Survey of Ophthalmology 69 (2024) 499–507

Contents lists available at ScienceDirect

Survey of Ophthalmology

journal homepage: www.elsevier.com/locate/survophthal

Review article

A systematic review of economic evaluation of artificial intelligence-based screening for eye diseases: From possibility to reality

Hongkang Wu[a,1], Kai Jin[a,1], Chee Chew Yip[b], Victor Koh[c], Juan Ye[a,*]

[a] Eye Center, The Second Affiliated Hospital, School of Medicine, Zhejiang University, Hangzhou, Zhejiang, China
[b] Department of Ophthalmology & Visual Sciences, Khoo Teck Puat Hospital, Singapore, Singapore
[c] Department of Ophthalmology, National University Hospital, National University of Singapore, Singapore

ARTICLE INFO

Keywords:
Ophthalmology
Health economic analysis
Artificial intelligence
Diabetic retinopathy
Retinoblastoma
Retinopathy of prematurity
Age-related macular degeneration
Glaucoma

ABSTRACT

Artificial Intelligence (AI) has become a focus of research in the rapidly evolving field of ophthalmology. Nevertheless, there is a lack of systematic studies on the health economics of AI in this field. We examine studies from the PubMed, Google Scholar, and Web of Science databases that employed quantitative analysis, retrieved up to July 2023. Most of the studies indicate that AI leads to cost savings and improved efficiency in ophthalmology. On the other hand, some studies suggest that using AI in healthcare may raise costs for patients, especially when taking into account factors such as labor costs, infrastructure, and patient adherence. Future research should cover a wider range of ophthalmic diseases beyond common eye conditions. Moreover, conducting extensive health economic research, designed to collect data relevant to its own context, is imperative.

1. Introduction

The wide range of digital technologies has created a unique opportunity in ophthalmology to embrace new models of care[28]. This can be achieved through the integration of telemedicine supported by digital innovation[31]. Artificial intelligence (AI) refers to the simulation of human intelligence in machines that are programmed to perform tasks autonomously[19], while deep learning (DL) is a subset of AI that utilizes neural networks with multiple layers to extract high-level representations and make complex decisions[30]. Digital ophthalmic diagnostics and therapeutics, integrating with AI and DL, reduces geographic disparities, improves ophthalmic care standards, and advances the prevention and treatment of visual impairment and blindness[52, 55]. Despite the significant potential of AI, its implementation in the real world is fraught with many pragmatic issues, including data sharing and privacy, transparency of AI algorithms, data standardization and interoperability across platforms, and patient safety[21]. In addition, the health economics of AI in ophthalmology is a critical consideration.

Health economic analysis (HEA) is a comparative analysis of the costs and consequences of a course of action[25]. Three types of analyses were utilized in the articles chosen for this paper. First, cost-minimization analysis is used to identify the least costly intervention when 2 interventions are assumed to have the same outcome. Second, cost-effectiveness analysis uses "natural units" (life years gained, symptom-free days, complications avoided, or cases diagnosed) to evaluate the costs and consequences of alternative interventions; that is, to maximize the health benefits to society within a limited budget. Third, cost-utility analysis is a form of affective expression that focuses on measuring patient preferences for particular health states; its main advantage is that it allows comparisons of outcomes across disease states[44].

HEA of different healthcare interventions including AI by compare costs with outcomes that are measured using natural methods. These outcomes can be quantified in terms of lives saved, life years gained, or days free of pain and symptoms[43]; however, despite the focus of research on the technical aspects of AI in recent years[3, 8, 21, 51], systematic evaluation of the health-economics of AI remains inadequate with scarce and fragmented data. We explore the health-economic implications of AI-based Eye screening in ophthalmology through a comprehensive review of relevant literature (Figure 1). Finally, we shall conclude with

Abbreviations: AI, artificial intelligence; AMD, age-related macular degeneration; DL, deep learning; DR, diabetic retinopathy; HEA, health economic analysis; ICER, incremental cost-effectiveness ratio; ICUR, incremental cost-utility ratio; QALY, quality-adjusted life years; ROP, retinopathy of prematurity.

* Corresponding author.
E-mail address: yejuan@zju.edu.cn (J. Ye).

[1] Both authors contributed equally to this article.

https://doi.org/10.1016/j.survophthal.2024.03.008
Received 29 August 2023; Received in revised form 4 March 2024; Accepted 12 March 2024
Available online 15 March 2024

pertinent questions and future perspectives in this field.

2. Method

To summarize the results of the different HEA studies, we conducted comprehensive literature following the PRISMA guidelines[41]. Searches were performed through the databases such as Pubmed, Google Scholar, and Web of Science from inception to July 7, 2023. We searched using the following keywords : ("Ophthalmology" OR "Eye" OR "Ophthalmology"[Mesh]) AND ("Deep Learning" OR "Artificial Intelligence" OR "Machine Learning" OR "Neural Network" OR "Artificial Intelligence"[Mesh] OR "Deep Learning"[Mesh]) AND ("Cost-Effectiveness Analysis" OR "Economic Evaluation" OR "Cost-Utility Analysis" OR "Marginal Analysis" OR "Cost-Benefit Data" OR "Cost-Effectiveness Analysis"[Mesh] OR "Cost-Benefit Analysis"[Mesh] OR "Health Care Economics and Organizations" [Mesh]). The terms from each category were cross-referenced independently with terms from the other category.

To conclude, we make a list of specific exclusion criteria in Figure 2. The publication must be the primary research, and the full text must be available online. It must analyze the patients who have an ophthalmologic diagnosis or who are currently under evaluation for an ophthalmologic diagnosis. The included research must also be published in English. Meanwhile, studies that meet the following criteria will be excluded from the analysis:[1] duplicated literature already included in the review,[2] topics that are unrelated to the research subject,[3] conference abstracts, and[4] non-original research, including case reports, editorials, and commentaries.

We selected articles that utilized HEA system to evaluate the cost-effectiveness of using AI-based versus ophthalmologist-based eye screening in diagnosing ophthalmic conditions in different disease groups and populations. A total of 98 papers were independently screened for eligible studies (title and abstract, then full text) by 2 reviewers (Hongkang Wu, Kai Jin). Any disagreements were resolved by discussion with the third author (Juan Ye). The HEA evaluation results and present the results in Table 1.

The articles selected were those that used HEA. Among the initial 98 articles, 15 articles met the inclusion/exclusion criteria and we classified the findings according to the different ophthalmic disease groups which AI has been used: "diabetic retinopathy", "retinoblastoma", "retinopathy of prematurity" , "age-related macular degeneration", "glaucoma" and "other blinding eye diseases". We analyze the articles and present the results according to four cost-effectiveness indicators. Firstly, "Incremental cost", the additional cost associated with implementing a new treatment or intervention compared to the standard or existing method. Secondly, "Incremental Quality-Adjusted Life Years (incremental QALYs)", a way to quantify the additional benefit in terms of quality and length of life. Thirdly, "Incremental Cost-Utility Ratio (ICUR)", a statistic used in health economics to assess the cost-effectiveness of a health intervention. Fourthly, "Incremental Cost-Effectiveness Ratio (ICER)", similar to ICUR, a statistic used in health economics to compare the relative costs and outcomes (effects) of two or more interventions. To facilitate meaningful comparisons, the financial figures in the articles from non-USA countries will be converted to USD at the year of article publication.

3. Results

3.1. Diabetic retinopathy

Diabetic retinopathy (DR) remains a significant cause of vision loss and preventable blindness, particularly in the 20–74 age group, with a

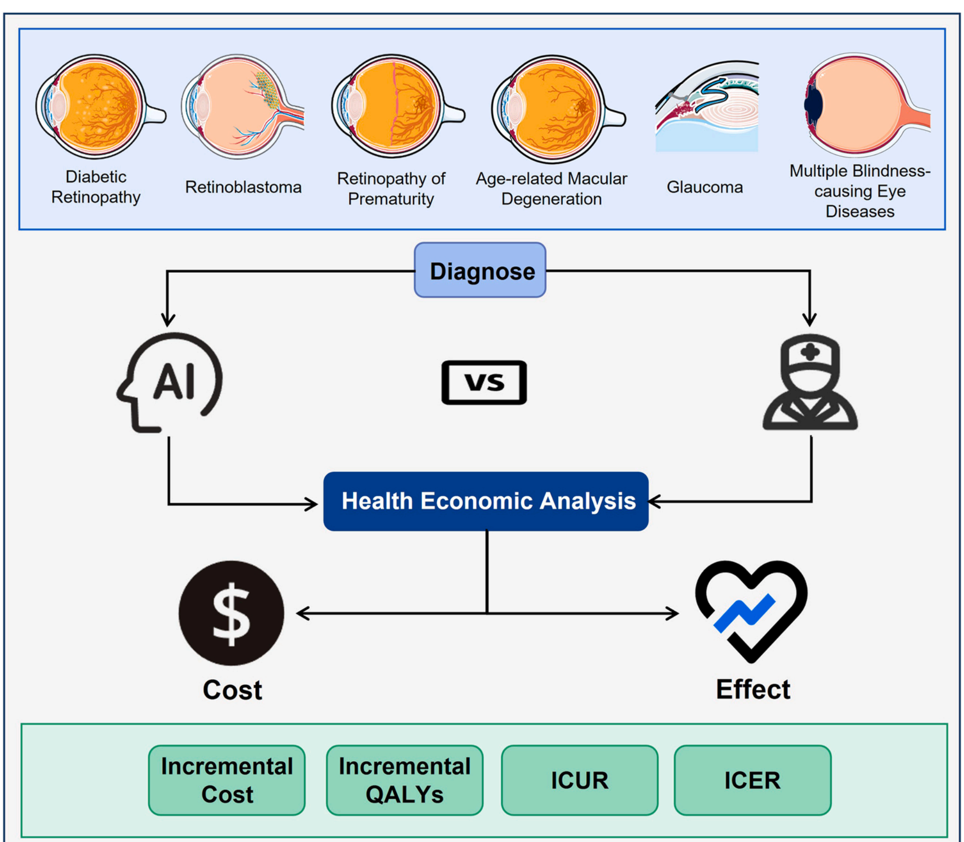

Fig. 1. The flow of economic evaluation of artificial intelligence-based screening for eye diseases. NA : QALYs, quality-adjusted life years ; ICUR, Incremental cost-utility ratios; ICER, Incremental cost-effectiveness ratio.

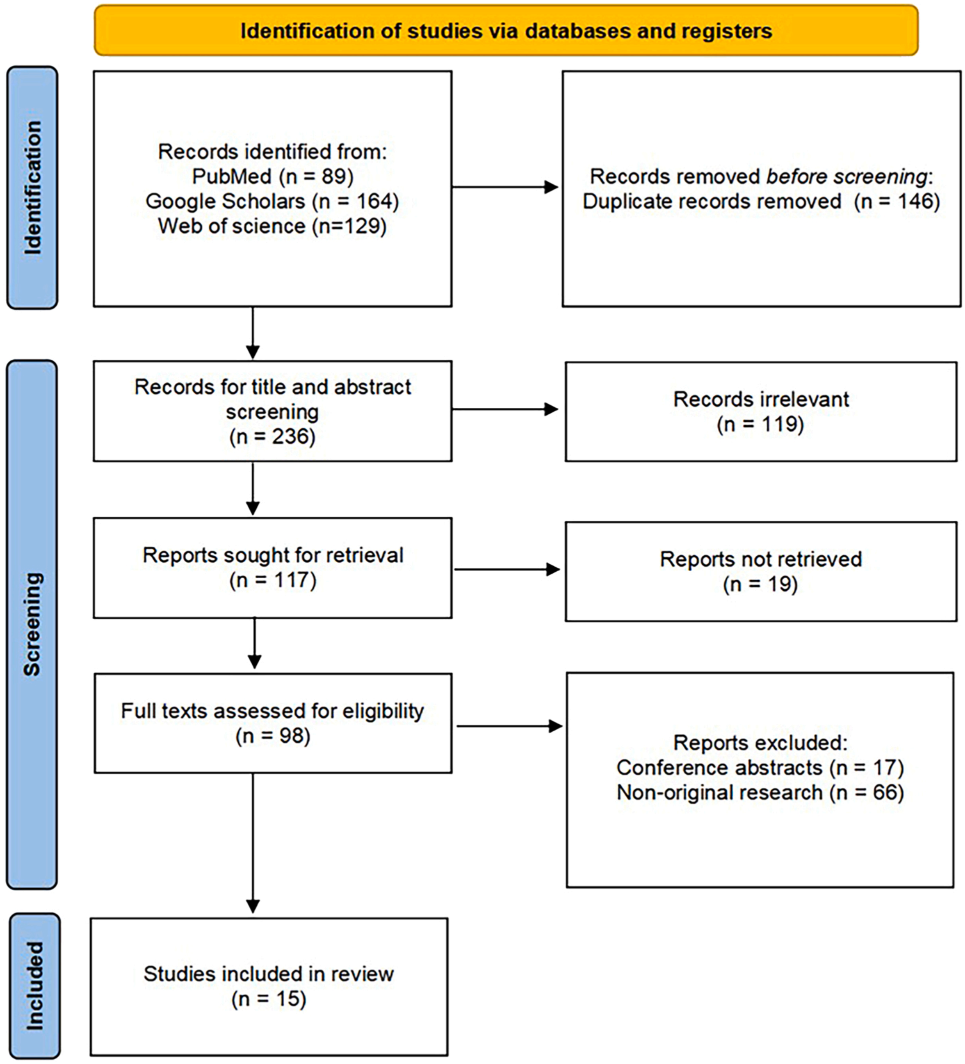

Fig. 2. PRISMA 2020 flow diagram for the systematic reviews which included searches of databases and registers only.

higher prevalence observed in middle and high income countries[7]. Therefore, the diagnosis and treatment of DR have received considerable attention in the field of -AI. A number of studies have demonstrated the safety and accuracy of AI for the screening of DR using digital fundus images[2, 26, 53]; however, there is much variability in the reported cost-effectiveness of AI-based DR screening in different countries.

3.2. *Cost-effectiveness in high-income countries*

Scotland and coworkers first compared the cost-savings of automated AI-based grading versus human grading of DR from fundus images of 6722 patients by 14,406 graders[46]. Automated grading was found potentially to save the National Health Service of Scotland $403, 200 per year. The additional cost for each additional reference case detected by manual grading was also computed which amounted to $8176[46].

Fuller and coworkers investigated the cost-effectiveness of AI-based grading of fundal images to detect DR over a 5-year period in America using two FDA-approved AI systems, IDx-DR and EyeArt, in comparison with human screening by clinicians[14]. Both AI systems achieved a 23.3% cost reduction per patient compared with human grading[14]. The analysis showed that AI was cost-effective up to a combined AI screening cost of $161.14 per screening and up to a standard eye exam adherence rate of 71.5%. Compared to the current practice, the ICUR is calculated to be $258,721.81, which significantly exceeds the willingness-to-pay threshold of $100,000. Furthermore, the incremental cost of AI compared to current practice is $485.92, which means AI saves that much money[14].

Similarly, Xie and coworkers conducted an analysis in Singapore to assess the feasibility of AI in DR screening, to assess the potential savings associated with using two AI models compared to human graders: a semi-automated AI as a filter (filtering for no DR and mild non-proliferative DR) prior to a second human assessment; and a fully automated model that completely replaces all human assessments. The results showed cost reductions of 19.5% and 14.3% for semiautomated and fully automated AI, respectively. In the fully automated model, a higher rate of false positives leads to an increase in unnecessary specialist visits, which results in higher overhead costs. The estimated use of semi-automated AI indicates that it could save the Singapore health system $361,000, which is about 20% of the current annual cost of screening[60]. Nonetheless, the cost estimates may have limited generalizability as they were based on average Singaporean wages[60]. Another study by Tufail and coworkers (2017) in the United Kingdom evaluated the cost-effectiveness of two AI systems, Retmarker and EyeArt, in adult DR screening[54]. This study aims to assess the cost-effectiveness of EyeArt and Retmarker by comparing the following two strategies:[1] using AI instead of the initial manual scoring and[2] using AI as a filter before manual scoring. Strategy 1 involves using AI instead of the initial manual scoring, while Strategy 2 involves using AI as a filter before manual scoring. The results indicate that for EyeArt, the

Table 1
Results of available health economic analysis in ophthalmology.

Reference	Journal/Year	Disease/Sample	Method	Model performance	Incremental costs	Incremental QALYs	Incremental cost-utility ratios (ICUR)	Incremental cost-effectiveness ratios (ICER)	Whether cost-effective
Zhang, R.[62]	Br J Cancer 2023	Retinoblastoma (36623)	Cost-utility analysis	Sensitivity: 97.80% Specificity: 80.00%	Screening: $– 1363 Surveillance: $– 633	Screening: 0.00,529 Surveillance: 0.00 091	Screening: $– 257,645 Surveillance: $– 694,323	*	Positive
Srisubat, A.[47]	Ophthalmol Ther 2023	DR	Cost-effectiveness analysis	Sensitivity: 95.00% Specificity: 98.00%	Societal perspective: $– 2.6 Provider perspective: $66.7	0.0043	*	Societal perspective: dominant Provider perspective: 16,020	Positive
Liu, H.[35]	Lancet Glob Health 2023	Multiple disease	Cost–utility analysis Cost-effectiveness analysis	*	Rural setting: $– 13,004,906 Urban setting: $4145,752 (per 100,000 people)	Rural setting: 3203 Urban setting: 16,959 (per 100,000 people)	Rural setting: dominating Urban setting: $244 (−315 to 1073, 95% CI)	Rural setting: dominating Urban setting:2567 (−4111 to 15 389, 95%CI)	Positive
Lin, S.[33]	JMIR Public Health Surveill 2023	DR	Cost–utility analysis and cost-effectiveness analysis	Sensitivity: 80.47% Specificity: 97.96%	$– 8289,840.65 per 100,000 people	-544.78 per 100,000 people	2553	15,216.96	Positive
Morrison, S. L.[38]	JAMA Ophthalmol 2022	Retinopathy of prematurity (52000)	Cost-effectiveness analysis	Sensitivity: 75-99.9% Specificity: 74%	*	*	*	*	Positive
Huang, X. M.[24]	BMC Health Serv Res 2022	DR (1000)	Cost-effectiveness analysis	Sensitivity: 90.79% Specificity: 98.50%	Health system perspective: $– 34.86 Societal perspective: $– 92.25	0.16	*	Health system perspective: 1107.63 Societal perspective: 10,347.12	Positive
Gomez Rossi, J.[15]	JAMA Netw Open 2022	DR	Cost-effectiveness analysis	*	$25.82	0.04	*	38,848	Negative
Fuller, S·D· (14)	J Diabetes Sci Technol 2022	DR (179)	Cost-utility analysis	*	$– 485.92	0.00188	258,721.81	*	Positive
Tamura, H.[50]	Jpn J Ophthalmo 2022	AMD (500000)	Cost-effectiveness analysis	Sensitivity: 98.20% Specificity: 91.20%	$597.20	0.0064	*	92,890	Negative
Wang, Y.[56]	Front Med (Lausanne) 2021	DR (88363)	Cost-minimization analysis	Sensitivity: 97.00% Specificity: 87.90%	*	*	*	*	Positive
Xiao, X.[59]	BMC Public health (19395)	Glaucoma	Cost-offset analysis	Sensitivity: 95.90% Specificity: 96.10%	$1464.3	*	*	*	Negative
Wolf, R. M.[57]	JAMA Ophthalmol 2020	DR (39006)	Cost-effectiveness analysis	Sensitivity: 87.00% Specificity: 91.00%	Type 1 diabetes: $0.61 Type 2 diabetes:$2.65	Type 1 diabetes: 0.02 Type 2 diabetes: 0.03	*	Type 1 diabetes:$31 Type 2 diabetes:$95	Negative
Xie, Y.[60]	Lancet Digit Health 2020	DR	Cost-minimization analysis	Sensitivity: 89.90% Specificity: 81.80%	$– 15	*	*	*	Positive
Tufail, A.[54]	Ophthalmology 2017	DR (20258)	Cost-effectiveness analysis	Retmarker: 73.00%/ 52.3% EyeArt: 94.70%/20%	Retmarker: strategy1: $– 101,820 strategy2: $– 63,063 EyeArt: strategy1: $– 167,250 strategy2: $– 137,152	*	*	Retmarker: strategy1: $7.14 strategy2: $4.43 EyeArt: strategy1: $18.69 strategy2: $15.36	Positive

(continued on next page)

Table 1 (*continued*)

Reference	Journal/Year	Disease/Sample	Method	Model performance	Incremental costs	Incremental QALYs	Incremental cost-utility ratios (ICUR)	Incremental cost-effectiveness ratios (ICER)	Whether cost-effective
Scotland, G. S.[46]	Br J Ophthalmol 2007	DR (6722)	Cost-effectiveness analysis	*	*	*	*	*	Positive

NA: AMD, age-related macular degeneration; QALYs, quality-adjusted life years; DR, diabetic retinopathy; DLA, deep learning algorithm; *, not mentioned.

ICER is $7.14 for Strategy 1 and $4.43 for Strategy 2. For Retmarker, the ICER is $18.69 for Strategy 1 and $15.36 for Strategy 2. The study concludes that utilizing AI as a substitute or precursor for human grading represents a more cost-effective approach.

3.3. Cost-effectiveness in low and middle-income countries

Srisubat and coworkers developed a cost-utility analysis model tailored to middle-income countries in their study published in 2023 using Thailand as a reference model. The study focused on a hypothetical cohort of 40-year-old patients with type 2 diabetes who had not previously been screened for DR. Both AI-based grading and human graders were found to be similarly cost-effective. From a societal perspective, AI-based DR screening had a slightly lower incremental cost of $2.6, positioning it as the preferred choice over human grading[47]. Conversely, from a provider perspective, deep learning was associated with a higher incremental cost of $66.7, resulting in an ICER of $16,020 per QALY gained[47]. The study suggests that implementing screening strategies using either human grading or deep learning can result in substantial cost savings if adherence to referral for treatment is high. Improving treatment adherence can help maximize the cost-effectiveness of these screening approaches[47].

Huang and coworkers examined the cost-effectiveness of ophthalmologist-based and AI-based DR screening in rural China. The researchers developed a hybrid Markov model-based decision tree to assess the long-term cost-effectiveness of these strategies over a 35-year period[24]. The ICER of AI screening strategies was calculated by comparing AI-based screening, ophthalmologist-based screening and no screening strategies. To determine whether AI screening was cost-effective, the ICER was compared to the World Health Organization threshold of 1–3 times gross domestic product per capita[24]. The study results showed that AI screening was associated with higher costs (an incremental cost of $180.19) but greater effectiveness (an incremental gain of 0.16 QALYs) compared with no screening. On the other hand, AI-based DR screening was found to be less expensive (an incremental cost of -$34.86) and more effective (an increase of 0.04 QALYs) than ophthalmologist-based DR screening[24]. The ICER for the AI-based DR screening group compared with the unscreened group was $1107.63 per QALY gained. Importantly, this ICER was below the willingness-to-pay threshold of $30,765.09, which is 3 times the per capita gross domestic product of China in 2019[24]. Consequently, based on this analysis, AI-based DR screening was considered cost-effective[24].

Also in China, Lin and coworkers conducted cost-effectiveness and cost-utility analyses to compare two scenarios of telemedicine screening for DR. The analyses used decision-analytic Markov models over a 30-year period, taking into account the societal impact perspective. The cost-effectiveness and cost-utility analyses showed that for a community resident with diabetes who underwent human graded medical screening, the total cost was $3265.40, which included screening, referral for confirmation, and necessary treatment[33]. This scenario resulted in 9.83 years without blindness and a QALY value of 6753. In contrast, the total cost of 9.80 years of blindness and 6.748 years of QALYs for a community member with diabetes in the AI-based eye screening scenario was $3182.50[33]. Thus, compared to using human graders, the AI-based telemedicine screening model reduced costs by 2.5%, blindness years by 0.3%, and QALYs by 0.1%[33]. Without improved adherence, AI-based telemedicine DR screening would not be more cost-effective than human grader-based telemedicine DR screening, primarily because of the limited direct healthcare cost savings of AI replacing manual grading when labor costs are low and the effectiveness of screening declines[33].

Furthermore, in a 2021 study, Wang and coworkers evaluated the cost-effectiveness of using semi-automated AI in a Chinese DR screening program. Their analysis, based on time and cost metrics, showed that semi-automated AI significantly reduced screening time per case compared to human screening, resulting in a time savings of 75.6%[56]. From a cost perspective, AI grading of retinal images was 90.1% less expensive than human grading[56]. The study included a large sample of 88,363 images, providing evidence that AI implementation in ophthalmology is economically viable and offers comparable efficacy to human grading[56].

3.4. Not cost-effectiveness or uncertain in demonstrating differences in cost-effectiveness

The economic impact of AI in healthcare, including its application in ophthalmology, is multifaceted and complex. They are often influenced by a variety of factors, one of which is the labor cost specific to the region where the technology is implemented. This is a crucial aspect that can have a significant impact on the cost-effectiveness of AI in healthcare.

This concept is further substantiated by the research conducted by Gomez and coworkers. They carried out a comprehensive study in Brazil, meticulously analyzing the cost-effectiveness of AI in diagnosing DR. According to their detailed analysis, the mean cost from a payer perspective for AI-based diagnosis of DR was R$1321, which is approximately US$559. In contrast, the cost without the use of AI was slightly lower, at R$1260, or approximately US$533[15]. Interestingly, both strategies yielded a similar mean benefit of 8.4 QALY;[15] however, the introduction of AI increased costs by R$ 61, or approximately US$ 25.82. The ICER was calculated at R$91,760, or approximately US$ 38, 848[15].

The study by Wolf and coworkers focused on the cost-effectiveness of using AI for DR screening in a pediatric cohort. Patients manually screened for type 1 diabetes incurred an average out-of-pocket cost of $7.91. For type 2 diabetes, the cost was $8.20. In contrast, patients with type 1 and type 2 diabetes who were screened with AI had average out-of-pocket costs of $8.52 and $10.85, respectively[57]. These results were based on an assumed average screening rate of 20% of the US population. The ICER for type 1 diabetes is $31, whereas for type 2 diabetes, it is $95[57]. Research has demonstrated that an AI screening must reach a screening rate of 23% to be more cost-effective than manual screening[57]. This increase is relatively small compared to the current rate, suggesting that AI could be a more cost-effective solution in the near future.

Both studies emphasize that although there may be specific situations where AI implementation leads to higher patient care expenses, the general trend suggests that AI has the potential to improve efficiency and save healthcare costs in most cases. These results align with the broader view that properly implemented and optimized, AI can streamline processes, improve accuracy, and enable better resource allocation in healthcare. The potential for higher efficiency and cost savings is a compelling argument for the continued research and implementation of AI in healthcare.

In summary, the majority of studies[14, 24, 33, 46, 47, 54, 56, 60] suggest that AI in ophthalmology can lead to cost reductions and improved efficiency in DR screening. Interestingly, screening strategies for the identical condition (DR) yield disparate results across different healthcare settings (high-income or low- and middle-income countries). This underscores the significance of considering additional factors including indirect costs to patients, the specific healthcare context (encompassing labor costs and medical equipment), patient adherence, and patient volume, in the economic evaluation of AI implementation. These factors influence outcomes including the ICER, QALY, and patient willingness-to-pay thresholds, which in turn lead to variations in HEA outcomes.

3.5. Retinoblastoma

Retinoblastoma, a frequently occurring ocular malignancy in pediatric patients, poses a substantial global health challenge. Regrettably, the latest global statistics indicate a notable fatality rate among the roughly 9000 newly diagnosed cases each year[10]. Nonetheless, significant breakthroughs in the comprehensive treatment of retinoblastoma, which involve establishing specialized facilities, enhancing infrastructure, and launching awareness campaigns, have resulted in nearly 100% survival rates in developed nations, with successful preservation of visual function in numerous instances[11]. In most developing nations, retinoblastoma patients have limited access to care, which is primarily offered by experienced ocular oncologists at tertiary eye centers. As a result of this uneven allocation of healthcare resources, there might be delays in diagnosis and increased treatment costs when patients are referred. Recently, DL algorithms have demonstrated remarkable performance in automatically detecting various ocular diseases, such as DR, age-related macular degeneration, glaucoma, and other diseases, by screening digital images from fundus screenings[9, 23, 53]. It also enables the detection and diagnosis of retinoblastoma, which reduces costs for the patient[29].

Zhang and coworkers[62] compared the cost-effectiveness of a conventional diagnostic model using ophthalmologists with an AI-based diagnostic model for retinoblastoma (DLA-RB) to monitor the clinical course of retinoblastoma from a societal perspective. The study employed 2 binary models. The first model simulated a hypothetical cohort of newborns with retinoblastoma over 5 1-year Markov cycles. The second model simulated a hypothetical cohort of 2-year-old retinoblastoma patients who underwent routine treatment and remained inactive over 3 1-year Markov cycles. During the study, patients received regular 3-monthly reviews. Those who have active lesions were given additional treatment. To identify active retinoblastoma tumors among all clinically suspected cases, the cumulative costs were $4882 and $3519 for the traditional ophthalmologist-based diagnostic mode and the DLA-RB-based diagnostic mode, respectively. The total QALYs for both diagnostic modes were 2.87,769 and 2.8724, respectively. In contrast, the traditional ophthalmologist-based diagnostic mode was deemed not cost-effective since it yielded only a gain of 1 QALY at a cost of $257,645 when compared to the DLA-RB-based diagnosis mode. To differentiate active retinoblastoma from stable retinoblastoma, the traditional diagnostic mode based on ophthalmic centers incurred cumulative costs and QALY values of $7614 and 5.46,436, respectively. In comparison, the same measures for the DLA-RB-based diagnostic mode were $6981 and 5.46,345[62].

DLA-RB has high accuracy and sensitivity in differentiating active retinoblastoma from normal and stable fundus retinoblastoma. It can be used to monitor retinoblastoma activity and to screen high-risk offspring during follow-up. Compared to referral to ophthalmic centers, screening and monitoring based on DLA-RB is cost-effective and can be included in telemedicine plans.

3.6. Retinopathy of prematurity

Retinopathy of prematurity (ROP) is a common and preventable cause of visual impairment in premature infants worldwide[45]. This condition results from abnormal development of retinal blood vessels and, if not detected and treated promptly, can lead to retinal detachment and profound and permanent visual impairment[22]. Each year, 20,000 to 30,000 infants lose their vision from ROP, with the majority of these infants born in areas where healthcare resources and ROP screening are inadequate[6, 12, 20]. Ensuring adequate ROP screening globally is an important task in reducing ROP-associated blindness, especially in areas with limited medical resources.

Morrison and coworkers did a cost-effectiveness analysis[38] of different screening modalities on a simulated cohort of 52,000 newborns. The study compared mean outcomes, costs, effectiveness, and ICER using a willingness-to-pay threshold of $100,000 per QALY. In the simulated cohort, autonomous AI was found to outperform three other modalities such as telemedicine, assistive AI, and ophthalmoscopy in terms of incremental cost-effectiveness. AI-based screening for ROP was shown to be cost-effective at thresholds of $7 for assistive screening and $34 for autonomous screening, relative to telemedicine. Furthermore, in comparison to ophthalmoscopy, the cost-effectiveness thresholds were $64 and $91, respectively. According to the probabilistic sensitivity analysis, autonomous AI screening demonstrated a probability of more than 60% of cost-effectiveness across all willingness-to-pay levels compared to other modalities. In a second simulated cohort with a sensitivity of 99%, utilization of autonomous AI reduced the number of instances of delayed treatment for ROP from 265 (with ophthalmoscopy screening) to 40. It was at least as effective but less costly than the other 3 options. AI had lower mean costs compared to telemedicine and ophthalmoscopy, while assistive AI had higher mean costs than telemedicine. Importantly, AI resulted in an increase in the number of newborns treated on time and a decrease in the number treated late compared to telemedicine or ophthalmoscopy.

3.7. Age-related macular degeneration

Age-related macular degeneration (AMD) is the primary cause of blindness in developed countries[39]. The global number of people with AMD is expected to reach 288 million by 2040[58]. Typically, the disease affects individuals aged 55 and above. Initially, AMD often presents as asymptomatic, eventually leading to vision distortions or central visual field defects in the mid to late stages. These complications significantly impede tasks that require clear vision, such as reading, driving, recognizing faces[17]. Early detection and intervention prove significantly beneficial for the treatment outcomes and prognosis of the disease[48].

Tamura and colleagues[50] utilized a Markov model to simulate a cohort of 500,000 Japanese individuals aged 40 years, estimating a 3.85% prevalence of monocular early AMD. Patients were categorized into a screening group and a non-screening group, with the former undergoing triennial screenings involving three protocols: AI fundus photography, AI optical coherence tomography (AI-OCT), and ophthalmologist-administered OCT. It was envisioned that individuals with AMD in the non-screening group would seek consultation at the ophthalmology department due to presbyopia-related symptoms. The findings indicated that the ICER for the AI fundus photography, the AI-OCT, and the ophthalmologist-led OCT screenings were $99,283.05, $98,465.69, and $98,974.20, respectively, all of which exceeded the willingness-to-pay threshold of $47,286[50]. Consequently, while ophthalmic screening for AMD proves clinically beneficial over the absence of screening, its cost-effectiveness is questionable, suggesting that the early adoption of a screening program may enhance its economic viability.

3.8. Glaucoma

Glaucoma is recognized as the second leading cause of blindness globally, following cataracts[42]. It principally inflicts damage on the optic nerve, leading to a progressive loss of retinal ganglion cells[27]. Given the irrevocable nature of visual field impairment in glaucoma, early detection and timely intervention are crucial for treatment. In recent years, AI has achieved remarkable advancements in the field of glaucoma detection, which not only offers a streamlined and rapid process, but also attains precision at the level of retinal specialists[32, 34], effectively addressing the resource constraints in primary care settings.

Xiao and coworkers[59] conducted a glaucoma screening initiative for individuals aged 65 years or older (approximately 19,395) residing in Changjiang County, China. The screening protocol comprised AI-assisted diagnosis, hospital referrals, and community follow-ups. Subsequently, Markov models were utilized to simulate the course of disease progression in patients over a period of 15 years. The projected incremental costs accumulated to $396,362.80, $424,907.90, and $434, 903.20 for 5, 10, and 15 years, respectively, compared to a no-screening scenario. When considered alongside the disease prevalence, the average incremental cost per diagnosed case over 15 years is $1464.3, which suggests the programme may not be cost-saving[59].

3.9. Other blinding eye diseases

In 2020, the global population with low vision was estimated to be 596 million, of which 43 million would be classified as blind. A further 510 million people will have uncorrected visual impairment because of uncorrected refractive errors. Notably, the vast majority of those affected (90%) live in low- and middle-income countries[5]. In China, the incidence of blinding eye diseases such as cataract, glaucoma, age-related macular degeneration, DR, and pathological myopia has increased with an ageing population and urbanization[61]. Early detection of asymptomatic patients and timely referral for treatment can significantly improve visual prognosis and reduce the burden of many eye disease[49]. Routine eye screening methods can be limited by time, transportation, and financial constraints, making it difficult to implement large-scale screenings in low- and middle-income countries[4]; however, the development of AI is providing new tools for ocular disease screening, solving the problem of inequitable distribution of ophthalmic healthcare resources, reducing the burden of patient travel, and bringing great benefits to ophthalmology.

Liu and coworkers[35] developed a decision-analytic Markov model consisting of a hypothetical set of individuals aged 50 years or older, with a total of 30 Markov cycles simulating the developmental process of individuals after screening. The results of the study indicate that integrating eye disease screening is a highly cost-effective health intervention in both rural and urban areas in China. In rural areas, non-telemedicine screening and non-AI telemedicine screening were found to be highly cost-effective compared to no screening, with incremental ICUR of $2494 and $2326, respectively. Similarly, in urban areas, non-telemedicine screening and non-AI telemedicine screening resulted in one QALY gained, with ICURs of $624 and $581, respectively. AI-based telemedicine screening was found to be superior to no screening in rural areas, while in urban areas the ICUR was $244. This form of screening could prevent one year of blindness at a ICER of $2567 in urban areas, and was also superior to no screening in rural areas. The study also evaluated the health and economic indicators of acceptable screening intervals, ranging from every 1 to 5 years, for both settings. More frequent screening intervals increased costs but also prevented more years of blindness. The study concluded that annual AI screening was the most cost-effective screening scenario in both rural and urban settings[35].

4. Discussion

This review identified 15 studies that conducted health economic analyses of AI in ophthalmology. The majority of these studies[14, 24, 33, 35, 38, 46, 47, 54, 56, 60, 62] (11 out of 15) concluded that the implementation of AI in ophthalmology is cost-effective or cost-saving. This suggests that AI has the potential to provide economic benefits in the field; however, the number of studies conducting health economic analyses of AI in ophthalmology is relatively limited compared to other AI application-related literature. The quality of such studies is variable, with the majority of published studies on the economic benefits of ophthalmic AI primarily focusing on provider perspectives; however, a societal perspective may prove more beneficial in furnishing policymakers with a more comprehensive understanding of the economic impact of AI deployment, thereby yielding macro-level and comprehensive data for robust health economic assessments. Furthermore, numerous HEA studies suffer from a paucity of cost and outcomes data, thereby failing to ascertain the economic impact and cost-effectiveness of ocular care technologies with precision[37].

Various obstacles limit the availability of data in health economic evaluation research. First, the bioethics must be addressed as it touches on many issues such as patient privacy, data security, and the potential for bias in decision-making algorithms[1, 36]. Second, patient compliance with new AI-based ophthalmic technology and eye care protocols is another important factor that may impact data availability on treatment effectiveness and costs for analysis[40]. Third, there is need for clear guidelines and regulations in the implementation of new AI-based eye care to ensure the responsible and ethical use of AI technologies. Lack of solid accountability systems may cause healthcare providers and institutions to be reluctant to adopt AI solutions[18]. Fourth, the cost of hardware, software, and personnel required to implement AI can be a significant barrier. The initial investment and ongoing maintenance costs associated with AI technologies can deter healthcare organizations from adopting them[15, 57]. There is a pressing need to overcome these barriers through the strategic and concerted efforts of researcher and policymakers to facilitate the collection of meaningful data for evaluation of the health economic impact of AI technologies, and make well-informed decisions regarding their implementation in ophthalmic care.

While current research focuses primarily on DR, further studies are needed to evaluate the economic feasibility of using AI technology to manage other ophthalmic conditions such as childhood myopia[13], as well as cataract[16]. By conducting comprehensive health economic evaluations in these areas, we can better understand the cost-effectiveness and potential benefits of integrating AI into the screening and management of ophthalmology, ultimately improving patient outcomes and optimizing resource allocation in ophthalmic care.

Finally, regional differences in labor costs, levels of economic development, and distribution of medical resources can significantly affect the applicability of research data. Findings from developed countries may not be directly applicable to developing countries because of these differences. In the case of China, there is a need for more relevant research that specifically addresses the country's unique context. Conducting further research in China and other developing countries will help generate data that is more representative of local conditions, allowing for more informed decision-making and the development of tailored strategies to address the specific challenges and needs of these regions.

5. Conclusion

We provide an overview of health evaluation research on AI-based eye care models on a few blinding diseases. Future research should focus on the cost-benefit analysis of AI, taking into account differences in labor costs, technology levels, and distribution of medical resources among countries. Collaboration of major eye care centers may accelerate the clinical application of AI and produce informative data for health economic evaluation.

Funding

This work has been financially supported by Natural Science Foundation of China (grant number 82201195) , and Clinical Medical Research Center for Eye Diseases of Zhejiang Province (grant number 2021E50007).

CRediT authorship contribution statement

Chee Chew Yip: Writing – review & editing. **Kai Jin:** Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Software, Resources, Project administration, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Hongkang Wu:** Writing – original draft, Visualization, Validation, Supervision, Software, Resources, Project administration, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Juan Ye:** Writing – review & editing, Supervision, Resources, Methodology, Funding acquisition. **Victor Koh:** Writing – review & editing.

Declaration of Competing Interest

This manuscript has not been published or presented elsewhere and is not under consideration by another journal. We have read and understood your journal's policies, and we believe that neither the manuscript nor the study violates any of these. There are no conflicts of interest to declare.

Acknowledgements

Thanks to all the peer reviewers and editors for their opinions and suggestions.

Study approval

This study was approved by the Medical Ethics Committee of the Second Affiliated Hospital, Zhejiang University, and complied with the Declaration of Helsinki.

References

1. Beets B, Newman TP, Howell EL, Bao L, Yang S. Surveying public perceptions of artificial intelligence in health care in the united states: systematic review. *J Med Internet Res*. 2023;25, e40337.
2. Bellemo V, Lim ZW, Lim G, 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(1):e35–e44.
3. Biousse V, Danesh-Meyer HV, Saindane AM, Lamirel C, Newman NJ. Imaging of the optic nerve: technological advances and future prospects. *Lancet Neurol*. 2022;21 (12):1135–1150.
4. Bourne RRA, Flaxman SR, Braithwaite T, 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(9): e888–e897.
5. Burton MJ, Ramke J, Marques AP, et al. The lancet global health commission on global eye health: vision beyond 2020. *Lanct Glob Health*. 2021;9(4):e489–e551.
6. Chan-Ling T, Gole GA, Quinn GE, Adamson SJ, Darlow BA. Pathophysiology, screening and treatment of ROP: a multi-disciplinary perspective. *Prog Retin Eye Res*. 2018;62:77–119.
7. Cheung N, Mitchell P, Wong TY. Diabetic retinopathy. *Lancet*. 2010;376(9735): 124–136.
8. De Fauw J, Ledsam JR, Romera-Paredes B, et al. Clinically applicable deep learning for diagnosis and referral in retinal disease. *Nat Med*. 2018;24(9):1342–1350.
9. Deng Y, Qiao L, Du M, et al. Age-related macular degeneration: epidemiology, genetics, pathophysiology, diagnosis, and targeted therapy. *Genes Dis*. 2022;9(1): 62–79.
10. Dimaras H, Kimani K, Dimba EA, et al. Retinoblastoma. *Lancet*. 2012;379(9824): 1436–1446.
11. Fabian ID, Onadim Z, Karaa E, et al. The management of retinoblastoma. *Oncogene*. 2018;37(12):1551–1560.
12. Fierson WM. Screening examination of premature infants for retinopathy of prematurity. *Pediatrics*. 2018;142(6).
13. 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(5):413–424.
14. Fuller SD, Hu J, Liu JC, et al. Five-year cost-effectiveness modeling of primary care-based, nonmydriatic automated retinal image analysis screening among low-income patients with diabetes. *J Diabetes Sci Technol*. 2022;16(2):415–427.
15. Gomez Rossi J, Rojas-Perilla N, Krois J, Schwendicke F. Cost-effectiveness of artificial intelligence as a decision-support system applied to the detection and grading of melanoma, dental caries, and diabetic retinopathy. *JAMA Netw Open*. 2022;5(3), e220269.
16. Gutierrez L, Lim JS, Foo LL, et al. Application of artificial intelligence in cataract management: current and future directions. *Eye Vis (Lond)*. 2022;9(1):3.
17. Guymer RH, Campbell TG. Age-related macular degeneration. *Lancet*. 2023;401 (10386):1459–1472.
18. Habli I, Lawton T, Porter Z. Artificial intelligence in health care: accountability and safety. *Bull World Health Organ*. 2020;98(4):251–256.
19. Hamet P, Tremblay J. Artificial intelligence in medicine. *Metabolism*. 2017, 69s:S36-s40.
20. Hartnett ME, Penn JS. Mechanisms and management of retinopathy of prematurity. *N Engl J Med*. 2012;367(26):2515–2526.
21. He J, Baxter SL, Xu J, Xu J, Zhou X, Zhang K. The practical implementation of artificial intelligence technologies in medicine. *Nat Med*. 2019;25(1):30–36.
22. Hellström A, Smith LE, Dammann O. Retinopathy of prematurity. *Lancet*. 2013;382 (9902):1445–1457.
23. Hood DC, La Bruna S, Tsamis E, et al. Detecting glaucoma with only OCT: implications for the clinic, research, screening, and AI development. *Prog Retin Eye Res*. 2022;90, 101052.
24. Huang XM, Yang BF, Zheng WL, et al. Cost-effectiveness of artificial intelligence screening for diabetic retinopathy in rural China. *BMC Health Serv Res*. 2022;22(1): 260.
25. Husereau D, Drummond M, Augustovski F, et al. Consolidated health economic evaluation reporting standards 2022 (CHEERS 2022) statement: updated reporting guidance for health economic evaluations. *Value Health*. 2022;25(1):3–9.
26. Ipp E, Liljenquist D, Bode B, et al. Pivotal evaluation of an artificial intelligence system for autonomous detection of referrable and vision-threatening diabetic retinopathy. *JAMA Netw Open*. 2021;4(11), e2134254.
27. Jayaram H, Kolko M, Friedman DS, Gazzard G. Glaucoma: now and beyond. *Lancet*. 2023;402(10414):1788–1801.
28. Jin K, Ye J. Artificial intelligence and deep learning in ophthalmology: current status and future perspectives. *Adv Ophthalmol Pract Res*. 2022;2(3).
29. Kaliki S, Vempuluru VS, Ghose N, Patil G, Viriyala R, Dhara KK. Artificial intelligence and machine learning in ocular oncology: retinoblastoma. *Indian J Ophthalmol*. 2023;71(2):424–430.
30. LeCun Y, Bengio Y, Hinton G. Deep learning. *nature*. 2015;521(7553):436–444.
31. Li JO, Liu H, Ting DSJ, et al. Digital technology, tele-medicine and artificial intelligence in ophthalmology: a global perspective. *Prog Retin Eye Res*. 2021;82, 100900.
32. Li Z, Guo C, Lin D, et al. Deep learning for automated glaucomatous optic neuropathy detection from ultra-widefield fundus images. *Br J Ophthalmol*. 2021; 105(11):1548–1554.
33. Lin S, Ma Y, Xu Y, et al. Artificial intelligence in community-based diabetic retinopathy telemedicine screening in Urban China: cost-effectiveness and cost-utility analyses with real-world data. *JMIR Public Health Surveill*. 2023;9, e41624.
34. Liu H, Li L, Wormstone IM, et al. Development and validation of a deep learning system to detect glaucomatous optic neuropathy using fundus photographs. *JAMA Ophthalmol*. 2019;137(12):1353–1360.
35. Liu H, Li R, Zhang Y, et al. Economic evaluation of combined population-based screening for multiple blindness-causing eye diseases in China: a cost-effectiveness analysis. *Lancet Glob Health*. 2023;11(3):e456–e465.
36. Maassen O, Fritsch S, Palm J, et al. Future medical artificial intelligence application requirements and expectations of physicians in german university hospitals: web-based survey. *J Med Internet Res*. 2021;23(3), e26646.
37. Marques AP, Ramke J, Cairns J, et al. The economics of vision impairment and its leading causes: a systematic review. *EClinicalMedicine*. 2022;46, 101354.
38. Morrison SL, Dukhovny D, Chan RVP, Chiang MF, Campbell JP. Cost-effectiveness of Artificial intelligence-based retinopathy of prematurity screening. *JAMA Ophthalmol*. 2022;140(4):401–409.
39. Muñoz B, West SK, Rubin GS, et al. Causes of blindness and visual impairment in a population of older Americans: the salisbury eye evaluation study. *Arch Ophthalmol*. 2000;118(6):819–825.
40. Nagendran M, Chen Y, Lovejoy CA, et al. Artificial intelligence versus clinicians: systematic review of design, reporting standards, and claims of deep learning studies. *Bmj*. 2020;368:m689.
41. Page MJ, Moher D, Bossuyt PM, et al. PRISMA 2020 explanation and elaboration: updated guidance and exemplars for reporting systematic reviews. *Bmj*. 2021;372: n160.
42. Pascolini D, Mariotti SP. Global estimates of visual impairment: 2010. *Br J Ophthalmol*. 2012;96(5):614–618.
43. Robinson R. Cost-effectiveness analysis. *Bmj*. 1993;307(6907):793–795.
44. Rudmik L, Drummond M. Health economic evaluation: important principles and methodology. *Laryngoscope*. 2013;123(6):1341–1347.
45. Sabri K, Ells AL, Lee EY, Dutta S, Vinekar A. Retinopathy of prematurity: a global perspective and recent developments. *Pediatrics*. 2022;150(3).

46. Scotland GS, McNamee P, Philip S, et al. Cost-effectiveness of implementing automated grading within the national screening programme for diabetic retinopathy in Scotland. *Br J Ophthalmol*. 2007;91(11):1518–1523.
47. Srisubat A, Kittrongsiri K, Sangroongruangsri S, et al. Cost-utility analysis of deep learning and trained human graders for diabetic retinopathy screening in a nationwide program. *Ophthalmol Ther*. 2023;12(2):1339–1357.
48. Stahl A. The diagnosis and treatment of age-related macular degeneration. *Dtsch Arztebl Int*. 2020;117(29-30).
49. Steinmetz JD, Bourne RR, Briant PS, et al. 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*. 2021;9(2):e144–e160.
50. Tamura H, Akune Y, Hiratsuka Y, et al. Real-world effectiveness of screening programs for age-related macular degeneration: amended Japanese specific health checkups and augmented screening programs with OCT or AI. *Jpn J Ophthalmol*. 2022;66(1):19–32.
51. Tan TE, Anees A, Chen C, 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(5): e317–e329.
52. Ting DSW, Pasquale LR, Peng L, et al. Artificial intelligence and deep learning in ophthalmology. *Br J Ophthalmol*. 2019;103(2):167–175.
53. Tsiknakis N, Theodoropoulos D, Manikis G, et al. Deep learning for diabetic retinopathy detection and classification based on fundus images: a review. *Comput Biol Med*. 2021;135, 104599.
54. Tufail A, Rudisill C, Egan C, et al. Automated diabetic retinopathy image assessment software: diagnostic accuracy and cost-effectiveness compared with human graders. *Ophthalmology*. 2017;124(3):343–351.
55. Vujosevic S, Aldington SJ, Silva P, et al. Screening for diabetic retinopathy: new perspectives and challenges. *Lancet Diabetes Endocrinol*. 2020;8(4):337–347.
56. Wang Y, Shi D, Tan Z, et al. Screening referable diabetic retinopathy using a semi-automated deep learning algorithm assisted approach. *Front Med (Lausanne)*. 2021; 8, 740987.
57. Wolf RM, Channa R, Abramoff MD, Lehmann HP. Cost-effectiveness of autonomous point-of-care diabetic retinopathy screening for pediatric patients with diabetes. *JAMA Ophthalmol*. 2020;138(10):1063–1069.
58. Wong WL, Su X, Li X, et al. Global prevalence of age-related macular degeneration and disease burden projection for 2020 and 2040: a systematic review and meta-analysis. *Lancet Glob Health*. 2014;2(2):e106–e116.
59. Xiao X, Xue L, Ye L, Li H, He Y. Health care cost and benefits of artificial intelligence-assisted population-based glaucoma screening for the elderly in remote areas of China: a cost-offset analysis. *BMC Public Health*. 2021;21(1):1065.
60. Xie Y, Nguyen QD, Hamzah H, et al. Artificial intelligence for teleophthalmology-based diabetic retinopathy screening in a national programme: an economic analysis modelling study. *Lancet Digit Health*. 2020;2(5):e240–e249.
61. Xu T, Wang B, Liu H, et al. Prevalence and causes of vision loss in China from 1990 to 2019: findings from the Global Burden of Disease Study 2019. *Lancet Public Health*. 2020;5(12):e682–e691.
62. Zhang R., Dong L., Li R., Zhang K., Li Y., Zhao H., et al. Automatic retinoblastoma screening and surveillance using deep learning. Br J Cancer. 2023.