.86%
	p-value	0.006**	0.000***	0.000***	0.012*	0.000***	0.035*	0.000***

*p value: * < 0.05, ** < 0.01, *** < 0.001.*

TABLE 8 | The comparison and test results in multiclass classification.

	Item	Accuracy	Recall (weighted)	PPV (weighted)	NPV (weighted)	F1-score (weighted)	Specificity (weighted)	AUC (macro)
Student 1	Metric	67.61%	67.61%	70.60%	41.54%	64.92%	36.98%	52.29%
	p-value	0.001**	0.001**	0.012*	0.000***	0.001**	0.000***	0.000***
Student 2	Metric	47.89%	47.89%	60.86%	38.91%	53.59%	50.25%	49.07%
	p-value	0.000***	0.000***	0.000***	0.000***	0.000***	0.001**	0.000***
OPH 1	Metric	53.52%	53.52%	63.76%	41.16%	57.21%	53.00%	53.26%
	p-value	0.000***	0.000***	0.001**	0.000***	0.000***	0.003**	0.000***
OPH 2	Metric	66.20%	66.20%	72.85%	50.61%	69.04%	59.85%	63.03%
	p-value	0.001**	0.001**	0.025*	0.000 ***	0.004**	0.029*	0.000***

*p value: * < 0.05, ** < 0.01, *** < 0.001.*

66 years, and most of our participants had various severities of cataract; many of the participants had received cataract surgery. Because SE may be influenced by lens condition and cylinder, which are not directly related to retinal or choroidal condition, we considered AXL as a more accurate feature to reflect a patient's myopic condition. Because some of our patients had retinal disease, such as AMD, DME, retinoschisis, or myopic CNV, the retinal condition may vary between patients and even among multiple visits for the same patient. Thus, we recorded the CT to avoid the potential variation of retinal condition. Dong et al. (38) predicted AXL and subfoveal CT from a color fundus image with high accuracy (39). In their heat map analysis, they demonstrated that different areas of the macula on the fundus image were used to predict various AXL. The entire macular region, foveal region, and extrafoveal region were used to predict AXL $<$ 22 mm, from 22 to 26 mm, and $>$ 26 mm, respectively. In our study, we measured the CT at fovea and perifovea, and predicted AXL from the features. Among all of the CTs at various positions, the nasal side had the highest correlation with AXL, followed by the central part, and the result was unanimous in binary and multiclass classifications. Compatible with that of Dong et al. our results also demonstrated that the CT at the fovea and perifoveal region can predict AXL in various classifications.

In classifier construction, feature scaling is an essential preprocessing step in AI. Before one evaluates and selects the features, all features must be standardized to prevent redundancy or duplication. We used Pearson, VIF, and Lasso analyses to select the proper features. Pearson's r indicates the linear relationship between a given feature and class label. The p value indicates the probability that a feature is uncorrelated with the class label, per the method of Kowalski (40). The VIF measures how substantially the variance of an independent feature is influenced by other independent features. If the VIF of the target feature was $>$ 10, we eliminated the target feature. The Lasso method was proposed by Santosa and Symes (41) and popularized by Tibshirani (42). Lasso performs covariate selection by forcing the sum of the absolute value of the regression coefficients to be less than a fixed value, which forces certain coefficients to be 0.

In hyperparameter optimization, Random Search is more effective than Grid Search for a fixed search number; (36) Hyperopt obtains superior values of hyperparameters within the same executing time (43). Our data set was imbalanced both in binary and multiclass classifications, and the use of such imbalanced data to train the model may yield a biased result. The method we used, oversampling, is a popular technique for treating imbalanced data to avoid the aforementioned problems.

The ensemble-learning approach has also been used in ophthalmology to diagnose DR and interpret OCT imaging (44, 45). In our study, we used 5 algorithms (BNN, SVM, RF, AdaBoost, and XGBoost) to construct the classifiers of axial length through CTs. The classifiers constructed by ensemble approach (RF, AdaBoost, and XGBoost) outperformed those constructed by single machine learning approach (BNN and SVM). Those constructed by AdaBoost (Classifiers 1, 2, 4, and 5) and XGBoost (Classifiers 3 and 6) had the most optimal performance. Essentially, AdaBoost and XGBoost features the combination of several weak classifiers to form a strong one with a boosting approach. For the 2 algorithms (AdaBoost and XGBoost), the boosting approach plays a crucial role in dealing with the bias-variance tradeoff, and the boosting approach is considered more effective.

In this study, we successfully conducted AXL classification at the accuracies of 94.34 and 88.73% for binary and multiclass classifications by hyperparameter optimization, oversampling, and boosting algorithms. The high prediction accuracy in our binary and multiclass classification could be attributed to two main reasons. First, all of our imputed CT were repeated measured and rechecked by an ophthalmologist familiar with OCT images to ensure the accuracy of each measurement and avoid segmentation errors. Second, the seven final classifiers were chosen from 8,518,200 candidates [from 425,910 searches, each went through 20 (4 oversampling and five cross-validation) complete experiments], thus enabled our model to have high accuracies.

This study has several limitations. First, our sample size was relatively small, especially those with AXL $<$ 22 mm or $>$ 26 mm; the distribution of AXLs is also relatively imbalanced. Second, the process of collecting the 10 features was time consuming. Considerable time and effort were required to measure the thickness of the choroid of 6 positions from 2 OCT images of each eye and to collect data on sex, age, weight, and height of each of the participants. Among the features we recorded, AXL did not increase with age after adulthood, and weight may change without change of AXL; these potentially induced bias in our results. Third, the manual measurement of CT may cause bias or inconsistency. Future studies should address these limitations, and we expect to conduct more investigations using a larger data set on the classification and diagnosis of eye diseases which may be revealed by SD-OCT.

We demonstrated the effectiveness of the proposed classifiers in classification prediction from medical data and provided an assistance tool for physicians.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Fu Jen Catholic University Hospital, Fu Jen Catholic University. Written informed consent for participation was not required for this study in accordance with the national legislation and the institutional requirements.

AUTHOR CONTRIBUTIONS

CY-T and H-CL: conceptualization and writing – original preparation. H-CL: methodology and software. C-YT, H-YC, and C-JH: data collection. H-CL and P-HC: formal analysis. C-YT, H-YC, L-SW, and P-HC: writing – review and edition. C-YT: supervision. H-CL, L-SW, and P-HC: funding acquisition.

All authors have read and agreed to the published version of the manuscript.

FUNDING

This study was supported by the Ministry of Science and Technology of Taiwan under the grants MOST 108-2410-H-030-078-MY2 and MOST 110-2410-H-182-008-MY3, grants from Fu Jen Catholic University (No. 910I112) and Fu Jen Catholic University Hospital (No. PL-20200800X).

ACKNOWLEDGMENTS

We thank Jia-Hao Jhan and Rui-Xuan Su for their assistance in data collection and analysis.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmed.2022.850284/full#supplementary-material

REFERENCES

1. Holden BA, Fricke TR, Wilson DA, Jong M, Naidoo KS, Sankaridurg P, et al. Global prevalence of myopia and high myopia and temporal trends from 2000 through 2050. *Ophthalmology.* (2016) 123:1036–42.
2. Tsai TH, Liu YL, Ma IH, Su CC, Lin CW, Lin LL, et al. Evolution of the prevalence of myopia among Taiwanese schoolchildren: a review of survey data from 1983 through 2017. *Ophthalmology.* (2021) 128:290–301. doi: 10.1016/j.ophtha.2020.07.017
3. Duan F, Yuan Z, Deng J, Wong YL, Yeo AC, Chen X. Choroidal thickness and associated factors among adult myopia: a baseline report from a medical university student cohort. *Ophthalmic Epidemiol.* (2019) 26:244–50. doi: 10.1080/09286586.2019.1597899
4. Tideman JW, Snabel MC, Tedja MS, van Rijn GA, Wong KT, Kuijpers RW, et al. Association of axial length with risk of uncorrectable visual impairment for Europeans with myopia. *JAMA Ophthalmol.* (2016) 134:1355–63. doi: 10.1001/jamaophthalmol.2016.4009
5. Saw SM, Gazzard G, Shih-Yen EC, Chua WH. Myopia and associated pathological complications. *Ophthalmic Physiol Opt.* (2005) 25:381–91. doi: 10.1111/j.1475-1313.2005.00298.x
6. Morgan IG, Ohno-Matsui K, Saw SM. Myopia. *Lancet.* (2012) 379:1739–48.
7. Bartol-Puyal FA, Isanta C, Ruiz-Moreno O, Abadia B, Calvo P, Pablo L. Distribution of choroidal thinning in high myopia, diabetes mellitus, and aging: a swept-source OCT study. *J Ophthalmol.* (2019) 2019:3567813. doi: 10.1155/2019/3567813
8. Ikuno Y, Kawaguchi K, Nouchi T, Yasuno Y. Choroidal thickness in healthy Japanese subjects. *Invest Ophthalmol Vis Sci.* (2010) 51:2173–6.
9. El-Shazly AA, Farweez YA, ElSebaay ME, El-Zawahry WMA. Correlation between choroidal thickness and degree of myopia assessed with enhanced depth imaging optical coherence tomography. *Eur J Ophthalmol.* (2017) 27:577–84. doi: 10.5301/ejo.5000936
10. Wang S, Wang Y, Gao X, Qian N, Zhuo Y. Choroidal thickness and high myopia: a cross-sectional study and meta-analysis. *BMC Ophthalmol.* (2015) 15:70. doi: 10.1186/s12886-015-0059-2
11. Usui S, Ikuno Y, Miki A, Matsushita K, Yasuno Y, Nishida K. Evaluation of the choroidal thickness using high-penetration optical coherence tomography with long wavelength in highly myopic normal-tension glaucoma. *Am J Ophthalmol.* (2012) 153:10–6.e1. doi: 10.1016/j.ajo.2011.05.037
12. Flores-Moreno I, Lugo F, Duker JS, Ruiz-Moreno JM. The relationship between axial length and choroidal thickness in eyes with high myopia. *Am J Ophthalmol.* (2013) 155:314–9.e1. doi: 10.1016/j.ajo.2012.07.015
13. Gupta P, Saw SM, Cheung CY, Girard MJ, Mari JM, Bhargava M, et al. Choroidal thickness and high myopia: a case-control study of young Chinese men in Singapore. *Acta Ophthalmol.* (2015) 93:e585–92. doi: 10.1111/aos.12631
14. Gupta P, Cheung CY, Saw SM, Koh V, Tan M, Yang A, et al. Choroidal thickness does not predict visual acuity in young high myopes. *Acta Ophthalmol.* (2016) 94:e709–15. doi: 10.1111/aos.13084
15. Wei WB, Xu L, Jonas JB, Shao L, Du KF, Wang S, et al. Subfoveal choroidal thickness: the Beijing eye study. *Ophthalmology.* (2013) 120:175–80.
16. Yang WH, Zheng B, Wu MN, Zhu SJ, Fei FQ, Weng M, et al. An evaluation system of fundus photograph-based intelligent diagnostic technology for diabetic retinopathy and applicability for research. *Diabetes Ther.* (2019) 10:1811–22. doi: 10.1007/s13300-019-0652-0
17. Takahashi H, Tampo H, Arai Y, Inoue Y, Kawashima H. Applying artificial intelligence to disease staging: deep learning for improved staging of diabetic retinopathy. *PLoS One.* (2017) 12:e0179790. doi: 10.1371/journal.pone.0179790
18. Grassmann F, Mengelkamp J, Brandl C, Harsch S, Zimmermann ME, Linkohr B, et al. A deep learning algorithm for prediction of age-related eye disease study severity scale for age-related macular degeneration from color fundus photography. *Ophthalmology.* (2018) 125:1410–20. doi: 10.1016/j.ophtha.2018.02.037
19. Dong L, Yang Q, Zhang RH, Wei WB. Artificial intelligence for the detection of age-related macular degeneration in color fundus photographs: a systematic review and meta-analysis. *EClinicalMedicine.* (2021) 35:100875. doi: 10.1016/j.eclinm.2021.100875
20. Omodaka K, An G, Tsuda S, Shiga Y, Takada N, Kikawa T, et al. Classification of optic disc shape in glaucoma using machine learning based on quantified ocular parameters. *PLoS One.* (2017) 12:e0190012. doi: 10.1371/journal.pone.0190012
21. Asaoka R, Murata H, Iwase A, Araie M. Detecting preperimetric glaucoma with standard automated perimetry using a deep learning classifier. *Ophthalmology.* (2016) 123:1974–80. doi: 10.1016/j.ophtha.2016.05.029
22. Hemelings R, Elen B, Blaschko MB, Jacob J, Stalmans I, De Boever P. Pathological myopia classification with simultaneous lesion segmentation using deep learning. *Comput Methods Programs Biomed.* (2021) 199:105920. doi: 10.1016/j.cmpb.2020.105920
23. Israelsen NM, Petersen CR, Barh A, Jain D, Jensen M, Hannesschläger G, et al. Real-time high-resolution mid-infrared optical coherence tomography. *Light Sci Appl.* (2019) 8:11. doi: 10.1038/s41377-019-0122-5
24. Waldstein SM, Vogl WD, Bogunovic H, Sadeghipour A, Riedl S, Schmidt-Erfurth U. Characterization of drusen and hyperreflective foci as biomarkers for disease progression in age-related macular degeneration using artificial intelligence in optical coherence tomography. *JAMA Ophthalmol.* (2020) 138:740–7. doi: 10.1001/jamaophthalmol.2020.1376
25. Kim YC, Chang DJ, Park SJ, Choi Y, Gong Y, Kim H, et al. Machine learning prediction of pathologic myopia using tomographic elevation of the posterior sclera. *Sci Rep.* (2021) 11:6950. doi: 10.1038/s41598-021-85699-0
26. Choi KJ, Choi JE, Roh HC, Eun JS, Kim JM, Shin YK, et al. Deep learning models for screening of high myopia using optical coherence tomography. *Sci Rep.* (2021) 11:21663. doi: 10.1038/s41598-021-00622-x
27. Li Y, Feng W, Zhao X, Liu B, Zhang Y, Chi W, et al. Development and validation of a deep learning system to screen vision-threatening conditions in high myopia using optical coherence tomography images. *Br J Ophthalmol.* (2020) 106:633–9. doi: 10.1136/bjophthalmol-2020-317825
28. Waldstein S, Faatz H, Szimacsek M, Glodan AM, Podkowinski D, Montuoro A, et al. Comparison of penetration depth in choroidal imaging using swept source vs spectral domain optical coherence tomography. *Eye (Lond).* (2015) 29:409–15. doi: 10.1038/eye.2014.319
29. Yilmaz I, Ozkaya A, Kocamaz M, Ahmet S, Ozkaya HM, Yasa D, et al. Correlation of choroidal thickness and body mass index. *Retina.* (2015) 35:2085–90. doi: 10.1097/IAE.0000000000000582
30. Wei WB, Zu L, Jonas JB, Shao L, Du KF, Wang S, et al. Subfoveal choroidal thickness: the Beijing eye study. *Ophthalmology.* (2013) 120:175–80.

31. Lee MW, Lee SE, Lim HB, Kim JY. Longitudinal changes in axial length in high myopia: a 4-year prospective study. *Br J Ophthalmol.* (2020) 104:600–3. doi: 10.1136/bjophthalmol-2019-314619
32. Lai TT, Yang CM. Lamellar hole-associated epiretinal proliferation in lamellar macular hole and full-thickness macular hole in high myopiA. *Retina.* (2018) 38:1316–23. doi: 10.1097/IAE.0000000000001708
33. Li M, Jin E, Dong C, Zhang C, Zhao M, Qu J. The repeatability of superficial retinal vessel density measurements in eyes with long axial length using optical coherence tomography angiography. *BMC Ophthalmol.* (2018) 18:326. doi: 10.1186/s12886-018-0992-y
34. Yang QH, Chen B, Peng GH, Li ZH, Huang YF. Accuracy of axial length measurements from immersion B-scan ultrasonography in highly myopic eyes. *Int J Ophthalmol.* (2014) 7:441–5.
35. Friedman J, Hastie T, Tibshirani R. Additive logistic regression: a statistical view of boosting (with discussion and a rejoinder by the authors). *Ann Statist.* (2000) 28:337–407.
36. Bergstra J, Bengio Y. Random search for hyper-parameter optimization. *J Mach Learn Res.* (2012) 13:281–305.
37. Varadarajan AV, Poplin R, Blumer K, Angermueller C, Ledsam J, Chopra R, et al. Deep learning for predicting refractive error from retinal fundus images. *Invest Ophthalmol Vis Sci.* (2018) 59:2861–8. doi: 10.1167/iovs.18-23887
38. Shi Z, Wang T, Huang Z, Xie F, Song G. A method for the automatic detection of myopia in Optos fundus images based on deep learning. *Int J Numer Method Biomed Eng.* (2021) 37:e3460. doi: 10.1002/cnm.3460
39. Dong L, Hu XY, Yan YN, Zhang Q, Zhou N, Shao L, et al. Deep learning-based estimation of axial length and subfoveal choroidal thickness from color fundus photographs. *Front Cell Dev Biol.* (2021) 9:653692.
40. Kowalski CJ. On the effects of non-normality on the distribution of the sample product-moment correlation coefficient. *J R Stat Soc Ser C Appl Stat.* (1972) 21:1–12.
41. Santosa F, Symes WW. Linear inversion of band-limited reflection seismograms. *J Sci Stat Comput.* (1986) 7:1307–30.
42. Tibshirani RJ. Regression shrinkage and selection via the lasso. *J R Stat Soc Ser B Appl Stat.* (1996) 58:267–88.
43. Bergstra J, Yamins D, Cox D. Making a science of model search: hyperparameter optimization in hundreds of dimensions for vision architectures. *Proc Int Conf Machine Learn.* (2013) 28:115–23.
44. Sikder N, Masud M, Bairagi AK, Arif ASM, Nahid A-A, Alhumyani HA. Severity classification of diabetic retinopathy using an ensemble learning algorithm through analyzing retinal images. *Symmetry.* (2021) 13:670.
45. Anoop BN, Pavan R, Girish GN, Kothari AR, Rajan J. Stack generalized deep ensemble learning for retinal layer segmentation in optical coherence tomography images. *Biocybern Biomed Eng.* (2020) 40:1343–58.

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers | Frontiers in Medicine

TYPE Original Research
PUBLISHED 25 August 2022
DOI 10.3389/fmed.2022.914504

OPEN ACCESS

EDITED BY
David Madrid-Costa,
Complutense University of Madrid, Spain

REVIEWED BY
Pablo De Gracia,
Midwestern University, United States
Neslihan Dilruba Koseoglu,
Cankiri State Hospital, Turkey

*CORRESPONDENCE
Tae-im Kim
TIKim@yuhs.ac

SPECIALTY SECTION
This article was submitted to Ophthalmology, a section of the journal Frontiers in Medicine

RECEIVED 06 April 2022
ACCEPTED 30 June 2022
PUBLISHED 25 August 2022

CITATION
Ahn H, Jun I, Seo KY, Kim EK and Kim T-i (2022) Femtosecond laser-assisted arcuate keratotomy for the management of corneal astigmatism in patients undergoing cataract surgery: Comparison with conventional cataract surgery. *Front. Med.* 9:914504. doi: 10.3389/fmed.2022.914504

Femtosecond laser-assisted arcuate keratotomy for the management of corneal astigmatism in patients undergoing cataract surgery: Comparison with conventional cataract surgery

Hyunmin Ahn[1], Ikhyun Jun[1,2], Kyoung Yul Seo[1], Eung Kweon Kim[2,3] and Tae-im Kim[1,2]*

[1]Department of Ophthalmology, Yonsei University College of Medicine, Seoul, South Korea, [2]Corneal Dystrophy Research Institute, Yonsei University College of Medicine, Seoul, South Korea, [3]Saevit Eye Hospital, Goyang, South Korea

Purpose: To assess the effects of femtosecond laser arcuate keratotomy with femtosecond laser-assisted cataract surgery in the management of corneal astigmatism, compared with conventional phacoemulsification cataract surgery.

Design: Retrospective comparative interventional case series.

Methods: A total of 2,498 eyes of consecutive patients who presented with 3.00 diopters (D) or under of astigmatism were included. The patients were treated with conventional phacoemulsification cataract surgery (conventional group) and femtosecond laser arcuate keratotomy with femtosecond laser-assisted cataract surgery (femtosecond group).

Results: Surgically induced astigmatism (SIA) was higher in the femtosecond group than the conventional group (0.215, $p < 0.001$). Difference vector (DV) was lower in the femtosecond group (-0.136, $p < 0.001$). The cut-off value of the overcorrection in the femtosecond group was 0.752 D of target induced astigmatism (TIA). For patients with TIA 0.75 D or under, DV and the value of index of success (TIA into DV) were significantly higher in the femtosecond group ($p = 0.022$ and < 0.001). The overcorrection ratios were 48.8% in the conventional and 58.9% in the femtosecond group. ($p < 0.001$). For patients with TIA over 0.75 D, SIA and correction index (TIA into SIA) was higher in femtosecond group (0.310 and 0.250, $p < 0.001$ and < 0.001, respectively). Absolute angle of error was 20.612 ± 18.497 in the femtosecond group and higher than the conventional group (2.778, $p = 0.010$).

Conclusion: Femtosecond laser arcuate keratotomy in cataract surgery was effective in SIA between 0.75 to 3.00 D of corneal astigmatism. However,

the overcorrection in the lower astigmatism and angle of error in the higher astigmatism were due to the postoperative corneal astigmatism not decreasing as much as SIA. Overcoming these challenges will lead to better management of corneal astigmatism.

KEYWORDS

arcuate keratotomy, corneal astigmatism, cataract surgery, femtosecond (fs) laser, keratotomy

Introduction

Cataract is the leading cause of blindness worldwide (1), and cataract surgery is among the most common procedures performed in the United States, with more than 30 million patients undergoing surgery each year (2). Phacoemulsification surgery is currently the standard method of treatment for patients with cataracts (1); however, advanced techniques involving the use of multifocal/toric intraocular lenses and other technologies, such as femtosecond laser-assisted cataract surgery (FLACS), have recently become commercially available (3–5).

Corneal astigmatism is a common consideration in cataract surgery, with approximately 40% of patients having astigmatism of more than 1.0 diopter (D) (6). Research has indicated that correction of corneal astigmatism yields better refractive outcomes following cataract surgery (7). Typically, surgical correction of corneal astigmatism involves toric intraocular lens implantation and the creation of corneal arcuate/limbal relaxing incisions, which can be performed either manually or using a femtosecond laser (i.e., FLACS) (4, 8, 9).

Although most previous studies have demonstrated that femtosecond laser-assisted arcuate keratotomy (FL-AK) is effective for correcting corneal astigmatism, there have been several inconsistencies in their results (**Supplementary Table 1**). Moreover, these studies had insufficient cohorts to conduct a detailed analysis of group differences or determine the factors influencing these differences. As the small sample sizes resulted in a comparative analysis without adjustment for confounders (i.e., independent *t*-test), discrepancies between the results of these previous studies and the effects of corneal astigmatism reduction during FL-AK observed in actual clinical settings are possible.

To overcome these limitations, we performed a detailed and well-controlled analysis using a massive real-world dataset and a statistical method that considered relevant confounders. In this study, we aimed to (a) assess the effectiveness of FL-AK for the management of corneal astigmatism in patients undergoing cataract surgery when compared with conventional phacoemulsification cataract surgery, (b) determine the degrees of preoperative corneal astigmatism for which FL-AK is indicated, and (c) identify the factors related to the effectiveness of FL-AK.

Abbreviations: absAE, absolute value of angle of error; AE, angle of error; ATR, against-the-rule astigmatism; CI, correction index; D, diopters; DV, difference vector; FLACS, femtosecond laser-assisted cataract surgery; FL-AK, femtosecond laser-assisted arcuate keratotomy; IOS, index of success; MofE, magnitude of error; OBL, oblique astigmatism; SIA, surgically induced astigmatism; TIA, target induced astigmatism; WTR, with-the-rule astigmatism.

Materials and methods

Study design

This retrospective study was performed at Severance Hospital, Yonsei University College of Medicine, between January 2018 and June 2021. The Severance Hospital Clinical Research Ethics Committee approved the protocol of the study (IRB protocol number-4-2021-0525), which was conducted in accordance with the tenets of the Declaration of Helsinki.

Participants

The study included 2,498 eyes of 1,767 consecutive patients aged $\geq$ 45 years diagnosed with age-related cataracts and corneal astigmatism $\leq$ 3.00 D. The surgical method was determined by the patient's decision-making after the consent of the surgery. Among them, 1,325 eyes of 922 patients were treated with conventional phacoemulsification cataract surgery (conventional group), while 1,173 eyes of 845 patients were treated with FLACS combined with FL-AK (femtosecond group). Patients exhibiting poor compliance during examination and eyes with irregular corneal astigmatism, corneal opacities, previous corneal surgery (including corneal refractive surgery), acute or chronic ophthalmic diseases of the anterior segment, or intraoperative complications were excluded.

Measurements

All patients underwent a detailed preoperative ophthalmological evaluation, including slit-lamp and fundus examinations. Calculations of intraocular lens power were performed using optical biometry results (IOLMaster 700®; Carl Zeiss Meditec AG, Jena, Germany). Corneal measurements and planning of the arcuate keratotomy procedure were based on autokeratometry (Topcon KR-800A; Topcon Corporation, Tokyo, Japan) and Scheimpflug topography (Pentacam®; Oculus Inc., Wetzlar, Germany) findings within 2 weeks before surgery. Patients exhibiting a more than 0.3-D difference in mean preoperative corneal astigmatism and those exhibiting axis measurements with differences of 5° or more between autokeratometry and Scheimpflug topography were excluded. Corneal astigmatism was measured *via* autokeratometry 3 months after surgery to assess postoperative outcomes.

Vector analysis

Vector analysis for corneal astigmatism was conducted using the results of autokeratometry, in accordance with the Alpins method (10). Considering the changes in the astigmatic axis, three vectors were measured: the target-induced astigmatism vector (TIA), defined the astigmatic change that the surgery was intended to induce; the surgically induced astigmatism vector (SIA), defined as the geographic change in corneal astigmatism actually induced by the surgery; and the difference vector (DV), defined as the induced astigmatic change that would enable the initial surgery to achieve its intended target. The magnitude of error (MofE; SIA minus TIA), angle of error (AE, the degree of angle between the TIA and SIA vectors), absolute value of AE (absAE), correction index (CI, SIA divided by TIA), and index of success (IOS; DV divided by TIA) were also measured. The CI > 1 means that the SIA is greater than the TIA, overcorrection. The IOS > 1 means that the corneal astigmatism is increased after surgery, compared to before surgery.

The preoperative corneal astigmatism axis was converted to a range of 0–90 degrees ($Axis_{90}$) and used to classify eyes into three groups: with-the-rule (WTR) astigmatism (0–30 degrees), oblique (OBL) astigmatism (30–60 degrees), and against-the-rule (ATR) astigmatism (60–90 degrees).

Femtosecond laser system and surgical technique

In the femtosecond group, FLACS and FL-AK were performed using the LenSx® femtosecond system (Alcon laboratory, Inc., Texas, United States). Intra-operative alignment of the corneal astigmatism axis was performed using a corneal topography system (Verion® image guide system; Alcon laboratory, Inc.; Texas, United States) immediately before surgery. A single arcuate keratotomy incision was paired in the opposite meridian. The depth of the astigmatic keratotomy was set at 80% corneal thickness according to a modified Donnenfeld limbal relaxing nomogram (11–14). The diameter of the optical zone was set to 8.0 mm. For FLACS, the laser was also used to perform a 5.0–5.3-mm capsulotomy and lens fragmentation. The femtosecond system was not used to make the main phacoemulsification incision or the peripheral incision.

In both groups, the clear corneal incision at the temporal side was created using a 2.80-mm keratome, and the anterior capsule button was removed. In the conventional group, capsulotomy was performed using forceps or a capsulotomy needle. Phacoemulsification was performed under local anesthesia in both groups using an Infiniti® system (Alcon laboratory, Inc.; Texas, United States). All operations were performed by an experienced surgeon (T.I.K). Patients were instructed to perform 1 month of postoperative care, which included instillation of eye drops, in accordance with the standard protocol for cataract surgery. Patients who did not follow the surgical instructions were excluded.

Statistical analysis

The data were analyzed using descriptive statistics, and the mean values and standard deviations were computed for

TABLE 1 Characteristics of patients in the conventional and femtosecond groups.

	Conventional group (n = 1,325)	Femtosecond group (n = 1,173)	p
Age	66.47 ± 10.94	66.71 ± 12.78	0.623
Sex (female)	824 (62.2%)	769 (65.6%)	0.087
Laterality (right, %)	674 (50.9%)	599 (51.1%)	0.936
Preoperative corneal refractive power (Km, D)	44.14 ± 1.88	44.20 ± 1.75	0.328
Preoperative corneal astigmatism	0.85 ± 0.58	0.85 ± 0.58	0.983
Preoperative corneal astigmatism axis ($Axis_{90}$)	43.45 ± 31.78	45.49 ± 31.90	0.085
Preoperative corneal astigmatism group			0.404
Against-the-rule	580 (43.8%)	487 (41.5%)	–
Oblique	228 (17.2%)	198 (16.9%)	–
With-the-rule	517 (39.0%)	488 (41.6%)	–

each variable. Differences between the groups were initially assessed using independent *t*-tests, which were adjusted for age, sex, laterality, preoperative corneal refractive power (Km), preoperative corneal astigmatism axis, and TIA. Linear regression analyses were used examine the association between TIA and SIA in each group. Using the cut-off value for overcorrection (i.e., SIA > TIA) identified in the linear regression analysis, stratified analyses were conducted for patients in the femtosecond group. Lower and higher TIA were defined as TIA values under and over the cut-off value, respectively. *P*-values < 0.05 were considered statistically significant.

TABLE 2 Vector analysis of postoperative corneal astigmatism.

A. Overall.

	Conventional (*n* = 1,325)		Femtosecond (*n* = 1,173)				Adjusted†	Adjusted†
	Mean	*SD*	Mean	*SD*	Diff	*p*	Diff	*p*
TIA	0.845	0.579	0.846	0.578	0.001	0.983	-	-
SIA	0.631	0.494	0.886	0.819	0.254	<0.001*	0.215	<0.001*
DV	0.913	0.823	0.803	0.772	-0.110	<0.001*	-0.136	<0.001*
AE	0.434	27.771	1.063	30.158	0.628	0.596	0.192	0.903
absAE	21.041	18.120	22.673	19.903	1.632	0.036*	1.634	0.033*
MofE	-0.135	0.777	0.041	0.862	0.175	<0.001*	0.215	<0.001*
CI	0.748	0.823	1.037	1.075	0.289	<0.001*	0.319	<0.001*
IOS	1.082	1.052	0.951	1.012	-0.131	0.002*	-0.047	0.668

B. Preoperative astigmatism ≤ 0.75 D.

	Conventional (*n* = 813)		Femtosecond (*n* = 693)				Adjusted†	Adjusted†
	Mean	*SD*	Mean	*SD*	Diff	*p*	Diff	*p*
TIA	0.486	0.214	0.485	0.219	–0.001	0.998	–	–
SIA	0.607	0.658	0.728	0.688	0.121	0.001*	0.234	<0.001*
DV	0.700	0.676	0.713	0.638	0.012	0.719	0.099	0.022*
AE	1.021	30.733	0.229	31.790	–0.792	0.623	–1.209	0.311
absAE	24.356	18.961	24.098	20.714	–0.258	0.378	–0.864	0.534
MofE	0.122	0.684	0.266	0.702	0.145	<0.001*	0.234	<0.001*
CI	1.418	1.685	1.776	2.018	0.358	<0.001*	0.316	<0.001*
IOS	1.598	1.685	1.744	1.965	0.146	0.132	0.249	<0.001*

C. Preoperative astigmatism > 0.75 D.

	Conventional (*n* = 512)		Femtosecond (*n* = 480)				Adjusted†	Adjusted†
	Mean	*SD*	Mean	*SD*	Diff	*p*	Diff	*p*
TIA	1.395	0.420	1.400	0.461	–0.004	0.879	–	–
SIA	0.790	0.591	1.114	0.932	0.324	<0.001*	0.310	<0.001*
DV	1.170	0.631	0.933	0.842	–0.237	<0.001*	-0.203	<0.001*
AE	0.242	24.321	2.268	27.618	2.026	0.133	0.954	0.670
absAE	18.224	16.087	20.612	18.497	2.388	0.016*	2.778	0.010*
MofE	–0.608	0.993	–0.286	0.963	–0.322	<0.001*	–0.310	<0.001*
CI	0.596	0.462	0.840	0.761	0.244	0.001*	0.250	<0.001*
IOS	0.848	0.442	0.668	0.599	–0.182	<0.001*	–0.186	<0.001*

*Statistically significant.
†Adjusted for age, sex, laterality, surgeon, Km, preoperative corneal astigmatism axis, and TIA.
absAE, absolute angle of error; AE, angle of error; CI, correction index; DV, difference vector; IOS, index of success; MofE, magnitude of error; SIA, surgically induced astigmatism; TIA, target induced astigmatism.

Results

The two groups exhibited no significant differences in baseline demographic or ophthalmological characteristics, including age, sex, laterality distribution, preoperative corneal refractive power (Km), preoperative corneal astigmatism, and axis (Table 1).

TIA was 0.845 ± 0.579 D in the conventional group and 0.846 ± 0.578 D in the femtosecond group (p = 0.983). However, SIA was significantly higher in the femtosecond group than in the conventional group (adjusted difference = 0.215, $p < 0.001$), whereas DV was significantly lower in the femtosecond group (adjusted difference = -0.136, $p < 0.001$). IOS did not significantly differ between the groups (adjusted difference = -0.047, $p = 0.668$) (Table 2A).

The cut-off TIA value for overcorrection in the femtosecond group was 0.752 (95% confidence interval: 0.512–0.992). The linear regression Equation between SIA and TIA in the femtosecond group was as follows (Supplementary Image 1):

$$SIA = 0.457 + 0.392\ TIA\ (R = 0.272, p < 0.001)$$

SIA and CI values were higher in the femtosecond group than in the conventional group among patients with both lower and higher TIA (all $p < 0.001$). In patients with lower TIA, the mean DV value was higher in the femtosecond group than in the conventional group (adjusted difference = 0.099, p = 0.022). The MofE values were 0.122 ± 0.684 D and 0.266 ± 0.702 D in the conventional and femtosecond groups, respectively (adjusted difference = 0.145, $p < 0.001$) (Table 2B). In patients with higher TIA, the DV value was significantly lower in the femtosecond group than in the conventional group (adjusted difference = -0.203, $p < 0.001$). The femtosecond group also had higher absolute AE values than the conventional group (adjusted difference = 2.778, p = 0.010) and an IOS of 0.668 ± 0.599 (Table 2C).

For both patients with lower and higher TIA, the overcorrection ratio was significantly higher in the femtosecond group than in the conventional group ($p < 0.001$ and < 0.001, respectively) (Table 3A). However, the overcorrection ratio did not significantly differ among the preoperative corneal astigmatism axis subgroups (ATR vs. OBL vs. WTR) when the analysis was restricted to patients of the femtosecond group with lower TIA (p = 0.643) (Table 3B).

A linear regression analysis adjusted for TIA and SIA indicated that absolute AE was significantly associated with DV among patients in the femtosecond group with higher TIA [B = 0.014 (95% confidence interval: 0.011–0.018), $p < 0.001$] (Table 4).

Overall, 72.0% of the patients in the femtosecond group exhibited decreased corneal astigmatism postoperatively. When compared with the preoperative values, 8.4% exhibited a decrease of more than 75%, while 22.9% exhibited a decrease of more than 50% (Supplementary Image 2). This Venn diagram shows the distribution of DV by TIA, SIA, and AE. Depending on AE, even with sufficient SIA compared to TIA, DV cannot reach TIA, and can even be greater than TIA.

Discussion

In this study, we investigated the effect of FL-AK in managing corneal astigmatism during cataract surgery using real-world data for 2,498 eyes. In the overall cohort, our findings indicate that the femtosecond group had higher SIA but

TABLE 3 Overcorrection ratio in stratified analyses for target induced astigmatism ≤ 0.75 diopters.

A. Treatment groups.

	Conventional	Femtosecond	Total
Overcorrection	397 (48.8%)	408 (58.9%)	805 (53.5%)
No overcorrection	416 (51.2%)	285 (41.1%)	701 (46.5%)
Total	813 (54.0%)	693 (46.0%)	1,506 (100%)
			$p < 0.001$*

B. Femtosecond subgroups based on preoperative corneal astigmatism axis.

	With-the-rule	Oblique	Against-the-rule	Total
Overcorrection	98 (40.2%)	58 (38.4%)	108 (36.2%)	285 (41.4%)
No overcorrection	146 (59.8%)	93 (61.6%)	190 (63.8%)	408 (58.9%)
Total	244 (35.2%)	151 (21.8%)	298 (43.0%)	693 (100%)
				$p = 0.643$

*Statistically significant.

TABLE 4 Linear regression analysis of the difference vector in patients of the femtosecond group with target induced astigmatism over 0.75 D (adjusted R^2 = 0.406, p < 0.001*).

	B	ß	95% Confidence Interval Lower	Upper	P	VIF
(Constant)	−0.054	–	−0.263	0.154	0.609	–
absAE	0.014	0.315	0.011	0.018	<0.001*	1.011
SIA	0.461	0.510	0.397	0.525	<0.001*	1.035
TIA	0.340	0.186	0.210	0.470	<0.001*	1.046

*Statistically significant.
absAE, absolute angle of error; SIA, surgically induced astigmatism; TIA, target induced astigmatism.

lower postoperative corneal astigmatism than the conventional group. However, the MofE (TIA subtracted from SIA) was positive, and the CI value (SIA divided by TIA) was over 1, indicating that overcorrection was common in the femtosecond group. The difference in SIA between the conventional and femtosecond groups was 0.254, but the absolute difference in postoperative corneal astigmatism (0.110. SIA) was higher in the femtosecond group; however, this effect was not fully reflected in the degree of postoperative corneal astigmatism. Linear regression analysis between TIA and SIA also indicated that the TIA cut-off value for overcorrection was 0.752 D, which is within the range reported in previous studies (15, 16).

In patients with TIA values of 0.75 D or under, FL-AK induced significant overcorrection and reductions in corneal astigmatism when compared with the conventional method. In a previous study that also included patients with relatively lower preoperative corneal astigmatism, overcorrection was related to the preoperative corneal astigmatism axis (17), as observed in the current study. However, the overcorrection rates did not significantly differ between the groups (p = 0.643). In another study, the authors reported that, in patients with a relatively lower degree of corneal astigmatism, FL-AK outcomes were influenced by preoperative corneal astigmatism and uncorrected visual acuity (18). The authors of that study utilized a novel formula to reduce the corneal incision arc by 20–30%, and the novel formula was more effective in correcting low astigmatism than the pre-existing method. Thus, reducing the corneal incision arc indicated by the existing formula may aid in lowering preoperative astigmatism in these patients.

In patients with TIA values > 0.75 D, the goal indicated by the FL-AK nomogram was 70–80% correction of preoperative corneal astigmatism, and SIA values indicated that 84% correction had been achieved. However, approximately 67% of preoperative corneal astigmatism remained in the FL-AK group, similar to findings reported in previous studies (0.47–0.71)

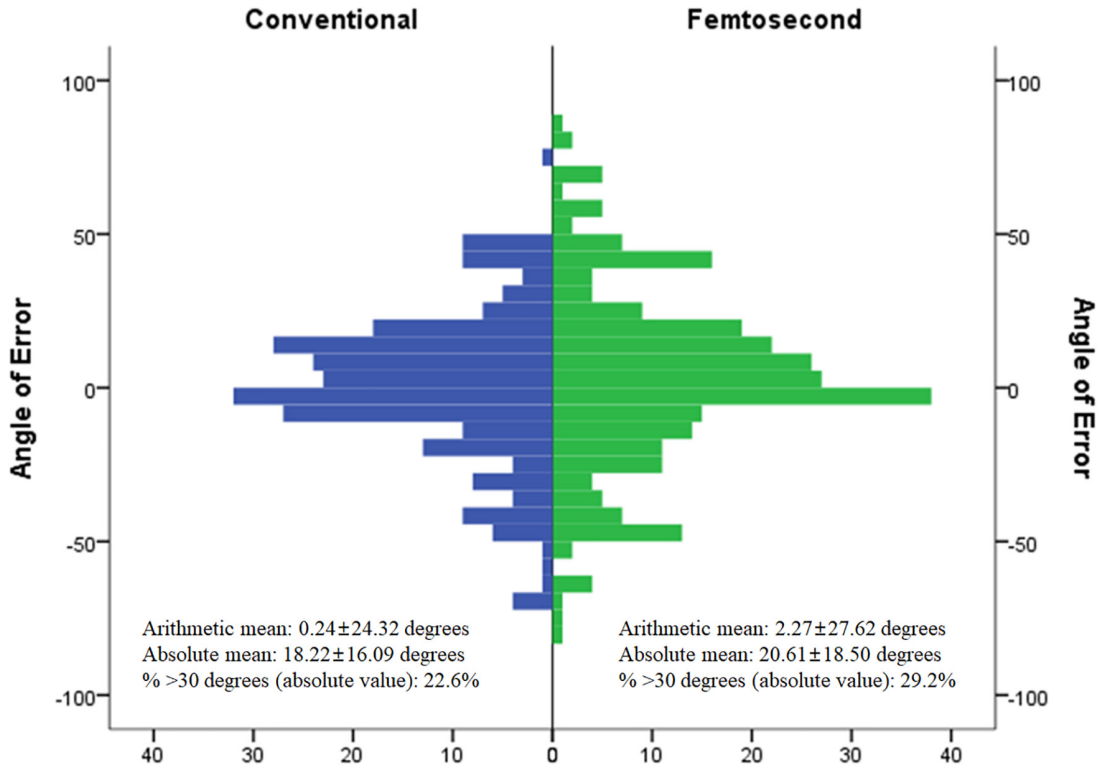

FIGURE 1
Angle of error between conventional group and femtosecond group in target induced astigmatism over 0.75 D.

(11, 15, 19, 20). IOS reflects the amount postoperative corneal astigmatism remaining when compared with the preoperative state. We focused on IOS because we assumed it to underlie the central conflict between the effects of FL-AK reported in previous studies (which focused on SIA) and the actual clinical situation. In the current study, postoperative corneal astigmatism was not defined based on the arithmetic difference between preoperative corneal astigmatism and SIA (21). Our findings suggest that AE (i.e., the angle between TIA and SIA) was the primary cause of these discrepancies (**Figure 1**). The absAE in the femtosecond group was 20.6, which is within the range of 17.5–25.1 reported in previous studies (9, 15). Considering the relationships among the vectors (TIA, SIA, and DV), the evidence indicates that AE is among the major factors influencing postoperative corneal astigmatism in patients treated with FL-AK during cataract surgery (10, 21).

In previous studies, torque was regarded as an ineffective component of the SIA vector (21–23), and the direction of SIA was tilted to the induced direction. Moreover, several studies have highlighted the importance of the reference axis in arcuate keratotomy (24, 25). Further studies are required to determine the precise factors affecting AE and to develop a novel nomogram incorporating AE, as this will help to improve the effectiveness of FL-AK in patients undergoing FLACS.

In conclusion, our findings indicate that cataract surgery with FL-AK resulted in significantly increased SIA but that it was effective in correcting preoperative corneal astigmatism > 0.75 D when compared with conventional phacoemulsification cataract surgery. However, overcorrection in patients with a lower degree of astigmatism and the angle of error in patients with higher astigmatism may have inhibited improvements in postoperative corneal astigmatism. Future studies should aim to overcome these challenges to achieve better efficacy in managing corneal astigmatism.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethics statement

The studies involving human participants were reviewed and approved by the Severance Hospital Clinical Research Ethics Committee. Written informed consent for participation was not required for this study in accordance with the national legislation and the institutional requirements.

Author contributions

T-IK conceived and designed the analysis. HA collected the data, performed the analysis, and wrote the manuscript. IJ, KS, EK, and T-IK contributed to the data. All authors contributed to the article and approved the submitted version.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmed.2022.914504/full#supplementary-material

References

1. Bourne RR, Stevens GA, White RA, Smith JL, Flaxman SR, Price H, et al. Causes of vision loss worldwide, 1990–2010: a systematic analysis. *Lancet Glob Health.* (2013) 1:e339–49. doi: 10.1016/S2214-109X(13)70113-X

2. Centers for Medicare and Medicaid Services. *CY 2019 Top 200 Level II HCPCS Codes Ranked by Charges (PDF).* (2020). Available online at: https://www.cms.gov/files/document/cy-2019-top-200-level-ii-hcpcs-codes-ranked-charges.pdf (accessed May 3, 2020).

3. Khandelwal SS, Jun JJ, Mak S, Booth MS, Shekelle PG. Effectiveness of multifocal and monofocal intraocular lenses for cataract surgery and lens replacement: a systematic review and meta-analysis. *Graefes Arch Clin Exp Ophthalmol.* (2019) 257:863–75. doi: 10.1007/s00417-018-04218-6

4. Kessel L, Andresen J, Tendal B, Erngaard D, Flesner P, Hjortdal J. Toric intraocular lenses in the correction of astigmatism during cataract surgery: a systematic review and meta-analysis. *Ophthalmology.* (2016) 123:275–86. doi: 10.1016/j.ophtha.2015.10.002

5. Day AC, Burr JM, Bennett K, Bunce C, Doré CJ, Rubin GS, et al. Femtosecond laser-assisted cataract surgery versus phacoemulsification cataract surgery (FACT): a randomized noninferiority trial. *Ophthalmology.* (2020) 127:1012–9.

6. Khan MI, Muhtaseb M. Prevalence of corneal astigmatism in patients having routine cataract surgery at a teaching hospital in the United Kingdom. *J Cataract Refract Surg.* (2011) 37:1751–5. doi: 10.1016/j.jcrs.2011.04.026

7. Wolffsohn JS, Bhogal G, Shah S. Effect of uncorrected astigmatism on vision. *J Cataract Refract Surg.* (2011) 37:454–60. doi: 10.1016/j.jcrs.2010.09.022

8. Roberts HW, Wagh VK, Sullivan DL, Archer TJ, O'Brart DP. Refractive outcomes after limbal relaxing incisions or femtosecond laser arcuate keratotomy to manage corneal astigmatism at the time of cataract surgery. *J Cataract Refract Surg.* (2018) 44:955–63. doi: 10.1016/j.jcrs.2018.05.027

9. Schwarzenbacher L, Schartmüller D, Röggla V, Meyer E, Leydolt C, Menapace R. One-year results of arcuate keratotomy in patients with low to moderate corneal astigmatism using a low-pulse-energy femtosecond laser. *Am J Ophthalmol.* (2021) 224:53–65. doi: 10.1016/j.ajo.2020.11.018

10. Alpins N. Astigmatism analysis by the alpins method. *J Cataract Refract Surg.* (2001) 27:31–49. doi: 10.1016/S0886-3350(00)00798-7

11. Ganesh S, Brar S, Arra RR. Comparison of astigmatism correction between anterior penetrating and intrastromal arcuate incisions in eyes undergoing femtosecond laser–assisted cataract surgery. *J Cataract Refract Surg.* (2020) 46:394–402. doi: 10.1097/j.jcrs.0000000000000069

12. Hiep NX, Khanh PTM, Quyet D, Van Thai T, Nga VT, Dinh TC, et al. Correcting corneal astigmatism with corneal arcuate incisions during femtosecond laser assisted cataract surgery. *Open Access Maced J Med Sci.* (2019) 7:4260.

13. Lopes D, Loureiro T, Carreira R, Barros SR, Cardoso JN, Campos P, et al. Transepithelial or intrastromal femtosecond laser arcuate keratotomy to manage corneal astigmatism at the time of cataract surgery. *Arch Soc Esp Oftalmol (English Edition).* (2021) 96:408–14. doi: 10.1016/j.oftale.2020.09.008

14. Donnenfeld E, Rosenberg E. Assisting femto incisions with nomograms. Assisting femto incisions with nomograms. *Ophthalmol Manage.* (2015) 19:48–52.

15. Chan TC, Cheng GP, Wang Z, Tham CC, Woo VC, Jhanji V. Vector analysis of corneal astigmatism after combined femtosecond-assisted phacoemulsification and arcuate keratotomy. *Am J Ophthalmol.* (2015) 160:250.–255. doi: 10.1016/j.ajo.2015.05.004

16. Visco DM, Bedi R, Packer M. Femtosecond laser–assisted arcuate keratotomy at the time of cataract surgery for the management of preexisting astigmatism. *J Cataract Refract Surg.* (2019) 45:1762–9. doi: 10.1016/j.jcrs.2019.08.002

17. Baharozian CJ, Song C, Hatch KM, Talamo JH. A novel nomogram for the treatment of astigmatism with femtosecond-laser arcuate incisions at the time of cataract surgery. *Clin Ophthalmol (Auckland NZ).* (2017) 11:1841. doi: 10.2147/OPTH.S141255

18. Wortz G, Gupta PK, Goernert P, Hartley C, Wortz B, Chiu J, et al. Outcomes of femtosecond laser arcuate incisions in the treatment of low corneal astigmatism. *Clin Ophthalmol (Auckland NZ).* (2020) 14:2229. doi: 10.2147/OPTH.S264370

19. Chan TC, Ng AL, Cheng GP, Wang Z, Woo VC, Jhanji V. Corneal astigmatism and aberrations after combined femtosecond-assisted phacoemulsification and arcuate keratotomy: two-year results. *Am J Ophthalmol.* (2016) 170:83–90. doi: 10.1016/j.ajo.2016.07.022

20. Chan TC, Ng AL, Wang Z, Chang JS, Cheng GP. Five-year changes in corneal astigmatism after combined femtosecond-assisted phacoemulsification and arcuate keratotomy. *Am J Ophthalmol.* (2020) 217:232–9. doi: 10.1016/j.ajo.2020.05.004

21. Alpins NA. Vector analysis of astigmatism changes by flattening, steepening, and torque. *J Cataract Refract Surg.* (1997) 23:1503–14. doi: 10.1016/S0886-3350(97)80021-1

22. Borasio E, Mehta JS, Maurino V. Torque and flattening effects of clear corneal temporal and on-axis incisions for phacoemulsification. *J Cataract Refract Surg.* (2006) 32:2030–8. doi: 10.1016/j.jcrs.2006.09.010

23. Alpins N. How to get accuracy in intraocular lens calculation in normal and extreme cases. *Eur Ophthal Rev.* (2017) 11:25–7. doi: 10.17925/EOR.2017.11.01.25

24. Geggel HS. Arcuate relaxing incisions guided by corneal topography for postkeratoplasty astigmatism: vector and topographic analysis. *Cornea.* (2006) 25:545–57. doi: 10.1097/01.ico.0000214222.13615.b6

25. Rückl T, Dexl AK, Bachernegg A, Reischl V, Riha W, Ruckhofer J, et al. Femtosecond laser–assisted intrastromal arcuate keratotomy to reduce corneal astigmatism. *J Cataract Refract Surg.* (2013) 39:528–38. doi: 10.1016/j.jcrs.2012.10.043

frontiers | Frontiers in Medicine

TYPE Original Research
PUBLISHED 13 September 2022
DOI 10.3389/fmed.2022.967710

OPEN ACCESS

EDITED BY
Yu-Chi Liu,
Singapore Eye Research Institute (SERI), Singapore

REVIEWED BY
Gilbert Yong San Lim,
SingHealth, Singapore
Dong Hui Lim,
Sungkyunkwan University, South Korea

*CORRESPONDENCE
Tae-im Kim
TIKim@yuhs.ac

SPECIALTY SECTION
This article was submitted to Ophthalmology, a section of the journal Frontiers in Medicine

RECEIVED 13 June 2022
ACCEPTED 17 August 2022
PUBLISHED 13 September 2022

CITATION
Ahn H, Jun I, Seo KY, Kim EK and Kim T-i (2022) Artificial intelligence approach for recommendation of pupil dilation test using medical interview and basic ophthalmologic examinations. *Front. Med.* 9:967710. doi: 10.3389/fmed.2022.967710

Artificial intelligence approach for recommendation of pupil dilation test using medical interview and basic ophthalmologic examinations

Hyunmin Ahn[1], Ikhyun Jun[1,2], Kyoung Yul Seo[1], Eung Kweon Kim[2,3] and Tae-im Kim[1,2]*

[1]Department of Ophthalmology, Yonsei University College of Medicine, Seoul, South Korea, [2]Corneal Dystrophy Research Institute, Yonsei University College of Medicine, Seoul, South Korea, [3]Saevit Eye Hospital, Goyang, South Korea

Purpose: To evaluate the value of artificial intelligence (AI) for recommendation of pupil dilation test using medical interview and basic ophthalmologic examinations.

Design: Retrospective, cross-sectional study.

Subjects: Medical records of 56,811 patients who visited our outpatient clinic for the first time between 2017 and 2020 were included in the training dataset. Patients who visited the clinic in 2021 were included in the test dataset. Among these, 3,885 asymptomatic patients, including eye check-up patients, were initially included in test dataset I. Subsequently, 14,199 symptomatic patients who visited the clinic in 2021 were included in test dataset II.

Methods: All patients underwent a medical interview and basic ophthalmologic examinations such as uncorrected distance visual acuity, corrected distance visual acuity, non-contact tonometry, auto-keratometry, slit-lamp examination, dilated pupil test, and fundus examination. A clinically significant lesion in the lens, vitreous, and fundus was defined by subspecialists, and the need for a pupil dilation test was determined when the participants had one or more clinically significant lesions in any eye. Input variables of AI consisted of a medical interview and basic ophthalmologic examinations, and the AI was evaluated with predictive performance for the need of a pupil dilation test.

Main outcome measures: Accuracy, sensitivity, specificity, and positive predictive value.

Results: Clinically significant lesions were present in 26.5 and 59.1% of patients in test datasets I and II, respectively. In test dataset I, the model performances were as follows: accuracy, 0.908 (95% confidence interval (CI): 0.880–0.936); sensitivity, 0.757 (95% CI: 0.713–0.801); specificity, 0.962 (95% CI: 0.947–0.977); positive predictive value, 0.878 (95% CI: 0.834–0.922); and F1 score, 0.813. In test dataset II, the model had an accuracy of 0.949 (95% CI: 0.934–0.964), a sensitivity of 0.942 (95% CI: 0.928–956), a specificity of 0.960 (95% CI: 0.927–0.993), a positive predictive value of 0.971 (95% CI: 0.957–0.985), and a F1 score of 0.956.

Conclusion: The AI model performing a medical interview and basic ophthalmologic examinations to determine the need for a pupil dilation test had good sensitivity and specificity for symptomatic patients, although there was a limitation in identifying asymptomatic patients.

KEYWORDS

artificial intelligence, machine learning, medical interview, ophthalmologic examination, pupil dilation test

Introduction

In today's era, artificial intelligence (AI) is one of the hottest topics in all fields worldwide. Digital device, marketing, education, and AI itself are the target of AI development (1–4). The medical field also cannot escape from this trend. However, only very specific settings in clinical practice, such as the detection of arterial fibrillation, epilepsy seizure, and hypoglycemia, or the diagnosis of disease based on histopathological examination or medical imaging benefit from the application of medical AI (5). Recent research in ophthalmology showed that AIs with deep learning algorithms had an acceptable performance in ophthalmic imaging data, such as fundus photography and topography (6). However, there are various challenges in the application of AI in actual clinical practice, even with AI with good performance for imaging analysis (7). Considering the flow of medical services from patients to doctors (Figure 1), tremendous applications of AI are possible.

Patient visit time for outpatient clinic is one of the key issues to address in order to improve not only the quality of medical services but also the clinic efficiency (8). Minimizing the medical process reduces the patient's waiting time and medical costs while improving the satisfaction of service providers and beneficiaries. Pupil dilation test and fundus examination is performed to differentiate between intraocular diseases. The majority of anterior segment diseases are diagnosed using slit-lamp biomicroscopic examination. In contrast, diseases in the lens, optic nerve, vitreous, and chorio-retina are basically diagnosed using pupil dilation test (9). However, after pupil dilation, some important examinations such as near vision test, pupillary light reflex, and visual field examination have a limitation or bias. Moreover, considering the dilation time after discontinuing mydriatics, fundus examination is a turning point in the process of medical service in ophthalmology, from visit to treatment (10). Pupil dilation test can be performed for a patient with symptoms and signs that suggest the possibility of an intraocular disease after a medical interview and basic ophthalmologic examinations. However, in many cases, because these processes have a limitation to presume some intraocular diseases, this test is performed after an additional process that an ophthalmologist conducts directly, such as slit-lamp biomicroscopy. Moreover, many intraocular diseases are asymptomatic and are detected incidentally (11–13).

We considered using AI to simplify the medical service process through the automatic determination of pupil dilation test. There is no study on AI that recommends pupil dilation test. This study aimed to determine whether AI can recommend pupil dilation test appropriately when only basic ophthalmologic information is provided, as in our clinical situation.

Abbreviations: AI, artificial intelligence; CI, confidence interval; ICD-10, International Statistical Classification of Diseases and Related Health Problems, 10th revision; IOP, intraocular pressure; IRB, institutional review board.

Materials and methods

The study was conducted in accordance with the tenets of the Declaration of Helsinki, and ethical approval for each follow-up was obtained from the institutional review board (IRB) of Yonsei University College of Medicine. All participants for prospective validation provided written informed consent before participating. For retrospective data, patient consent was waived after IRB approval (Protocol number 4-2022-0326).

Participants

The study was conducted at Severance Hospital, Yonsei University College of Medicine, Republic of Korea. Medical records from 2017 to 2021 were analyzed. All patients who visited the outpatient clinic of Severance Eye hospital for the first time were included the study. The medical service process of the first visiting outpatient is presented in the flowchart in Figure 1. The key inclusion criteria were as follows: (1) completed medical interview and basic ophthalmologic examinations such as slit-lamp biomicroscopy, pupil dilation test, and fundus examination, (2) communicated directly (for children, including parents), and (3) medical records confirmed by a subspecialist. Patients who did not complete all examinations were excluded. The diagnosis, treatment, and follow-up plan were confirmed to exclude unspecified disease.

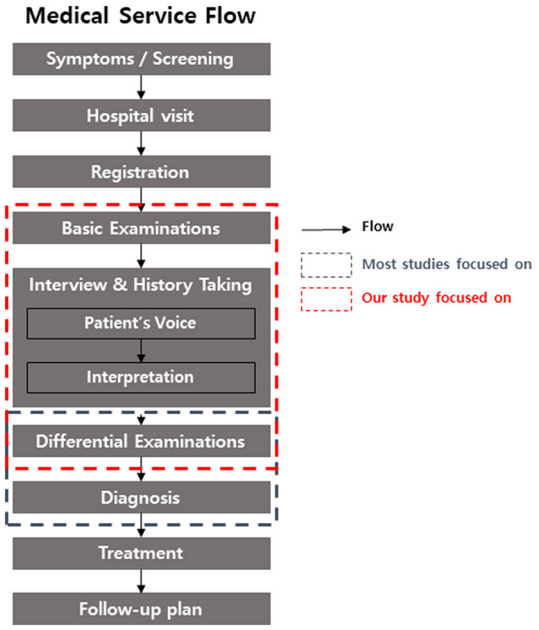

FIGURE 1
The flow of medical services from to doctors in this study.

Basic protocol for first-visiting patients

Medical interview and basic ophthalmologic examinations

Patients who visited the outpatient clinic for the first time were interviewed by ophthalmologists and experienced paramedics who had been trained for at least 2 years with confirmed hospital protocols. In the interview, chief complaint, comorbid symptoms, duration, systemic/ophthalmic history, and familial history were collected (see also Supplementary material 1). Systemic and ophthalmic diseases were categorized by the International Statistical Classification of Diseases and Related Health Problems, tenth revision (ICD-10) classification. After the medical interview, all patients underwent basic ophthalmologic examinations such as uncorrected distance visual acuity, corrected distance visual acuity if wearing glasses and contact lenses, autokeratometry, corrected distance visual acuity with autokeratometry, and intraocular pressure (IOP) with non-contact tonometry. When the IOP was under 7 mmHg or over 21 mmHg, the measurements were repeated twice. When the initial measurement of refraction or keratometric power by autokeratometry failed, a repeat measurement was performed.

Pupil dilation test and fundus examination

All new patients who visited our clinic were required to undergo a pupil dilation test and fundus examination.

Further processes

After medical interview and basic ophthalmologic examination, all new patients underwent additional examinations, or treatments after referral to subspecialist. All the contents of the medical processes were saved in the electronic medical record.

AI modeling

The overall process of AI modeling is described in Figure 2. AI modeling was constructed based on the electronic medical record by Python 3.8 program.

Training dataset

Prior to AI modeling, patients who first visited the outpatient clinic between 2017 and 2020 were included in the training dataset.

Test dataset

Patients who first visited the clinic in 2021 were included in the test dataset. First, patients who underwent an eye screening test for their systemic disease and treatment, as well as asymptomatic patients including consultation cases from other medical parts, were included in test dataset I. Subsequently, test patients that were not included in test dataset I, were included in test dataset II.

Input variables

The input variables were as follows: (1) General patient information, including age, sex, systemic/ophthalmologic history, and family history, (2) symptoms and events, and (3) results of the basic ophthalmologic examination (see also Supplementary material 1). Patients' symptoms were sorted based on the list in the website of American Academy of Ophthalmology and our previous study (14, 15). Characteristics, duration, time aspect, related events, and other purposes of visiting (i.e., health check-up and screening ophthalmic complications of systemic diseases and treatments) were also interpreted (see also Supplementary material 1). For model training, training dataset was split into training and validation data in a 3: 7 ratios. Standard scaler was used for visual acuities, IOP, and the values of autokeratometry.

Output variables

The output variable was set to the binary value of the need for a pupil dilation test (yes or no). The need for a pupil dilation test was determined when there were clinically significant lesions in any of the eyes. A clinically significant lesion for the pupil

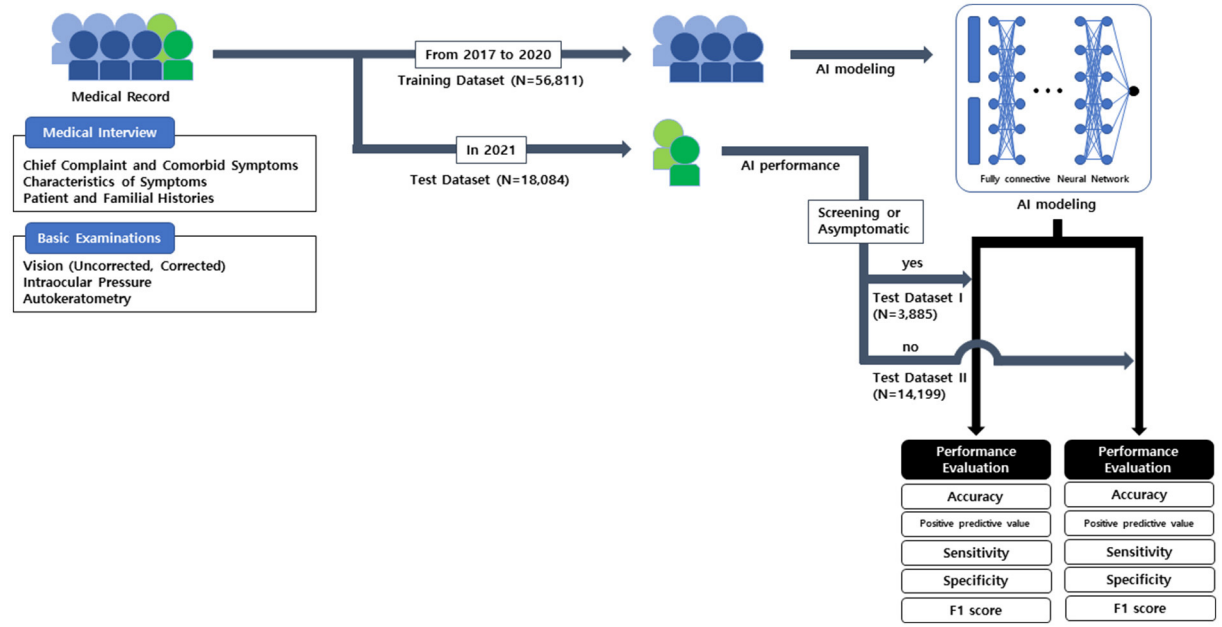

FIGURE 2
The process used in this study and artificial intelligence modeling for recommending pupil dilation test.

dilation test was defined as a lesion in the lens, vitreous, retina, and optic disc area. It was confirmed by each subspecialist when **one** or more of the following criteria were met: (1) clearly explained patient's subjective symptoms, (2) required additional follow-up with the possibility of exacerbation of disease and/or intervention within 3 months, and (3) required additional detailed examinations for treatment plan (see also Supplementary material 2). Lesions such as asymptomatic mild macular drusen, lattice degeneration without retinal break or vitreous traction, simple retinal pigmentation and chronic scars, low-risk glaucoma suspect with long-term follow-up over 6 months, and non-vision impairing cataract were not deemed clinically significant by subspecialists (16–19). Functional disorders and extraocular disorders that did not require a pupil dilation test were also not deemed clinically significant.

Model construction

AI modeling was conducted with fully connected deep neural network. The activation function for hidden layers was rectified linear (ReLU) function. Adam optimization was used. Accuracy was used as a metric. The depth of the hidden layer and the nodes in each hidden layer were automatically modulated with network topology. Batch and epoch size were automatically modulated. Dropout 0.5 and L2 regularization were used to prevent overfitting. Performance and loss were surveilled to prevent underfitting. The early stopping method was used.

Statistical analysis

Statistical analysis was conducted with Python 3.8 program. Model performance was evaluated with accuracy, sensitivity (also called recall), specificity, and positive predictive value (also called precision); 95% confidence intervals (CIs) were used. The false-positive and false-negative cases in each test dataset were descriptively analyzed with the location of the lesion based on the ICD-10 classification. If the locations overlapped, all locations were considered.

Results

In the training dataset, 56,811 patients were enrolled, with women accounting for 54.1%. The mean age of the patients was 57.5 ± 18.9 years. The clinically significant lesions for pupil dilation test were present in 65.1% of the patients. Of the clinically significant lesions, 28.9% were in the lens, 7.8% in the vitreous, 38.6% in the macular area, 12.6% in the peripheral retina, and 20.1% in the optic disc. A total of 3,885 patients were enrolled in test dataset I, and 14,199 patients were enrolled in test dataset II. The clinically significant lesions were present in 26.5% of patients in test dataset I and 59.1% of patients in test dataset II (Table 1).

In test datasets I and II, the AI recommendation for pupil dilation test had an accuracy of 0.908 (95% CI: 0.880–0.936) and 0.949 (95% CI: 0.934–0.964), respectively. The sensitivity,

TABLE 1 Characteristics of study patients in training dataset and test dataset I and II.

Characteristics	Training ($N = 56,811$)	Test I ($N = 3,885$)	Test II ($N = 14,199$)
Age (years, mean±SD)	57.5 ± 18.9	48.5 ± 22.3	60.0 ± 17.8
Sex (proportion of female, %)	54.1	60.2	54.2
Uncorrected distance visual acuity (logMAR, mean±SD)	0.57 ± 0.51	0.48 ± 0.52	0.57 ± 0.50
Corrected distance visual acuity (logMAR, mean±SD)[y]	0.19 ± 0.27	0.07 ± 0.08	0.19 ± 0.29
Intraocular pressure (mmHg, mean±SD)[z]	15.3 ± 3.5	14.7 ± 3.4	15.2 ± 3.3
Spherical equivalent (diopters, mean±SD)[y]	−1.65 ± 3.06	−1.68 ± 2.81	−1.60 ± 3.45
Corneal power (diopters, mean±SD)[y]	43.33 ± 2.02	42.19 ± 3.75	43.30 ± 2.02
Clinically significant lesion (% of eyes)	65.1	26.5	59.1
Lens (% of clinically significant lesion)	28.9	35.3	23.5
Vitreous (% of clinically significant lesion)	7.8	2.5	7.0
Macula (% of clinically significant lesion)	38.6	14.8	40.5
Peripheral retina (% of clinically significant lesion)	12.6	30.0	15.3
Optic disc (% of clinically significant lesion)	20.1	25.5	19.7

[y]Measured by auto-keratometry.
[z]Measured by non-contact tonometry.

TABLE 2 The performance of AI for recommendation pupil dilation test in test dataset I and II.

	Test dataset I ($n = 3,885$)		Test dataset II ($n = 14,199$)	
	Estimate	95% CI	Estimate	95% CI
Accuracy	0.908	0.880–0.936	0.949	0.934–0.964
Sensitivity (Recall)	0.757	0.713–0.801	0.942	0.928–0.956
Specificity	0.962	0.947–0.977	0.960	0.927–0.993
Positive predictive value (Precision)	0.878	0.834–0.922	0.971	0.957–0.985
F1 score	0.813	-	0.956	-

specificity, and positive predictive value in test dataset I were 0.757 (95% CI: 0.713–0.801), 0.962 (95% CI: 0.947–0.977), and 0.878 (95% CI: 0.834–0.922), respectively, and those in test dataset II were 0.942 (95% CI: 0.928–0.956), 0.960 (95% CI: 0.927–0.993), and 0.971 (95% CI: 0.957–0.985), respectively (Table 2). F1 score was 0.813 in test dataset I and 0.956 in test dataset II.

Table 3 shows the proportion of the locations of the clinically significant lesions in the false-negative and false-positive categories in the entire test dataset. In the false-negative category, 37% of the lesions were in the macular area, 28.1% in the optic disc, 20.1% in the peripheral retina, 10.3% in the lens, and 5.3% in the vitreous. In the false-positive category, 73% of the lesions were in the anterior segment, including the cornea and anterior chamber, and 10.7% were in the eyelid and extra-orbital area. Further, 17% of the false-positive cases had non-ophthalmologic causes.

TABLE 3 Locations of clinically significant lesions in false-negative and false-positive categories with overall test dataset.

Locations	Proportions (%)
False-negative	
Lens	10.3
Vitreous	5.3
Macula	37.2
Peripheral retina	20.1
Optic disc	28.1
False-positive	
Cornea and Anterior chamber	73.1%
Eyelid and Extra-orbital area	10.7%
Non-ophthalmologic	17.3%

Discussion

The performance of the AI in recommending pupil dilation test using a medical interview and basic ophthalmologic examinations was good, with ∼95% accuracy in symptomatic patients. However the AI had a limitation in detecting asymptomatic lesions in the lens, vitreous, chorio-retina, and

optic nerve, and the sensitivity and positive predictive value of test dataset I was ∼76 and 88%.

In the ophthalmologic service process, examination of vision and IOP are generally recommended at a visiting eye clinic (10). Ophthalmologists select additional differential and detailed examinations based on the information obtained from a medical interview and basic ophthalmologic examinations. AI automation is expected to improve the efficiency of the medical process (20). We aimed to develop a decision-making AI for ophthalmologic examinations as a type of AI that will help to reduce the time and cost of medical services.

Because the characteristics of study population in test dataset II were more similar to those in training dataset than in test dataset I, this AI model might perform better in test dataset II than in test dataset I. This study was conducted at a tertiary medical service institution, and the number of asymptomatic patients was smaller than the number of symptomatic patients as confirmed by the sample size of test dataset I which was smaller than the sample size of test dataset II. Analyzing AI model performance in detail, test dataset I had lower sensitivity and positive predictive value, which was attributed to the lower true positive ratio. According to previous studies, fundus examination is important to detect asymptomatic diseases (16). Glaucoma, diabetic retinopathy, and age-related macular degeneration are well-known diseases that are asymptomatic in the early phase (21–23). Ocular symptoms are not the only reason for a pupil dilation test; the patient's ophthalmologic history, systemic disease, and familial history are also considered (24–27). The importance of pupil dilation test and fundus examination in asymptomatic patients is a contrary evidence that the investigator cannot predict the disease of the posterior segment of the eye from symptoms alone. Moreover, the past history and familial history of the patient may be unclear or unrevealed. These problems were also reflected in the results of our study, especially in asymptomatic patients.

In this study, information obtained from the medical interview and basic ophthalmologic examinations used as the input dataset were limited in determining whether to conduct a pupil dilation test. The performance of the AI can be improved by changing the AI model or using a large sample size (28). We used several methods to overcome the technical problem. First, the hyperparameters, especially the number of hidden layers, nodes, batch size, and epochs were modulated automatically with surveillance of overfitting and underfitting. Increasing the number of hyperparameters does not always increase the performance of AI (29, 30). In this study, because there was no continuous performance improvement with the additional training process, and the plateau phenomenon was detected in all of the sequences with hyperparameter modulation, the possibility of underfitting was carefully estimated to be minimized. Second, we evaluated the performance with two validation datasets. We tried to determine whether the lower performance was due to a technical issue or a limitation of clinical factor. The results of this study suggest that insufficient performance in test dataset I of the AI model was caused by asymptomatic lesions, limitation of clinical factor, and aleatoric uncertainty. In order to improve the performance of the AI in cases with asymptomatic lesions, completely new input information is needed rather than simply increasing the sample size or changing the AI model.

This study has some limitations. First, the dataset was collected from a tertiary care hospital. The proportion of patients with clinically significant lesions in tertiary care hospitals is different from that in a primary care service. The performance of AI may vary in a primary care setting, depending on the application area, such as telemedicine. Prospective applicable research in various clinical settings is needed. Second, the result of this study is applicable only for patients visiting the clinic for the time. AI in patients with previous visiting history is different, and it could be considered with other AI models such as recurrent neural network. Third, the patients' symptoms were interpreted by medical personnel and did not directly reflect the patients' expression. This study did not evaluate the use of AI by patients. An advanced AI using dataset directly expressed by patients in ways such as speech or writing is now being planned. Finally, the definition of "need for pupil dilation test" was determined by each subspecialist in our hospital. The definition might be clinically acceptable and the controversial cases between subspecialists which were <0.1% in this study were excluded in this study. However, bias from individual cases could not be completely excluded. Perhaps this issue depends on the protocol guidelines within hospital or group.

In conclusion, the AI recommending pupil dilation test had a good performance with only basic ophthalmologic information for symptomatic lesions, although there was a limitation of the performance for asymptomatic lesions.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethics statement

The studies involving human participants were reviewed and approved by the Institutional Review Board (IRB) of Yonsei University College of Medicine. Written informed consent from the participants' legal guardian/next of kin was not required to participate in this study in accordance with the national legislation and the institutional requirements.

Author contributions

HA and T-iK designed the study and wrote the original draft of the manuscript. HA collected patient data, performed the artificial intelligence modeling, and analysis. IJ, KS, and EK are responsible for reviewing the manuscript. All authors provided critical review and approved the version for publication.

Acknowledgments

We thank the subspecialists of cornea, glaucoma, neuro-ophthalmology, oculoplastics, retina/uvea, and strabismus/pediatric ophthalmology in Severance Hospital for academic supports.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmed.2022.967710/full#supplementary-material

References

1. Borges AF, Laurindo FJ, Spínola MM, Gonçalves RF, Mattos CA. The strategic use of artificial intelligence in the digital era: Systematic literature review and future research directions. *Int J Infm Manag.* (2021) 57:102225. doi: 10.1016/j.ijinfomgt.2020.102225

2. Verma S, Sharma R, Deb S, Maitra D. Artificial intelligence in marketing: Systematic review and future research direction. *Int J Infm Manag Data Insight.* (2021) 1:100002. doi: 10.1016/j.jjimei.2020.100002

3. Chen L, Chen P, Lin Z. Artificial intelligence in education: a review. *IEEE Access.* (2020) 8:75264–78. doi: 10.1109/ACCESS.2020.2988510

4. Thiagarajan JJ, Venkatesh B, Sattigeri P, Bremer PT. Building calibrated deep models via uncertainty matching with auxiliary interval predictors. In: *Proceedings of the AAAI Conference on Artificial Intelligence*, Vol. 34. AAAI (2020). p. 6005–12.

5. Briganti G, Le Moine O. Artificial intelligence in medicine: today and tomorrow. *Front Med.* (2020) 7:27. doi: 10.3389/fmed.2020.00027

6. Ting DSW, Pasquale LR, Peng L, et al. Artificial intelligence and deep learning in ophthalmology. *Br J Ophthalmol.* (2019) 103:167–75. doi: 10.1136/bjophthalmol-2018-313173

7. Gunasekeran DV, Wong TY. Artificial intelligence in ophthalmology in 2020: a technology on the Cusp for translation and implementation. *Asia Pac J Ophthalmol (Phila).* (2020) 9:61–6. doi: 10.1097/01.APO.0000656984.56467.2c

8. Ciulla TA, Tatikonda MV, ElMaraghi YA, Hussain RM, Hill AL, Clary JM, et al. Lean six sigma techniques to improve ophthalmology clinic efficiency. *Retina.* (2018) 38:1688–98. doi: 10.1097/IAE.0000000000001761

9. Peng J, Zhang Q, Jin HY, Lu WY, Zhao PQ. Ultra-wide field imaging system and traditional retinal examinations for screening fundus changes after cataract surgery. *Int J Ophthalmol.* (2016) 9:1299. doi: 10.18240/ijo.2016.09.11

10. Rupp JD. The 8-Point Eye Exam. *American Academy of Ophthalmology.* (2016). Available online at: http://www.aao.org/young-ophthalmologists/yo-info/article/how-to-conduct-eight-point-ophthalmology-exam (accessed September 10, 2021).

11. Kohner EM. Diabetic retinopathy. *BMJ.* (1993) 307:1195–9. doi: 10.1136/bmj.307.6913.1195

12. Mukesh BN, McCarty CA, Rait JL, Taylor HR. Five-year incidence of open-angle glaucoma: the visual impairment project. *Ophthalmology.* (2002) 109:1047–51. doi: 10.1016/S0161-6420(02)01040-0

13. Weih LM, Nanjan M, McCarty CA, Taylor HR. Prevalence and predictors of open-angle glaucoma: results from the visual impairment project. *Ophthalmology.* (2001) 108:1966–72. doi: 10.1016/S0161-6420(01)00799-0

14. Anonymous. Eye Symptoms. *American Academy of Ophthalmology*. Available online at: https://www.aao.org/eye-health/symptoms-list#l (accessed September 12, 2021).

15. Ahn H. Artificial intelligence method to classify ophthalmic emergency severity based on symptoms: a validation study. *BMJ Open.* (2020) 10:e037161. doi: 10.1136/bmjopen-2020-037161

16. Pollack AL, Brodie SE. Diagnostic yield of the routine dilated fundus examination. *Ophthalmology.* (1998) 105:382–6. doi: 10.1016/S0161-6420(98)93718-6

17. Wong TY, Mitchell P. Hypertensive retinopathy. *N Engl J Med.* (2004) 351:2310–7. doi: 10.1056/NEJMra032865

18. Chew EY, Clemons TE, Agrón E, Sperduto RD, Sangiovanni JP, Davis MD, et al. Ten-year follow-up of age-related macular degeneration in the age-related eye disease study: AREDS report no. 36. *JAMA Ophthalmol.* (2014) 132:272–7. doi: 10.1001/jamaophthalmol.2013.6636

19. Lewis H. Peripheral retinal degenerations and the risk of retinal detachment. *Am J Ophthalmol.* (2003) 136:155–60. doi: 10.1016/S0002-9394(03)00144-2

20. Yarlagadda RT. ai automation and it's future in the United States. *Int J Creat Res Thoughts.* (2017) 5. Available online at: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3798873

21. Cheung N, Mitchell P, Wong T. Diabetic retinopathy. *Lancet.* (2010) 376:124–36. doi: 10.1016/S0140-6736(09)62124-3 Available online at: http://www.ijcrt.org/papers/IJCRT1133935.pdf

22. Guymer R, Wu Z. Age-related macular degeneration (AMD): more than meets the eye. The role of multimodal imaging in today's management of AMD—a review. *Clin Exper Ophthalmol.* (2020) 48:983–95. doi: 10.1111/ceo.13837

23. Heijl A, Bengtsson B, Hyman L, Leske MC, Group EMGT. Natural history of open-angle glaucoma. *Ophthalmology.* (2009) 116:2271–6. doi: 10.1016/j.ophtha.2009.06.042

24. Williams GA, Scott IU, Haller JA, Maguire AM, Marcus D, McDonald HR. Single-field fundus photography for diabetic retinopathy screening: a report

by the American Academy of Ophthalmology. *Ophthalmology*. (2004) 111:1055–62. doi: 10.1016/j.ophtha.2004.02.004

25. Komulainen R, Tuulonen A, Airaksinen PJ. The follow-up of patients screened for glaucoma with non-mydriatic fundus photography. *Int Ophthalmol*. (1992) 16:465–9. doi: 10.1007/BF00918438

26. Gillies MC, Zhu M, Chew E, Barthelmes D, Hughes E, Ali H, et al. Familial asymptomatic macular telangiectasia type 2. *Ophthalmology*. (2009) 116:2422–9. doi: 10.1016/j.ophtha.2009.05.010

27. Khurram Butt D, Gurbaxani A, Kozak I. Ultra-wide–field fundus autofluorescence for the detection of inherited retinal disease in difficult-to-examine children. *J Pediatric Ophthalmol Strabismus*. (2019) 56:383–7. doi: 10.3928/01913913-20190925-03

28. Markham IS, Rakes TR. The effect of sample size and variability of data on the comparative performance of artificial neural networks and regression. *Comput Oper Res*. (1998) 25:251–63. doi: 10.1016/S0305-0548(97)00074-9

29. Yu C, Manry MT, Li J, Narasimha PL. An efficient hidden layer training method for the multilayer perceptron. *Neurocomputing*. (2006) 70:525–35. doi: 10.1016/j.neucom.2005.11.008

30. Hutter F, Hoos H, Leyton-Brown K. An efficient approach for assessing hyperparameter importance. In *International conference on machine learning. In: Proceedings of Machine Learning Research*, Vol. 32. PMLR (2014). p. 754–62.

frontiers | Frontiers in Medicine

TYPE Original Research
PUBLISHED 06 October 2022
DOI 10.3389/fmed.2022.1008950

OPEN ACCESS

EDITED BY
Tyler Hyungtaek Rim,
Duke-NUS Medical School, Singapore

REVIEWED BY
Min-Yen Hsu,
Chung Shan Medical University, Taiwan
Thanapong Somkijrungroj,
Chulalongkorn University, Thailand

*CORRESPONDENCE
De-Kuang Hwang
m95gbk@gmail.com

SPECIALTY SECTION
This article was submitted to Ophthalmology, a section of the journal Frontiers in Medicine

RECEIVED 01 August 2022
ACCEPTED 16 September 2022
PUBLISHED 06 October 2022

CITATION
Lin T-Y, Chen H-R, Huang H-Y, Hsiao Y-I, Kao Z-K, Chang K-J, Lin T-C, Yang C-H, Kao C-L, Chen P-Y, Huang S-E, Hsu C-C, Chou Y-B, Jheng Y-C, Chen S-J, Chiou S-H and Hwang D-K (2022) Deep learning to infer visual acuity from optical coherence tomography in diabetic macular edema.
Front. Med. 9:1008950.
doi: 10.3389/fmed.2022.1008950

Deep learning to infer visual acuity from optical coherence tomography in diabetic macular edema

Ting-Yi Lin[1], Hung-Ruei Chen[2], Hsin-Yi Huang[3,4], Yu-Ier Hsiao[2], Zih-Kai Kao[5], Kao-Jung Chang[2,3], Tai-Chi Lin[2,6], Chang-Hao Yang[7], Chung-Lan Kao[2,5,8,9], Po-Yin Chen[5,8,10,11,12], Shih-En Huang[5,8], Chih-Chien Hsu[2,3,6], Yu-Bai Chou[2,3,6], Ying-Chun Jheng[3,6,13,14], Shih-Jen Chen[2,3,6], Shih-Hwa Chiou[2,3,6,13] and De-Kuang Hwang[2,3,6]*

[1]Doctoral Degree Program of Translational Medicine, National Yang Ming Chiao Tung University and Academia Sinica, Taipei, Taiwan, [2]School of Medicine, National Yang Ming Chiao Tung University, Taipei, Taiwan, [3]Institute of Clinical Medicine, National Yang Ming Chiao Tung University, Taipei, Taiwan, [4]Taipei Veterans General Hospital Biostatistics Task Force, Taipei, Taiwan, [5]Department of Physical Medicine and Rehabilitation, Taipei Veterans General Hospital, Taipei, Taiwan, [6]Department of Ophthalmology, Taipei Veterans General Hospital, Taipei, Taiwan, [7]Department of Ophthalmology, National Taiwan University, Taipei, Taiwan, [8]Department of Physical Therapy and Assistive Technology, National Yang Ming Chiao Tung University, Taipei, Taiwan, [9]Center for Intelligent Drug Systems and Smart Bio-devices (IDS2B), National Yang Ming Chiao Tung University, Hsinchu, Taiwan, [10]School of Gerontology and Long-Term Care, College of Nursing, Taipei Medical University, Taipei, Taiwan, [11]Master Program in Long-Term Care, College of Nursing, Taipei Medical University, Taipei, Taiwan, [12]International Ph.D. Program in Gerontology and Long-Term Care, College of Nursing, Taipei Medical University, Taipei, Taiwan, [13]Big Data Center, Department of Medical Research, Taipei Veterans General Hospital, Taipei, Taiwan, [14]Center for Quality Management, Taipei Veterans General Hospital, Taipei, Taiwan

Purpose: Diabetic macular edema (DME) is one of the leading causes of visual impairment in diabetic retinopathy (DR). Physicians rely on optical coherence tomography (OCT) and baseline visual acuity (VA) to tailor therapeutic regimen. However, best-corrected visual acuity (BCVA) from chart-based examinations may not wholly reflect DME status. Chart-based examinations are subjected findings dependent on the patient's recognition functions and are often confounded by concurrent corneal, lens, retinal, optic nerve, or extraocular disorders. The ability to infer VA from objective optical coherence tomography (OCT) images provides the predicted VA from objective macular structures directly and a better understanding of diabetic macular health. Deviations from chart-based and artificial intelligence (AI) image-based VA will prompt physicians to assess other ocular abnormalities affecting the patients VA and whether pursuing anti-VEGF treatment will likely yield increment in VA.

Materials and methods: We enrolled a retrospective cohort of 251 DME patients from Big Data Center (BDC) of Taipei Veteran General Hospital (TVGH) from February 2011 and August 2019. A total of 3,920 OCT images, labeled as "visually impaired" or "adequate" according to baseline VA, were grouped into training (2,826), validation (779), and testing cohort (315). We applied confusion matrix and receiver operating characteristic (ROC) curve to evaluate the performance.

Results: We developed an OCT-based convolutional neuronal network (CNN) model that could classify two VA classes by the threshold of 0.50 (decimal notation) with an accuracy of 75.9%, a sensitivity of 78.9%, and an area under the ROC curve of 80.1% on the testing cohort.

Conclusion: This study demonstrated the feasibility of inferring VA from routine objective retinal images.

Translational relevance: Serves as a pilot study to encourage further use of deep learning in deriving functional outcomes and secondary surrogate endpoints for retinal diseases.

KEYWORDS

treatment response, diabetic macular edema (DME), medical image, visual acuity, deep learning

Introduction

The best-corrected visual acuity (BCVA) exam is the most popular test to reflect the condition of the central fovea and the severity of many ocular diseases. Introduced in 1862 by Herman Snellen, the visual chart remained the gold standard for visual acuity (VA) clinical measurement. Visual charts rely on the ability of the patient to identify rows of letters at a fixed distance as each row (line) appears increasingly smaller in size. Although the chart performance depends on the subjective nature of the human response, chances in the correct guessing, or human learning from routine follow-up, the chart remained the basis for VA assessment in clinics and clinical trials. Traditional examinations such as the Early Treatment Diabetic Retinopathy Study (ETDRS) grading scale are usually considered more preferential than other modalities as ETDRS is associated with an escalated risk for vision-threatening retinopathy and serves as a grading scale for retinopathy (1). However, the clinical relevance of the ETDRS grading scale of diabetic retinopathy and other chart-based examinations has been challenged by the difficulty to implement in real-world settings and the technological advances in image acquisition. Thus, the ability to easily derive VA surrogate from routine image modalities provides significant clinical insights throughout the clinical trajectory of macular diseases.

Since the introduction of intravitreal injections (IVI) anti-VEGF, physicians are able to treat exudative macular diseases and recover VA (2–5). In clinical practice, ophthalmologists rely on multiple information, accumulated experience, and intuitive predictions to predict diabetic macular edema (DME) treatment response and whether the treatment is worth pursuing based on an individual's response (6, 7). In daily clinical practice, clinicians often encounter DME patients with concurrent ocular diseases (**Figure 1**) . Therefore, traditional VA examinations based on charts may not wholly reflect DME status or be accurately quantified. For this reason, we aimed to provide surrogate VA based on optical coherence tomography (OCT) that depict macular structural health directly.

Optical coherence tomography is routinely used to screen patients with macular disease where the technology depicts the structural retinal health via scans of retinal cross layers (8–10). Besides, the popularity of OCT across medical settings (i.e., optic glass store, non-ophthalmic clinics) makes the utility practical for disease screening and earlier referral. The wealth of information generated via non-invasive retinal scans makes the technology ideal to distinguish baseline status and treatment response (11–14). The ability to infer surrogate VA from OCT and by assisting physicians in detecting OCT-VA and chart-based VA mismatch will allow the physician to derive treatment strategies taking account of concurrent ocular disease to maximize VA recovery.

To evaluate the potential of deep learning in predicting VA outcomes from structural and functional assessments in the early stages of the diagnosis, we built an SD-OCT-based deep learning model using real-world data to infer the VA cut-off value of 0.50, consistent with the minimal requirement for referral by the AAO (15). To our knowledge, this is the first study to implement deep learning in inferring VA from OCT images in DME patients.

Methods and materials

Ethical approval and data source

This study was approved by the Institutional Review Board (IRB) of Taipei Veterans General Hospital (TVGH) and written informed consent was signed. This study does not include minors, or minorities. Optical coherence tomography (SD-OCT) B-Scans were selected as the primary input information

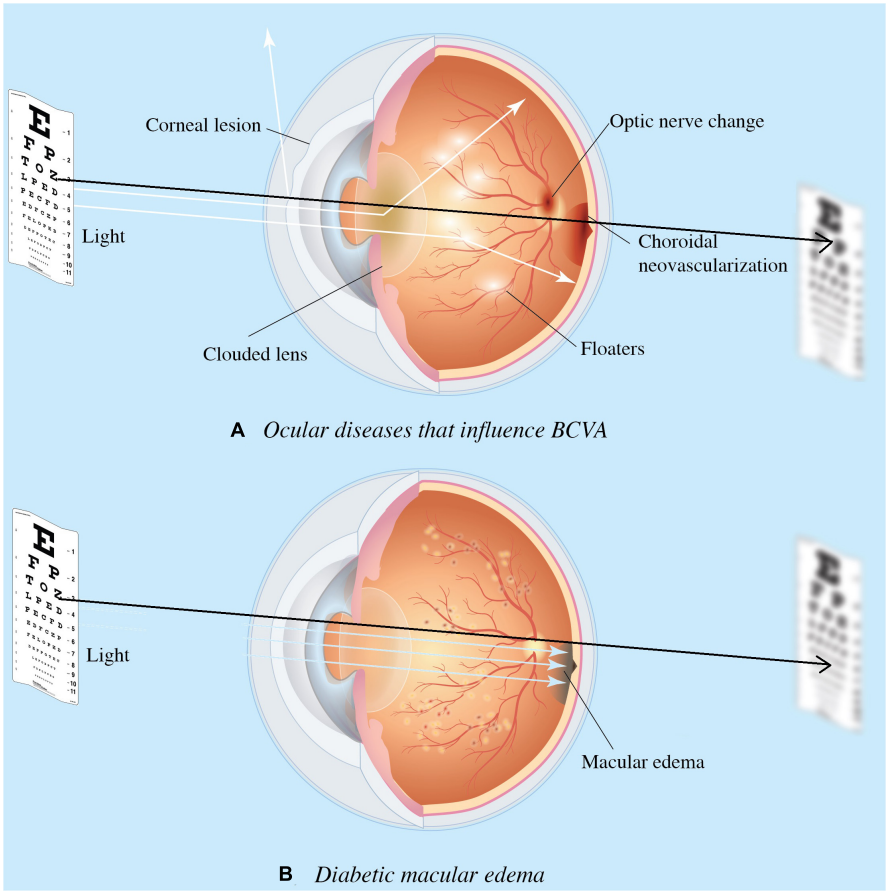

FIGURE 1
Ocular diseases that influence visual acuity (VA). Diseases impact the visual axis, such as corneal lesions caused by degeneration, clouded lens by cataract, floaters by uveitis, choroidal neovascularization due to age-related macular degeneration (AMD); and optic nerve neuropathy by glaucoma. Such impact obscures diabetic macular edema's (DME) involvement in the functional outcome of treatment response, and the need for DME treatment. Diagram **(A)** presents concurrent ocular disorders that impact VA measurement, while **(B)** demonstrates the VA directly measures macular health when isolate DME is present. Black arrow denotes the visual axis.

to establish the computer-assisted visual acuity diagnosis system. All OCT images and subjective, objective, assessment, and plan (SOAP) notes between February 2011 and August 2019 were retrieved from the databank in the big data center (BDC) of TVGH. This dataset consists of de-identified secondary data released for retrospective research purposes. In addition, the OCT images were collected from the patients diagnosed with diabetic macular edema (DME) who sought medical help in the TVGH's Department of Ophthalmology and received an ophthalmology image inspection using the RTVue XR AngioVue OCT device (Optovue Inc., Fremont, CA, USA).

Study participants

Patients were enrolled based on the following inclusion criteria: (1) age above 20 years old, (2) diagnosis of diabetes mellitus (I or II), (3) diagnosis of DME with available baseline OCT image and VA, (4) BCVA measured by Snellen chart from 0.05 to 1.50 (decimal), (5) central-involved macular edema defined by the retinal thickness of >250 μm in the central subfield based on Optovue's automated quantification and the presence of intraretinal fluid (IRF) and subretinal fluid (SRF) seen on SD-OCT, Exclusion criteria were as follows: the presence of cataract or clouded lens, without cataract surgery records. The ocular conditions were obtained from the clinical charts documented by ophthalmologists on the same day when OCT images were taken. In addition, patient charts were reviewed for demographic data, hemoglobin A1C (HbA1C) values, and BCVA.

Clinical labeling

Best-corrected visual acuity of both eyes was measured on the same day when OCT images were acquired in the

Department of Ophthalmology, TVGH. Physicians obtained each OCT scan with a ground truth BCVA and documented it on chart review in each visit. We excluded patients with unspecified BCVA or profound visual impairment defined by the International Classification of Diseases, 11th Revision (ICD-11) as BCVA of decimal notation less than 0.05. Our study employed the cut-off value of 0.50, consistent with the minimal requirement for referral by the American Association of Opthalmology (AAO) (15). We defined BCVA values greater than or equal to 0.50 labeled as "adequate" and those less than 0.50 as "impaired" (Figures 2, 3). The same 0.50 thresholds to discriminate against patients with adequate and impaired vision is consistently used in the literature (16–18).

Datasets and image pre-processing

All participants in this retrospective study were selected based on a comprehensive ophthalmic examination. OCT is accessed via the Big Data Center where reports containing horizontal scan and vertical scan of mid-foveal position is uploaded as PDF reports by Optometrist to the institutions medical image storage PACS (Picture archiving communication system). We cropped the region of interest (ROI) from both vertical and horizontal scans and saved the image in png format (resolution 1960 × 645, bit depth 8) for subsequent model development. The ROI is extractracted from the middle one third of scan areas and downsized them to 224 × 224 pixels resolution by bicubic interpolation. The images were divided into training, validation, and testing groups (Figures 2, 3). First, 70% of the images were incorporated into the training group to train and generate the model parameters. Then, the model's performance was checked by evaluating an independent validation group (20%). The model that generated the smallest error was designated as the final model. Finally, the test group was composed of the remaining dataset (10%) independent of the training. This group was used to appraise the accuracy rate of the final model. To improve deep learning DL efficiency, we conducted data augmentation by horizontal and vertical translation, zooming, Gaussian blurring of the additional noise, horizontal flipping, and random rotation within 30° translation, zooming, Gaussian blurring of the additional noise, horizontal

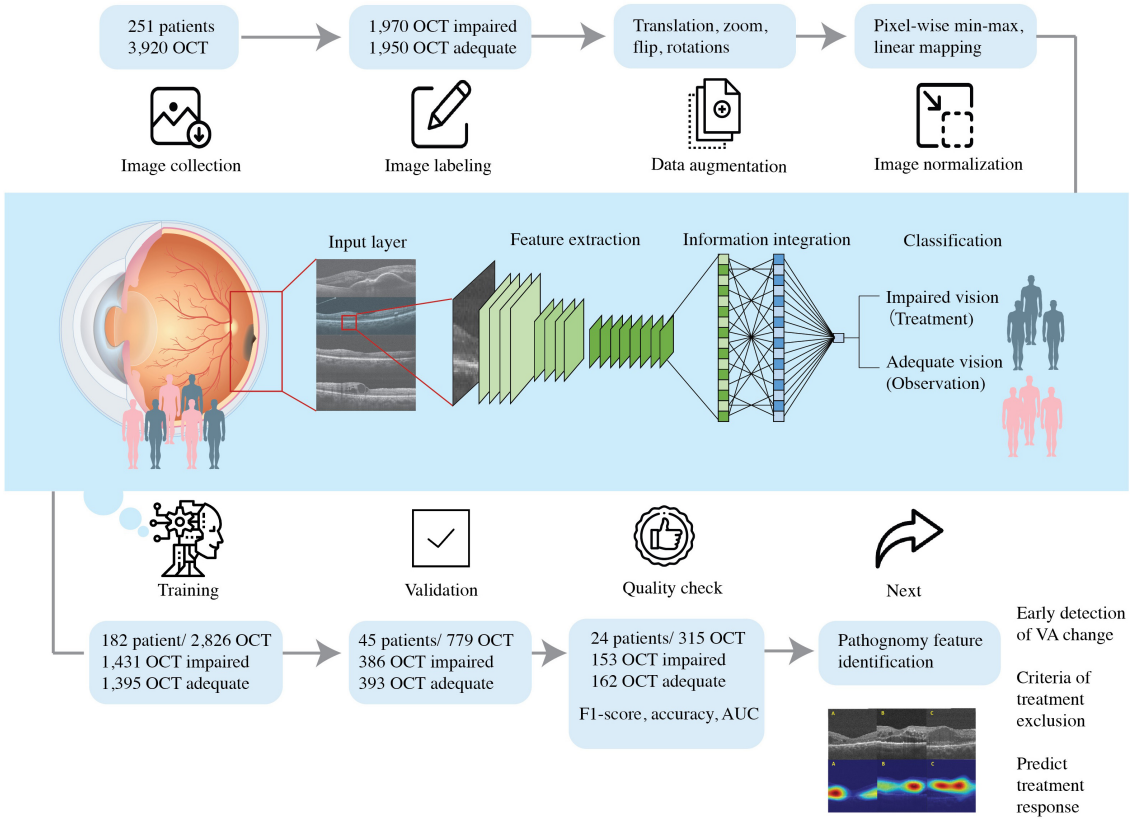

FIGURE 2
Schematic diagram showing the flow of this study. We included patients diagnosed with DME with the best- BCVA between 0.05 and 1.50 and collected the optical coherence tomography (OCT) dataset. The dataset was labeled accordingly with BCVA obtained and DME OCT features by experienced ophthalmologists. The pre-processed OCT database trained the convolutional neural network, so the artificial intelligence algorithm could predict VA and guide therapeutic strategy.

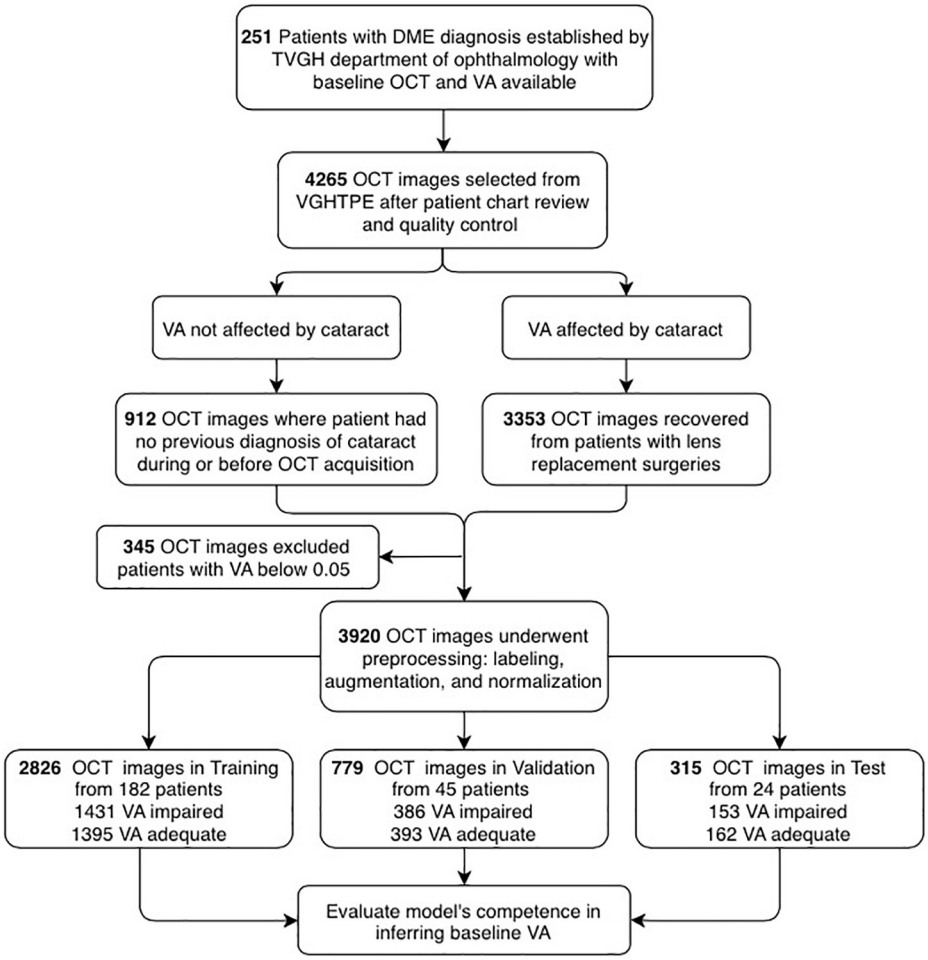

FIGURE 3
Flowchart showing the selection of optical coherence tomography (OCT) images and their analysis. OCT images and patient clinical information were de-identified secondary data released for retrospective research purposes (*N* = 4,265). The OCT images were collected from the patients diagnosed with DME with clear lens (*N* = 912) or artificial intraocular lens (*N* = 3,353), best-corrected visual acuity (BCVA) measured from 0.05 to 1.5 (decimal), excluding those with lower than 0.05 (*N* = 3,920). Afterimage preprocessing, the dataset was categorized into training (*N* = 2,926), validation (*N* = 779), and test (*N* = 315) for the establishment of the AI platform.

flipping, and random rotation within 30°. The augmented dataset was used only for training and not validation or testing. The resized or augmented images then underwent pixel-wise min-max normalization, linear mapping of pixel intensities to the range [−1, 1]. We then used the F1-score, accuracy, and area under the curve (AUC) to evaluate the AI model's performance. F1-score evaluates the test's accuracy calculated from the test's precision and recall (sensitivity) (**Figure 2**).

TABLE 1 The details of the final trained models.

Parameters	Setting
Architecture	EfficientNet
Optimizer	SGD
Loss function	Binary cross-entropy
Learning rate	1e-4 and 1e-5
Batch size	32
Total number of epochs run during training	310

The final model parameters showed the most superior performance where we also compared transfer learning models of the different network architectures, VGG11, VGG16, and ResNet34.

Establishing the artificial intelligence models

An efficient recognition algorithm, convolutional neural network (CNN), is frequently used in image processing and pattern recognition (19, 20). We used EfficientNet-B0 deep neural network architecture to classify OCT images in this study (21). Employing transfer learning, we compared

EfficientNet-B0 with models of different network architectures, VGG11, VGG16, and ResNet34, which were pre-trained for different tasks, converged them for considerably faster steady value, and reduced training time. Furthermore, the AI models were established using the Google cloud platform with two-core vCPU, 7.5 GB RAM and an NVIDIA Tesla K80 GPU card; the software used was CentOS7 with Keras 2.2.4 and TensorFlow-GPU 1.6.0 for training and validation. Because of the retina's size and shape variations, a stochastic gradient descent (SGD) algorithm trained the computational layers with a relatively small batch size (32 images). The total training iteration was 310 epochs; the learning rate was le-4 in the first ten epochs, and the learning rate was downgraded to le-5 in the successive epochs. The training for all categories was performed for 310 epochs, and the loss was calculated using the binary cross-entropy loss function.

To prevent overestimating and overfitting our model's performance, we ensured that both the previous train-test split and the subdivision of the training set were done "patient-dependent" to ensure that no images from a single patient could appear in training corresponding validation sets. The final model parameters, listed in **Table 1**, were selected based on the validation set's accuracy (**Figure 4**) and used for the testing set.

Final test and clinical evaluation

To evaluate the final AI model's performance, we used the confusion matrix and the receiver operating characteristic curve (ROC curve) (22, 23). The confusion matrix, comprising four parameters such as true positive (TP), true negative (TN), false positive (FP), and false-negative (FN), was used to evaluate the accuracy, precision, recall (sensitivity), and F1-score. The ROC curve evaluated the false-negative performance with both continuous and ordinal scales (24). Negatives were summarized with a graphical plot of 1-specificity against the sensitivity and the area under the ROC curve (AUC). Attempting to fathom which pathognomy features were critical in associating with BCVA, we used the Grad-CAM technique to visualize the heat map of AI's recognitions (25–27).

Results

Image collection

A total of 259 patients with DME were recruited, and eight patients with visual acuity of decimal notation less than 0.05 were excluded. The participants were mostly over 60 years old, with an average age of 63 years. The ratio of males was 130 (51.8%). While 17.5% of patients had clear lenses, the remaining 82.5% had undergone intraocular lens (IOL) surgery. The database contained 3,920 images. Images from 24 randomly selected patients (9.6% of 251 patients) were preserved as the final test set, and the rest of the images constituted the training and validation sets. A total of 182 and 45 patients have been assigned to training and validation datasets, respectively. Therefore, a total of 1,431 OCT images labeled as "impaired vision" and 1,395 OCT images labeled as "adequate vision" constituted the training set (70% of all enrolled images), the validation dataset (20% of all enrolled images) contained 386 OCT images with "impaired vision" label and 393 OCT images with "adequate vision" label. The test dataset (10% of all enrolled images) was composed of 315 OCT images, which contained 162 images with an "adequate vision" label and 153 images with an "impaired vision label," as shown in **Table 2**. Besides, BCVA values of the impaired and adequate groups dataset were similar in each dataset (the visual acuity of the impaired group and adequate group in each dataset was close to 0.22 and 0.68, respectively) (**Table 2**).

Model development

The CNN model EfficientNet achieved superior performance during the training process and was selected as the final model for subsequent verifications. The training process's detailed learning curve revealed that iterations attained lower loss and higher accuracy as the model underwent successive iterations (**Figure 4**). Finally, the validation accuracy curve achieved a testable level, and the training accuracy was higher than the validation accuracy, which meant that the training process was finished. The 232^{nd} epoch represented the best performance of the validation accuracy (76.1%). Hence, this trained AI model has been selected as the final model to execute the final test.

The final test of the trained artificial intelligence model

Finally, the final trained AI model was verified by the final test dataset to evaluate its realistic performance. The test dataset contained 162 images with the "adequate vision" label and 153 images with the "impaired vision" label. Our AI model's accuracy, precision, recall, and F1-score were 75.9, 68.6, 78.9, and 73.4%, respectively (**Figure 5A**). As was calculated from the receiver operator characteristic (ROC) curve, the area under the curve (AUC) was 0.801, with the confidence interval (CI) from 0.751 to 0.851 (**Figure 5B**). Furthermore, we applied heat map visualization to identify OCT image areas recognized by the AI to discriminate between BCVA classes (**Figure 6**). The heat maps highlighted a more extensive area covering nearly the entire retinal layer instead of specific smaller lesions in some cases. The more extensive coverage of heat maps identified by AI to be critical for the determination of BCVA could be related

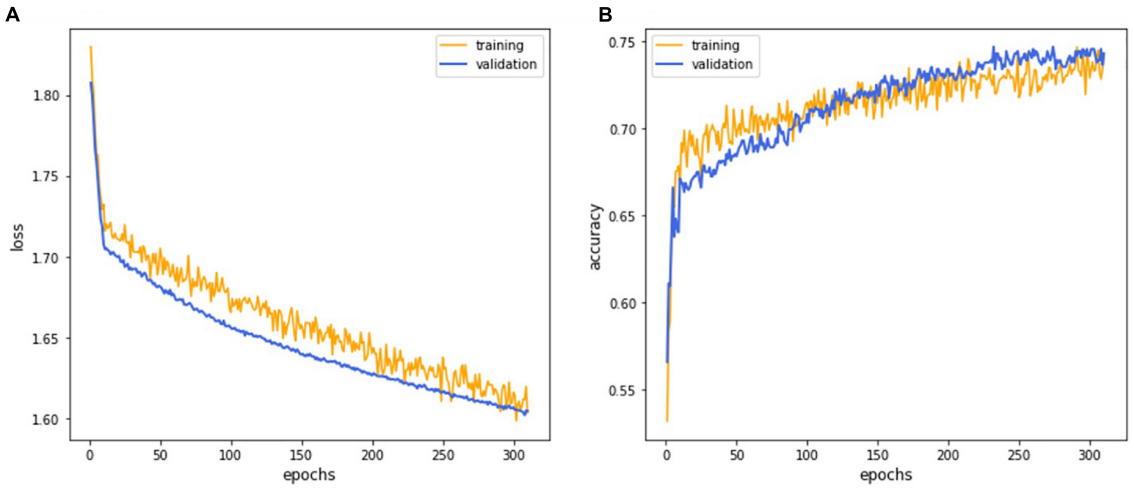

FIGURE 4
The deep learning model training curve. The CNN model EfficientNet training process revealed that iterations attained lower loss **(A)** and higher accuracy **(B)** as the model underwent successive iterations with the 252nd epoch representing the best performance.

TABLE 2 The details of the training, validation, and final test datasets list the numbers of allocated patients and optical coherence tomography (OCT) images and average BCVA values of patients.

Dataset	Training	Validation	Final test
Number of patients	182	45	24
Number of images	2,826 (1,431 impaired, 1,395 normal)	779 (386 impaired, 393 normal)	315 (153 impaireds, 162 normal)
BCVA (SD)	Impaired 0.22 (0.12)	Impaired 0.23 (0.13)	Impaired 0.22 (0.14)
	Normal 0.68 (0.17)	Normal 0.69 (0.20)	Normal 0.66 (0.14)

BCVA, best-corrected visual acuity; SD, standard deviation.

to the multiple microstructural changes and the thickness of the retina.

Discussion

Artificial intelligence in DR screening and referral decisions has achieved clinical reality. The first FDA-approved AI system (2018), IDx-DR (Iowa, USA), can analyze the digital fundus photograph (FP) in DR screening to provide referral suggestions (23). Apart from using the FP-based AI model, researchers have also developed an OCT-based algorithm, Notal OCT analyzer (NOATM, Notal Vision, Israel), which uses deep learning algorithms to detect the retinal fluid in AMD patients (24). However, up-to-date, in our literature review, there is no study focusing on an AI-based model to evaluate visual acuity in DR nor DME. Thus, our study developed an OCT-based AI model that could infer the binary VA status separated by the threshold of 0.5 (decimal notation) and attained an accuracy of 75.9%, a sensitivity of 78.9% with AUC = 80.1% based on OCT images only (Figure 4). Furthermore, to verify that our deep learning model indeed analyzed the exacted structural features of DME, we applied the heat map visualization to graphically show the different weighted values in pixel matrices of OCT images (Figure 6).

While visual acuity is a clinical measurement of changes in visual function as a primary endpoint, FDA recommends that retinal imaging technologies help determine anatomic markers for clinical progression of the disease. With the advent of imaging technologies such as the color fundus, angiography, and OCT, clinicians can observe the structural health of the neurosensory retina and generate new endpoints not previously accessible. For example, current technologies can identify onset, or progressions before symptom occurrence, leading to smaller marginal changes for earlier intervention and better visual outcomes. The use of OCT images was therefore incorporated into the work routine of the ophthalmologist to quantify the structural changes in individual patients' retinal pathological and topographic profiles (28). In addition, the ease of use and adoption into routine clinical practice makes the technology powerful to derive surrogate endpoints that change along with clinical endpoints and represent the disease status.

There are several limitations to our study. Rather than inferring the continuous VA variable, we only employed a binary classification of "impaired" and "normal" VA. Some may argue that the grade of impairment is essential as we may evaluate

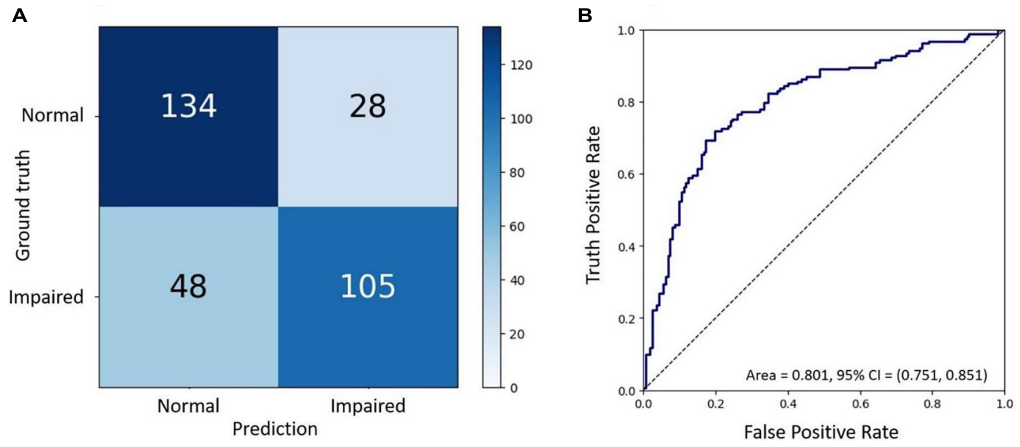

FIGURE 5
The final test of the trained AI model. **(A)** Confusion matrix demonstrating the accuracy of prediction of two visual acuity classes based on the validation dataset of OCT images. **(B)** Receiver operating characteristic (ROC) curve showing the accuracy of prediction with the area under the curve (AUC) = 0.801.

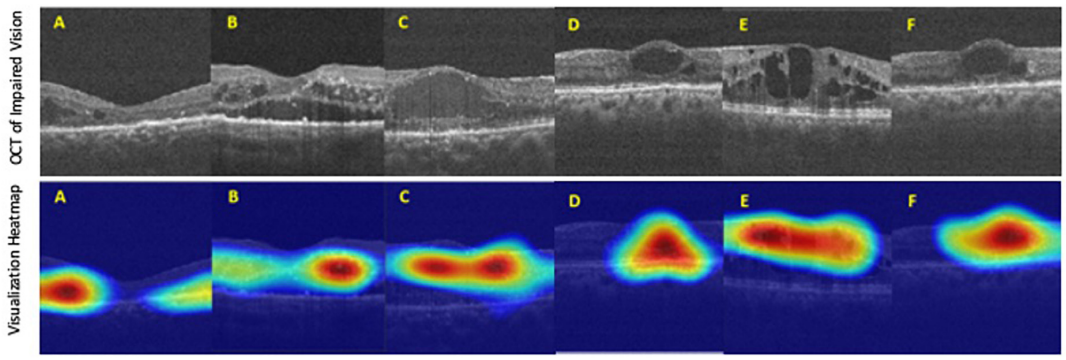

FIGURE 6
The heat map visualization of representative six **(A–F)** OCT images recognized by our AI model as predictors of best-corrected visual acuity (BCVA)-defined impaired vision. *Top panel*: original input images from the final test dataset. *Bottom panel*: heat map visualization of areas used by our AI model to discriminate between BCVA classes.

whether the patient is close to the treatment threshold or far away. Linear regression was not performed as we face small sample size that does not follow the assumption of normality, constant variance, and independent sampling, to construct a robust model in predicting visual function status at the decimal-level. Besides, our small sample size coupled with real-world heterogeneity caused our standard deviation of VA relatively large and BCVA measurements in the clinic may not be recorded as vigorously in controlled trials with EDTRS logmar standard. Our deep learning model may assist in evidence-based assistance to the physician, alleviating their burden in determining those with impaired vision (less than 0.5 baseline VA). Moreover, we only excluded patients with cataract diagnoses without cataract surgery. To achieve a better yield, we ideally have to impose exclusion criteria such as (1) prior history of choroidal neovascularization due to AMD, retinal vein occlusion, uveitis, or any other inflammatory disease, (2) presence of cataract or clouded lens, (3) glaucoma or any other neuropathy, (4) epiretinal membrane, vitreomacular traction disease, or any other maculopathy, and (5) corneal disease or degeneration. By only excluding cataracts, we obtain broader inclusion criteria that allow this AI model to closely imitate real-world settings and be expanded to accommodate most DME patients. Our AI model may be extended to serve a wider population by not excluding patients who underwent previous treatment and can be used for screening, referral, and monitoring. Finally, our model is constructed with horizontal and vertical scans of the mid-foveal position and not OCT volume. Therefore, we cannot analyze the concordance of the binary outcome of several OCT slices of the same patient and quantify their contradicting outcomes.

In the future, inferring VA based on imaging may be considered as quasi-functional surrogate endpoints for interventional clinical trials. By doing so, clinical trials can

enroll a larger set of patients that resemble those in the real world and provide treatment recommendations that can be implemented in the clinic (29–31). Furthermore, DME results in loss of visual function long before visual acuity is impaired as central acuity is not always affected. Herein, only a subgroup of DME fits the standard. Clinically significant macular edema (CSME) is defined as a lesion within 500 μm of the foveal center and center involved macular edema as central subfield retinal thickness of >250 μm in central 1 mm ETDRS grid (foveal thickness). Much macular retinal health recovery is not reflected in visual acuity. Visual acuity measured by visual charts (EDTRS, Snellen test) measures the photopic function of the central retina and is not reflective or sensitive to gain of retinal health or therapeutic benefits. Therefore, it is proposed that patient-reported outcome measures assess impairment of visual function in more detail. Redefine investigation of treatment effects superior to standard visual acuity testing without the need for extensive psychophysical examination. The European Medicines Agency (EMA) and FDA now demand the employment of patient-reported outcome measure (PROM) as functional endpoints in clinical trials (NEI-VFQ-25) are now routinely used as a valid and reliable measure of patients' vision-related quality of life. However, these tests are time-consuming, demanding for the elderly patient, and present significant inter-interpreter variability. In addition, rather than inferring function in a cross-sectional time manner for baseline VA, another interesting aspect is to predict VA in the future – what are the estimated letter gains after IVI-VEGF for my disease status? These algorithms inform patients about treatment prognosis and give patients the power to self-assess the cost-benefit of pursuing the IVI-VEGF. Overall, sensitive and robust outcome measures of retinal function are pivotal for measuring the clinical trial primary endpoint of VA and reinforce patient autonomy in the decision-making process.

Conclusion

This study built an OCT-based deep learning model that inferred VA status based on OCT and was correlated with the concurrent BCVA measured by standard visual charts. We achieved an accuracy of 75.9%, sensitivity of 78.9%, and a ROC AUC of 80.1%. This demonstrated the feasibility of predicting the functional outcome VA from routine ophthalmic images and served as a pilot study to develop further surrogate markers that can better represent the disease.

Data availability statement

The datasets presented in this article are not readily available due to legal, ethical and privacy restrictions of hospital data. Further inquiries can be directed to the corresponding author.

Ethics statement

The studies involving human participants were reviewed and approved by Institutional Review Board (IRB) of Taipei Veterans General Hospital (TVGH). Written informed consent for participation was not required for this study in accordance with the national legislation and the institutional requirements.

Author contributions

Y-CJ, Y-BC, S-JC, and D-KH contributed to the conception and design of the study. T-CL, C-HY, C-LK, P-YC, S-EH, S-HC, C-CH, and Z-KK organized the database. T-YL, H-RC, and H-YH performed the statistical analysis. T-YL and Y-IH wrote the first draft of the manuscript. K-JC wrote sections of the manuscript. All authors contributed to the manuscript revision, read, and approved the submitted version.

Funding

This study was funded by Academia Sinica (AS-TM-110-02-02), Technology (MOST 108-2314-B-010-042-MY3, 109-2321-B-010-006, and 109-2811-B-010-523) and the Ministry of Health and Welfare (MOHW108-TDU-B-211-133001 and MOHW109-TDU-B-211-114001). Taipei Veterans General Hospital (V111C-209, V107E-002-2, and V108D46-004-MY2-1), VGH, TSGH, NDMC, AS Joint Research Program (VTA107/108-V1-5-1), VGH, NTUH Joint Research Program (VN106-02, VN107-16), and the "Center for Intelligent Drug Systems and Smart Bio-devices (IDS2B)" from The Featured Areas Research Center Program within the framework of the Higher Education Sprout Project by the Ministry of Education (MOE) in Taiwan.

Acknowledgments

Institutional Review Board (IRB) of Taipei Veterans General Hospital (TVGH) approved the study under IRB accession number 2021-08-003AC.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

References

1. Early Treatment Diabetic Retinopathy Study Research Group. ETDRS report number 10. Early treatment diabetic retinopathy study research group. *Ophthalmology.* (1991) 98:786–806.

2. Brown DM, Nguyen QD, Marcus DM, Boyer DS, Patel S, Feiner L, et al. Long-term outcomes of ranibizumab therapy for diabetic macular edema: the 36-month results from two phase III trials: RISE and RIDE. *Ophthalmology.* (2013) 120:2013–22.

3. Heier JS, Korobelnik JF, Brown DM, Schmidt-Erfurth U, Do DV, Midena E, et al. Intravitreal aflibercept for diabetic macular edema: 148-week results from the VISTA and VIVID studies. *Ophthalmology.* (2016) 123:2376–85.

4. Ishibashi T, Li X, Koh A, Lai TY, Lee FL, Lee WK, et al. The REVEAL study: ranibizumab monotherapy or combined with laser versus laser monotherapy in Asian patients with diabetic macular edema. *Ophthalmology.* (2015) 122:1402–15.

5. Schmidt-Erfurth U, Lang GE, Holz FG, Schlingemann RO, Lanzetta P, Massin P, et al. Three-year outcomes of individualized ranibizumab treatment in patients with diabetic macular edema: the RESTORE extension study. *Ophthalmology.* (2014) 121:1045–53.

6. Diabetic Retinopathy Clinical Research Network, Elman MJ, Qin H, Aiello LP, Beck RW, Bressler NM, et al. Intravitreal ranibizumab for diabetic macular edema with prompt versus deferred laser treatment: three-year randomized trial results. *Ophthalmology.* (2012) 119:2312–8.

7. Elman MJ, Ayala A, Bressler NM, Browning D, Flaxel CJ, Glassman AR, et al. Intravitreal ranibizumab for diabetic macular edema with prompt versus deferred laser treatment: 5-year randomized trial results. *Ophthalmology.* (2015) 122:375–81.

8. Schlegl T, Waldstein SM, Bogunovic H, Endstraßer F, Sadeghipour A, Philip AM, et al. Fully automated detection and quantification of macular fluid in OCT using deep learning. *Ophthalmology.* (2018) 125:549–58.

9. Deak GG, Schmidt-Erfurth UM, Jampol LM. Correlation of central retinal thickness and visual acuity in diabetic macular edema. *JAMA Ophthalmol.* (2018) 136:1215–6.

10. Zur D, Iglicki M, Busch C, Invernizzi A, Mariussi M, Loewenstein A, et al. OCT biomarkers as functional outcome predictors in diabetic macular edema treated with dexamethasone implant. *Ophthalmology.* (2018) 125:267–75.

11. Wells JA, Glassman AR, Jampol LM, Aiello LP, Antoszyk AN, Baker CW, et al. Association of baseline visual acuity and retinal thickness with 1-year efficacy of aflibercept, bevacizumab, and ranibizumab for diabetic macular edema. *JAMA Ophthalmol.* (2016) 134:127–34.

12. Schmidt-Erfurth U, Vogl WD, Jampol LM, Bogunovic H. Application of automated quantification of fluid volumes to anti-VEGF therapy of neovascular age-related macular degeneration. *Ophthalmology.* (2020) 127:1211–9.

13. Roberts PK, Vogl WD, Gerendas BS, Glassman AR, Bogunovic H, Jampol LM, et al. Quantification of fluid resolution and visual acuity gain in patients with diabetic macular edema using deep learning: a post hoc analysis of a randomized clinical trial. *JAMA Ophthalmol.* (2020) 138:945–53.

14. Sophie R, Lu N, Campochiaro PA. Predictors of functional and anatomic outcomes in patients with diabetic macular edema treated with ranibizumab. *Ophthalmology.* (2015) 122:1395–401.

15. Wong TY, Sun J, Kawasaki R, Ruamviboonsuk P, Gupta N, Lansingh VC, et al. Guidelines on diabetic eye care: the international council of ophthalmology recommendations for screening, follow-up, referral, and treatment based on resource settings. *Ophthalmology.* (2018) 125:1608–22.

16. Dandona L, Dandona R. Revision of visual impairment definitions in the international statistical classification of diseases. *BMC Med.* (2006) 4:7. doi: 10.1186/1741-7015-4-7

17. Nguyen QD, Brown DM, Marcus DM, Boyer DS, Patel S, Feiner L, et al. Ranibizumab for diabetic macular edema: results from 2 phase III randomized trials: RISE and RIDE. *Ophthalmology.* (2012) 119:789–801.

18. Korobelnik JF, Do DV, Schmidt-Erfurth U, Boyer DS, Holz FG, Heier JS, et al. Intravitreal aflibercept for diabetic macular edema. *Ophthalmology.* (2014) 121:2247–54.

19. Suzuki K. Overview of deep learning in medical imaging. *Radiol Phys Technol.* (2017) 10:257–73.

20. Alqudah AM. AOCT-NET: a convolutional network automated classification of multiclass retinal diseases using spectral-domain optical coherence tomography images. *Med Biol Eng Comput.* (2020) 58:41–53.

21. Tan M, Le QV. EfficientNet: rethinking model scaling for convolutional neural networks. In: *Proceedings of the 36th International Conference on Machine Learning, ICML 2019.* Long Beach, CA (2019). p. 6105–14.

22. Hwang DK, Hsu CC, Chang KJ, Chao D, Sun CH, Jheng YC, et al. Artificial intelligence-based decision-making for age-related macular degeneration. *Theranostics.* (2019) 9:232–45.

23. Wang S, Dong L, Wang X, Wang X. Classification of pathological types of lung cancer from CT images by deep residual neural networks with transfer learning strategy. *Open Med (Wars).* (2020) 15:190–7.

24. Hoo ZH, Candlish J, Teare D. What is an ROC curve? *Emerg Med J.* (2017) 34:357–9.

25. Cheng CT, Ho TY, Lee TY, Chang CC, Chou CC, Chen CC, et al. Application of a deep learning algorithm for detection and visualization of hip fractures on plain pelvic radiographs. *Eur Radiol.* (2019) 29:5469–77.

26. Iizuka T, Fukasawa M, Kameyama M. Deep-learning-based imaging-classification identified cingulate island sign in dementia with Lewy bodies. *Sci Rep.* (2019) 9:8944.

27. Zhang Y, Hamada M. DeepM6ASeq: prediction and characterization of m6A-containing sequences using deep learning. *BMC Bioinformatics.* (2018) 19:524. doi: 10.1186/s12859-018-2516-4

28. Piech C, Malik A, Topol EJ. Digitising the vision test. *Lancet.* (2021) 398:1296.

29. Baker, CW, Glassman AR, Beaulieu WT, Antoszyk AN, Browning DJ, Chalam KV, et al. Effect of initial management with aflibercept vs laser photocoagulation vs observation on vision loss among patients with diabetic macular edema involving the center of the macula and good visual acuity: a randomized clinical trial. *JAMA.* (2019) 321:1880–94.

30. Busch C, Iglicki M, Fraser-Bell S, Zur D, Rodríguez-Valdés PJ, Cebec Z, et al. Observation versus treatment in diabetic macular edema with very good visual acuity – the OBTAIN study. *Invest Ophthalmol Vis Sci.* (2019) 60:2600.

31. Bressler NM, Varma R, Doan QV, Gleeson M, Danese M, Bower JK, et al. Underuse of the health care system by persons with diabetes mellitus and diabetic macular edema in the United States. *JAMA Ophthalmol.* (2014) 132:168–73.

frontiers | Frontiers in Medicine

ORIGINAL RESEARCH
published: 13 October 2022
doi: 10.3389/fmed.2022.875242

Acceptance and Perception of Artificial Intelligence Usability in Eye Care (APPRAISE) for Ophthalmologists: A Multinational Perspective

Dinesh V. Gunasekeran[1,2,3], Feihui Zheng[1], Gilbert Y. S. Lim[1,3], Crystal C. Y. Chong[1], Shihao Zhang[1], Wei Yan Ng[1], Stuart Keel[4], Yifan Xiang[5], Ki Ho Park[6], Sang Jun Park[6,7], Aman Chandra[8], Lihteh Wu[9], J. Peter Campbel[10], Aaron Y. Lee[11], Pearse A. Keane[12], Alastair Denniston[13,14], Dennis S. C. Lam[15,16], Adrian T. Fung[17,18], Paul R. V. Chan[19], SriniVas R. Sadda[20], Anat Loewenstein[21], Andrzej Grzybowski[22,23], Kenneth C. S. Fong[24], Wei-chi Wu[25], Lucas M. Bachmann[26], Xiulan Zhang[5], Jason C. Yam[27], Carol Y. Cheung[27], Pear Pongsachareonnont[28], Paisan Ruamviboonsuk[29], Rajiv Raman[30], Taiji Sakamoto[31], Ranya Habash[32], Michael Girard[1,3], Dan Milea[1,33], Marcus Ang[1,2,3], Gavin S. W. Tan[1,2,3], Leopold Schmetterer[1,2,3], Ching-Yu Cheng[1,2,3], Ecosse Lamoureux[1,2,3], Haotian Lin[5], Peter van Wijngaarden[4], Tien Y. Wong[1,2,3,34] and Daniel S. W. Ting[1,2,3]

[1] Singapore Eye Research Institute (SERI), Singapore National Eye Center (SNEC), Singapore, Singapore, [2] School of Medicine, National University of Singapore (NUS), Singapore, Singapore, [3] Duke-NUS Medical School, Singapore, Singapore, [4] Department of Ophthalmology, University of Melbourne, Melbourne, VIC, Australia, [5] State Key Laboratory of Ophthalmology, Zhongshan Ophthalmic Center (ZOC), Sun Yat-sen University, Guangzhou, China, [6] Department of Ophthalmology, Seoul National University College of Medicine, Seoul, South Korea, [7] Department of Ophthalmology, Seoul National University Bundang Hospital, Seongnam-si, South Korea, [8] Department of Ophthalmology, Southend University Hospital, Southend-on-Sea, United Kingdom, [9] Asociados de Macula, Vitreo y Retina de Costa Rica, San José, Costa Rica, [10] Casey Eye Institute, Oregon Health and Science, Portland, OR, United States, [11] Department of Ophthalmology, University of Washington, Seattle, WA, United States, [12] Moorfields Eye Hospital, London, United Kingdom, [13] Department of Ophthalmology, University Hospitals Birmingham NHS Foundation Trust, Birmingham, United Kingdom, [14] Institute of Ophthalmology, University College London (UCL), London, United Kingdom, [15] International Eye Research Institute of the Chinese University of Hong Kong (Shenzhen), Shenzhen, China, [16] C-MER International Eye Research Center of the Chinese University of Hong Kong (Shenzhen), Shenzhen, China, [17] Specialty of Clinical Ophthalmology and Eye Health, Faculty of Medicine and Health, Westmead Clinical School, The University of Sydney, Sydney, NSW, Australia, [18] Department of Ophthalmology, Faculty of Medicine, Health and Human Sciences, Macquarie University Hospital, Sydney, NSW, Australia, [19] Department of Ophthalmology, University of Illinois College of Medicine, Chicago, IL, United States, [20] Department of Ophthalmology, Doheny Eye Institute, Los Angeles, CA, United States, [21] Department of Ophthalmology, Tel Aviv Medical Center, Tel Aviv, Israel, [22] Department of Ophthalmology, University of Warmia and Mazury, Olsztyn, Poland, [23] Institute for Research in Ophthalmology, Ponzan, Poland, [24] OasisEye Specialists, Kuala Lumpur, Malaysia, [25] Department of Ophthalmology, Chang Gung Memorial Hospital, Taoyuan, Taiwan, [26] Oculocare Medical AG, Zurich, Switzerland, [27] Department of Ophthalmology and Visual Sciences, The Chinese University of Hong Kong (CUHK), Hong Kong, Hong Kong SAR, China, [28] Vitreoretinal Research Unit, Department of Ophthalmology, Chulalongkorn University and King Chulalongkorn Memorial Hospital, Bangkok, Thailand, [29] Department of Ophthalmology, College of Medicine, Rangsit University, Rajavithi Hospital, Bangkok, Thailand, [30] Vitreo-Retinal Department, Sankara Nethralaya, Chennai, India, [31] Department of Ophthalmology, Kagoshima University, Kagoshima, Japan, [32] Bascom Palmar Eye Institute, Miami, FL, United States, [33] Copenhagen University Hospital, Copenhagen, Denmark, [34] Tsinghua Medicine, Tsinghua University, Beijing, China

OPEN ACCESS

Edited by:
Tyler Hyungtaek Rim, Duke-NUS Medical School, Singapore

Reviewed by:
Saif Aldeen AlRyalat, The University of Jordan, Jordan
Qingsheng Peng, Guangdong Provincial People's Hospital, China

***Correspondence:**
Daniel S. W. Ting
daniel.ting.s.w@singhealth.com.sg

Specialty section:
This article was submitted to Ophthalmology, a section of the journal Frontiers in Medicine

Received: *14 February 2022*
Accepted: *29 March 2022*
Published: *13 October 2022*

Citation:
Gunasekeran DV, Zheng F, Lim GYS, Chong CCY, Zhang S, Ng WY, Keel S, Xiang Y, Park KH, Park SJ, Chandra A, Wu L, Campbel JP, Lee AY, Keane PA, Denniston A, Lam DSC, Fung AT, Chan PRV, Sadda SR, Loewenstein A, Grzybowski A, Fong KCS, Wu W-c, Bachmann LM, Zhang X, Yam JC, Cheung CY, Pongsachareonnont P, Ruamviboonsuk P, Raman R, Sakamoto T, Habash R, Girard M, Milea D, Ang M, Tan GSW, Schmetterer L, Cheng C-Y, Lamoureux E, Lin H, van Wijngaarden P, Wong TY and Ting DSW (2022) Acceptance and Perception of Artificial Intelligence Usability in Eye Care (APPRAISE) for Ophthalmologists: A Multinational Perspective. Front. Med. 9:875242. doi: 10.3389/fmed.2022.875242

Background: Many artificial intelligence (AI) studies have focused on development of AI models, novel techniques, and reporting guidelines. However, little is understood about clinicians' perspectives of AI applications in medical fields including ophthalmology, particularly in light of recent regulatory guidelines. The aim for this study was to evaluate the perspectives of ophthalmologists regarding AI in 4 major eye conditions: diabetic retinopathy (DR), glaucoma, age-related macular degeneration (AMD) and cataract.

Methods: This was a multi-national survey of ophthalmologists between March 1st, 2020 to February 29th, 2021 disseminated *via* the major global ophthalmology societies. The survey was designed based on microsystem, mesosystem and macrosystem questions, and the software as a medical device (SaMD) regulatory framework chaired by the Food and Drug Administration (FDA). Factors associated with AI adoption for ophthalmology analyzed with multivariable logistic regression random forest machine learning.

Results: One thousand one hundred seventy-six ophthalmologists from 70 countries participated with a response rate ranging from 78.8 to 85.8% per question. Ophthalmologists were more willing to use AI as clinical assistive tools (88.1%, $n = 890/1,010$) especially those with over 20 years' experience (OR 3.70, 95% CI: 1.10–12.5, $p = 0.035$), as compared to clinical decision support tools (78.8%, $n = 796/1,010$) or diagnostic tools (64.5%, $n = 651$). A majority of Ophthalmologists felt that AI is most relevant to DR (78.2%), followed by glaucoma (70.7%), AMD (66.8%), and cataract (51.4%) detection. Many participants were confident their roles will not be replaced (68.2%, $n = 632/927$), and felt COVID-19 catalyzed willingness to adopt AI (80.9%, $n = 750/927$). Common barriers to implementation include medical liability from errors (72.5%, $n = 672/927$) whereas enablers include improving access (94.5%, $n = 876/927$). Machine learning modeling predicted acceptance from participant demographics with moderate to high accuracy, and area under the receiver operating curves of 0.63–0.83.

Conclusion: Ophthalmologists are receptive to adopting AI as assistive tools for DR, glaucoma, and AMD. Furthermore, ML is a useful method that can be applied to evaluate predictive factors on clinical qualitative questionnaires. This study outlines actionable insights for future research and facilitation interventions to drive adoption and operationalization of AI tools for Ophthalmology.

Keywords: ophthalmology, artificial intelligence (AI), regulation, implementation, translation

INTRODUCTION

Aging populations are fueling an exponential growth in the demand for eye care and insufficient capacity of eye care services in many health systems (1–3). This has created mounting pressure to develop solutions that optimize existing resources, facilitate the triage of patients, and expand the surge capacity of health systems (4, 5). These constraints were heightened by clinical service disruptions during the coronavirus disease 2019 (COVID-19) outbreak, ranging from operational reorganization for pandemic responses as well as a mounting backlog of postponed elective services (5, 6). In response, the medical community has identified artificial intelligence (AI) as a potential solution to mitigate these pressures. A mature implementable AI digital solution could provide scalable automation, alleviate resource bottlenecks and expedite treatment process. This is particularly relevant for Ophthalmology, where extensive use of digital sensors and image-acquisition technologies provide a strong foundation for AI deployment (7).

Currently, AI for automated classification in ophthalmic imaging has been validated with clinically acceptable performance and evaluated in many studies (8–15), including clinical trials (7, 16), health economic analyses (17), reporting standards such as CONSORT-AI, SPIRIT-AI, and STARD-AI (18–22), AI ethics, trust, reproducibility, and explainability (23, 24). However, expert consensus for the acceptable forms of clinical AI applications have not been established. In recent clinical AI implementation studies, a range of barriers were reported to hinder successful clinical translation, for example lack of trust amongst stakeholders, organizational lack of capacity, and system limitations in necessary supporting infrastructure (25).

Earlier studies have surveyed general perceptions of AI among different users (e.g., medical students, radiologists) (26–28), although sample sizes were relatively small and limited to specific society or geographical location. Furthermore, none have evaluated the entire healthcare ecosystem from the microsystem (e.g., individual practitioners) (29, 30), to mesosystem (e.g., specific organizations) and macrosystem (e.g., system-level policies and population screening services) (31). These are crucial steps to determine practical requirements for effective clinical

implementation at each level of the health system, and to inform initiatives to facilitate sustained adoption (32). The objective of this study is to evaluate the acceptance and perception of AI applications among ophthalmologists for the leading causes of preventable blindness including diabetic retinopathy (DR), glaucoma, age-related macular degeneration (AMD), and/or cataract, using the United States Food and Drug Administration (US FDA) software as medical device (SaMD) guideline as a reference framework.

METHODS

This was an expert survey investigation of eye care practitioners regarding their perspectives for clinical artificial intelligence (AI) solutions in Ophthalmology. Responses from ophthalmologists to this anonymous web-based electronic survey are investigated in partnership with professional associations through convenient selection to reflect the spectrum of geographical regions and subspecialties across the Ophthalmology medical field. The temporal proximity of the study period (1 March 2020–1 March 2021) to the COVID-19 outbreak (declared a pandemic by the World Health Organization on 11 March 2020) also enabled collection of data regarding its impact on provider perspectives of AI applications. This research adhered to the tenets of the declaration of Helsinki, and Singhealth Institutional Review Board (IRB) approval was obtained with waiver of the need for informed consent (CIRB Ref 2020/2219).

Survey Development

The study survey was iteratively refined through literature review to develop semi-structured dichotomous and Likert questions (**Appendix 1**). This was followed by a pilot exercise with 6 clinical and academic Ophthalmology experts in Singapore, China, and Australia who have extensive experience in the conduct of AI-related research and recently published an AI-related peer-reviewed manuscript. Based on the results of the pilot exercise, the survey was finalized with optional responses programmed for individual qualitative questions. This was to avoid forced responses in the event a question was irrelevant for a given participants' practice setting [e.g., for **Supplementary Tables 1A–C**, regions with a lack of trained allied primary eye care services (PECS) or primary care provider (PCP) with eye care services]. Research was conducted remotely during the COVID-19 pandemic. It was hosted on an online survey platform (SurveyMonkey, San Mateo, USA) and designed to assess ophthalmologists' perspectives regarding their own organizations willingness to adopt AI as well as their own professional acceptance of various clinical AI applications for eye care.

First, professional acceptance of various clinical AI applications for eye care services was evaluated based on the regulatory guidance outlined in the SAMD document prepared by the International Medical Device Regulators Forum (IMDRF) working group chaired by the US FDA (33). A risk-based approach is applied accordingly, with ophthalmologists responding about their acceptance of AI applications in a matrix questionnaire based on the intended user, clinical context, and significance of the information provided to the healthcare decision based on the SaMD framework.

Intended users included ophthalmologists, primary eye care providers (PECPs, such as optometrists and opticians) and primary care providers (PCPs) with eye care services. Clinical contexts evaluated include the detection of common eye diseases DR, glaucoma, AMD and/or cataract.

Significance of information provided to the healthcare decision were classified based on the SaMD framework for intended uses to inform clinical management, drive clinical management, or diagnose eye diseases, as assistive tool, clinical decision support (CDS) tool or diagnostic tool, respectively (33). Applications of AI as assistive tools to inform clinical management include highlighting areas of interest in ophthalmic images for the practitioners' consideration to arrive at a diagnosis and treatment plan. Applications of AI as CDS tools to drive clinical management include providing possible provisional diagnoses based on areas of interest in ophthalmic images for the practitioners' consideration to develop a treatment plan. Applications of AI as diagnostic tools include providing a clinical diagnosis including stage of disease based on ophthalmic images, with or without management recommendations.

Next, Ophthalmologists' views on factors contributing to AI acceptance were evaluated considering all levels of the healthcare ecosystem from the microsystem to the macrosystem. First, the factors contributing to technology acceptance at the level of the healthcare microsystem including professional acceptance of clinical AI applications, acceptable level of error, perceived impact on professional roles, and potential barriers/enablers for adoption were explored (31). Second, factors contributing to technology acceptance at the level of the healthcare mesosystem were explored, including perceived willingness to adopt AI for clinical services within the organizations they practice in, anticipated organizational impact of clinical AI adoption, and likelihood of organizational facilitation of its adoption. At the mesosystem level, participants were also asked about the perceived willingness of their organizations to adopt AI for screening or diagnosis of the four major contributors to avoidable blindness, namely, diabetic retinopathy (DR), glaucoma, age-related macular degeneration (AMD) or cataract (34). Furthermore, participants were asked to report any anticipated organizational impact of the adoption of clinical AI for eye care services. Third, participants were asked to report their perspectives on the potential value of AI at the level of the macrosystem for eye care services. Finally, given the proximity of survey dissemination with the onset of the coronavirus disease 2019 (COVID-19) pandemic, participants were also surveyed about their perceptions regarding its impact acting at the level of the healthcare macrosystem, on future meso- and micro-system priorities for adoption.

Survey Dissemination

The web-based survey was disseminated through snowball sampling of professional Ophthalmology associations. Collaborating associations were selected to represent participants from a breadth of clinical Ophthalmology and imaging subspecialties as well as geographical regions of practice

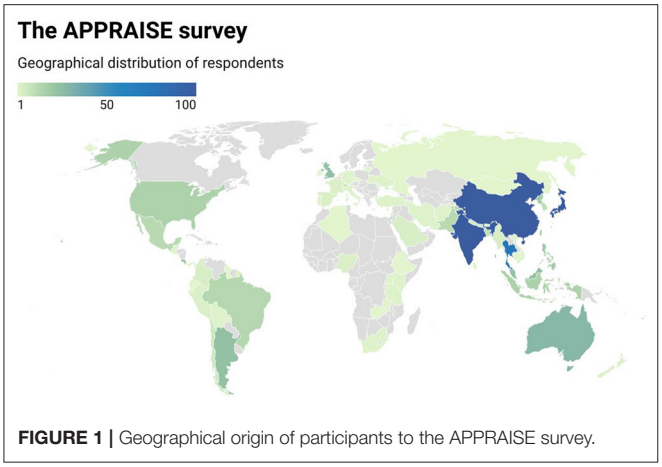

FIGURE 1 | Geographical origin of participants to the APPRAISE survey.

(Acknowledgment). Study recruitment was conducted using standardized invitations sent by the associations *via* their official established channels with all actively enrolled members. Recruitment was led by a study team member that was a member in good standing in each participating professional association.

The initial invitation to participate was sent to all actively enrolled members within each association. All invitations were sent by email and supplemented by regional practices based on the societies established channels with their members, such as WeChat in China. Invitations included the unique uniform resource locator (URL) of the web-based survey, that was programmed to restrict entries to one per participant-device to avoid duplication of entries from providers enrolled in multiple associations and/or receiving invites from multiple channels. Invitations were followed by 3 reminders at ∼2-week intervals, coordinated by the study team members.

Statistical Analysis

Responses were described with valid percentages for categorical variables as well as mean and standard deviation (SD) for continuous variables, with response rate tabulated for each question. The geographical origin of participants is classified based on the World Bank (WB) classification for 7 global regions (35). The economic background of participants is categorized using the 2017 International Council of Ophthalmology (ICO) classifications for low/intermediate and high resource settings, whereby countries grouped under resource-constrained settings were those classified by the WB as low- to upper-middle- income economies, and countries under resource-abundant settings were those classified by the WB as high-income economies (35, 36).

Quantitative analysis of any associations between provider acceptance and demographic information are reported. Multivariable logistic regression was performed to investigate any linear associations between provider acceptance of AI application in Ophthalmology and demographic information including age, gender, country (region of practice), economic background, experience, and self-rated understanding of AI for participants. To obtain a 95% confidence interval with 5% for the margin of error and 50% response distribution, a minimum sample size of 385 was calculated for the outcome of willingness to adopt AI in the next 5 years. Statistical significance was set at a *p*-value of 0.05. Analysis was performed using SPSS (IBM, SPSS Inc, USA).

In addition, machine learning (ML) analysis of survey responses was conducted using six selected input variables (clinical practice experience, World Bank geographical region, 2017 ICO classification for resource availability, gender, age, and self-reported AI understanding), to predict a total of 15 outcomes (output variables). An independent random forest model was trained to predict each outcome from the input variables in an exploratory analysis to assess for any non-linear associations between provider acceptance of AI and demographic information. The training dataset was randomly divided into

TABLE 1 | Demographics of participants.

Question	Responses	*N* (%)
Gender	Female	430 (36.6)
	Male	597 (50.8)
	Prefer not to say	149 (12.7)
Age (years)	Mean, Standard deviation (SD)	46.84, 10.936
Clinical practice experience in eye care services (eye screening, optometry, ophthalmology, etc)	Not practicing in eye care	13 (1.1)
	Currently in training (students in Ophthalmology and/or Optometry)	28 (2.4)
	<5 years clinical practice experience	73 (6.2)
	5–10 years clinical practice experience	183 (15.6)
	10–20 years clinical practice experience	323 (27.5)
	20–30 years clinical practice experience	337 (28.7)
	>30 years clinical practice experience	219 (18.6)
Participant region of practice using world bank classification (All: Missing 2)	East Asia & Pacific	870 (74.0)
	Europe & Central Asia	51 (4.3)
	Latin America & the Caribbean	87 (7.4)
	Middle East & North Africa	20 (1.7)
	North America	13 (1.1)
	South Asia	128 (10.9)
	Sub-Saharan Africa	7 (0.6)
How would you rate your understanding about deep learning, machine learning, and AI?	Excellent	66 (5.6)
	Above average	226 (19.2)
	Average	639 (54.3)
	Below average	196 (16.7)
	Very poor	49 (4.2)

1,000 subjects for training, and 176 subjects for validation. To train each random forest model, five-fold cross-validation was first performed on the training dataset, to optimize four hyperparameters: the entropy criterion, the maximum depth of the random forest trees, the maximum number of features, and the number of tree estimators. The optimal hyperparameters thus found were then used on train the final model on the full training dataset, and subsequently applied to the validation dataset to evaluate the area under the receiver operating curve (AUC).

The Breiman-Cutler permutation importance measure was used to determine the most important input variable(s) in predicting each outcome (37). The permutation importance measure was computed by permuting the column values of a single input variable, and calculating the drop in overall accuracy caused by the permutation. For the outcome variables with six initial options (Strongly Agree, Agree, Neutral, Disagree, Strongly Disagree, Unsure), Strongly Agree and Agree were grouped together as positive outcomes, with the remaining options considered negative outcomes. For the outcome variables with three initial options (Yes, No, Unsure), Yes was considered a positive outcome, and No/Unsure as negative outcomes.

RESULTS

A total of 1,176 ophthalmologists from 70 countries responded to the survey with representation from all 7 world bank geographical regions (**Figure 1**), although a majority practice in the East Asia & Pacific (74.0%), South Asia (10.9%) and Latin America & Caribbean (7.4%) regions. Participants had a mean age of 46.7 +/– 10.9 years. There was a slightly increased number of 597 male (50.8%) compared to 430 female (36.6%) participants, whereby 149 (12.7%) participants opted not to disclose their gender.

Participants reported a spectrum of clinical experience mostly between 10 and 30 years, whereby 323 participants had 10–20 years (27.5%) and 337 participants 20–30 years (28.7%) of experience. When asked to rate their understanding about machine learning (ML), deep learning (DL), and artificial intelligence (AI), a majority self-rated their understanding as average (54.3%, n = 639/1,176). Participant demographics are detailed in **Table 1**.

Microsystem—Professional Acceptance of Clinical AI Applications for Eye Care

Participants were asked about their acceptance of various applications of AI for eye care services based on the solutions' intended user and clinical application in accordance with the SaMD regulatory framework. Assistive tools to inform clinical management were the most acceptable form of clinical AI application in ophthalmology, with applications designed for use by ophthalmologists (89.2%, n = 901/1,010) receiving higher acceptance than those intended for use by Primary Eye Care Providers (88.1%, n = 890/1,010). Professional acceptance of AI applications as CDS tools to drive clinical management received lower acceptance. Diagnostic tools intended for use by ophthalmologists received the lowest (59.1%, n = 597/1,010) acceptance among the 6 categories (**Figure 2A, Supplementary Table 1A**).

Multivariate analysis was also conducted for the professional acceptance of clinical AI applications based on their intended users with demographic factors included for adjustment (**Table 2**). In this model, the odds of professional acceptance of AI applications for PECPs as assistive tools was lower among participants practicing in Latin America and the Caribbean (OR 0.42, 95% CI: 0.19–0.90, p = 0.025) than those practicing in East Asia and the Pacific. However, acceptance of AI applications for PECPs as diagnostic tools was relatively higher among participants that self-rated their understanding of AI as average (OR 1.06, 95% CI: 0.01–0.84, p = 0.033) or above average (OR 1.21, 95% CI: 0.02–0.95, p = 0.044).

On the other hand, the odds of professional acceptance of AI applications for ophthalmologists as assistive tools was relatively higher among those with clinical experience of 20 or more years (OR 3.70, 95% CI: 1.10–12.5, p = 0.035). Similarly, acceptance of AI applications for ophthalmologists as diagnostic tools was relatively higher among participants with increasing age (OR 1.03, 95% CI: 1.00–1.05, p = 0.019), male gender (OR 1.40, 95% CI: 1.02–1.91, p = 0.037), and participants practicing in the Latin America and the Caribbean region (OR 2.20, 95% CI: 1.14–4.24, p = 0.019), although it was lower among resource-abundant practice settings (OR 1.36, 95% CI: 0.37–0.80, p = 0.002). No demographic variables had statistically significant associations with professional acceptance of AI applications as CDS tools for either group of intended users.

Next, participants reported the acceptable level of error for the various applications of AI for eye care services based on the intended user and application, when level of error was benchmarked against various practitioners (PECPs, general practitioners, general, and subspecialty-trained ophthalmologists). Overall, participants had greater expectations for the performance of clinical AI applications intended for use by ophthalmologists as opposed to that for use by PECPs (**Figure 2B, Supplementary Table 1B**). Specifically, the acceptable levels of error for Assistive and CDS applications of AI that were reported most frequently was the level of error equivalent to the intended user (whether PECP or ophthalmologist). On the other hand, the acceptable level of error for AI applications as diagnostic tools varied based on the intended user. The acceptable level of error for diagnostic tool applications of AI intended to be used by PECPs that was most frequently reported was a level of error equivalent to a general ophthalmologist (38.5%, n = 389/1,010). Participants were divided about the acceptable level of error for diagnostic tool applications of AI intended to be used by ophthalmologists, with equivalent to a general ophthalmologist (35.5%, n = 359/1,010) or subspecialty-trained ophthalmologist (36.0%, n = 364/1,010) being the most frequent responses.

Next, participants were surveyed about the potential impact of clinical AI on their professional roles and responsibilities

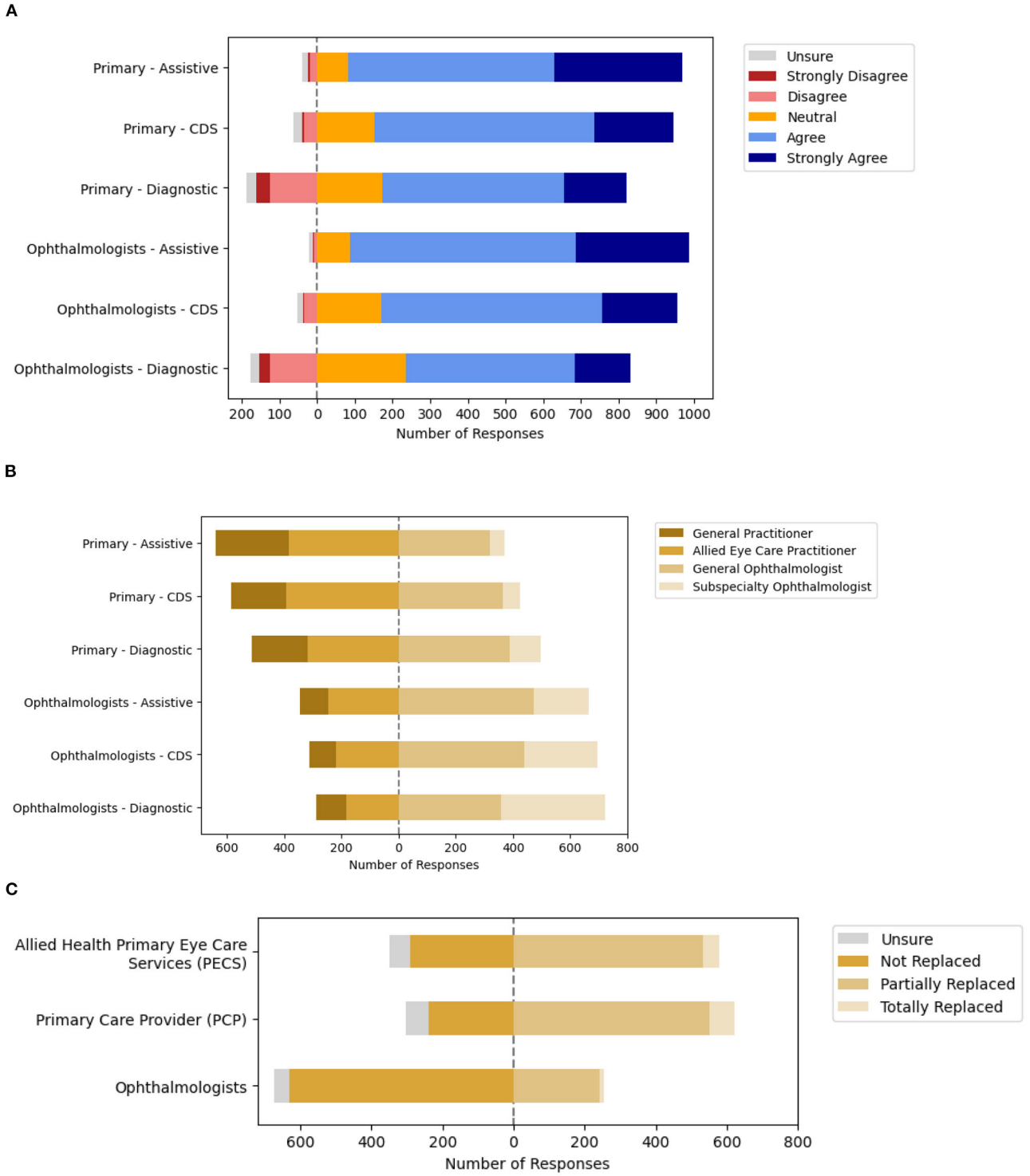

FIGURE 2 | Ophthalmologists acceptance of artificial intelligence (AI) based on the software as a medical device (SaMD) regulatory framework. **(A)** Acceptance of clinical AI based on significance of information and intended user. **(B)** Acceptable level of error based on significance of information and intended user. **(C)** Anticipated impact of clinical AI on professional roles.

at the level of the healthcare microsystem. The majority of participants that responded indicated that the eye care roles of ophthalmologists are not likely to be replaced (68.2%, n = 632/927), although those of others may be partially replaced including allied primary eye care service (PECS) providers (57.6%, n = 534/927) and primary care providers (PCP) with

TABLE 2 | Multivariate analysis for professional acceptance of artificial intelligence (AI) applications in Ophthalmology.

		Assistive tool for PECPs				CDS tool for PECPs				Diagnostic tool for PECPs				Assertive tool for ophthalmologists				CDS tool for ophthalmologists				Diagnostic tool for ophthalmologists			
		OR	95% CI		*p*-value	OR	95% CI		*p*-value	OR	95% CI		*p*-value	OR	95% CI		*p*-value	OR	95% CI		*p*-value	OR	95% CI		*p*-value
Age		0.982	0.953	1.012	0.236	1.002	0.977	1.027	0.884	1.020	0.999	1.042	0.065	1.002	0.971	1.035	0.892	1.017	0.992	1.044	0.188	1.026	1.004	1.049	0.019
Gender	Female	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref
	Male	1.253	0.794	1.976	0.332	1.191	0.830	1.711	0.343	1.089	0.798	1.486	0.590	1.316	0.817	2.121	0.259	1.434	0.991	2.076	0.056	1.398	1.020	1.914	0.037
Clinical experience	Currently in training	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref
	<20 years	2.158	0.674	6.910	0.195	0.806	0.287	2.263	0.683	1.018	0.446	2.323	0.966	2.450	0.956	6.280	0.062	0.951	0.389	2.325	0.912	0.747	0.328	1.700	0.487
	>20 years	1.676	0.437	6.423	0.451	0.700	0.214	2.291	0.556	1.007	0.382	2.654	0.988	3.702	1.096	12.510	0.035	1.165	0.395	3.436	0.782	1.037	0.394	2.727	0.941
Geographical region	East Asia and Pacific	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref
	Europe and Central Asia	1.200	0.450	3.204	0.716	1.000	0.475	2.103	0.999	1.334	0.700	2.543	0.380	0.638	0.274	1.486	0.298	1.292	0.574	2.908	0.536	1.068	0.565	2.018	0.840
	Latin America & the Caribbean	0.415	0.192	0.898	0.025	0.722	0.381	1.369	0.318	1.010	0.568	1.797	0.973	3.595	0.809	15.987	0.093	1.297	0.599	2.809	0.509	2.195	1.137	4.236	0.019
	Middle east and North Africa	0.578	0.153	2.187	0.419	0.832	0.254	2.721	0.761	0.833	0.290	2.393	0.735	1.793	0.224	14.323	0.582	0.996	0.266	3.728	0.995	0.740	0.250	2.188	0.586
	North America	1.000				2.965	0.372	23.658	0.305	1.869	0.488	7.160	0.361	1.174	0.145	9.542	0.881	0.983	0.205	4.703	0.983	2.210	0.567	8.615	0.253
	South Asia	0.874	0.381	2.006	0.751	1.008	0.552	1.840	0.980	0.987	0.595	1.637	0.960	0.598	0.287	1.249	0.171	0.785	0.432	1.428	0.428	0.604	0.364	1.000	0.050
	Sub-Saharan Africa	0.455	0.049	4.244	0.490	0.480	0.083	2.777	0.413	0.661	0.107	4.090	0.656	1.000				0.554	0.095	3.245	0.513	2.061	0.222	19.089	0.524
Income level	Resource-constrained	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref
	Resource-abundant	0.630	0.345	1.150	0.132	0.914	0.583	1.434	0.696	0.738	0.501	1.086	0.124	0.776	0.422	1.428	0.416	0.799	0.502	1.270	0.342	0.543	0.367	0.805	0.002
Self-rated understanding of AI	Very poor	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref
	Below average	0.713	0.078	6.553	0.765	1.286	0.315	5.261	0.726	0.418	0.050	3.489	0.420	0.611	0.070	5.293	0.655	1.058	0.295	3.798	0.931	0.851	0.210	3.450	0.821
	Average	0.587	0.073	4.714	0.616	0.840	0.226	3.125	0.795	0.107	0.014	0.837	0.033	0.608	0.076	4.865	0.639	1.499	0.441	5.094	0.517	0.325	0.086	1.226	0.097
	Above average	0.575	0.071	4.632	0.603	1.061	0.284	3.968	0.930	0.121	0.016	0.949	0.044	0.670	0.083	5.379	0.706	1.712	0.502	5.842	0.391	0.324	0.086	1.223	0.096
	Excellent	0.620	0.069	5.557	0.669	1.196	0.284	5.030	0.807	0.137	0.017	1.124	0.064	0.643	0.072	5.766	0.693	1.525	0.400	5.814	0.537	0.418	0.102	1.709	0.225

**Wherein "ref" denotes the reference category.*

The color values are added to draw attention of readers to analyses for which p-value was < 0.05.

eye care services (59.3%, n = 553/927). Detailed responses are included in **Figure 2C** and **Supplementary Table 1C**.

Finally, participants were surveyed about their perceptions of potential advantages and disadvantages of clinical AI for Ophthalmology to identify potential barriers and enablers for clinical AI adoption. Overall, the perceived advantages of clinical AI for Ophthalmology that were most frequently reported include improved patient access to disease screening (94.5%, n = 876/927), more targeted referrals to specialist care (87.1%, n = 807/927), and reduced time spent by specialists on monotonous tasks (82.7%, n = 767/927). The disadvantages of clinical AI for Ophthalmology that were most frequently reported include concerns over medical liability due to machine error (72.5%, n = 672/927), data security & privacy concerns (64.9%, n = 602/927), and concerns over the divestment of healthcare to large technology and data companies (64.1%, n = 594/927). Further detailed responses are depicted in **Figure 3** and **Supplementary Tables 1D,E**.

Mesosystem—Organizational Adoption of AI for Clinical and Eye Care Services

Participants were asked about the willingness to adopt clinical artificial intelligence (AI) in their organizations. Six hundred four participants (51.4%) reported that their organizations were willing to adopt AI for clinical practice in general within the next 5 years. A multivariate logistic regression model was applied to evaluate associations with participant demographics.

In this model, the odds of participants indicating organizational willingness to adopt AI within 5 years was higher among ophthalmologists of male gender (OR 1.58, 95% CI: 1.18–2.10) and those practicing in the North American region (OR 8.54, 95% CI: 1.86–39.5, $p = 0.006$) compared to the East Asia and Pacific region. However, the odds of organizational willingness to adopt AI was lower among participants from resource-abundant regions (OR 0.39, 95% CI: 0.27–0.56, $p < 0.001$). There were no significant associations between the odds of organizational willingness to adopt AI and the remaining demographic factors, including age, clinical experience, and self-rated understanding of AI (**Table 3A**).

Next, participants were asked about the perceived willingness of their organizations to adopt AI for leading validated applications in eye care services (**Figure 4A**). For screening applications of AI, most participants indicated organizational willingness to adopt AI including 920 participants for DR screening (78.2%), 832 for glaucoma screening (70.7%), 786 for AMD screening (66.8%), and 604 for cataract screening (51.4%). A multivariate logistic regression model was applied to evaluate associations between willingness to adopt AI for screening applications reported by participants, with their demographic factors (**Table 3B**).

In this model, the odds of organizational willingness to adopt AI being reported by ophthalmologists practicing in South Asia was relatively higher for DR screening (OR 2.07, 95% CI: 1.04–4.13, $p = 0.039$) compared to East Asia and the Pacific. On the other hand, that for glaucoma screening reported by

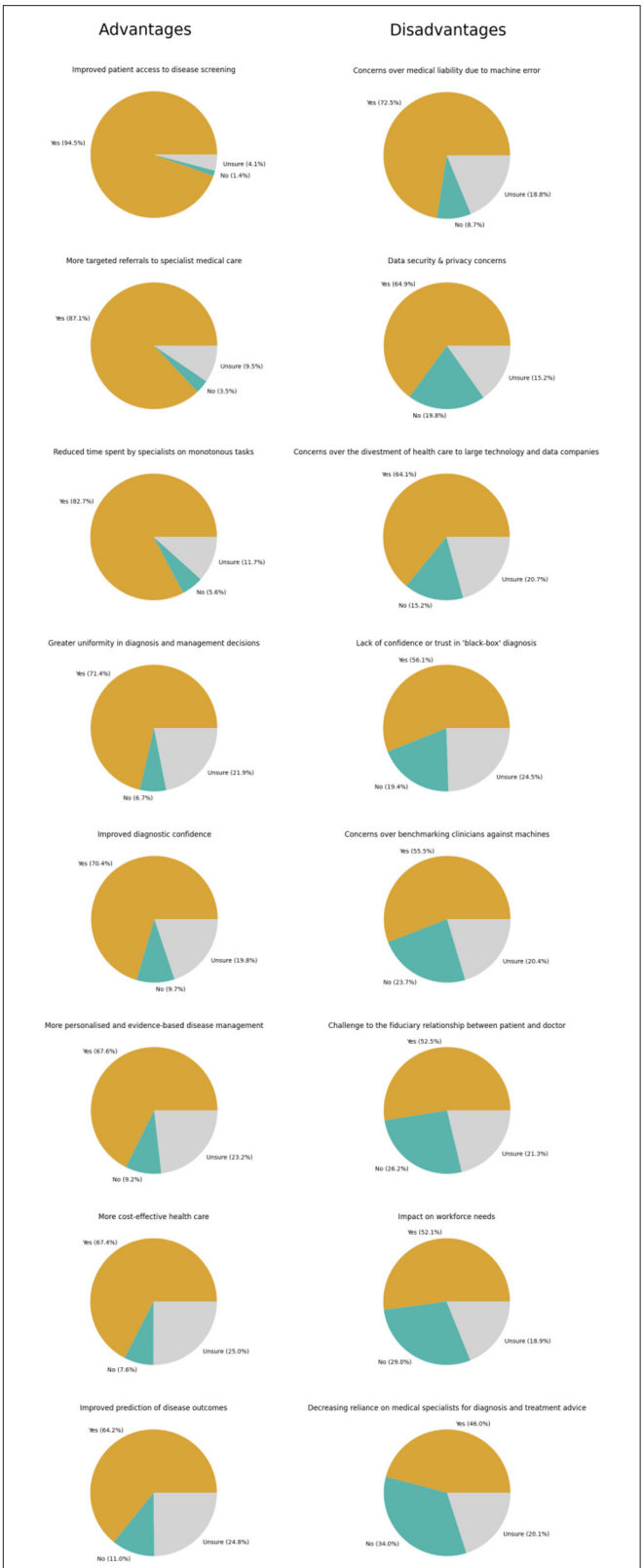

FIGURE 3 | Perceptions regarding advantages and disadvantages of artificial intelligence (AI) of ophthalmologists.

TABLE 3A | Organizational willingness to adopt AI for general eye practice.

		Odds ratio	95% confidence interval	*P*-value	
Age		1.007923	0.9889255	1.027285	0.416
Gender	Female	Ref	Ref	Ref	Ref
	Male	1.575258	1.181954	2.099436	0.002
Years of clinical experience	Student	Ref	Ref	Ref	Ref
	<20 years	0.7718551	0.3476733	1.713564	0.525
	20 or more years	0.7120689	0.2867984	1.767939	0.464
Geographical region	East Asia and Pacific	Ref	Ref	Ref	Ref
	Europe and Central Asia	1.471252	0.8191418	2.642501	0.196
	Latin America & the Caribbean	1.220085	0.7132367	2.087116	0.468
	Middle east and North Africa	0.6701729	0.252713	1.77724	0.421
	North America	8.540024	1.84594	39.50941	0.006
	South Asia	1.586754	0.9715878	2.591415	0.065
	Sub-Saharan Africa	0.4363696	0.0930749	2.045863	0.293
Income level	Resource-constrained	Ref	Ref	Ref	Ref
	Resource-abundant	0.390912	0.2737758	0.5581654	<0.001
Participant self-rated understanding of artificial intelligence (AI)	Very Poor	Ref	Ref	Ref	Ref
	Below Average	1.105834	0.3658182	3.342835	0.859
	Average	0.8993603	0.3182563	2.541501	0.841
	Above Average	0.6707563	0.2368241	1.899781	0.452
	Excellent	0.7322512	0.2382404	2.250634	0.586

**Wherein "ref" denotes the reference category.*
The color values are added to draw attention of readers to analyses for which p-value was < 0.05.

ophthalmologists in the Middle east and North Africa was relatively lower for glaucoma screening (OR 0.365, 95% CI: 1.04–4.13, $p = 0.033$) compared to East Asia and the Pacific. Finally, the odds of organizational willingness to adopt AI for cataract screening being reported by ophthalmologists of older age was higher (OR 1.02, 95% CI: 1.00–1.04, $p = 0.030$), while that by ophthalmologists of male gender was lower (OR 0.67, 95% CI: 0.50–0.88, $p = 0.04$). That for cataract screening was similarly lower among those practicing in the Europe and Central Asia region (OR 0.44, 95% CI: 0.24–0.84, $p = 0.012$) and North American region (OR 0.174, 95% CI: 0.04–0.81, $p = 0.026$).

Notably, the perception of organizational willingness to adopt diagnostic applications of AI was lower than screening applications (**Figure 4A**). Fewer participants indicated organizational willingness to adopt AI for diagnostic applications: positive responses were recorded from 638 participants for DR (54.3%), 490 for glaucoma (41.7%), 497 for AMD (42.3%) and 435 for cataract diagnosis (37.0%). A multivariate logistic regression model was applied to evaluate associations between willingness to adopt AI for diagnostic applications reported by participants, with their demographic factors (**Table 3C**).

In this model, the odds of organizational willingness to adopt AI being reported for DR diagnosis was higher for participants of older age (OR 1.02, 95% CI: 1.00–1.04, $p = 0.020$). That for DR diagnosis was also higher among participants practicing in the Latin America & the Caribbean region (OR 1.80, 95% CI: 1.04–3.11, $p = 0.035$) as well as the South Asian region (OR 1.69, 95% CI: 1.05–2.72, $p = 0.032$) relative to the East Asia and Pacific region. Similarly, the odds of organizational willingness to adopt AI being reported for Glaucoma diagnosis was higher for participants of older age (OR 1.04, 95% CI: 1.02–1.06, $p < 0.001$) and male gender (OR 1.53, 95% CI: 1.15–2.03, $p = 0.004$). However, that for Glaucoma diagnosis was lower for participants in the Middle east and North African region (OR 0.31, 95% CI: 0.10–0.95, $p = 0.040$).

Furthermore, the odds of organizational willingness to adopt AI being reported for AMD diagnosis was higher for participants of older age (OR 1.02, 95% CI: 1.00–1.04, $p = 0.029$) and male gender (OR 1.40, 95% CI: 1.06–1.87, $p = 0.020$). In addition, that for AMD diagnosis was also higher for participants practicing in the Latin America and the Caribbean region (OR 1.83, 95% CI: 1.08–3.11, $p = 0.026$) compared to the East Asia and the Pacific. On the other hand, the odds of organizational willingness to adopt AI being reported for Cataract screening were lower for participants practicing in the Europe and Central Asia region (OR 0.486, 95% CI: 0.24–1.00, $p = 0.049$) compared to the East Asia and the Pacific region.

Next, participants were asked about the anticipated organizational impact of the adoption of clinical AI for eye care services (**Figure 4B, Supplemental Table 2A**). Interestingly, some 55 participants indicated AI was already adopted for eye care services in their organizations (4.7%). Most of these participants had self-rated their understanding of clinical AI as excellent (16.4%, 9/55) or above average (27.3%, 15/55). These included 17 of the participants from the South Asian region (13.3%), 2 of the participants from Europe and central Asia region (3.9%), 33 of the participants from the East Asia and

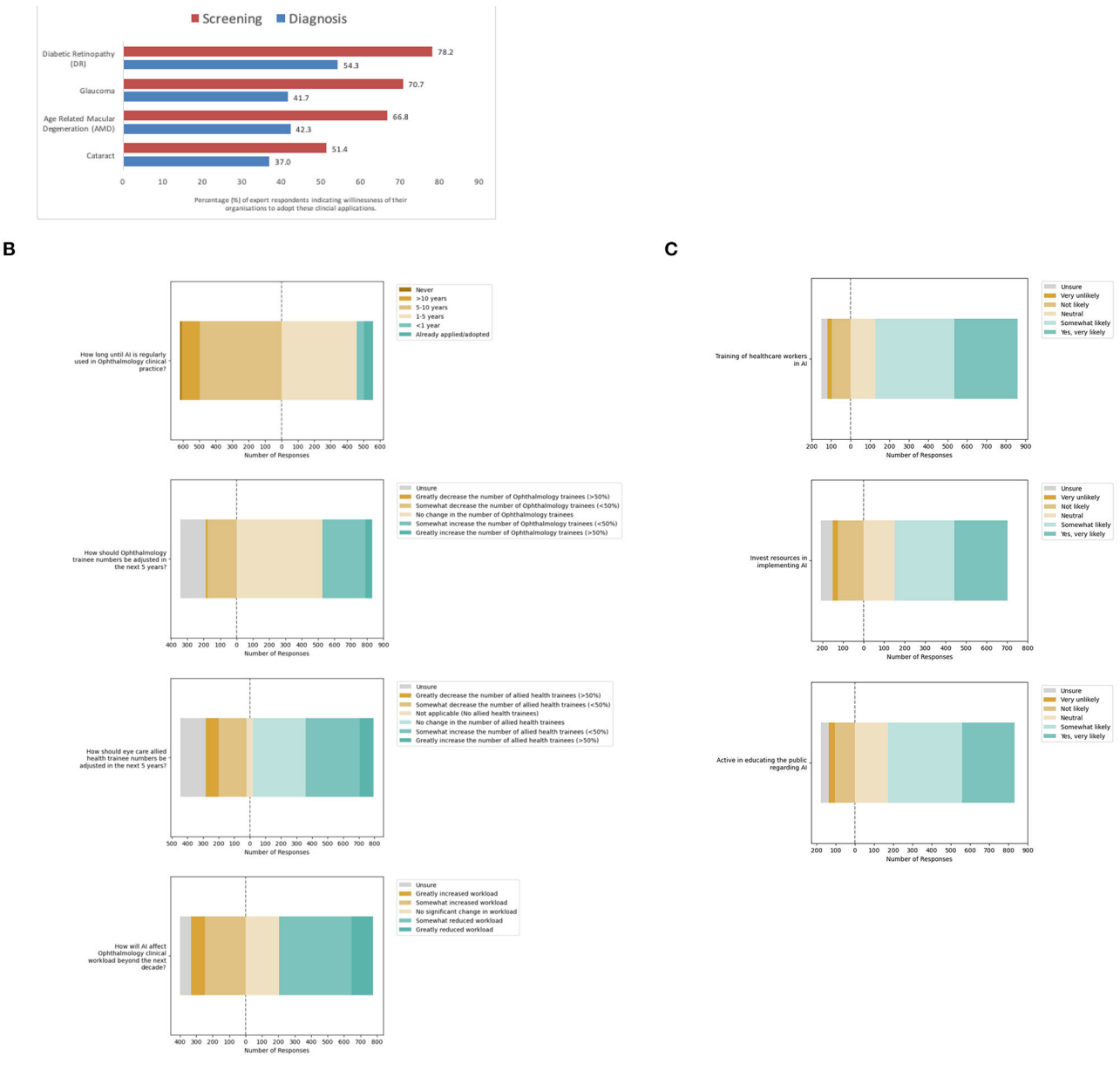

FIGURE 4 | Meso-system—Organizational adoption of clinical artificial intelligence (AI). **(A)** Organizational willingness to adopt AI for specific eye care services. **(B)** Anticipated organizational impact from adoption. **(C)** Perceived likelihood of organizational facilitation of adoption.

Pacific region (3.8%), and 3 of the participants from the Latin America and the Caribbean region (3.4%).

Despite the current progress in validation and implementation of AI for eye care services, less than half of all participants in this survey felt that AI would be regularly used in clinical practice within the next 5 years (47.4%, n = 558/1,176). Furthermore, participants had mixed views regarding the impact of AI on ophthalmology clinical workload, with some anticipating reduced workload (48.6%, n = 572/1,176) and others instead anticipating increased workload (28.2%, n = 332/1,176). When asked if trainee numbers should be increased, decreased or kept the same, most participants indicated that ophthalmology trainee numbers should not be adjusted (44.8%, n = 527/1,176), although some indicated allied health eye care trainee numbers should be increased (37.0%, n = 435/1,176).

With respect to their organizational willingness to facilitate adoption of AI tools for eye care services, study participants were optimistic overall. Many participants indicated that their

TABLE 3B | Organizational willingness to adopt applications of AI in Screening for eye diseases.

		Diabetic retinopathy (DR)				Glaucoma				Age related macular degeneration (AMD)				Cataract			
		OR	**95% CI**		***p*-value**	**OR**	**95% CI**		***p*-value**	**OR**	**95% CI**		***p*-value**	**OR**	**95% CI**		***p*-value**
Age		1.023	0.999	1.047	0.056	1.020	0.999	1.041	0.066	1.018	0.998	1.039	0.075	1.021	1.002	1.041	0.03
Gender	Female	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref
	Male	1.080	0.765	1.525	0.66	1.015	0.743	1.387	0.923	1.023	0.759	1.378	0.883	0.667	0.505	0.882	0.004
Clinical experience	Currently in training	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref
	<20 years	1.237	0.511	2.998	0.637	1.150	0.519	2.550	0.73	1.068	0.485	2.351	0.871	0.724	0.329	1.593	0.422
	>20 years	0.972	0.349	2.709	0.957	1.172	0.463	2.971	0.737	1.047	0.419	2.613	0.922	0.748	0.303	1.846	0.529
Geographical region	East Asia and Pacific	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref
	Europe and Central Asia	1.669	0.757	3.676	0.204	0.559	0.306	1.021	0.058	1.495	0.758	2.948	0.246	0.444	0.236	0.836	0.012
	Latin America & the Caribbean	1.631	0.767	3.470	0.204	1.695	0.899	3.195	0.103	1.691	0.931	3.073	0.085	0.863	0.515	1.446	0.576
	Middle east and North Africa	0.515	0.188	1.413	0.198	0.365	0.144	0.922	0.033	0.570	0.224	1.450	0.238	0.644	0.254	1.635	0.355
	North America	1.471	0.314	6.889	0.624	0.333	0.106	1.046	0.06	2.095	0.451	9.736	0.346	0.174	0.037	0.812	0.026
	South Asia	2.068	1.037	4.125	0.039	1.719	0.998	2.960	0.051	1.207	0.737	1.976	0.455	1.249	0.784	1.992	0.35
	Africa	1.180	0.137	10.176	0.88	2.248	0.263	19.214	0.459	0.373	0.080	1.732	0.208	1.597	0.296	8.615	0.586
Income level	Resource-constrained	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref
	Resource-abundant	0.729	0.472	1.126	0.154	1.205	0.821	1.769	0.34	1.193	0.825	1.726	0.349	0.755	0.533	1.069	0.113
Self-rated understanding of AI	Very poor	ref	ref	ref	ref	ref	ref	ref	ref	ref	ref	ref	ref	ref	ref	ref	ref
	Below average	0.927	0.235	3.659	0.914	1.201	0.371	3.887	0.759	0.842	0.263	2.694	0.772	1.778	0.594	5.317	0.303
	Average	1.151	0.313	4.240	0.832	1.274	0.422	3.846	0.668	1.038	0.344	3.130	0.948	1.224	0.442	3.388	0.698
	Above average	0.859	0.234	3.156	0.818	1.090	0.361	3.295	0.878	0.855	0.283	2.582	0.782	1.126	0.406	3.123	0.82
	Excellent	0.539	0.137	2.127	0.378	0.658	0.202	2.141	0.486	0.718	0.220	2.340	0.583	1.129	0.376	3.389	0.828

**Wherein "ref" denotes the reference category.*

The color values are added to draw attention of readers to analyses for which p-value was < 0.05.

TABLE 3C | Organizational willingness to adopt applications of AI for Diagnosis of eye diseases.

		Diabetic retinopathy (DR)			Glaucoma			Age related macular degeneration (AMD)			Cataract					
		OR	**95% CI**	***p*-value**	**OR**	**95% CI**	***p*-value**	**OR**	**95% CI**	***p*-value**	**OR**	**95% CI**	***p*-value**			
Age		1.023	1.004	1.042	0.02	1.036	1.016	1.056	0	1.021	1.002	1.041	0.029	1.018	0.998	1.038	0.072
Gender	Female	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref
	Male	1.158	0.877	1.529	0.301	1.527	1.148	2.031	0.004	1.403	1.055	1.865	0.02	0.828	0.619	1.107	0.202
Clinical experience	Currently in training	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref
	<20 years	1.369	0.628	2.982	0.429	1.143	0.493	2.651	0.755	1.667	0.705	3.944	0.245	0.771	0.343	1.732	0.528
	>20 years	1.221	0.501	2.976	0.66	0.860	0.334	2.212	0.754	1.904	0.729	4.972	0.188	0.936	0.372	2.357	0.888
Geographical region	East Asia and Pacific	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref
	Europe and Central Asia	1.245	0.693	2.234	0.464	0.968	0.528	1.775	0.916	1.656	0.915	2.996	0.095	0.486	0.237	0.997	0.049
	Latin America & the Caribbean	1.801	1.043	3.109	0.035	1.242	0.735	2.098	0.418	1.829	1.077	3.106	0.026	0.737	0.427	1.272	0.273
	Middle east and North Africa	0.721	0.288	1.806	0.485	0.306	0.099	0.948	0.04	0.530	0.185	1.521	0.238	0.347	0.099	1.221	0.099
	North America	0.734	0.237	2.277	0.593	0.633	0.196	2.046	0.445	0.809	0.252	2.594	0.722	0.307	0.066	1.433	0.133
	South Asia	1.688	1.046	2.724	0.032	1.102	0.687	1.766	0.687	1.295	0.810	2.071	0.28	1.118	0.697	1.794	0.643
	Africa	1.733	0.324	9.272	0.521	2.734	0.504	14.826	0.243	0.453	0.084	2.454	0.358	3.202	0.594	17.247	0.176
Income level	Resource-constrained	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref
	Resource-abundant	0.772	0.545	1.095	0.147	0.864	0.604	1.236	0.423	0.712	0.498	1.018	0.063	0.723	0.502	1.040	0.08
Self-rated understanding of AI	Very poor	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref
	Below average	0.669	0.210	2.129	0.497	0.924	0.307	2.781	0.889	0.706	0.233	2.139	0.538	1.507	0.494	4.598	0.471
	Average	0.535	0.179	1.600	0.263	0.637	0.226	1.793	0.393	0.508	0.179	1.442	0.203	0.985	0.344	2.818	0.978
	Above average	0.436	0.145	1.305	0.138	0.573	0.203	1.618	0.293	0.475	0.167	1.351	0.163	1.028	0.359	2.948	0.959
	Excellent	0.628	0.194	2.030	0.437	0.656	0.215	2.001	0.458	0.755	0.246	2.322	0.624	1.214	0.391	3.763	0.737

**Wherein "ref" denotes the reference category.*

The color values are added to draw attention of readers to analyses for which p-value was < 0.05.

organizations were very likely or somewhat likely to specifically train healthcare workers in the use and understanding of AI (72.3%, n = 730/1,010), invest resources for implementation (54.5%, n = 550/1,010), and actively educate the public regarding the use of AI in Ophthalmology (65.3%, n = 660/1,010). Detailed responses are included in **Figure 4C** and **Supplemental Table 2B**.

Macrosystem—Value of Clinical AI Applications for Eye Care Across the Health System

Many participants indicated that they strongly agree or agree that clinical AI will facilitate improvements in accessibility (84.7%, n = 785/927), affordability (61.9%, n = 574/927), and quality (69.4%, n = 643/927) in eye care services. Detailed responses are depicted in **Figure 5A** and **Supplemental Table 3A**. Next, participants were surveyed about their perceptions regarding the impact of COVID-19 acting at the level of the healthcare macrosystem (**Figure 5B, Supplemental Table 3B**). Notably, many participants were optimistic regarding the potential for AI to reduce non-essential contact between providers and patients (80.9%, n = 750/927). However, participants were closely divided regarding whether COVID-19 increased the likelihood of organizational AI adoption (50.2%, n = 465/927) as well as organizational facilitation. Participants remained divided when the likelihood of organizational facilitation was explored in greater detail in terms of investing resources to implement AI (51.1%, n = 474/927), training healthcare workers in AI (52.4%, n = 486/927), and educating the public regarding AI (54.2%, n = 502/927).

Machine Learning Analysis for Clustering of Survey Responses

On analysis of the survey responses using machine learning (ML) models, predictive AUCs of between 0.52 and 0.83 were obtained in predicting binary outcomes with corresponding permutation importance depicted in **Appendix 2**. The outcome variable predicted with the highest AUC of 0.83 was on whether AI could be an acceptable assistive tool for ophthalmologists, whereas the AUC for predicting the application of AI as a diagnostic tool for ophthalmologists had relatively low values of 0.59 or below. Finally, the model achieved an AUC of 0.65 in predicting organizational willingness to adopt AI in clinical practice in 5 years, whereby the variables that had the greatest predictive value were those for self-reported AI understanding and resource availability, with clinical practice experience having low predictive value. Detailed results are demonstrated in the **Appendix 2**.

DISCUSSION

To our knowledge, this is the first study providing an in-depth evaluation of ophthalmologists acceptance of clinical AI for Ophthalmology that incorporates the relevant medical device regulatory framework. Provider perspectives on professional and organizational acceptance of clinical AI tools for eye care services are evaluated in this study involving participants from a spectrum of geographies and clinical subspecialties. A machine learning (ML) approach was applied to highlight the clustering of responses, illustrating the relevance of individual demographic and attitude variables on professional acceptance and likelihood of adoption. Overall, participants indicated high levels of professional and organizational acceptance of AI for eye care services. Potential important barriers and enablers for the implementation of these tools in clinical practice were also highlighted. Furthermore, the impact of COVID-19 on clinical AI adoption in Ophthalmology was assessed.

Healthcare Micro-System Considerations for the Implementation of Clinical AI

The results of this study suggest several considerations for facilitating adoption of clinical AI at the level of healthcare microsystems. Participants were more accepting of clinical AI applications as assistive tools rather than CDS or diagnostic tools, based on the software as a medical device (SaMD) regulatory framework for clinical intended uses of AI technology to inform clinical management by highlighting areas of interest, drive clinical management by initiating referrals, or diagnose eye diseases to recommend management, respectively. The process of forming a clinical diagnosis is a fundamental role of healthcare practitioners. It is a complexed art based on probabilistic, causal or deterministic reasoning, often without the availability of complete information (38). The practitioner has to identify patterns in clinical information about each individual patient in the context of their prior medical and contextual knowledge to form an impression, then validate it through trial of treatment or investigations (39). This may explain why assistive tools received the greatest acceptance as opposed to CDS or diagnostic tools that suggest or provide a diagnosis, given the inability to incorporate additional contextual and non-verbal information in AI for a holistic approach to evaluating patients (**Figure 2A**).

The perceived enablers of improved accessibility and optimized referrals from screening as well as acceptance of these tools when designed for used by allied PECPs also suggests avenues to optimize solution development and deployment. Applying design considerations to facilitate operationalising clinical AI used by PECPs within the community to filter out patients with advanced illness requiring tertiary care are more acceptable to stakeholders. This may require embedded systems to facilitate referrals to ophthalmologists using the sorting conveyor or pyramid operational models where required (31).

Most ophthalmologists were not concerned about the threat of being replaced by AI (**Figure 2C**). The views of participants in this study are consistent with that from studies in other fields including Pathology and Radiology, 2 other medical fields with leading applications of AI for classification of medical imaging (11, 40). Many Diagnostic Pathologists in recent study reported that a negative impact of clinical AI on their professional compensation was unlikely (65.6%) and displacement or negative career impacts were limited (38.0%). They instead anticipated an increase in employment prospects (42.4%) (28). Similarly, among European Radiologist participants, most anticipated an increase in job opportunities (58%), with increased clinical roles

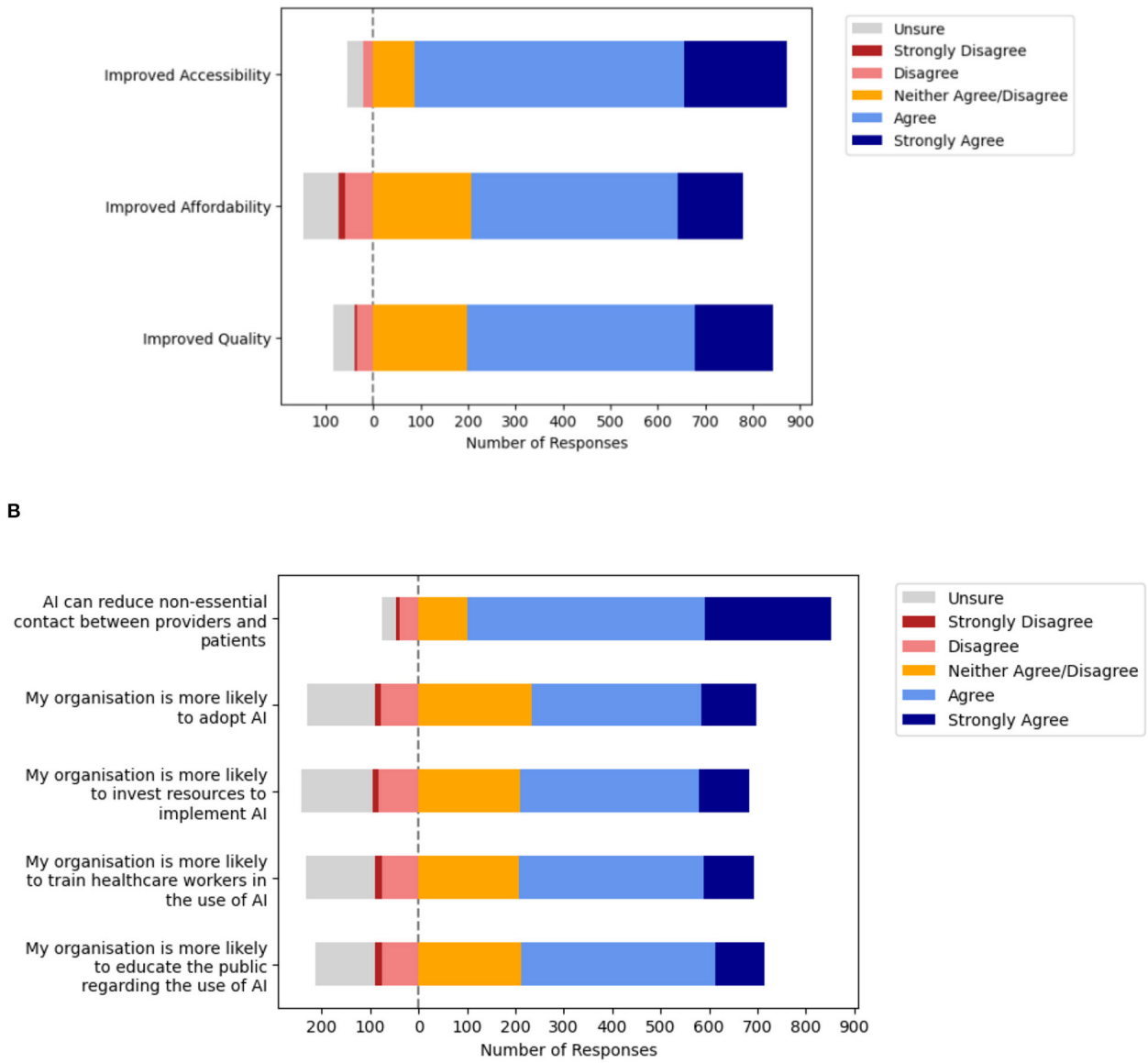

FIGURE 5 | Macrosystem—Ophthalmologists perceptions about artificial intelligence (AI) and the impact of the pandemic on adoption within health systems. **(A)** Value of clinical artificial intelligence (AI) perceived by Ophthalmologists. **(B)** Impact of the coronavirus disease 2019 (COVID-19) on health system adoption.

(54%) and decreased administrative roles such as reporting (75%) (27), in keeping with the advantage of reduced monotonous tasks perceived by participants.

Therefore, we find that participants are confident in their clinical roles and do not perceive AI to be a major threat to professional roles. This is consistent with the lower agreement reported for relevant potential disadvantages of AI (**Figure 3**), such as decreased reliance of medical specialists for diagnosis and treatment advice. These results are also consistent with an acceptance survey conducted across 22 provinces in China, whereby few healthcare workers anticipate replacement of clinical activities with AI (6.0%), while being receptive to applications that assist diagnosis (40.0%) and treatment (39.2%) (41).

The major disadvantages of clinical AI that participants agreed upon include potential medical liability from machine error, data security, privacy, and potential divestment of healthcare to corporate entities (**Figure 3**). Yet, despite these limitations, it has been successfully trained and validated for classification tasks of medical imaging for screening and diagnosis with clinically acceptable performance (42). The progress in this technology is reflected in the high rates of provider acceptance for the various abovementioned clinical applications. Furthermore, participants largely agreed on the advantages of AI (**Figure 3**), including

improved patient access to disease screening, targeted referrals to specialist medical care, and reduce time spent on monotonous tasks. Decentralized and improved access to screening has increased relevance today given widespread fear of viral exposure within hospitals that has prompted many patients to post-pone regular eye screening and monitoring (43, 44). However, fewer participants agreed that AI would improve care by making it more personalized, cost-effective, or predictive to pre-empt the clinical needs of patients.

These findings highlight the need for greater stakeholder engagement to emphasize advantages of AI in tandem with research to address disadvantages perceived by experts (31, 45). For example, participants flagged up lack of trust and confidence in the "black-box" diagnosis inherent with existing solutions, which could be addressed by emerging solutions such as saliency maps to improve algorithmic transparency. More pragmatic and qualitative investigations of AI implementation to address these potential barriers and enablers of adoption are needed to facilitate successful implementation of clinical AI in practice (46).

Healthcare Meso-System Considerations for the Implementation of Clinical AI

The analysis of healthcare meso-system considerations for clinical AI implementation highlight several trends in receptiveness to the adoption of clinical AI at the level of specific eye care services and within healthcare organizations. Overall, participants reported greater likelihood of organizational willingness to adopt screening applications rather than diagnostic applications. There was greater acceptance of applications for detection of DR, particularly in South Asia, and lower acceptance of applications for detection of Cataract, particularly in Europe and Central Asia (**Tables 3B,C**). This may relate to the importance of symptoms in the clinical evaluation of Cataracts, whereby screening models that incorporate AI screening or diagnostic applications within telemedicine platforms may facilitate real-world operational adoption (47).

Interestingly, the odds of reporting organizational willingness to adopt AI for certain applications were higher among participants with advanced age, including applications for detection of DR, glaucoma and AMD, as well as screening for cataracts (**Tables 3B,C**). This is congruent with the results of prior studies that have suggested increasing age may not be negatively correlated with health technology acceptance (41, 48). Furthermore, the odds of reporting organizational willingness to adopt AI were higher among participants with male gender for detection of glaucoma and AMD, although they were lower for cataract. These findings for AI adoption reflect the facilitating conditions, subjective norm, and social influence factors required for successful technology adoption from established theoretical models such as the technology acceptance model (TAM) and unified theory of acceptance and use of technology (UTAUT) (49, 50).

Yet, despite all the progress in the field of AI for ophthalmology, less than half of participants felt that AI is likely to be implemented in the next 5 years nor likely to reduce clinical workload (**Figure 4**). It follows that participants felt ophthalmology trainee numbers should not be adjusted (44.8%, $n = 527/1{,}176$). In ophthalmology, confidence in professional responsibilities likely stems from the procedural and surgical roles of professionals that cannot be replaced by AI. This interventional workload will likely increase with enhanced detection of eye diseases through the use of clinical AI to scale-up screening services, as reflected in the advantages of clinical AI anticipated by participants (**Figure 3**) including improved patient access to eye screening (94.5%) and targeted referrals to specialists (87.1%).

Furthermore, current evidence supports the improved cost-effectiveness of AI for eye care when applied in semi-autonomous models due to lower false positive referrals (17), highlighting that AI applied in partnership with healthcare practitioners will likely result in superior outcomes. In addition, apart from interventions, provision of clinical care also requires considerable management of technology for operational and administrative requirements. Earlier studies have highlighted that implementation of new technology such as electronic medical records (EMRs) can lead to delays and reduced efficiency in eye care services (51), requiring added time to review and interpret information (52). This highlights the need for design thinking approaches in the development of these tools, to streamline the aggregation and visualization of clinically relevant information from AI that can be conveniently interpreted by practitioners and applied in clinical practice (30).

Notably, most participants reported that their organizations are currently likely to facilitate adoption (**Figure 4**) through training healthcare workers in AI, investing resources for implementation, and actively educating the public. Based on the disadvantages reported in this study (**Figure 3**), remaining barriers that need to be addressed for adoption include potential medical liability arising from machine error (72.5%), data security & privacy (64.9%), as well as potential divestment of healthcare to large technology and data companies (64.1%). These can be addressed through the engagement of relevant stakeholders to develop medicolegal and cybersecurity guidelines, as well as co-development of these tools with the concerted involvement of relevant clinical participants (7, 16). Enablers to facilitate adoption include improving access to eye screening (94.5%), optimizing the flow of patients within eye care services for more targeted referrals to specialists (87.1%), and reducing the need for specialists to spend time on monotonous tasks (82.7%), that can be targeted as operational outcomes or goals during health services research and product development.

Health Macro-System Considerations and Impact of COVID-19 on Clinical AI Implementation

Finally, factors affecting AI adoption at the level of the healthcare macrosystem were evaluated including perceptions regarding value of clinical AI and the impact of the pandemic on the likelihood of adoption. Participants were generally positive regarding the value of clinical AI. Most agreed that clinical AI for eye care services will improve the accessibility of eye care services (84.7%), although they were less certain

regarding improvements in affordability (61.9%) and quality (69.4%). Moreover, having experienced the macrosystem changes brought about by the coronavirus disease 2019 (COVID-19) pandemic, participants were optimistic about the potential value of clinical AI to minimize non-essential patient contact (80.9%). However, participants were divided regarding whether COVID-19 increased the likelihood of organizational AI adoption (50.2%) and facilitation through investment (51.1%), training of healthcare workers (52.4%), and educating the public (54.2%). This suggests that ophthalmologists are unsure about the impact of the pandemic at the macrosystem level on influencing the willingness to implement clinical AI among service providers and organizations within the health system.

Machine Learning Analysis for Clustering of Survey Responses

Feature permutation importance estimates were obtained to estimate the contribution of each feature, within the random forest model trained to predict each question outcome. For example, for the question "Will your organization be willing to adopt AI in clinical practice in 5 years," the features corresponding to one's understanding of DL/ML, and ICO2017 income, were significantly more important than the others. Feature permutation importance involves randomly shuffling the values for a particular feature, and observing the decrease in model performance due to this shuffling. If the performance decreases appreciably, the feature is regarded as relatively predictive.

The AUCs of between 0.52 and 0.83 obtained in predicting binary outcomes with corresponding permutation importance suggest that eye care professionals' acceptance of AI and perceived likelihood of implementation can be anticipated from their demographics and attitudes toward AI to an extent (**Appendix 2**). The outcome variable predicted with the highest AUC of 0.83 was on whether AI could be an acceptable assistive tool for ophthalmologists, for which World Bank geographical region was the relatively most important predictor by permutation importance value. On the other hand, the AUC for the model in predicting acceptance for the application of AI as a diagnostic tool for ophthalmologists had relatively low values of 0.59 or below, which suggests the input variables in this study were not able to accurately predict ophthalmologists' acceptance of this application.

Responses to the APPRAISE survey indicated that participants were optimistic about the deployments of AI applications that reduce non-essential contact between patients and providers, thereby minimizing the risk of infectious disease transmission. This highlights the public health importance for further research and capacity building in this critical field to rapidly scale-up eye care services to meet the growing needs for eye screening in aging populations. Ophthalmologists will need to work closely with computer scientists to ensure that AI solutions for healthcare are appropriately designed to assimilate into clinical workflows and incorporate relevant considerations such as professional acceptance for specific AI applications based on the intended users for the given solution. Interestingly, participants from resource abundant settings reported potential barriers to adoption including lower odds of organizational willingness to adopt AI as well as lower acceptance of specific AI applications, including diagnostic tools for ophthalmologists that would need to be considered in developing specific applications. Although professional acceptance for AI solutions was relatively greater in resource-constrained settings, potential challenges with infrastructural availability such as intermittent electricity or internet access will need to be considered in the design of these solutions (7, 25).

Limitations and Strengths

Limitations of this study include that a majority of participants originated from Asia Pacific, with less representation of participants from the West. Although the timing of survey dissemination facilitated evaluation of professional perspectives on AI adoption during an ongoing public health emergency, regions with higher official burden of COVID-19 at the time of survey dissemination had less representation in our results, as ophthalmologists may have been occupied with related public health initiatives at the time. Moreover, statistical assessment of the survey was not conducted and responses for all qualitative questions were not made compulsory, whereby non-response rates are indicated in the study tables. Therefore, description and analysis was conducted based on the valid responses with non-responses programmed as "missing."

In addition, limitations of the snowball sampling method used for the purpose of hypothesis-generation in this study include the exclusion of other stakeholders such as primary care providers (PCPs). Furthermore, specific response rate calculation for different regions and channels for recruitment were not possible given privacy restrictions of the professional associations and inability to deconflict participants with membership in multiple associations. Although the survey was programmed to restrict one response per participant to avoid duplicate responses, there is potential selection bias for stakeholders that are more actively engaged in professional associations. These limitations may limit the generalisability of findings from this study. Future studies can address these limitations through survey validation for reliability and reproducibility, multiple testing correction for future hypothesis-testing research, probability sampling methods with inclusion of PCPs, stratified response rate tabulation based on individual channels of recruitment, and increased representation of participants from the West.

Finally, another limitation consistent with earlier survey investigations is the use of a logistic regression analytic approach to investigate associations between independent variables such as demographics and dependent variables such as acceptance, which assumes a linear relationship between them. However, unlike earlier investigations, one strength of this study is the use of a decision-tree based machine learning (ML) analytic approach called random forests to analyse responses in tandem with traditional logistic regression. The main distinction between the two analytic approaches rests in the transparency and underlying assumptions, whereby the flexibility of ML has allowed it to outperform the predictive accuracy of logistic regression in large empirical evaluations (53).

For logistic regression, a linear regression model on the input variables is transformed using the logistic function. It is therefore readily interpretable in terms of these input variables, assuming linearity of the input variables and log odds. However, logistic regression therefore remains a linear classifier, and non-linear relationships are not well-accounted for by the model. In contrast, the random forests ML classifier is able to model non-linear relationships in the data, with the trade-off of being less interpretable. This allows evaluation of a broader variety of potential relationships between the variables.

Additional strengths of this study include the consolidation of perspectives from a large and diverse spectrum of Ophthalmologists on the timely topic of clinical AI applications. The survey was also disseminated with close time-proximity to the COVID-19 outbreak, allowing assessment of the impact of a public health emergency on provider perspectives regarding clinical AI adoption. Finally, this study provides an in-depth investigation of professional acceptance of clinical AI solutions for automated classification of medical imaging in ophthalmology, incorporating a systematic approach to address factors affecting adoption at all levels from the micro-, meso-, and macrosystem. Furthermore, the intricacies of the latest regulatory guidance were applied in the evaluation of AI applications based on the intended user, significance of the information to the healthcare decision, and clinical context.

CONCLUSION

Artificial Intelligence (AI) has been established as a tool for health systems to improve the right-siting of patients. This study outlines several key considerations that inform future research, communication and facilitation interventions to drive effective adoption and operationalization of these tools in clinical practice. Actionable insights to facilitate AI adoption are also highlighted, including engagement of relevant stakeholders and operationalization based on the enablers of AI adoption identified in this study, as well as addressing perceived barriers through development of the technology and guidelines in collaboration with ophthalmologists.

DATA AVAILABILITY STATEMENT

The datasets presented in this article are not readily available because aggregated de-identified data for participant responses in this study are provided in the segment on **Supplementary Tables** and **Appendix**. Informed consent was not obtained for sharing of individual responses (study data) outside of the preparation of this article. Requests to access the datasets should be directed to DG, mdcdvg@nus.edu.sg.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by this research adhered to the tenets of the declaration of Helsinki, and Singhealth Institutional Review Board (IRB) approval was obtained with waiver of the need for informed consent (CIRB Ref 2020/2219). Written informed consent for participation was not required for this study in accordance with the national legislation and the institutional requirements.

AUTHOR CONTRIBUTIONS

DG, FZ, GL, CCYC, SZ, and WN contributed to the data analysis and drafting the initial manuscript. All listed authors contributed to study conceptualization, design, recruitment, data interpretation, manuscript preparation, and meet the criteria for authorship as agreed upon by the International Committee of Medical Journal Editors and are in agreement with the content of the manuscript. All authors contributed to the article and approved the submitted version.

ACKNOWLEDGMENTS

We would like to acknowledge the following ophthalmology societies for supporting the dissemination of the survey questionnaires to the society members: American Ophthalmology Society, Asia-Pacific Ocular Imaging Society (APOIS), Singapore College of Ophthalmology, Asia-Pacific Academy of Ophthalmology (APAO); Asia-Pacific Myopia Society (APMS); Asia-Pacific Vitreo-retina Society (APVRS), British and Eire Association of Vitreoretinal Surgeons (BEAVRS), China Ophthalmology Society, Chinese American Association of Ophthalmology, European Academy of Ophthalmology, European Association for Vision and Eye Research, Glaucoma Research Society, The Hong Kong Ophthalmological Society, International Retinal Imaging Symposium (IntRIS), Israel Ophthalmology Society, Japanese Vitreo-retinal Society, Korean Ophthalmology Society, Malaysian Society of Ophthalmology, Pan American Association of Ophthalmology, Switzerland Ophthalmology Society, The Royal Australian and New Zealand College of Ophthalmologists (RANZCO), The Royal College of Ophthalmologists of Thailand, Vitreoretinal Society of India (VRSI), All India Ophthalmological Society (AIOS).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmed.2022.875242/full#supplementary-material

REFERENCES

1. Resnikoff S, Felch W, Gauthier TM, Spivey B. The number of ophthalmologists in practice and training worldwide: a growing gap despite more than 200,000 practitioners. *Br J Ophthalmol.* (2012) 96:783–7. doi: 10.1136/bjophthalmol-2011-301378

2. Bourne RRA, Flaxman SR, Braithwaite T, Cicinelli MV, Das A, Jonas JB, et al. Magnitude, temporal trends, and projections of the global prevalence of blindness and distance and near vision impairment: a systematic review and meta-analysis. *Lancet Glob Health.* (2017) 5:e888–97. doi: 10.1016/S2214-109X(17)30293-0

3. Taylor HR. Global blindness: the progress we are making and still need to make. *Asia Pac J Ophthalmol.* (2019) 8:424–8. doi: 10.1097/APO.0000000000000264
4. Ting DS, Gunasekeran DV, Wickham L, Wong TY. Next generation telemedicine platforms to screen and triage. *Br J Ophthalmol.* (2019) 104:299–300. doi: 10.1136/bjophthalmol-2019-315066
5. Gunasekeran DV, Ting DSW, Tan GSW, Wong TY. Artificial intelligence for diabetic retinopathy screening, prediction and management. *Curr Opin Ophthalmol.* (2020) 31:357–65. doi: 10.1097/ICU.0000000000000693
6. Ting DSW, Carin L, Dzau V, Wong TY. Digital technology and COVID-19. *Nat Med.* (2020) 26:459–61. doi: 10.1038/s41591-020-0824-5
7. Gunasekeran DV, Wong TY. Artificial intelligence in ophthalmology in 2020: a technology on the cusp for translation and implementation. *Asia Pac J Ophthalmol.* (2020) 9:61–6. doi: 10.1097/01.APO.0000656984.56467.2c
8. 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) 137:987–93. doi: 10.1001/jamaophthalmol.2019.2004
9. Abràmoff MD, Tobey D, Char DS. Lessons learned about autonomous ai: finding a safe, efficacious, and ethical path through the development process. *Am J Ophthalmol.* (2020) 214:134–42. doi: 10.1016/j.ajo.2020.02.022
10. Wong TY, Bressler NM. Artificial intelligence with deep learning technology looks into diabetic retinopathy screening. *JAMA*. (2016) 316:2366–7. doi: 10.1001/jama.2016.17563
11. Ting DSW, Liu Y, Burlina P, Xu X, Bressler NM, Wong TY. AI for medical imaging goes deep. *Nat Med.* (2018) 24:539–40. doi: 10.1038/s41591-018-0029-3
12. Ruamviboonsuk P, Cheung CY, Zhang X, Raman R, Park SJ, Ting DSW. Artificial intelligence in ophthalmology: evolutions in Asia. *Asia Pac J Ophthalmol.* (2020) 9:78–84. doi: 10.1097/01.APO.0000656980.41190.bf
13. Milea D, Najjar RP, Zhubo J, Ting D, Vasseneix C, Xu X, et al. Artificial intelligence to detect papilledema from ocular fundus photographs. *N Engl J Med.* (2020) 382:1687–95. doi: 10.1056/NEJMoa1917130
14. Kugelman J, Alonso-Caneiro D, Read SA, Hamwood J, Vincent SJ, Chen FK, et al. Automatic choroidal segmentation in OCT images using supervised deep learning methods. *Sci Rep.* (2019) 9:13298. doi: 10.1038/s41598-019-49816-4
15. He M, Li Z, Liu C, Shi D, Tan Z. Deployment of artificial intelligence in real-world practice: opportunity and challenge. *Asia Pac J Ophthalmol.* (2020) 9:299–307. doi: 10.1097/APO.0000000000000301
16. Li JO, Liu H, Ting DSJ, Jeon S, Chan RVP, Kim JE, et al. Digital technology, tele-medicine and artificial intelligence in ophthalmology: a global perspective. *Prog Retin Eye Res.* (2020) 82:100900. doi: 10.1016/j.preteyeres.2020.100900
17. Xie Y, Nguyen QD, Hamzah H, Lim G, Bellemo V, Gunasekeran DV, et al. Artificial intelligence for teleophthalmology-based diabetic retinopathy screening in a national programme: an economic analysis modelling study. *Lancet Dig Health.* (2020) 2:e240–9. doi: 10.1016/S2589-7500(20)30060-1
18. Sounderajah V, Ashrafian H, Aggarwal R, De Fauw J, Denniston AK, Greaves F, et al. Developing specific reporting guidelines for diagnostic accuracy studies assessing AI interventions: The STARD-AI Steering Group. *Nat Med.* (2020) 26:807–8. doi: 10.1038/s41591-020-0941-1
19. Group D-AS. DECIDE-AI: new reporting guidelines to bridge the development-to-implementation gap in clinical artificial intelligence. *Nat Med.* (2021) 27:186–7. doi: 10.1038/s41591-021-01229-5
20. Liu X, Cruz Rivera S, Moher D, Calvert MJ, Denniston AK, SPIRIT-AI and CONSORT-AI Working Group. Reporting guidelines for clinical trial reports for interventions involving artificial intelligence: the CONSORT-AI extension. *Lancet Dig Health*. (2020) 2:e537–48. doi: 10.1136/bmj.m3164
21. Cruz Rivera S, Liu X, Chan AW, Denniston AK, Calvert MJ, SPIRIT-AI and CONSORT-AI Working Group, et al. Guidelines for clinical trial protocols for interventions involving artificial intelligence: the SPIRIT-AI extension. *Lancet Dig Health*. (2020) 2:e549–60. doi: 10.1136/bmj.m3210
22. Collins GS, Moons KGM. Reporting of artificial intelligence prediction models. *Lancet.* (2019) 393:1577–9. doi: 10.1016/S0140-6736(19)30037-6
23. Amann J, Blasimme A, Vayena E, Frey D, Madai VI, Precise QC. Explainability for artificial intelligence in healthcare: a multidisciplinary perspective. *BMC Med Inform Decis Mak*. (2020) 20:310. doi: 10.1186/s12911-020-01332-6
24. Vollmer S, Mateen BA, Bohner G, Király FJ, Ghani R, Jonsson P, et al. Machine learning and artificial intelligence research for patient benefit: 20 critical questions on transparency, replicability, ethics, and effectiveness. *BMJ.* (2020) 368:l6927. doi: 10.1136/bmj.l6927
25. Beede E, Baylor E, Hersch F, Iurchenko A, Wilcox L, Raumviboonsuk P, Vardoulakis L. 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.* ACM (2020). p. 1–12. doi: 10.1145/3313831.3376718
26. Sit C, Srinivasan R, Amlani A, Muthuswamy K, Azam A, Monzon L, et al. Attitudes and perceptions of UK medical students towards artificial intelligence and radiology: a multicentre survey. *Insights Imaging.* (2020) 11:14. doi: 10.1186/s13244-019-0830-7
27. European Society of Radiology. Impact of artificial intelligence on radiology: a EuroAIM survey among members of the European society of radiology. *Insights Imaging.* (2019) 10:105. doi: 10.1186/s13244-019-0798-3
28. Sarwar S, Dent A, Faust K, Richer M, Djuric U, Van Ommeren R, et al. Physician perspectives on integration of artificial intelligence into diagnostic pathology. *NPJ Digit Med.* (2019) 2:28. doi: 10.1038/s41746-019-0106-0
29. Chew AMK, Ong R, Lei HH, Rajendram M, K V G, Verma SK, et al. Digital health solutions for mental health disorders during COVID-19. *Front Psychiatry.* (2020) 11:582007. doi: 10.3389/fpsyt.2020.582007
30. Gunasekeran DV. Technology and chronic disease management. *Lancet Diabetes Endocrinol.* (2018) 6:91. doi: 10.1016/S2213-8587(17)30441-2
31. Gunasekeran DV, Tham YC, Ting DSW, Tan GSW, Wong TY. Digital health during COVID-19: lessons from operationalising new models of care in ophthalmology. *Lancet Digit Health.* (2021) 3:e124–34. doi: 10.1016/S2589-7500(20)30287-9
32. Greenhalgh T, Wherton J, Shaw S, Morrison C. Video consultations for covid-19. *BMJ.* (2020) 368:m998. doi: 10.1136/bmj.m998
33. IMDRF Software as a Medical Device (SaMD) Working Group. *"Software as a Medical Device": Possible Framework for Risk Categorization and Corresponding Considerations 2014*. IMDRF/SaMD WG/N12FINAL. IMDRF Software as a Medical Device (SaMD) Working Group (2014). Available online at: https://www.imdrf.org/documents/software-medical-device-possible-framework-risk-categorization-and-corresponding-considerations
34. GBD 2019 Blindness and Vision Impairment Collaborators; Vision Loss Expert Group of the Global Burden of Disease Study. Causes of blindness and vision impairment in 2020 and trends over 30 years, and prevalence of avoidable blindness in relation to VISION 2020: the right to sight: an analysis for the global burden of disease study. *Lancet Glob Health.* (2020) 9:e144–60. doi: 10.1016/S2214-109X(20)30489-7
35. The World Bank (WB). *World Bank Country and Lending Groups*. The World Bank. Available online at: https://datahelpdesk.worldbank.org/knowledgebase/articles/906519-world-bank-country-and-lending-groups (accessed September 14, 2020).
36. Wong TY, Sun J, Kawasaki R, Ruamviboonsuk P, Gupta N, Lansingh VC, et al. Guidelines on diabetic eye care: the international council of ophthalmology recommendations for screening, follow-up, referral, and treatment based on resource settings. *Ophthalmology.* (2018) 125:1608–22. doi: 10.1016/j.ophtha.2018.04.007
37. Breiman L. Random forests. *Mach Learn.* (2001) 45:5–32. doi: 10.1023/A:1010933404324
38. Kassirer JP. Diagnostic reasoning. *Ann Intern Med.* (1989) 110:893–900. doi: 10.7326/0003-4819-110-11-893
39. Eddy DM, Clanton CH. The art of diagnosis: solving the clinicopathological exercise. *N Engl J Med.* (1982) 306:1263–8. doi: 10.1056/NEJM198205273062104
40. Scheetz J, Rothschild P, McGuinness M, Hadoux X, Soyer HP, Janda M, et al. A survey of clinicians on the use of artificial intelligence in ophthalmology, dermatology, radiology and radiation oncology. *Sci Rep.* (2021) 11:5193. doi: 10.1038/s41598-021-84698-5
41. Xiang Y, Zhao L, Liu Z, Wu X, Chen J, Long E, et al. Implementation of artificial intelligence in medicine: status analysis and development suggestions. *Artif Intell Med.* (2020) 102:101780. doi: 10.1016/j.artmed.2019.101780
42. Davenport T, Kalakota R. The potential for artificial intelligence in healthcare. *Fut Healthc J.* (2019) 6:94–8. doi: 10.7861/futurehosp.6-2-94

43. Tham YC, Husain R, Teo KYC, Tan ACS, Chew ACY, Ting DS, et al. New digital models of care in ophthalmology, during and beyond the COVID-19 pandemic. *Br J Ophthalmol.* (2021) 106:452–7. doi: 10.1136/bjophthalmol-2020-317683
44. Chew AMK, Gunasekeran DV. Social media big data: the good, the bad, and the ugly (un)truths. *Front Big Data.* (2021) 4:623794. doi: 10.3389/fdata.2021.623794
45. He J, Baxter SL, Xu J, Zhou X, Zhang K. The practical implementation of artificial intelligence technologies in medicine. *Nat Med.* (2019) 25:30–6. doi: 10.1038/s41591-018-0307-0
46. Gunasekeran DV, Tseng RMWW, Tham YC, Wong TY. Applications of digital health for public health responses to COVID-19: a systematic scoping review of artificial intelligence, telehealth and related technologies. *NPJ Digit Med.* (2021) 4:40. doi: 10.1038/s41746-021-00412-9
47. Ting DSJ, Ang M, Mehta JS, Ting DSW. Artificial intelligence-assisted telemedicine platform for cataract screening and management: a potential model of care for global eye health. *Br J Ophthalmol.* (2019) 103:1537–8. doi: 10.1136/bjophthalmol-2019-315025
48. Liu Z, Ng M, Gunasekeran DV, Li H, Ponampalam K, Ponampalam R. Mobile technology: usage and perspective of patients and caregivers presenting to a tertiary care emergency department. *World J Emerg Med.* (2020) 11:5–11. doi: 10.5847/wjem.j.1920-8642.2020.01.001
49. Venkatesh V, Morris MG, Davis GB, Davis FD. User acceptance of information technology: toward a unified view. *MIS Q.* (2003) 27 425–78. doi: 10.2307/30036540
50. Holden RJ, Karsh BT. The technology acceptance model: its past and its future in health care. *J Biomed Inform.* (2010) 43:159–72. doi: 10.1016/j.jbi.2009.07.002
51. Redd TK, Read-Brown S, Choi D, Yackel TR, Tu DC, Chiang MF. Electronic health record impact on productivity and efficiency in an academic pediatric ophthalmology practice. *J AAPOS.* (2014) 18:584–9. doi: 10.1016/j.jaapos.2014.08.002
52. Goldstein IH, Hribar MR, Reznick LG, Chiang MF. Analysis of total time requirements of electronic health record use by ophthalmologists using secondary EHR data. In: *AMIA Annual Symposium Proceedings.* American Medical Informatics Association (2018). p. 490–7. Available online at: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6371357/
53. Couronné R, Probst P, Boulesteix AL. Random forest versus logistic regression: a large-scale benchmark experiment. *BMC Bioinformatics.* (2018) 19:270. doi: 10.1186/s12859-018-2264-5

Conflict of Interest: DG reports appointment as Physician Leader (Telemedicine) for Raffles Medical Group (SGX:$BSL.SI) and investments in digital health start-ups AskDr, Doctorbell (acquired by MaNaDr), Shyfts, and VISRE. JC reports appointment as a consultant to Boston AI labs. AYL reports grants from Santen, personal fees from Genentech, US FDA, Johnson and Johnson, grants from Carl Zeiss Meditec, personal fees from Topcon, Gyroscope, non-financial support from Microsoft, grants from Regeneron, outside the submitted work; This article does not reflect the views of the US FDA. PK reports having acted as a consultant for DeepMind, Roche, Novartis, Apellis, and BitFount and is an equity owner in Big Picture Medical. He has received speaker fees from Heidelberg Engineering, Topcon, Allergan, and Bayer. AF reports honoraria, advisory board and grant funding from Alcon, Bayer, Novartis, Allergan, Roche, and Syneos Health. AL reports grants from Roche and Novartis, and appointment as a consultant to NotalVision, Allergan, Bayer, WebMD, and Beyeonics. AG reports appointment to provide lectures for Pfizer, Thea, and Polpharma. TS reports appointment as a consultant & advisory board for Bayer Yakuhin, Boehringer-Ingelheim, Novartis, Chugai, Senju, and Santen. DM reports funding support from the Singapore National Medical Research Council (NMRC-CIRG18Nov-0013), and the Duke-NUS Medical School, Singapore (ACP 05/FY2019/P2/06-A60). DM also reports appointment as consultant and Advisory Board Member of Optomed, Finland. TW reports appointment as the deputy group chief executive officer (research and education) of Singapore Health Services, a consultant & advisory board for Allergan, Bayer, Boehringer-Ingelheim, Genentech, Merck, Novartis, Oxurion (formerly ThromboGenics), Roche, and co-founder of Plano. DT reports funding from the following grants for research about AI in healthcare: National Medical Research Council, Singapore (NMRC/HSRG/0087/2018; MOH-000655-00), National Health Innovation Center, Singapore (NHIC-COV19-2005017), SingHealth Fund Limited Foundation (SHF/HSR113/2017), Duke-NUS Medical School, Singapore (Duke-NUS/RSF/2021/0018; 05/FY2020/EX/15-A58), and Agency for Science, Technology and Research (ASTAR), Singapore (A20H4g2141 and A20H4g2141). DT, GL, and TW also report being the co-inventors of a deep learning system for retinal diseases and co-founders of related start-up Eyris; potential conflicts of interests are managed according to institutional policies of the Singapore Health System (SingHealth) and the National University of Singapore (NUS).

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers | Frontiers in Medicine

TYPE Original Research
PUBLISHED 02 November 2022
DOI 10.3389/fmed.2022.1001673

OPEN ACCESS

EDITED BY
Darren Shu Jeng Ting,
University of Nottingham,
United Kingdom

REVIEWED BY
M. Ahmed,
Phcog.Net, India
Jiong Zhang,
University of Southern California,
United States

*CORRESPONDENCE
Youxin Chen
chenyx@pumch.cn

†These authors share first authorship

SPECIALTY SECTION
This article was submitted to Ophthalmology, a section of the journal Frontiers in Medicine

RECEIVED 23 July 2022
ACCEPTED 20 September 2022
PUBLISHED 02 November 2022

CITATION
Yang J, Wu S, Dai R, Yu W and Chen Y (2022) Publication trends of artificial intelligence in retina in 10 years: Where do we stand?
Front. Med. 9:1001673.
doi: 10.3389/fmed.2022.1001673

Publication trends of artificial intelligence in retina in 10 years: Where do we stand?

Jingyuan Yang[1,2†], Shan Wu[3†], Rongping Dai[1,2], Weihong Yu[1,2] and Youxin Chen[1,2*]

[1]Department of Ophthalmology, Peking Union Medical College Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing, China, [2]Key Laboratory of Ocular Fundus Diseases, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing, China, [3]Beijing Hospital, National Center of Gerontology, Institute of Geriatric Medicine, Chinese Academy of Medical Sciences, Beijing, China

Purpose: Artificial intelligence (AI) has been applied in the field of retina. The purpose of this study was to analyze the study trends within AI in retina by reporting on publication trends, to identify journals, countries, authors, international collaborations, and keywords involved in AI in retina.

Materials and methods: A cross-sectional study. Bibliometric methods were used to evaluate global production and development trends in AI in retina since 2012 using Web of Science Core Collection.

Results: A total of 599 publications were retrieved ultimately. We found that AI in retina is a very attractive topic in scientific and medical community. No journal was found to specialize in AI in retina. The USA, China, and India were the three most productive countries. Authors from Austria, Singapore, and England also had worldwide academic influence. China has shown the greatest rapid increase in publication numbers. International collaboration could increase influence in this field. Keywords revealed that diabetic retinopathy, optical coherence tomography on multiple diseases, algorithm were three popular topics in the field. Most of top journals and top publication on AI in retina were mainly focused on engineering and computing, rather than medicine.

Conclusion: These results helped clarify the current status and future trends in researches of AI in retina. This study may be useful for clinicians and scientists to have a general overview of this field, and better understand the main actors in this field (including authors, journals, and countries). Researches are supposed to focus on more retinal diseases, multiple modal imaging, and performance of AI models in real-world clinical application. Collaboration among countries and institutions is common in current research of AI in retina.

KEYWORDS

artificial intelligence, bibliometric, deep learning, retina, retinal diseases

Introduction

The application of artificial intelligence (AI) in retinal images has shown reliable performance as well as or better than human clinicians at some key medical care tasks, such as analysis of images, diagnosis, and prediction of prognosis (1). Currently ophthalmologists begin to embrace an age of AI-assistant ophthalmology. However, how to evaluating and applying AI technique in retina and to integrate AI technique into ophthalmic profession remain challenging. Analyzing previous works is helpful to identify current research situation in AI in retina.

Due to these issues, a comprehensive review of the publications in AI in retina is urgently needed. Bibliometric analyses are helpful to address these problems by describing distribution patterns of publications, geographical distribution of research, latest evolution in the field of AI in retina. Therefore, bibliometrics is helpful in understanding of a specific field and in governing policymaking (2). Bibliometric analysis is a sort of original studies and not a systematic review or meta-analysis. However, to our knowledge, no similar studies which focused on AI in retina have been specifically conducted. Consequently, there is a lack of knowledge about the research situation in the field of AI in retina. In this study, we investigated the frontier researches of and the trends within the fields of AI in retina across the international scientific literature. We also tried to predict trends for the next few years, noting that the increase of the amount of AI researches in retina is expected to lead to a better application of AI technique in real-world settings.

Materials and methods

Search methods

Web of Science (WOS) Core Collection is regarded as the most suitable database for bibliometric analysis. The search for papers to be included in the current study was carried on July 4, 2022, and all the included publications were published from January 1, 2012 to July 1, 2022. The search strategy was "(TS = artificial intelligence OR TS = deep learning) AND (TS = retina OR TS = vitreous OR TS = choroid)". 622 literatures were identified. 77 publications were excluded except articles and review articles according to the document type, and 6 non-English publications were also excluded. 599 publications were included ultimately.

Abbreviations: AI, artificial intelligence; HDI, Human development index; OCT, optical coherence tomography; RRI, Relative research interest; WOS: web of science.

Data collection

All the data were extracted and downloaded from WOS databases, including metrics of publication numbers, countries and regions, authors, citations, and H-indexes. The classification of countries and regions were defined according to the default classification in Web of Science. The literatures from Hong Kong were part of the literatures from People's Republic of China (China). We also investigated the relationship between global productivity of AI in retina and Human development index (HDI), which measures the level of human development based on knowledge, life expectancy, and income per capita indicators, rather than economic growth alone. Human development report 2020 was published by United Nations Development Programme (3). Countries and areas were divided into four categories based on HDI, including very high human development, high human development, medium human development, and low human development. The countries and regions classification system for Human development index, which were come up with by United Nations, were converted to the classification in Web of Science. Prism 9, R (R. app GUI 1.79), VOSviewer 1.6.18, and SPSS 26 were used to input and analyze data.

Bibliometric analysis

The descriptive indexes were extracted from WOS and calculated by SPSS. The co-occurrence networks of keywords, authors, and countries/regions were constructed by VOSviewer. The keywords were extracted from titles and abstracts. Frequencies over 20 was the criteria of the inclusion for analyses. Average appearing year was used to assess the novelty of keywords. For creating the wordcloud of keywords, Biblioshiny, an R tool, was used to generate wordcloud map of keywords. Frequencies over 20 was also the criteria of the inclusion for analyses. H-indexes were collected from WOS database, and can partially reflect the impact of researchers. Relative research interest (RRI) was defined as the number of publications in a specific field per year divided by all publications in all fields per year (formula in **Supplementary Table**). The value of this metric reflects the global attention and study interest in a specific field. A higher value of RRI for AI in retina represents more research interest and more research hotspots in this field. The third order polynomial method was used in the prediction model using Prism. In order to analyze the increasing trend of publication numbers, we calculated the average growth rate, compound average growth rate, relative growth rate, and doubling time (formulas in **Supplementary Table**). For investigating the degree of international collaboration, the degree of collaboration was calculated (formula in **Supplementary Table**),

and Pearson's correlation of publication numbers among countries was calculated.

Results

Productivity and collaboration between countries and regions

A total of 599 publications were analyzed. From 2012 to 2022, USA contributed to the most publications (171, 28.5%), followed by China (149, 24.9%) (**Figure 1A**). More than half publications were contributed by USA and China. Except USA and China, no countries published more than 100 publications. The total number of publications on AI in retina has maintained steady growth over the past 10 years, especially in the past 5 years (**Figure 1C**). USA and China were also the countries with highest H-index, and the top 5 counties with most publications also had the highest H-index. USA also published the most papers per year from 2012 until now, and published 45 papers. Additionally, RRI of AI in retina also increased from < 0.001% in 2012 to 0.008% in 2021 (**Figure 1C**), indicating that the research interest of AI in retina keep increasing worldwide over the past decade. According to the HDI category, we noticed that most publications were from very high human development countries or areas, and the numbers of publications of AI in retina were consistent with HDI classification on the whole (**Figure 1D**).

We analyzed the co-occurrence of 32 countries and regions (**Supplementary Figure 1**); the analysis suggested 6 clusters: 1. USA, and Taiwan, China; 2. China, and South Africa; 3. India, Suadi Arabia, Pakistan, U Arab Emirates, Egypt, Bangladesh, Poland, and Russia; 4. South Korea, Japan, France, Malaysia, and Vietnam; 5. England, Spain, Germany, Iran, Brazil, Italy, Turkey, and Israel; and 6. Australia, Singapore, Canada, Switzerland, and Austria, Netherlands.

The top 10 countries/regions of high degree of collaboration were Singapore (95.2%), England (87.1%), Germany (85.7%), Saudi Arabia (84.4%), Switzerland (84.2%), Pakistan (79.2%), Australia (75.0%), Canada (68.0%), South Korea (67.4%), and USA (63.7%). Although China contributed a large number of publications, the degree of collaboration of China was only 39.6%. We further investigated the correlation among countries and regions, and found no significant correlation was detected between countries (all P value > 0.1).

The publication rate of papers on AI in retina has continued to increase over the past 10 years; predictions for next 5 years show this increase continuing (**Figure 2**). China has shown the greatest rapid increase in publication numbers since 2012. USA is projected to maintain its leading position and steady growth, but has the possibility to publish less papers than China and India in 2025.

For global publication number from 2012 to 2021, the average growth rate was 79.4%, compound average growth rate was 67.9%, relative growth rate was 81.2%, and doubling time was 1.3 years.

Citations and H-index

WOS citation reports revealed a total of 6445 citations without self-citations of the 7267 relevant citations since 2012. Each paper was cited an average of 12.13 times. USA contributed to the most citations (2466 citations, 2366 without self-citations) and H-index (28) (**Figure 1A**) from 2012. China ranked second in both citations and H-index (1401 citations, 1339 without self-citations, H-index 20).

The most cited publication has been cited for a total of 481 times (**Figure 1B**). We divided these publications into three groups according to the frequencies, including high frequency (more than 100 citations), medium frequency (more than 50 citations and < 100 citations), and low frequency (< 50 citations). Most of the publications were in low frequency group, 24 papers were cited with a medium frequency and only 7 papers were cited with a high frequency.

To further explore the distribution of citation number in each year, we supplemented the heatmaps of each group of citation frequency (**Supplementary Figures 2B–D**). Every row in the heatmap represents a publication, the x axis means year, and the color represents the citation number. The time span of high frequency and medium frequency is similar, and is longer than that of most publications in low frequency group. Besides, we also analyzed the distribution of publication year of each group (**Supplementary Figure 2A**). Most of the publications in low frequency group were published in recent years, and that could possibly due to less citations.

The leading institutions, journals and authors

We examined the top institutions in this field and found that the University of College London (30, 5.01%), and the Moorfields Eye Hospital NHS Foundation Trust (27, 4.51%) in England published the most papers on AI in retina since 2012. University of California system (22, 3.67%), Johns Hopkins University (16, 2.67%), and Chinese Academy of Sciences (16, 2.67%) ranked third and fourth (**Figure 3A**).

About half (287, 47.91%) of the papers on AI in retina were published in 27 journals, including *IEEE Access*, which published the most relevant publications (45). *Translational Vision Science Technology* and *British Journal of Ophthalmology* and published the second- and third-most with both 21 publications (**Figure 3B**).

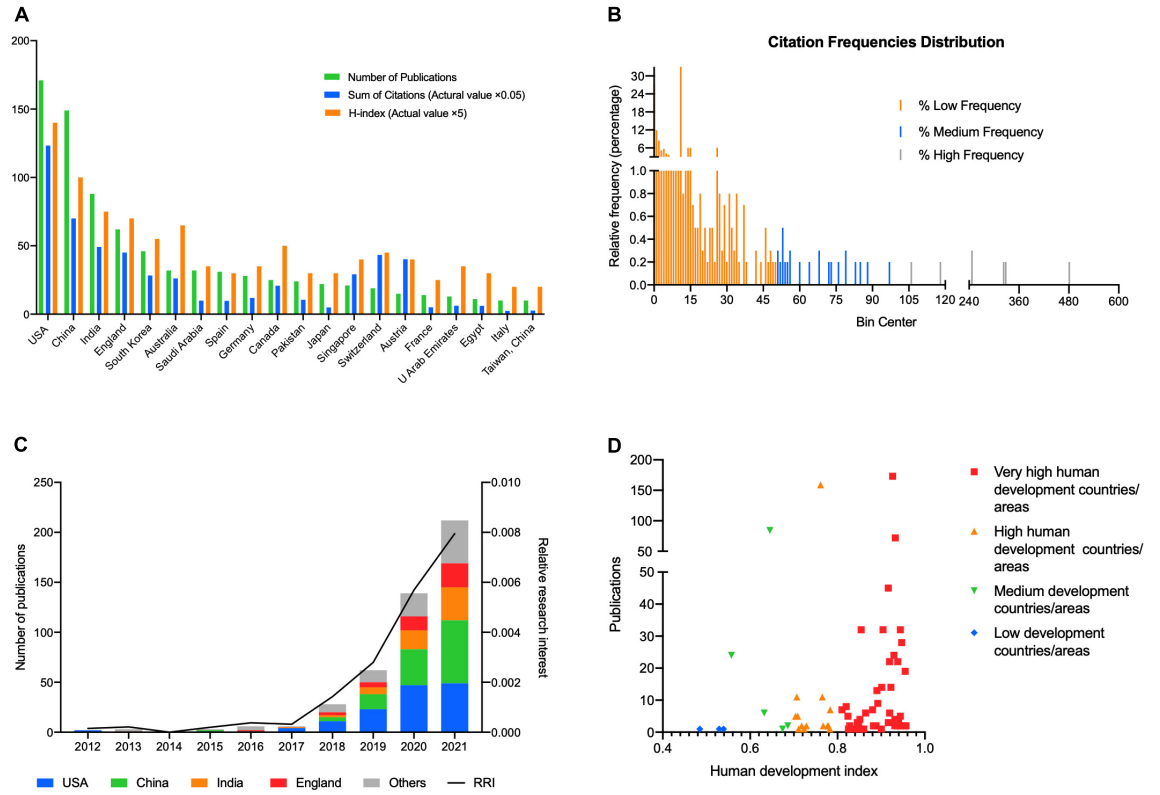

FIGURE 1
(A) Top 20 countries/regions in the publications of artificial intelligence in retina. The green bar shows the number of publications, the blue bar shows the sum of citations in total (actual value multiply by 0.05), and the orange bar shows the H-index (actual value multiply by 5). **(B)** The relative frequency (percentage) distribution of publications in various citation number. The publications were divided into three groups according to the total citation frequency, including high citation frequency (more than 100 citations) group, medium frequency (more than 50 citations and < 100 citations) group, and low frequency (<50 citations) group. **(C)** The proportion of publications of USA, China, India, England, and other countries/regions and relative research interest (RRI) in each year on the field of artificial intelligence in retina. **(D)** The number of publications of artificial intelligence in retina in countries or areas of various levels of human development. Very high human development countries or areas contributed to most publications.

The 10 papers with the most citations in total are displayed in Table 1. The most cited paper was published in *IEEE Transactions on Medical Imaging*, a classic and authoritative medical imaging periodical, and is called *Segmenting Retinal Blood Vessels With Deep Neural Networks*. The corresponding author was Pawel Liskowski. Most publications of AI in retina were published in journals focusing on engineering and computing (Table 2).

The top 10 authors in this field are listed in Table 3 according to the number of publications and citations, as well as their position in author orders. The works of Ursula Schmidt-Erfurth from Medical University of Vienna were published the most since 2012, with 9 papers and 458 citations (449 without self-citations). Hrvoje Bogunoviæ, also from Medical University of Vienna, ranked second with 8 publications and 464 citations (460 without self-citations). Pearse A. Keane also ranked third with 8 publications and 368 citations (364 without self-citations) (Table 3).

We also analyzed cooperation between investigators (Supplementary Figure 3); the node size within a collaboration network indicates the strength of the connections between every author. Several authors, including David Alonso-Caneiro, Michael J. Collins, Scott A Read, and Ursula Schmidt-Erfurth had close cooperation to other researchers and teams. The researchers in the figure usually lead large-scale research teams and maintain close connections to others in the field of AI in retina.

Research hotspots in artificial intelligence in retina

Keywords analysis defined the most frequently used words and their linkage within the field of AI in retina research. We analyzed the keywords that appeared over 20 times across the included publications. Merging repeated words excluded meaningless ones resulted in 41 total keywords

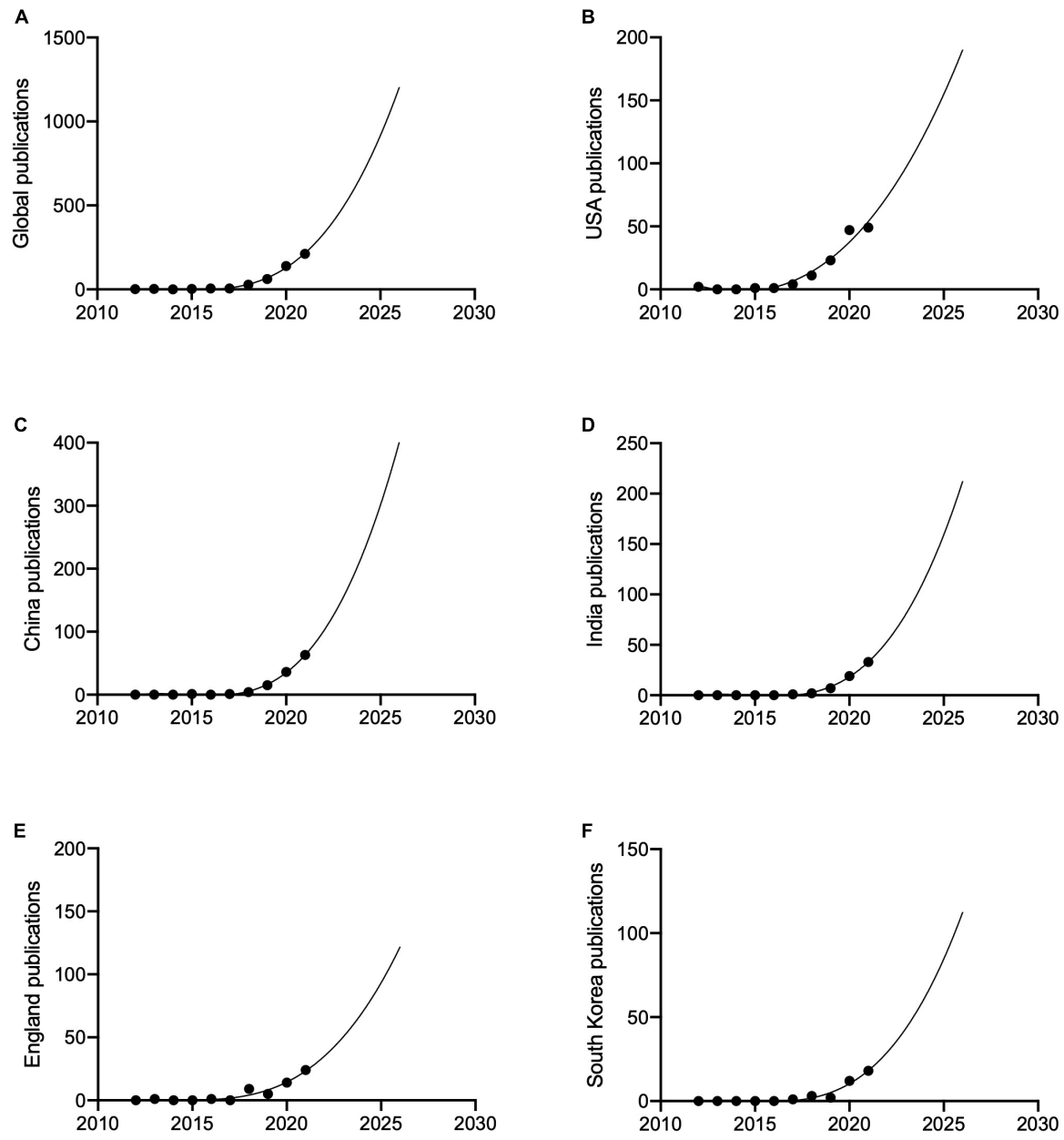

FIGURE 2
The publication trends and prediction curve of global and countries which had the most publications since 2012. **(A)** Global. **(B)** USA. **(C)** China. **(D)** India. **(E)** England. **(F)** South Korea.

that can be divided into three primary clusters by co-occurrence frequency, including diabetic retinopathy-related cluster (in red), optical coherence tomography (OCT) on multiple diseases-related cluster (in green), algorithm-related cluster (in blue) (Figure 4A).

We also color-coded the keywords by average time of appearance, and found that these keywords appeared in a short period (Figure 4B). And the words "retinal images" and "deep learning" were the most frequent words (Figure 4C).

Discussion

The current study aims to conduct a bibliometric analysis related to AI in retina. We found that the increasing trends in AI in retina since 2012. USA has published the most publications (45 publications), the most citations (1976 citations in total) and the highest H-index (24). China shows the greatest potential on publications in this field. The authors who published most on AI in retina come from Austria, England, USA,

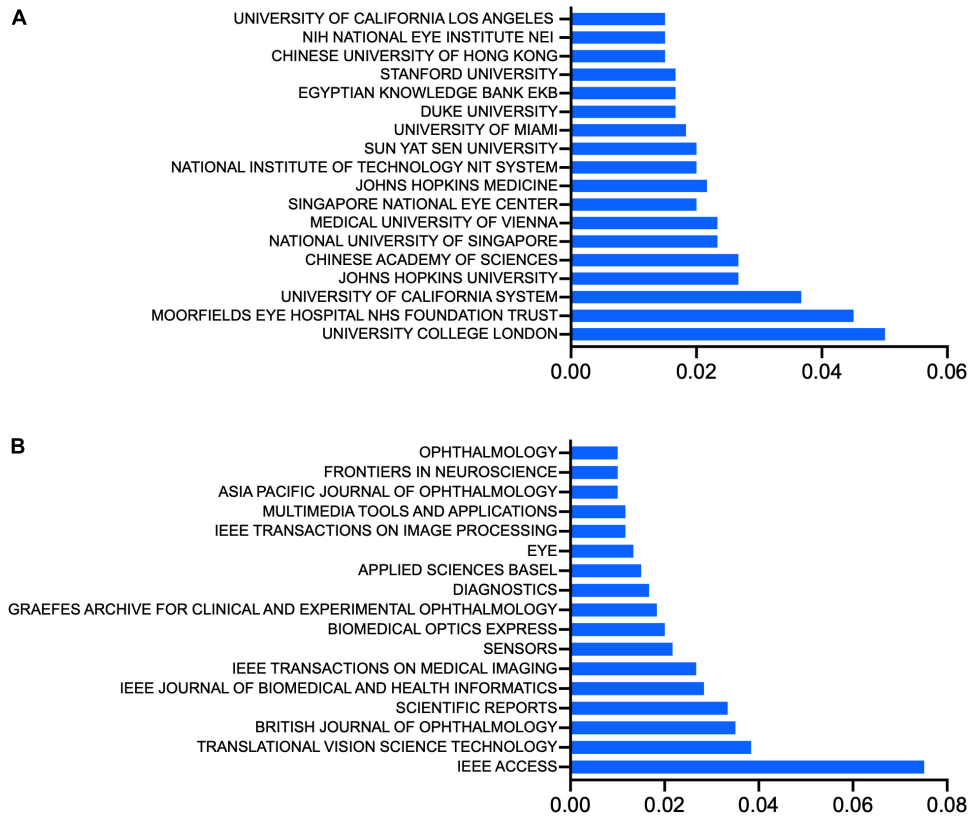

FIGURE 3
(A) Top institutions, ranked by the percentage of their publications in the total number of publications of artificial intelligence in retina. **(B)** Top journals, ranked by the percentage of their publications in the total number of publications of artificial intelligence in retina.

Singapore, South Korea, and China. By analyzing keywords, we summarized previous hotspots and predicted future hotspots. We found that AI algorithm studies, diabetic retinopathy-related AI studies, OCT-related AI studies, and other application studies of AI on retina are the current hotspots in the field of AI in retina. The average years of appearance for keywords are concentrated over a relatively short period, which suggests that AI in retina is an attractive topic in science. These results suggest the rapid progress made in AI in retina research, which might guide the research directions of future studies.

In the current study, we searched publications in the last decade, which were published since 2012. The current mainstream of artificial intelligence network was deep learning network (4), which was found to be most suitable for imaging data and replaced classic machine learning (1). Gulshan et al. and Ting et al. reported the application of deep learning network on detection of diabetic retinopathy based on fundus photographs achieved high sensitivity and specificity, respectively (5, 6). From then on, more and more papers focused on deep learning network on retina were published, which was accord with our results. The deep learning networks presented satisfied outcomes and potential to revolutionize how ophthalmology is practiced in the future (7). Various deep learning technologies have been applied in this field, such as generative adversarial networks and automated machine learning (8, 9). Current AI models mainly rely on cloud computing, which usually requires high bandwidth and low latency. With the development of communication technology, such as 5G, local edge models or offline AI models that could run on compact low power devices and do not require a continuous internet connection could help alleviate some of application challenges, and improve the quality of healthcare in underdeveloped areas (10–12). The trend of rapid development of AI in retina occurs in recent years. Therefore, we included publications in the latest decade (from only 2 publications in 2012 to 212 publications in 2021), rather than much earlier studies, whose methods might not suitable for current medical imaging tasks in most settings. And we used *deep learning* as topic words rather than other AI methods to search studies. Moreover, as this scientific research field is extremely dynamic (13), in order to include most recent publications, we included 138 publications in the first half of 2022, just before we wrote the manuscript. Therefore, considering the development of the field of AI in retina and its

TABLE 1 The top 10 papers with the most citations relevant to artificial intelligence in retina.

Title	Corresponding authors	Journal	Publication year	Total citations
Segmenting retinal blood vessels with deep neural networks	Liskowski, P	IEEE T MED IMAGING	2016	481
A reconfigurable on-line learning spiking neuromorphic processor comprising 256 neurons and 128K synapses	Indiveri, G	FRONT NEUROSCI-SWITZ	2015	328
Artificial intelligence and deep learning in ophthalmology	Ting, DSW	BRIT J OPHTHALMOL	2019	323
Artificial intelligence in retina	Schmidt-Erfurth, U	PROG RETIN EYE RES	2018	247
Real-time classification and sensor fusion with a spiking deep belief network	Pfeiffer, M	FRONT NEUROSCI-SWITZ	2013	247
Retinal vessel segmentation based on fully convolutional neural networks	Oliveira, A; Silva, CA	EXPERT SYST APPL	2018	118
Using a deep learning algorithm and integrated gradients explanation to assist grading for diabetic retinopathy	Peng, L	OPHTHALMOLOGY	2019	106
Multi-categorical deep learning neural network to classify retinal images: A pilot study employing small database	Yoo, TK; Rim, TH	PLOS ONE	2017	97
A deep learning ensemble approach for diabetic retinopathy detection	Shamshirband, S	IEEE ACCESS	2019	88
Choroid segmentation from optical coherence tomography with graph edge weights learned from deep convolutional neural networks	Zheng, YJ	NEUROCOMPUTING	2017	85

TABLE 2 Top 10 Web of Science categories of journals on artificial intelligence in retina research.

Web of Science categories	No. of publications (%)
Ophthalmology	133 (22.20)
Engineering Electrical Electronic	123 (20.53)
Computer Science Information Systems	93 (15.53)
Computer Science Artificial Intelligence	70 (11.69)
Radiology Nuclear Medicine Medical Imaging	60 (10.02)
Computer Science Interdisciplinary Applications	59 (9.85)
Telecommunications	56 (9.35)
Engineering Biomedical	53 (8.85)
Mathematical Computational Biology	41 (6.85)
Multidisciplinary Sciences	35 (5.84)

rapid progress, we included relevant publications only published in the latest decade.

In the field of AI in retina, the metric of RRI has maintained growth over the last 10 years from < 0.001% in 2001 to 0.004% in 2021, and the number of publications also increased rapidly. These indicate great interest in AI in retina on the part of the scientific and medical community. It is noticed that the number of publications in AI in retina grew more and more sharply after 2016 (total number of publications was more than 98% in the present study). In the year when Google published its influential work of AI in detection of diabetic retinopathy (5), Microsoft announced that 2016 would be the year of AI (14). This shows that the current wave of interest in AI covers many fields beyond ophthalmology. However, the current percentage of articles related to AI in retina published in the same journals which published most relevant papers was not high, and these papers were scattered in a great number of journals (more than 100 journals). The lack of a specialized journal could lead to difficulties for researchers to track latest developments in AI in retina.

When analyzing countries and authors, we found that the countries which published most papers were not consistent with the countries which had most influential authors. For example, USA (171 publications) and China (149 publications) contributed to the greatest number of publications, but only 1 author in the top 5 authors with greatest number of citations came from USA. On the contrary, authors from Austria and Singapore had the top worldwide influence according to the number of citations. On one hand, except benefiting from the bonus of number of researchers, the analysis of co-occurrence of countries and regions revealed that researchers in USA and China tend to collaborate with researchers from other countries (**Supplementary Figure 1**). On the other hand, this position of Singapore and Austria could be explained not only by the production of local institutions, such as National University of Singapore (14 publications) and Medical University of Vienna in Austria (14 publications), but also by the willing of collaboration (13, 15). Another possible reason is that researchers in small countries have much more cooperation partners outside of their countries than partners in

TABLE 3 Top 10 authors who published most and cited most in the field of artificial intelligence in retina.

Author	Country	Latest Affiliation	No. of publications	No. of citations
Schmidt-Erfurth U	Austria	Medical University of Vienna	9	455
Bogunovic H	Austria	Medical University of Vienna	8	464
Keane PA	England	Moorfields Eye Hospital NHS Foundation Trust	8	368
Balaskas K	England	Moorfields Eye Hospital NHS Foundation Trust	7	75
Gerendas BS	Austria	Medical University of Vienna	7	383
Lee AY	USA	University of Washington	7	410
Wong, Tien Y	Singapore	National University of Singapore	6	401
Yoo TK	South Korea	Aerosp Med Ctr	6	124
Cheung Carol Y	China	Chinese University of Hong Kong	6	99
Huang TJ	China	Peking University	6	63

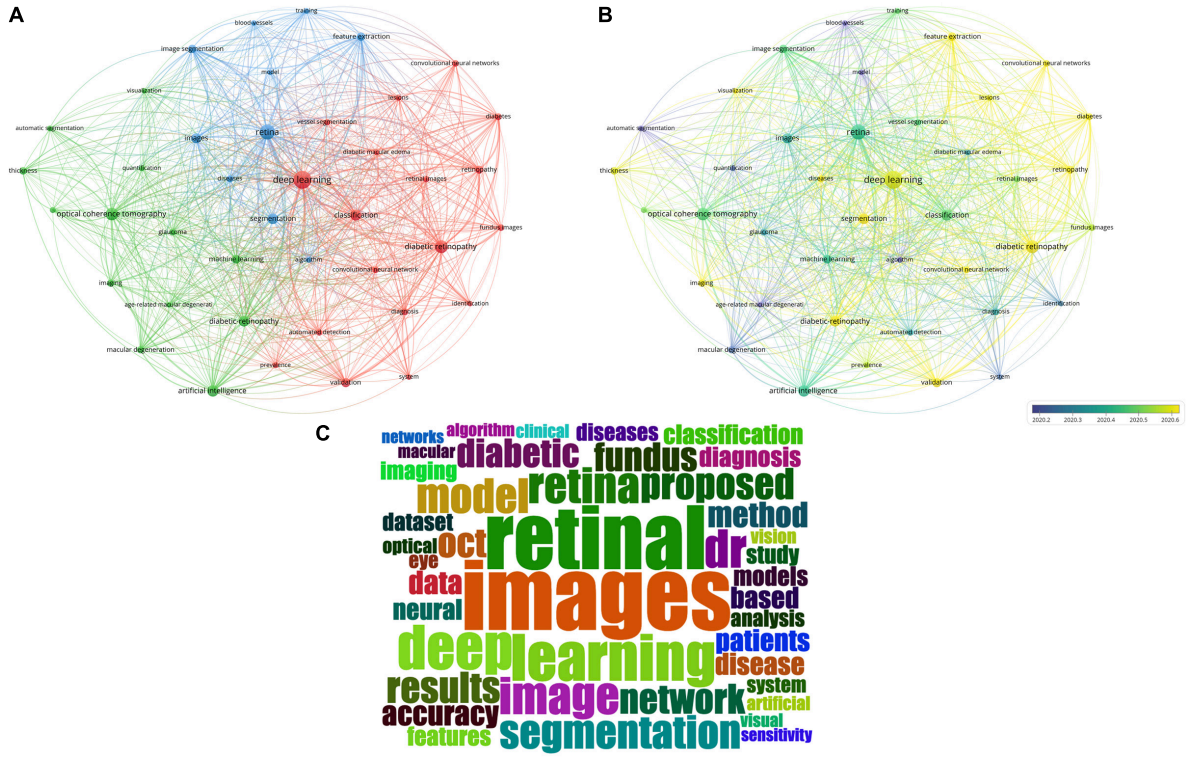

FIGURE 4
Keywords analysis by VOSviewer and R. **(A)** Co-occurrence map of keywords in titles and abstracts. Keywords were classified into three clusters by co-occurrence frequency, including diabetic retinopathy-related cluster (in red), optical coherence tomography on multiple diseases-related cluster (in green), algorithm-related cluster (in blue). **(B)** Color-coded map of keywords by the average time of appearance. Blue keywords appeared earlier, while yellow keywords appeared more recently. **(C)** The wordcloud map the most frequent keywords.

their countries (16). International collaboration on AI-related studies has shown its advantages on offering opportunities to countries regardless of country area (17). Furthermore, we applied a classification system based on HDI, and found that most of the top countries (8 of top 10 countries)and top authors (8 of top 10 authors) belongs to the level of very high human development countries. Therefore, cooperation and collaboration between countries and areas of various levels of human development are needed to promote the popularity of AI technique. However, the unavoidable filling of collaboration agreements and negotiation on intellectual property issues with partners on an institutional level might cause loss of interest and risk of postponing cooperation due to prolonged preparation before onset of AI researches (18). Researchers might need to investigate a new rapid pattern for international collaboration.

Besides the effort from medical researchers and AI engineers, a concerted effort from all stakeholders, including administrations, patients, and insurances, is needed. Sustainable models of AI applications in real world worth further investigation, taking the benefit of the patient, the health care provider, and the payer into consideration (18). Ethical and legal regulations for AI researches, which usually involve data of patient privacy (19), are also needed to cover the full workflow of AI studies and clinical trials.

We analyzed the top journals which published most papers and top articles which were most cited, and found that most publications of AI in retina are technical researches related to technology, engineering, and computing (6 of top 10 journals), not related to clinical practice. Similar to most AI researches in the category of ophthalmology, researches of AI in retina mainly focus on technical methods (7 of top 10 publications) (13). The analysis revealed that this research field involves more scientists rather than clinicians. In the future, more efforts are supposed to be paid on translation of the AI technique into a powerful tool in clinical practice on screening, triage, and decision-making process (20).

Based on the analysis of keywords in the current studies, most studies of AI in retina are correlated with diabetic retinopathy (200 relevant publications in the current study) (5, 6), followed by optical coherence tomography and associated technology (172 relevant publications). AI has also been used in diagnosing age-related macular degeneration (60 relevant publications) (21, 22), glaucoma (57 relevant publications) (23), and retinopathy of prematurity (15 relevant publications) (24). However, the technique of AI could be used for assistance for screening and triage of more retinal diseases, such as retinal detachment (14 relevant publications) and retinal vein occlusion (8 relevant publications), which are also vision-threatening diseases. More studies are needed for diagnosis of multiple retinal diseases simultaneously, and even general conditions (25), which meets the current clinical needs (26).

Moreover, multimodal imaging method rather than single modal imaging method is currently mainstream in real-world settings. However, most AI studies used single modal images of fundus photography or optical coherence tomography as training data, and the performance of trained models usually cannot overcome the challenges from numerous variabilities in reality, including field of view, image magnification, image quality, and race origin (7). Only limited researches investigated the application of multimodal imaging in AI researches, which not only improved the diagnostic performance, but also a more precise definition (27). Considering the growing demand of healthcare and limited number of retinal specialists, particularly in rural areas, the application of AI on more retinal diseases and multimodal imaging could help enhance medical care for multiethnic populations in real-world clinical practice.

We analyzed the keywords and abstracts of current studies of AI in retina, and found that most studies focused on imaging analysis (436 relevant publications), such as segmentation and detection of lesions, and diagnosis. Recently, Ting et al. found that using AI system as an assistive tool to screen for diabetic retinopathy is an economical method in real-world settings (28). The technique of AI could also be used for guidance of therapy with automated detection of lesion activity, quantitative analysis of therapeutic effects, and determination of recurrence (1). However, most studies of AI in retina were conducted in a well-prepared setting with selected data, and the performance of AI models in real world might showed unexpected outcomes. A team from Google used a deep learning system in clinics for detection of diabetic retinopathy (29), and failure to reproduce the performance they published in 2016 due to socio-environmental factors (5). Moreover, increase the diversity of dataset by applying training data from various sources helps enhance the generalization ability of AI models, and helps bring AI models to clinical practice (30). Therefore, the validation of application of AI technique in real world is an important topic in future researches.

Although studies of AI in retina is broad, it mainly focuses on detecting structures and lesions and diagnosing diseases from fundus photography and optical coherence tomography images through deep learning algorithms currently, trying to achieve expert-level performance in well-prepared experiment environment. Enhancing the performance of detecting retinal diseases in real world in encouraged (29). Firstly, additional structured clinic data (including but not limited to, age, sex, ethnicity, other examination results, medical history, clinical diagnosis, comorbidities, or genetic indicators, etc.) or multimodal images could be collected together. This would be helpful to build datasets with diversity and further to cover more application scene and more populations, such as screening and clinical diagnosis in countries and regions of various development levels. Secondly, more application requirements from clinical practice were put up with, including earlier detection and referral, personalized treatment based on guidelines and the reality of each patient and prediction of prognosis, covering more diseases for global automated screening, resulting in lower economic burden and better life quality for patients. AI has the potential to provide direct patient care after combining with mobile devices and communication technology, especially in under-resourced areas (31). However, there is a lot of work to be done to make this a reality, and the application of AI models in clinical practice needs clinical validation and regulatory requirements. Moreover, AI models are supposed to progress in parallel with the advance of new management strategies.

Our study has several limitations. The nature of selection bias existed in the methods, including in favor of English-language journal and only papers published on authoritative and influential journals which were listed in WOS Core Collection were included in our analysis. Although most high-quality papers are included by most databases, the results might be

partially affected. Moreover, using different searching method in WOS Core Collection could lead to different results, and the method we used did not include publications that were assigned same keywords by other databases, which would not result in unrepeatable searching results. Therefore, the total number of included publications using forementioned methods might be less than some other methods.

In summary, this study comprehensively analyzed most published researches on AI in retina, and presents a current view of mainstream studies on AI in retina. AI in retina is a very attractive topic in researches, and the relevant technique developed rapidly, although no journal specializes in AI in retina. International collaboration is important in conduction of influential researches. Researches are supposed to focus on more retinal diseases, multiple modal imaging, and performance of AI models in real world. This study may help clinicians, but also scientists understand the current trend of publications of AI in retina, know the main actors in the field, and predict and guide the future developments in this research field.

Data availability statement

The original contributions presented in this study are included in the article/**Supplementary material**, further inquiries can be directed to the corresponding author.

Author contributions

JY: design, definition of intellectual content, data acquisition, data analysis, manuscript preparation, and manuscript editing. SW: design, definition of intellectual content, data acquisition, data analysis, funding, manuscript preparation, and manuscript editing. RD and WY: manuscript review. YC: concepts and manuscript review. All authors contributed to the article and approved the submitted version.

Funding

This study was supported by the National High Level Hospital Clinical Research Funding (BJ-2021-208).

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmed.2022.1001673/full#supplementary-material

SUPPLEMENTARY FIGURE 1
The co-occurrence map of 32 countries and regions, which showed the international collaboration among countries/regions.

SUPPLEMENTARY FIGURE 2
(A) The distribution of publication year for publications of various citation frequency. **(B–D)** The heatmaps of high citation frequency (more than 100 citations) group, medium frequency (more than 50 citations and < 100 citations) group, and low frequency (<50 citations) group, respectively. Every row in the heatmap represents a publication. The color represents the total citation number in each year (x axis).

SUPPLEMENTARY FIGURE 3
The co-occurrence map of scholars who published papers of artificial intelligence in retina, which showed the cooperation among researchers.

References

1. Schmidt-Erfurth U, Sadeghipour A, Gerendas BS, Waldstein SM, Bogunoviæ H. Artificial intelligence in retina. *Prog Retin Eye Res.* (2018) 67:1–29. doi: 10.1016/j.preteyeres.2018.07.004

2. King DA. The scientific impact of nations. *Nature.* (2004) 430:311–6. doi: 10.1038/430311a

3. UNDP. *Human Development Report 2020*. New York, NY: UNDP (2020).

4. LeCun Y, Bengio Y, Hinton G. Deep learning. *Nature.* (2015) 521:436–44. doi: 10.1038/nature14539

5. 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. doi: 10.1001/jama.2016.17216

6. 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. doi: 10.1001/jama.2017.18152

7. Ting DSW, Pasquale LR, Peng L, Campbell JP, Lee AY, Raman R, et al. Artificial intelligence and deep learning in ophthalmology. *Br J Ophthalmol.* (2019) 103:167–75. doi: 10.1136/bjophthalmol-2018-313173

8. Liu Y, Yang J, Zhou Y, Wang W, Zhao J, Yu W, et al. Prediction of OCT images of short-term response to anti-VEGF treatment for neovascular age-related macular degeneration using generative adversarial network. *Br J Ophthalmol.* (2020) 104:1735–40. doi: 10.1136/bjophthalmol-2019-315338

9. Yang J, Zhang C, Wang E, Chen Y, Yu W. Utility of a public-available artificial intelligence in diagnosis of polypoidal choroidal vasculopathy. *Graefes Arch Clin Exp Ophthalmol.* (2020) 258:17–21. doi: 10.1007/s00417-019-04493-x

10. Merenda M, Porcaro C, Iero D. Edge machine learning for AI-enabled IoT devices: a review. *Sensors.* (2020) 20:2533. doi: 10.3390/s20092533

11. Keane PA, Topol EJ. Medicine and meteorology: cloud, connectivity, and care. *Lancet.* (2020) 395:1334. doi: 10.1016/s0140-6736(20)30813-8

12. Greco L, Percannella G, Ritrovato P, Tortorella F, Vento M. Trends in IoT based solutions for health care: moving AI to the edge. *Pattern Recognit Lett.* (2020) 135:346–53. doi: 10.1016/j.patrec.2020.05.016

13. Boudry C, Al Hajj H, Arnould L, Mouriaux F. Analysis of international publication trends in artificial intelligence in ophthalmology. *Graefes Arch Clin Exp Ophthalmol.* (2022) 260:1779–88. doi: 10.1007/s00417-021-05511-7

14. The New York Times blog. *Microsoft Reorganizes Its Research Efforts Around A.I.* New York, NY: The New York Times blog (2016).

15. Tran BX, Vu GT, Ha GH, Vuong QH, Ho MT, Vuong TT, et al. Global evolution of research in artificial intelligence in health and medicine: a Bibliometric study. *J Clin Med.* (2019) 8:360. doi: 10.3390/jcm8030360

16. Mindeli LE, Markusova VA. Bibliometric studies of scientific collaboration: international trends. *Autom Doc Math Linguist.* (2015) 49:59–64. doi: 10.3103/S0005105515020065

17. Narin F, Stevens K, Whitlow ES. Scientific co-operation in Europe and the citation of multinationally authored papers. *Scientometrics.* (1991) 21:313–23. doi: 10.1007/BF02093973

18. Ting DSW, Peng L, Varadarajan AV, Keane PA, Burlina PM, Chiang MF, et al. Deep learning in ophthalmology: the technical and clinical considerations. *Prog Retin Eye Res.* (2019) 72:100759. doi: 10.1016/j.preteyeres.2019.04.003

19. Price WN II, Cohen IG. Privacy in the age of medical big data. *Nat Med.* (2019) 25:37–43. doi: 10.1038/s41591-018-0272-7

20. Kras A, Celi LA, Miller JB. Accelerating ophthalmic artificial intelligence research: the role of an open access data repository. *Curr Opin Ophthalmol.* (2020) 31:337–50. doi: 10.1097/icu.0000000000000678

21. Burlina PM, Joshi N, Pekala M, Pacheco KD, Freund DE, Bressler NM. Automated grading of age-related macular degeneration from color fundus images using deep convolutional neural networks. *JAMA Ophthalmol.* (2017) 135:1170–6. doi: 10.1001/jamaophthalmol.2017.3782

22. Grassmann F, Mengelkamp J, Brandl C, Harsch S, Zimmermann ME, Linkohr B, et al. A deep learning algorithm for prediction of age-related eye disease study severity scale for age-related macular degeneration from color fundus photography. *Ophthalmology.* (2018) 125:1410–20. doi: 10.1016/j.ophtha.2018.02.037

23. Li Z, He Y, Keel S, Meng W, Chang RT, He M. 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

24. Brown JM, Campbell JP, Beers A, Chang K, Ostmo S, Chan RVP, et al. Automated diagnosis of plus disease in retinopathy of prematurity using deep convolutional neural networks. *JAMA Ophthalmol.* (2018) 136:803–10. doi: 10.1001/jamaophthalmol.2018.1934

25. Rim TH, Lee G, Kim Y, Tham YC, Lee CJ, Baik SJ, et al. Prediction of systemic biomarkers from retinal photographs: development and validation of deep-learning algorithms. *Lancet Digit Health.* (2020) 2:e526–36. doi: 10.1016/s2589-7500(20)30216-8

26. Li B, Chen H, Zhang B, Yuan M, Jin X, Lei B, et al. Development and evaluation of a deep learning model for the detection of multiple fundus diseases based on colour fundus photography. *Br J Ophthalmol.* (2022) 106:1079–86. doi: 10.1136/bjophthalmol-2020-316290

27. Morelle O, Wintergerst M, Finger RP. [Multimodal imaging and evaluation in the age of artificial intelligence] Multimodale bildgebung und -auswertung im zeitalter von künstlicher intelligenz. *Ophthalmologe.* (2020) 117:965–72. doi: 10.1007/s00347-020-01210-6

28. Xie Y, Nguyen QD, Hamzah H, Lim G, Bellemo V, Gunasekeran DV, et al. Artificial intelligence for teleophthalmology-based diabetic retinopathy screening in a national programme: an economic analysis modelling study. *Lancet Digit Health.* (2020) 2:e240–9. doi: 10.1016/s2589-7500(20)30060-1

29. Beede E, Baylor E, Hersch F, Iurchenko A, Wilcox L, Ruamviboonsuk P, et al. 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. Association for Computing Machinery*. Honolulu HI: (2020). p. 1–12.

30. Benet D, Pellicer-Valero OJ. Artificial intelligence: the unstoppable revolution in ophthalmology. *Surv Ophthalmol.* (2022) 67:252–70. doi: 10.1016/j.survophthal.2021.03.003

31. Bellemo V, Lim ZW, Lim G, Nguyen QD, Xie Y, Yip MYT, et al. Artificial intelligence using deep learning to screen for referable and vision-threatening diabetic retinopathy in Africa: a clinical validation study. *Lancet Digit Health.* (2019) 1:e35–44. doi: 10.1016/s2589-7500(19)30004-4

frontiers | Frontiers in Medicine

TYPE Original Research
PUBLISHED 18 January 2023
DOI 10.3389/fmed.2022.1069649

OPEN ACCESS

EDITED BY
Darren Shu Jeng Ting,
University of Nottingham,
United Kingdom

REVIEWED BY
Gilbert Yong San Lim,
SingHealth, Singapore
Jun Jiang,
Wenzhou Medical University, China

*CORRESPONDENCE
Xiyan Zhang
✉ xyzhang0220@foxmail.com
Jie Yang
✉ 51478536@qq.com

†PRESENT ADDRESSES
Jie Yang,
Jiangsu Provincial Center for Disease Control and Prevention, Nanjing, Jiangsu, China
Xiyan Zhang,
Jiangsu Provincial Center for Disease Control and Prevention, Nanjing, Jiangsu, China

SPECIALTY SECTION
This article was submitted to Ophthalmology, a section of the journal Frontiers in Medicine

RECEIVED 14 October 2022
ACCEPTED 02 December 2022
PUBLISHED 18 January 2023

CITATION
Zhang X, Zhou Y, Wang Y, Du W and Yang J (2023) Trend of myopia through different interventions from 2010 to 2050: Findings from Eastern Chinese student surveillance study. *Front. Med.* 9:1069649. doi: 10.3389/fmed.2022.1069649

Trend of myopia through different interventions from 2010 to 2050: Findings from Eastern Chinese student surveillance study

Xiyan Zhang[1,2*†], Yonlin Zhou[1], Yan Wang[1], Wei Du[3] and Jie Yang[1,2*†]

[1]Department of Child and Adolescent Health Promotion, Jiangsu Provincial Center for Disease Control and Prevention, Nanjing, China, [2]School of Public Health, Nanjing Medical University, Nanjing, China, [3]School of Public Health, Southeast University, Nanjing, China

Purpose: First, to investigate the utilization rate and effect of proven myopic interventions. Second, to predict the prevalence of myopia and high myopia, as well as Years Lived with Disability (YLD) caused by an uncorrected refractive error in children and teens in Eastern China from 2010 to 2050 under different interventions.

Methods: (1) The surveillance of common diseases among children and adolescents in Jiangsu Province from 2010 to 2021 provides the database for myopia screening and intervention utilization surveys. (2) The National Bureau of Statistics and the Global Burden of Disease Study 2016 (GBD2016) are the foundation for the estimated myopes and YLD. (3) A systematic review provides the strong or weak impact of intervention in the prediction model. (4) The trend of screening myopia from 2010 to 2050 under various treatments is predicted using a GM (1,1) model.

Results: By the year 2050, myopia is expected to affect 8,568,305 (7–12 years old) and 15,766,863 (13–18 years old) children and adolescents, respectively (95% CI: 8,398,977–8,737,633). The utilization prevalence of myopia-proven interventions for myopic children included outdoor activities, orthokeratology lenses, atropine treatment, contact lenses, frame glasses, and eye exercises, with respective rates of 31.9–33.1, 2.1–2.3, 6.0–7.5, 2.2–2.7, 60.4–62.2, and 64.7–72.5%. All interventions have substantial effects on myopia after parental myopia and behavior pattern adjustment, including physical activity, near work, dietary pattern, and sleep. Under strong intervention, the estimated reduced myopia prevalence by the year 2050 is 1,259,086 (95% CI: 1,089,758–1,428,414) for children aged 7–12, and 584,785 (95% CI: 562,748–606,823) for children aged 13–18, respectively.

Conclusion: Among myopic Chinese children and adolescents, the use rates and effects of proven myopia interventions vary. Under the present intervention strategy, the prevalence of myopia and high myopia will increase from 2010 to 2050. The overall number of myopic people can be greatly decreased by implementing timely, steady, comprehensive interventions.

KEYWORDS

myopia, prevalence trend, interventions, children and adolescents, Eastern China

Introduction

Myopia occurs because the cornea or lens is too powerful or because the eyeball is longer than normal (1). Uncorrected refractive error is the most common cause of distance vision impairment and the second most common cause of blindness (2). In myopia, distant objects are focused in front of the retina instead of on it, as occurs in non-myopic individuals. Holden estimated that by 2050, there will be 4,758 million people with myopia (49.8% of the world population) and 938 million people with high myopia (9.8% of the world population) (3).

The prevalence of myopia is severe in China, and visual impairment occurs with complications of high myopia, such as retinal detachment, cataracts, glaucoma, and blindness (4–7). Anhui province, Fujian province, Jiangsu province, Jiangxi province, Shandong province, Shanghai city, and Zhejiang province are the seven provinces that make up the eastern mainland of China (also known as Eastern China). An area of 798,300 km^2 in Eastern China is home to approximately 30% of the country's population and 40% of its GDP (8). The prevalence of myopia and vision impairment is high among school students in Eastern China, and the greater prevalence rate of myopia in Eastern China is significantly influenced by the increased academic load on students at younger ages (5, 9, 10).

The effects of myopia interventions on children and adolescents have been extensively studied worldwide. Pharmaceuticals, optical devices, and lifestyle changes are among the interventions that had proven effects (11). Increasing the amount of time that children spent outdoors at school resulted in statistically significant reductions in incident myopia and myopic shift, as shown by the previous studies (12). A study of 571 students aged 7–11 years in Taiwan reported a 1-year reduction in the incidence rate of myopia of 8.4% in the intervention group vs. 17.7% in the control group (13). In most cases, wearing quality eyeglasses at the correct time and properly could easily correct children's vision problems (14). Recent surveys conducted in rural China indicate that among myopic students, less than one-third of myopic students reported the use of glasses, and more than two-third of myopic students denied wearing them. The mean (SD) spherical equivalent refractive error of participants was −2.16 (1.12) D (range, −0.625 to −4.0 D) in the right eye. In three schools, the proportion of children with myopia (both eyes' spherical equivalent $\leq$ −0.5 D) ranged from 25 to 58%, whereas the proportion wearing glasses at the time of examination was between 8 and 30% (15). According to the results of a study by Yi et al., only one-sixth of myopic students in rural China used eyeglasses (16). Orthokeratology was becoming more and more popular especially in the Asia-Pacific region to control the progression of myopia in young children (17). Orthokeratology is defined as the "reduction, modification, or elimination of a refractive error by programmed application of contact lenses (18)." Also, previous reviews, meta-analyses, and clinical trials suggested that atropine eye drops conferred the best efficacy among all myopia prevention methods (19, 20).

The two main goals of this investigation are as follows: First, based on a repeated cross-sectional survey in 2019, 2020, and 2021, to examine the usage rate of myopia interventions (such as outdoor activities, orthokeratology lenses, atropine treatment, contact lenses, frame glasses, and eye exercises) among an annual random sampling of Chinese children and adolescents. To calculate the odd-ratio effects of interventions that have been modified for parental myopia and behaviors (such as physical activity, near work, dietary pattern, and sleep). Second, to predict the estimated myopia population, high myopia population, reduced myopia population, and Years Lived with Disability (YLD) of refraction and accommodation disorders (such as severe vision impairment due to uncorrected refractive error, moderate vision impairment due to uncorrected refractive error, and blindness due to uncorrected refractive error) under various intervention strategies (current intervention situation, strong intervention, and weak intervention) among children and adolescents from 2010 to 2050 in Eastern China.

Materials and methods

Profile of Eastern China

With seven provinces and cities, including Shanghai, Shandong, Jiangsu, Anhui, Jiangxi, Zhejiang, and Fujian,

East China is one of the most developed regions in China (Supplementary Figure 1). According to a 2010 National Bureau of Statistics report, this region is home to approximately 400 million people or almost one-third of all Chinese citizens. Supplementary Figure 2 provides comprehensive data for children aged 7–18 years.

Definition

Definition of screening Myopia (21, 22): (1) For children aged 7–12 years, the screening myopia is defined as uncorrected visual acuity (UCVA) < 0.5, as well as non-cycloplegic auto refraction (NCAR) < −0.50 D. (2) For children aged 13–18 years, the screening myopia is defined as UCVA < 0.5.

Years Lived with Disability: Long-term disability due to a given cause including blindness, moderate vision impairment and severe vision impairment (23).

Gray model GM (1,1) (24): The model is established based on the Gray System Theory using a time-series prediction realm. The Gray prediction model includes the classic univariate gray prediction model [GM (1,1) model] and the multi-variable gray prediction model [GM (1,N) model]. A group of new data series with the obvious trend is generated by accumulating a certain original data series, and the growth trend of the new data series is used to establish a model for prediction, and then the reverse calculation is performed by accumulating the new data series to recover the original data series, and finally, the prediction results are obtained. GM (1,1) equation:

$$\hat{x}_0(k+1) = \hat{x}_1(k+1) - \hat{x}_1(k)$$
$$= (1 - e^{\hat{a}})[x_0(1) - \frac{\hat{u}}{\hat{a}}]e^{-\hat{a}k}$$

Physical examination measurements

Myopia screening: In Jiangsu Province, a myopia screening was carried out since 2010. With no cycloplegia, an auto-refractor (Topcon KR-800; Topcon Co., Tokyo, Japan) was used. Children without myopia and the absence of any other major eye conditions, Chinese Han students, and parents or guardians who could give informed consent were the inclusion criteria for our subjects.

Utilization and effectiveness of interventions

Students were requested to complete a questionnaire about myopic information during 2019, 2020, and 2021 as part of an intervention usage study. **Generalized linear model (GLM) regression analysis:** GLM analysis is performed, and a log odds ratio with 95% CI is computed to assess the effect of intervention adjusted by parental myopia, physical activity, near work, dietary pattern, and sleep, and dependent variable are classified as non-myopia (P_1), low myopia (SE more than −3.00D,P_2), medium

TABLE 1 The demographic characteristics of the study participants from 2010 to 2021 aged 7–18 years.

	Number of survey population	Male/Female (P = 0.223)	Age-standardized height (cm) (P < 0.001)		Age-standardized weight (kg) (P < 0.001)		Age-standardized vision (P < 0.001)	
Year	7–18	7–18	7–12	13–18	7–12	13–18	7–12	13–18
2010	12,000	50.0/50.0	141.5 ± 11.8	165.7 ± 8.3	35.4 ± 10.3	58.7 ± 12.2	4.81 ± 0.29	4.54 ± 0.39
2011	12,000	50.0/50.0	141.6 ± 12.3	166.0 ± 8.5	36.8 ± 11.0	57.1 ± 11.8	4.89 ± 0.30	4.48 ± 0.38
2012	48,295	52.4/47.6	146.6 ± 15.0	165.4 ± 8.7	40.3 ± 13.5	57.7 ± 12.4	4.81 ± 0.38	4.44 ± 0.39
2013	57,500	52.6/47.4	141.2 ± 13.1	166.2 ± 9.1	36.7 ± 11.9	57.9 ± 12.9	4.72 ± 1.15	4.35 ± 0.90
2014	71,650	53.5/46.5	141.5 ± 12.4	166.5 ± 8.5	37.6 ± 11.7	58.4 ± 12.4	4.84 ± 0.31	4.52 ± 0.36
2015	73,687	52.6/47.4	140.7 ± 12.1	165.9 ± 8.7	36.4 ± 11.2	57.8 ± 12.0	4.88 ± 0.29	4.50 ± 0.36
2016	60,857	53.7/46.3	143.7 ± 10.6	166.2 ± 8.6	40.3 ± 9.8	58.3 ± 12.0	4.86 ± 0.29	4.48 ± 0.35
2017	30,067	50.9/49.1	141.4 ± 12.2	165.2 ± 8.5	37.9 ± 12.0	59.9 ± 12.8	4.85 ± 0.30	4.47 ± 0.38
2018	52,694	53.1/46.9	141.1 ± 12.4	165.3 ± 8.8	38.3 ± 12.5	60.2 ± 13.7	4.79 ± 0.32	4.39 ± 0.40
2019	48,649	52.7/47.3	141.8 ± 12.5	166.2 ± 8.4	38.1 ± 12.8	61.3 ± 14.3	4.79 ± 0.33	4.41 ± 0.39
2020	48,288	52.7/47.3	142.0 ± 12.6	166.3 ± 8.4	39.3 ± 13.2	62.0 ± 14.3	4.78 ± 0.31	4.41 ± 0.38
2021	47,498	52.4/47.6	141.9 ± 12.6	166.4 ± 8.3	39.2 ± 13.0	62.9 ± 14.9	4.76 ± 0.33	4.42 ± 0.38

myopia (SE more than −6.00D,P_3), and high myopia (SE less than −6.00D,P_4). The regression equation is listed as follows:

$$GLM_{(P1,P2,P3,P4)} = \beta_0 + \beta_{main-effect1(Intervention1)}X_1 + \beta_{main-effect2(Intervention2)}X_2 + \beta_{main-effect3(Intervention3)}X_3 + \beta_{main-effect4(Intervention4)}X_4 + \beta_{main-effect5(Intervention5)}X_5 + \beta_{main-effect6(Intervention6)}X_6 + \beta_{adjusted-parentmyopia}X_7 + \beta_{adjusted-physicalactivity}X_8 + \beta_{adjusted-nearwork}X_9 + \beta_{adjusted-dietarypattern}X_{10} + \beta_{adjusted-sleep}X_{11}$$

GM (1,1) model predicting the trend of screening myopia from 2010 to 2050 under current interventions

Data are used from 2010 to 2021 as the original data to establish a GM (1, 1) model (25) and to predict the number of myopic students from 2022 to 2050. **Related parameters are obtained from:** (1) surveillance of common diseases among students in Jiangsu Province from 2010 to 2021; (2) Global Burden of Disease Study 2016 (GBD2016) (26) and National Bureau of Statistics; and (3) a systematical review. Details on these parameters can be found in **Supplementary Table 1**. In this study, we also predict a reduced myopia population under weak intervention (R: 0 to 0.25 D/year, more outdoor activities, MOA, 0.14 D/year) and strong weak intervention [R: >0.50 D/year, high-dose atropine (1 or 0.5%), ATRH, 0.7 D/year] (27). (4) Projected demographic changes in the relevant age groups are shown in **Supplementary Figure 2**.

Statistical analysis

Continuous variable trend analysis is a one-way variance trend analysis, with mean ± SD as the description. For discontinuous variables, chi-square is used for trend analysis and percentage as the description. Indicators including weight, height, vision, and spherical equivalent (SE) were standardized using age. The standardization makes the proportion of the population per age as 1:1.

All data are analyzed using office software and SPSS V.20.0 software.

Ethics statement

The Institutional Review Board approved the Ethics Committee of Jiangsu Province CDC's study protocol. The students and their parents are informed about the survey's aim, and teachers obtained participants' and their parents' oral and written consent. Detailed information can be found in the previous article (28–30).

Results

Characteristics of the study participants from 2010 to 2021

A total of 563,185 students participated in this study from 2010 to 2021. The adjusted distribution of sex is not significant ($P > 0.05$). Age-standardized weight and height show upward trends, while age-standardized vision shows downward trends (trend $P < 0.01$) (**Table 1**). Changes in spherical equivalent (SE, D) from 2018 to 2021 are not significant (trend $P > 0.05$) (**Figure 1**).

Current myopic intervention situation

The utilization rate of interventions (outdoor activities/atropine treatment) reveals a decreasing trend from 2019 to 2021 (trend $P < 0.01$). Orthokeratology has the lowest utilization rate (2.1–2.3%) and eye exercises the highest (64.7–72.5%) (**Table 2**). Effects of all interventions on myopia adjusted by parental myopia and behavior pattern (including physical activity, near work, dietary pattern, and sleep) are significant. The odds ratio value for frame glass of <0.70 indicates a medium effect on the progression of myopia. While orthokeratology, atropine treatment, and contact lenses took up a relatively low utilization rate; protective effects on myopia could still be observed. Outdoor activities and eye exercise showed a weak protective effect on primary and middle/high school students, respectively (**Figure 2**).

Estimated population of myopia, high myopia, and YLD from 2010 to 2050 under the current situation

By the year 2050, the estimated population of myopia will be 8,568,305 (95% CI: 8,398,977–8,737,633) for 7–12 years old and 15,766,863 (95% CI: 15,744,826–15,788,900) for 13–18 years old, respectively. The number of high myopia cases is 205,639 (95% CI: 201,575–209,703) for 7–12 years old and 2,395,613 (95% CI: 2,392,264–2,398,961) for 13–18 years old, respectively. The prevalence of myopia will be 62.1% for 7–12 years old and 99.7% for 13–18 years old, respectively. The prevalence of high myopia will be 1.5% for 7–12 years old and 15.2% for 13–18 years old, respectively. The estimated YLD/1,000 person-year (blindness due to uncorrected refractive error) is 151.3 (95% CI: 100.3–210.4) among children and adolescents (**Figure 3**).

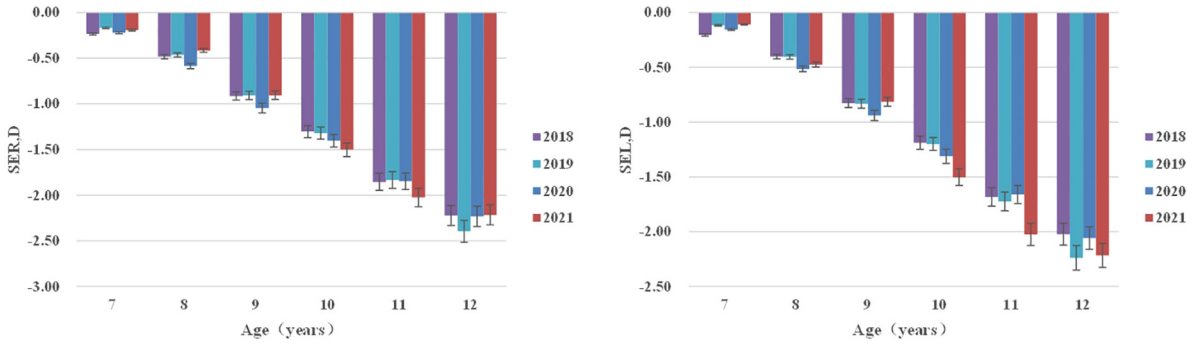

FIGURE 1
Spherical equivalent (SE, D) distribution of right/left eyes stratified by age from 2018 to 2021.

TABLE 2 Current myopic intervention utilization among myopic Chinese children and adolescents based on surveys from 2019, 2020, and 2021.

	2019	2020	2021	Absolute increase	*P*-value for trend
Intervention 1: Outdoor activities > 2 h/d					
Primary/middle and high school	38.7/31.3	39.4/30.7	39.0/29.4	−1.2	0.003
Male/female	37.0/29.1	36.1/29.5	35.3/28.5		
Urban/rural	34.0/31.8	32.8/33.0	32.4/31.3		
	33.1	32.9	31.9		
Intervention 2: Orthokeratology					
Primary/middle and high school	2.5/2.0	2.3/2.2	2.6/2.2	0.2	0.227
Male/female	2.1/2.1	2.2/2.2	2.3/2.2		
Urban/rural	2.2/2.0	2.5/1.8	2.6/1.8		
	2.1	2.2	2.3		
Intervention 3: Atropine treatment					
Primary/middle and high school	10.2/5.9	10.7/6.7	9.4/5.1	−0.8	0.001
Male/female	6.6/7.0	7.2/7.9	6.0/6.0		
Urban/rural	6.6/7.0	7.8/7.2	6.4/5.5		
	6.8	7.5	6.0		
Intervention 4: Contact lens					
Primary/middle and high school	1.5/3.0	1.2/2.5	1.0/2.9	−0.2	0.114
Male/female	1.8/3.6	1.4/3.1	1.5/3.5		
Urban/rural	3.3/1.8	2.7/1.6	3.1/1.6		
	2.7	2.2	2.5		
Intervention 5: Frame glasses					
Primary/middle and high school	40.9/65.3	44.2/66.9	45.1/67.1	1.7	0.000
Male/female	56.3/64.5	57.4/67.1	57.7/66.5		
Urban/rural	61.7/58.4	63.3/60.6	63.5/60.1		
	60.4	62.2	62.1		
Intervention 6: Eye exercise					
Primary/middle and high school	94.9/65.4	88.3/56.7	95.2/64.3	0.0	0.941
Male/female	71.9/73.2	64.7/64.7	72.0/73.0		
Urban/rural	66.7/81.1	58.0/74.7	65.6/82.4		
	72.5	64.7	72.5		

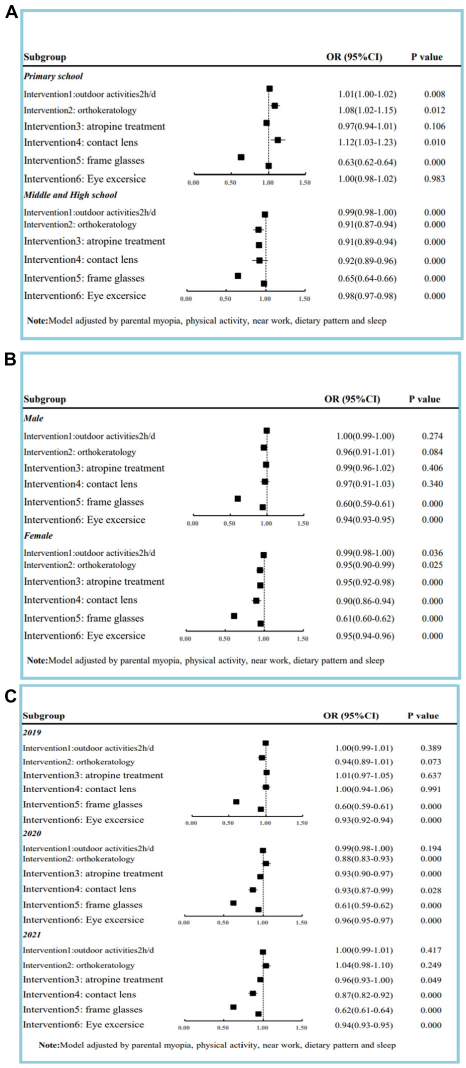

FIGURE 2
Effect of intervention on myopia adjusted by parental myopia and behavior pattern among children and adolescents based on survey in 2019, 2020, and 2021. **(A)** Effect of intervention on myopia classified by study period. **(B)** Effect of intervention on myopia classified by gender. **(C)** Effect of intervention on myopia classified by year.

Internal validation: The regressive performance of the GM (1,1) model is presented in **Supplementary Figure 3**, and the R^2 values are >0.900.

Estimation of the reduced number of the myopes with strong/weak intervention

(1) Weak effect intervention (MOV): By the year 2050, the estimated reduced population of myopia is 13,295 (95% CI: 0–182,623) for 7–12 years old and 41,731 (95% CI: 19,694–63,768) for 13–18 years old, respectively. For 7–12 years old, the reduced population of high myopia was 319 (95% CI: 0–4,383) and for 13–18 years old, it was 6,341 (95% CI: 2,992–9,689). **(2) Strong effect intervention (ATRH)**: The estimated reduced population of myopia by 2050 is 1,259,086 (95% CI: 1,089,758–1,428,414) for 7–12 years old and 584,785 (95% CI: 562,748–606,823) for 13–18 years old. The population of high myopia is 30,218 (95% CI: 26,154–34,282) for 7–12 years old and 88,852 (95% CI: 85,504–92,200) for 13–18 years old, respectively (**Figure 4**).

Discussion

According to our study, there will be roughly 24.3 and 2.6 million children and adolescents in Eastern China between the ages of 7 and 18 who have myopia or high myopia, respectively. Among myopic Chinese children and adolescents, varying percentages of effective myopia interventions (such as orthokeratology lenses, atropine therapy, contacts, glasses frames, and outdoor activities) were used. By 2050, myopia and high myopia are predicted to be reduced by nearly 55,027–1,843,872 and 6,660–119,070, respectively, according to estimates of weak or strong intervention effects. These have essential implications for myopia interventions, such as population-level and personal preventive strategies.

What are the main factors affecting our projects? First, lifestyle factors such as physical activity and close work have significantly impacted the onset and progression of myopia (31). High-pressure educational systems in Eastern China can be listed as a crucial factor (32). The human environment in this area supports culture and fosters education, and since ancient times education has received significant attention in this region. For instance, Confucius, who was born in the Eastern Chinese province of Shandong, is revered as a representative of Chinese culture and thinkers. Myopia growth may be influenced by "culture-gene." Jiangsu Province has been considered a representative area in Eastern China from the perspective of economic, educational, and cultural levels. Economic, educational, and cultural levels are highly correlated with myopia. Jiangsu Province is used in this study to approximate the myopia level in Eastern China. Japan, Singapore, and Hong Kong (China), which are part of the "Circle of Confucius-Culture" and share a lot of the same values, social structures, and cultural traditions, may be also impacted by this phenomenon.

Second, the effect of myopia controls the interventions. The prevalence of myopia is rapidly increasing globally, and this phenomenon has driven investment into myopia prevention and control (33). In this study, interventions include exhibited significant effects after adjusting parental myopia and behavior patterns among students sorted by age and gender. However, utilization rates of proven myopic interventions are still relatively low. Interventions that sufficiently slow down or delay

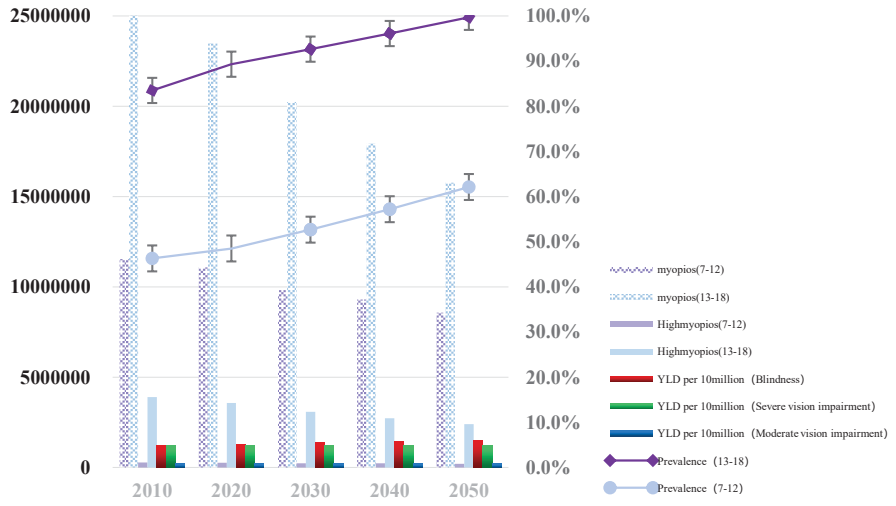

FIGURE 3
Graph depicting the estimated number of children and adolescents with myopia, high myopia, and Years Lived with Disability (YLD) from 2010 to 2050 under the current intervention situation. *YLD per 10 million (Blindness: Blindness due to uncorrected refractive error per 10 million children and adolescents); YLD per 10 million (Moderate vision impairment: Moderate vision impairment due to uncorrected refractive error per 10 million children and adolescents); YLD per 10 million (Severe vision impairment: Severe vision impairment due to uncorrected refractive error per 10 million children and adolescents).

myopia can potentially prevent an individual from developing high myopia provided treatment is started early enough (3). We estimated the reduced number of myopia, high myopia, and related YLD after strong/weak interventions. This indicated that a concerted effort by the government, education, and health systems should be undertaken to control myopia.

Different interventions have been attempted to reduce myopic progression, including increasing outdoor time (12), optical methods [orthokeratology (34, 35), contact lens (36, 37), and eyeglasses], and pharmacological methods including atropine eye drops (38). Spending 10–14 h per week in outdoor activities (MOA) as compared to "engaging in outdoor activities" only 0–5 h per week was associated with approximately half the risk of developing myopia (39). Atropine treatment shared a relatively lower usage rate in Eastern China. High-dose atropine (0.5–1%) is the most effective, but it has significant trade-offs with respect to the rebound of myopia on discontinuation and side effects. Low doses of atropine have been trialed and show a dose-dependent efficacy (40). Students receiving interventions (orthokeratology lenses, atropine treatment, contact lenses, and frame glasses) were more likely to have a higher level of myopia than those not receiving interventions (**Supplementary Table 2**). As explained in the introduction section, due to specific differences in the cognitive level of myopia, parents pay more attention to myopia intervention for children with high myopia. In contrast, parents ignore children with low myopia, leading to such a phenomenon in single-factor analysis. In the multivariate analysis, the addition of correction variables reduces the occurrence of such bias. Age, gender, level of myopia, family type, parent, region economic level, and active medical treatment record significantly impacted intervention utilization among Chinese children (**Supplementary Table 3**).

Our study design has some potential limitations. (1) The population predicted comes from the general education group, and the population's myopia from education diversion is not considered. (2) The definition of screening myopia was based on our previous studies; in 2018, we randomly selected 36 primary and secondary schools in 12 counties of Jiangsu Province and conducted myopia diagnosis and myopia screening for 7,441 students aged 7–18 years. We suggested that under the condition of non-mydriasis and large populations, we recommend that the myopia screening strategy for students aged 7–12 was UCVA $<$ 0.5 and NCAR $<$ −0.50D, and the myopia screening strategy for students aged 13–18 was UCVA $<$ 0.5. Myopia screening data are utilized, but changes in optical refraction before and after cyclic events are not considered. This could lead to an overestimation of myopic individuals in future. (3) Frame glasses and contact lenses, such as single and defocus, with different myopia control effects, were not differentiated in the study, which may yield bias on utilization and effect of intervention estimation.

Strengths can be listed as follows: (1) From 2010 to 2022, Eastern China's ongoing large-sample size surveillance of myopia screening can correctly describe the health trends of children and adolescents. (2) Considering the influence of

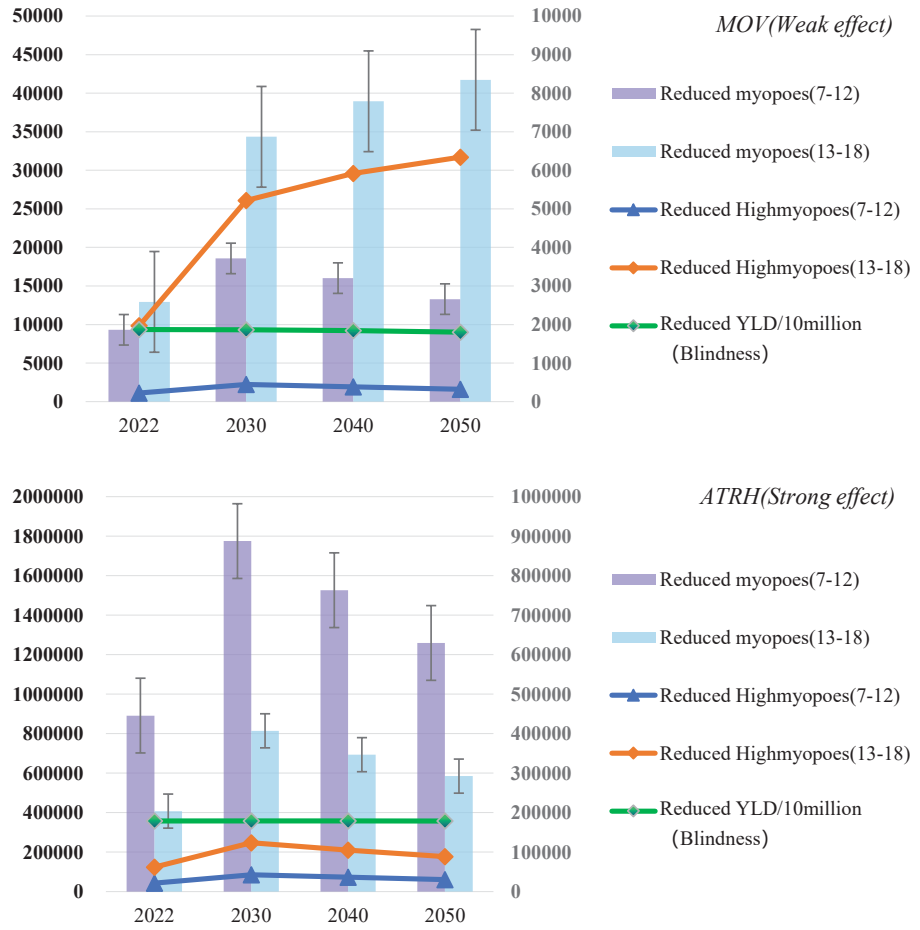

FIGURE 4
Estimates for 2022–2050 after MOV (weak effect) and ATRH (strong effect) are 100% covered: the number of myopia reductions, the number of high myopia reductions, and the rate of Years Lived with Disability (YLD) (blindness) reductions.

myopia control interventions and a thorough prediction of the decrease in myopes among children and adolescents, this study provides a research foundation for the burden of myopia disease in Eastern China.

In conclusion, our research offers predictions for myopia, high myopia, and YLD of refraction and accommodation disorders through various interventions by 2050. The usage and effects of myopia interventions (such as outdoor activities, orthokeratology lenses, atropine treatment, contact lenses, frame glasses, and eye exercises) vary among Chinese children and adolescents. The total number of children and adolescents with myopia in East China declined from 2010 to 2050 under the current intervention status, while the prevalence of myopia significantly increased. Myopia prevalence can be significantly decreased by increasing strong and weak interventions. A concerted campaign by the government, schools, and health systems should be made to prevent myopia, as Eastern China has long placed a high value on education.

Data availability statement

The original contributions presented in this study are included in the article/**Supplementary material**, further inquiries can be directed to the corresponding authors.

Ethics statement

The studies involving human participants were reviewed and approved by Ethics Committee of Jiangsu Province CDC. Written informed consent to participate in this study was provided by the participants' legal guardian/next of kin.

Author contributions

XZ and JY: data curation. YW and YZ: investigation and methodology. WD: project administration and supervision.

XZ: design, advice, and writing the original draft. WD and YZ: manuscript modification. All authors have read and approved the manuscript.

Funding

Annually Basic Public Health of Child and Adolescent Health (School Health) funding was covered by the Jiangsu Provincial Department of Finance and Jiangsu Provincial Health Committee for 2019–2020 (2019 64) and 2020–2021 (2020 124) (2020 52), respectively. The role of the funder(s) was not applicable in this study. This research was supported by a Demonstration Project of Comprehensive Prevention and Control of Emerging Infectious Diseases (BE2015714) http://www.jshealth.com/xxgk/sewcjz/xmzl/sbjkt/201708/t20170830_59379.html (Fengyun Zhang).

Acknowledgments

The authors thank Professor Fengyun Zhang and Dr. Dian Lu for constructive suggestions and help with the manuscript.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmed.2022.1069649/full#supplementary-material

References

1. Walline J, Lindsley K, Vedula S, Cotter S, Mutti D, Ng S, et al. Interventions to slow progression of myopia in children. *Cochrane Database Syst Rev.* (2020) 1:CD004916. doi: 10.1002/14651858.CD004916.pub4

2. Bourne R, Stevens G, White R, Smith J, Flaxman S, Price H, et al. Causes of vision loss worldwide, 1990-2010: a systematic analysis. *Lancet Glob Health.* (2013) 1:e339–49. doi: 10.1016/S2214-109X(13)70113-X

3. Holden B, Fricke T, Wilson D, Jong M, Naidoo K, Sankaridurg P, et al. Global prevalence of myopia and high myopia and temporal trends from 2000 through 2050. *Ophthalmology.* (2016) 123:1036–42. doi: 10.1016/j.ophtha.2016.01.006

4. Li Y, Liu J, Qi P. The increasing prevalence of myopia in junior high school students in the Haidian District of Beijing, China: a 10-year population-based survey. *BMC Ophthalmol.* (2017) 17:88. doi: 10.1186/s12886-017-0483-6

5. Chen M, Wu A, Zhang L, Wang W, Chen X, Yu X, et al. The increasing prevalence of myopia and high myopia among high school students in Fenghua city, eastern China: a 15-year population-based survey. *BMC Ophthalmol.* (2018) 18:159. doi: 10.1186/s12886-018-0829-8

6. Ohno-Matsui K. What is the fundamental nature of pathologic myopia? *Retina.* (2017) 37:1043–8. doi: 10.1097/IAE.0000000000001348

7. Sankaridurg P, Tahhan N, Kandel H, Naduvilath T, Zou H, Frick K, et al. IMI impact of myopia. *Invest Ophthalmol Vis Sci.* (2021) 62:2. doi: 10.1167/iovs.62.5.2

8. National Bureau of Statistics. (2020). Available online at: http://www.stats.gov.cn/tjsj/pcsj/

9. Xu C, Pan C, Zhao C, Bi M, Ma Q, Cheng J, et al. Prevalence and risk factors for myopia in older adult east Chinese population. *BMC Ophthalmol.* (2017) 17:191. doi: 10.1186/s12886-017-0574-4

10. Wang J, Ying G, Fu X, Zhang R, Meng J, Gu F, et al. Prevalence of myopia and vision impairment in school students in Eastern China. *BMC Ophthalmol.* (2020) 20:2. doi: 10.1186/s12886-019-1281-0

11. Mak C, Yam J, Chen L, Lee S, Young A. Epidemiology of myopia and prevention of myopia progression in children in East Asia: a review. *Hong Kong Med J.* (2018) 24:602–9. doi: 10.12809/hkmj187513

12. Mingguang H, Fan X, Yangfa Z, Mai J, Chen Q, Zhang J, et al. Effect of time spent outdoors at school on the development of myopia among children in China: a randomized clinical trial. *JAMA.* (2015) 314:1142–8. doi: 10.1001/jama.2015.10803

13. Wu P, Tsai C, Wu H, Yang Y, Kuo H. Outdoor activity during class recess reduces myopia onset and progression in school children. *Ophthalmology.* (2013) 120:1080–5. doi: 10.1016/j.ophtha.2012.11.009

14. Indian Journal of Medical Sciences. Sight test and glasses could dramatically improve the lives of 150 million people with poor vision. *Indian J Med Sci.* (2006) 60:485–6.

15. Li L, Lam J, Lu Y, Ye Y, Lam D, Gao Y, et al. Attitudes of students, parents, and teachers toward glasses use in rural China. *Arch Ophthalmol.* (2010) 128:759–65. doi: 10.1001/archophthalmol.2010.73

16. Yi H, Zhang L, Ma X, Congdon N, Shi Y, Pang X, et al. Poor vision among China's rural primary school students: prevalence, correlates and consequences. *China Econ Rev.* (2015) 33:247–62. doi: 10.1016/j.chieco.2015.01.004

17. Sun Y, Wang L, Gao J, Yang M, Zhao Q. Influence of overnight orthokeratology on corneal surface shape and optical quality. *J Ophthalmol.* (2017) 2017:3279821. doi: 10.1155/2017/3279821

18. Kerns R. Research in orthokeratology. Part I: introduction and background. *J Am Optom Assoc.* (1976) 47:1047–51.

19. Yam J, Jiang Y, Tang S, Law A, Chan J, Wong E, et al. Low-concentration atropine for myopia progression (LAMP) Study: a randomized, double-blinded, placebo-controlled trial of 0.05%, 0.025%, and 0.01% atropine eye drops in myopia control. *Ophthalmology.* (2019) 126:113–24. doi: 10.1016/j.ophtha.2018.05.029

20. Chia A, Chua W, Cheung Y, Wong W, Lingham A, Fong A, et al. Atropine for the treatment of childhood myopia: safety and efficacy of 0.5%, 0.1%, and 0.01% doses (atropine for the treatment of myopia 2). *Ophthalmology.* (2012) 119:347–54. doi: 10.1016/j.ophtha.2011.07.031

21. Yu XA, Zhang L, Zhang R, Bai X, Zhang Y, Hu Y, et al. Comparison of three schemes for myopia screening in children and adolescents. *Chin Sch Hyg.* (2019) 40:3.

22. National Health Commission. *Guidelines for appropriate techniques for myopia prevention and control in children and adolescents*. Beijing: National Health Commission (2019).

23. Tan W, Chen L, Zhang Y, Xi J, Hao Y, Jia F, et al. Regional years of life lost, years lived with disability, and disability-adjusted life-years for severe mental disorders in Guangdong Province, China: a real-world longitudinal study. *Glob Health Res Policy.* (2022) 7:17. doi: 10.1186/s41256-022-00253-3

24. Ju-Long D. Control problems of grey systems. *Syst Control Lett.* (1982) 1:288–94. doi: 10.1016/S0167-6911(82)80025-X

25. Yan J, Li Y, Zhou P. Impact of COVID-19 pandemic on the epidemiology of STDs in China: based on the GM (1,1) model. *BMC Infect Dis.* (2022) 22:519. doi: 10.1186/s12879-022-07496-y

26. Global Burden of Disease Study [GBD]. *Data Resources*. (2016). Available online at: ttps://ghdx.healthdata.org/gbd-2016

27. Huang J, Wen D, Wang Q, McAlinden C, Flitcroft I, Chen H, et al. Efficacy comparison of 16 interventions for myopia control in children: a network meta-analysis. *Ophthalmology.* (2016) 123:697–708. doi: 10.1016/j.ophtha.2015.11.010

28. Zhang X, Yang J, Wang Y, Liu W, Yang W, Gao L, et al. Epidemiological characteristics of elevated blood pressure among middle and high school students aged 12-17 years: a cross-sectional study in Jiangsu Province, China, 2017-2018. *BMJ Open.* (2019) 9:e27215. doi: 10.1136/bmjopen-2018-027215

29. Zhang X, Wang Y, Pan C, Yang W, Xiang Y, Yang J, et al. Effect of genetic-environmental interaction on Chinese childhood myopia. *J Ophthalmol.* (2020) 2020:6308289. doi: 10.1155/2020/6308289

30. Zhang X, Zhou Y, Yang J, Wang Y, Yang W, Gao L, et al. The distribution of refraction by age and gender in a non-myopic Chinese children population aged 6-12?years. *BMC Ophthalmol.* (2020) 20:439. doi: 10.1186/s12886-020-01709-1

31. Morgan I, Wu P, Ostrin L, Tideman J, Yam J, Lan W, et al. IMI risk factors for myopia. *Invest Ophthalmol Vis Sci.* (2021) 62:3. doi: 10.1167/iovs.62.5.3

32. Morgan I, Ohno-Matsui K, Saw S. Myopia. *Lancet.* (2012) 379:1739–48.

33. Cho P, Tan Q. Myopia and orthokeratology for myopia control. *Clin Exp Optom.* (2019) 102:364–77. doi: 10.1111/cxo.12839

34. Cho P, Cheung S. Retardation of myopia in orthokeratology (ROMIO) study: a 2-year randomized clinical trial. *Invest Ophthalmol Vis Sci.* (2012) 53:7077–85. doi: 10.1167/iovs.12-10565

35. Hiraoka T, Kakita T, Okamoto F, Takahashi H, Oshika T. Long-term effect of overnight orthokeratology on axial length elongation in childhood myopia: a 5-year follow-up study. *Invest Ophthalmol Vis Sci.* (2012) 53:3913–9. doi: 10.1167/iovs.11-8453

36. Lam C, Tang W, Tse D, Lee R, Chun R, Hasegawa K, et al. Defocus Incorporated Multiple Segments (DIMS) spectacle lenses slow myopia progression: a 2-year randomised clinical trial. *Br J Ophthalmol.* (2020) 104:363–8. doi: 10.1136/bjophthalmol-2018-313739

37. Kanda H, Oshika T, Hiraoka T, Hasebe S, Ohno-Matsui K, Ishiko S, et al. Effect of spectacle lenses designed to reduce relative peripheral hyperopia on myopia progression in Japanese children: a 2-year multicenter randomized controlled trial. *Jpn J Ophthalmol.* (2018) 62:537–43. doi: 10.1007/s10384-018-0616-3

38. Fan D, Lam D, Chan C, Fan A, Cheung E, Rao S. Topical atropine in retarding myopic progression and axial length growth in children with moderate to severe myopia: a pilot study. *Jpn J Ophthalmol.* (2007) 51:27–33. doi: 10.1007/s10384-006-0380-7

39. Jones L, Sinnott L, Mutti D, Mitchell G, Moeschberger M, Zadnik K, et al. Parental history of myopia, sports and outdoor activities, and future myopia. *Invest Ophthalmol Vis Sci.* (2007) 48:3524–32. doi: 10.1167/iovs.06-1118

40. Tran H, Tran Y, Tran T, Jong M, Coroneo M, Sankaridurg P. A Review of myopia control with atropine. *J Ocul Pharmacol Ther.* (2018) 34:374–9. doi: 10.1089/jop.2017.0144

Frontiers in Medicine

Translating medical research and innovation into improved patient care

A multidisciplinary journal which advances our medical knowledge. It supports the translation of scientific advances into new therapies and diagnostic tools that will improve patient care.

Discover the latest Research Topics

See more →

Frontiers

Avenue du Tribunal-Fédéral 34
1005 Lausanne, Switzerland
frontiersin.org

Contact us

+41 (0)21 510 17 00
frontiersin.org/about/contact