Efficacy of artificial intelligence in detecting diabetic retinopathy from retinal fundus images — a systematic review

Gabrielius Dailidė[1], Saulius Lagunavičius[1], Vilma Jūratė Balčiūnienė[2]

[1]Lithuanian University of Health Sciences, Kaunas, Lithuania
[2]Department of Ophthalmology, Lithuanian University of Health Sciences, Kaunas, Lithuania

ABSTRACT

BACKGROUND: Diabetic retinopathy (DR) is a microvascular disorder that damages the retina's blood vessels. This review aims to evaluate scientific literature about the efficacy of artificial intelligence (AI) in detecting diabetic retinopathy from retinal fundus images.

MATERIAL AND METHODS: Systematic literature review was carried out following preferred reporting items for systematic reviews and meta-analyses (PRISMA) guidelines. Selected Medical Subject Headings (MeSH) keywords were used to search "PubMed" and "ScienceDirect" databases. Inclusion and exclusion criteria were developed using the patient, intervention, comparison, outcome (PICO) tool. Quality assessment of selected studies was done using a modified seven-item checklist based on the methodological index for nonrandomized studies (MINORS) criteria.

RESULTS: 15 studies from 14 scientific publications were included in this systematic review. AI algorithms analyzed a total of 150179 images. The AI-based algorithm's average sensitivity (Se) was 92.58 %, ranging from 76.2% to 100%. The average specificity (Sp) was 87.22%, with the lowest of 53.16% and the highest of 98.5%. The average area under the receiver operating characteristic (AUROC) curve was 0.937, with a variation of 0.843 to 0.9905.

CONCLUSION: Our results show that AI-based algorithms can accurately detect DR in retinal fundus images. These systems should be considered of use in clinical practice to save time and reduce costs.

KEY WORDS: ophthalmology; diabetic retinopathy; artificial intelligence; machine learning; neural networks; deep learning

Ophthalmol J 2022; Vol. 7, 144–151

INTRODUCTION

Diabetic retinopathy (DR) is a microvascular disorder occurring due to the long-term effects of diabetes that can lead to vision-threatening damage of the retina [1]. According to the Global Burden of Disease study, DR is the leading cause of blindness in the world for those of working age [2]. Screening coupled with timely referral and treatment is a universally accepted strategy for blindness prevention [3]. Although there are several methods to diagnose DR, the most widely used one is slit lamp biomicroscopy — fundoscopy.

CORRESPONDING AUTHOR:
Gabrielius Dailidė, Neries krantinė 32–40, Kaunas, Lithuania, tel: +37067378295; e-mail: gabrieliusdail@gmail.com

The global prevalence of diabetes has nearly doubled in the last three decades, and 35 % of patients suffer from some form of DR. This makes it the most common complication of the disease [4, 5]. Epidemiologic studies show that 11% of the world's population is currently over 60 years of age, projected to double by 2050 [6]. Due to the rapidly aging population, the healthcare system is anticipated to struggle with timely diagnosis and treatment of patients. In fact, DR is the only cause of blindness with increasing prevalence over the past three decades, suggesting that ophthalmologists have already failed to meet the growing need [2]. Therefore, new methods to help streamline the process are being explored. One of the approaches that are starting to be established and integrated into global health care systems is artificial intelligence (AI).

AI is often used as an umbrella term to cover a computer science branch capable of analyzing complex data [7]. A convolutional neural network (CNN) is an AI-based system, most used to analyze visual images. It utilizes a set of algorithms that take an input image, evaluate the importance of various objects present, and classify it. The field of ophthalmology is well suited for AI studies due to numerous digital techniques widely used in clinical practice, such as color fundus photography, optical coherence tomography (OCT), and computerized visual field testing. Collected data is stored in publicly available databases and is being used to develop AI algorithms capable of doing image recognition diagnostic tasks [8]. In recent years many studies about the detection and screening of DR using different algorithms have been published.

This review aims to evaluate the efficacy of different AI algorithms used by researchers in detecting DR from retinal fundus images.

MATERIAL AND METHODS

This systematic literature review was carried out in accordance with Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines [9].

Information sources and search strategy

The literature search was performed using "PubMed" and "ScienceDirect" databases. A manual review of references within the included articles was carried out to ensure all relevant studies were assessed. Keywords and combinations of possible synonyms used for the search were selected using the Medical Subject Headings dictionary (MeSH). The terms used were "diabetic retinopathy" combined with "artificial intelligence", "deep neural network", "deep learning", "machine learning", and "neural networks". The databases were last searched in February 2022.

Study selection and eligibility criteria

Study selection criteria were developed using the PICO tool. PICO is an acronym for Patient, Intervention, Comparison, and Outcome. It is used to conceptualize the main research question and selection criteria by identifying an essential issue, intervention, and outcome for a diagnostic method. By clearly defining these terms, researchers can systematically examine potentially relevant studies according to their stated selection criteria [10]. This review aimed to analyze articles describing the efficacy of AI models detecting DR from retinal fundus images. Studies had to be written in English, published in the last 10 years, and have open access to the full text. Research papers using either a publicly available or their own dataset were included. The performance of AI-based algorithms had to be compared with that of an ophthalmologist. Outcome measures used to evaluate the performance of AI had to include sensitivity, specificity, and area under the receiver operating characteristic curve (AUROC). Inclusion and exclusion criteria are described in Table 1.

Selection process and data collection process

After removing the duplicates, two reviewers independently screened the titles and abstracts of the remaining studies. Studies that failed to meet the inclusion criteria were excluded, and full-text articles of relevant studies were assessed and evaluated, conforming with the eligibility criteria. Mutual consensus on disagreements was reached after a discussion or a consultation with a third reviewer. The PRISMA flow diagram displays specific information about the selection process.

Data extraction and outcome measures

Two reviewers extracted data into an electronic database using a standardized form. Collected data include the following:
1. Title, author, and year of publication.
2. Study design.

Table 1. Inclusion and exclusion criteria	
Inclusion criteria	**Exclusion criteria**
Publications written in English with open access to full text	Literature reviews and case reports
Published in the last 10 years	
Used AI to detect DR	
Diagnosis was made solely based on retinal images captured by fundus photography	
Index test was compared with ophthalmologist diagnosis	
Efficacy of method evaluated by sensitivity, specificity and AUROC	

AI — artificial intelligence; DR — diabetic retinopathy; AUROC — area under the receiver operating characteristic curve

3. A number of photos analyzed, and the dataset used.
4. Outcomes that include diagnostic accuracy measures of AI (sensitivity, specificity, and area under the receiver operating characteristic (AUROC) curve).

Sensitivity and specificity are the two fundamental performance measures of a diagnostic test. They describe the actual number of positive and negative cases, respectively in a dataset. The more accurately a true positive and a true negative are correctly identified, the higher sensitivity and specificity are achieved [11]. However, there is significant variation in diagnostic confidence when making decisions based on image findings. One of the most popular measures that can more accurately convey the overall performance of a diagnostic test is the area under the receiver operating characteristic (AUROC) curve [12]. AUROC curve represents the selected method's ability to correctly distinguish whether a condition is present or not. It is the combined measure of sensitivity and specificity and can range between 0 and 1. The better an AI-based algorithm can correctly identify true positives and negatives, the higher the AUROC curve will be. An AUROC of 0.5 implies the prediction is no better than chance, while an AUROC of 1 implies perfect prediction [13]. Evaluating these three outcome measures allows us to objectively compare the efficacy of different diagnostic tests.

Quality assessment

Two independent reviewers used a modified seven-item checklist based on the methodological index for nonrandomized studies (MINORS) criteria to assess the quality of selected studies. The criteria included disclosure, study aim, input features, determination of ground truth labels, dataset distribution, performance metric, and explanation of the used AI model. Modified MINORS criteria have already been used in previously conducted systematic reviews for diagnostic studies [14].

RESULTS
Review of literature

The systematic search resulted in a total of 2322 scientific publications collected. 1363 potentially relevant studies were left after the elimination of duplicates. A brief analysis of titles and abstracts was conducted, after which 168 studies remained. Full texts of the remaining articles were further examined for final inclusion, and 15 studies from 14 scientific publications were included in this systematic review. Reasons for exclusion are described in the PRISMA flow chart (Fig. 1).

Study characteristics

Pei et al. and Ming et al. used EyeWisdom system [15,16], Zhang et al. and Li et al. used the Inception V3 system [17,18], Arenas-Cavalli et al. used the DART system [19], Roychowdhury et al. used DREAM system [20], Wang et al. used both DeepDR and Lesion-Net in different reports [21, 22]. In contrast, others didn't specify the market name for their systems. Eight of the studies included used CNN to detect DR [19, 21–27], while others did not specify or used another kind of architecture for their AI systems. Part of the studies was done using publicly available datasets, while others collected their own data. Eight studies used images from public datasets (Kaggle, MESSIDOR, MESSIDOR-2, E-Ophtha, DIARETDB1, EyePACS) [17, 20, 24–28], five studies used images from local hospitals [15, 19, 21, 23] and three studies used a combination of images gathered locally and from the datasets [16, 18, 22]. The study sample size varied from 321 to 75137 images, with a total sam-
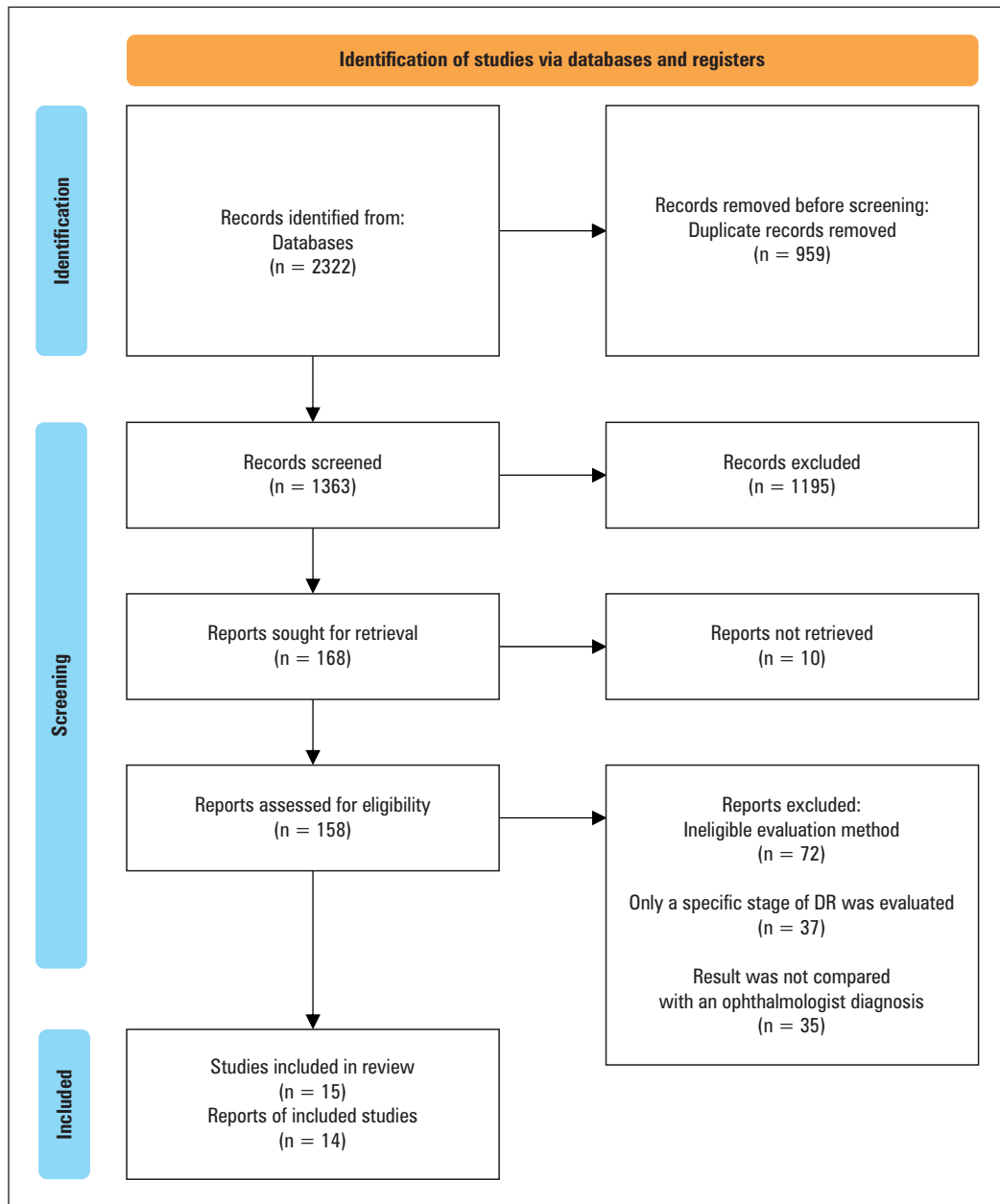
FIGURE 1. PRISMA flow diagram

ple size of 150179 images and an average of 9895 images per study. The full characteristics of included studies are presented in Table 2.

Results of individual studies

The average sensitivity of the AI-based algorithm was 92.58 % (95% CI: 89.40% to 95.70%), highest Se of 100% was achieved by Roychowdhury et al. [20] using the DREAM system, with an AUROC of 0,904 and the single lowest Sp of all the studies included — 53.2%. The system with the lowest Se of 76.2% was EyeWisdom MCS used by Pei

et al. [15]. It achieved Sp of 92.4% and AUROC of 0,843, which was the lowest amongst all studies. Mean Sp was 87.22% (95% CI: 80.36% to 94.10%), highest Sp was 98.5%, achieved by He et al. [23] using CNN with Se of 90.79% and AUROC of 0.946. Average AUROC was 0.937 (95% CI: 0.913–0.960), highest AUROC of 0.991 was achieved by Li et al. [18], with Se of 96.93% and Sp of 93.45%. Results suggest that it was harder for researchers to achieve high specificity than high sensitivity. The full results of included studies are presented in Table 2.

Table 2. Characteristics and results of included studies

Authors	Name of the system used	Study sample	Source of images	Sensitivity, %	Specificity, %	AUROC
Arenas-Cavali et al. 2022 [19]	DART	1123	Locally gathered	94.6	74.3	0.915
He et al. 2020 [23]	Not specified	3556	Locally gathered	90.8	98.5	0.946
Zhang et al. 2021 [17]	Inception V3	7025	Kaggle dataset	92.5	90.7	0.968
Gargeya et al. 2017 [24]	Not specified	75137	MESSIDOR-2, E-Ophtha datasets	94.0	98.0	0.97
Li et al. 2019 [18]	Inception-v3 (deep transfer learning)	19233	Locally gathered	96.9	93.5	0.991
Wang et al. 2020 [21]	DeepDR, CNN	6788	Locally gathered	93.5	77.1	0.93
Wang et al. 2021 [22]	Lesion Net, CNN	12252	Kaggle dataset and locally gathered	90.5	78.5	0.938
Roychowdhury et al. 2014 [20]	DREAM	1200	DIARETDB1, MESSIDOR datasets	100.0	53.2	0.904
Cao et al. 2018 [28]	Not specified	1200	MESSIDOR dataset	92.4	91.5	0.939
Ming et al. 2021 [16]	EyeWisdom	321	Locally gathered	90.0	96.6	0.933
Pei et al. 2022 [15]	EyeWisdomDSS	1768	Locally gathered	91.0	81.3	0.862
Pei et al. 2022 [15]	EyeWisdomMCS	1768	Locally gathered	76.2	92.4	0.843
Saxena et al. 2020 [25]	Not specified	1200	EyePACS, MESSIDOR, MESSIDOR-2 datasets	88.8	89.9	0.958
Baget-Bernaldiz et al. 2021 [26]	Not specified	14327	MESSIDOR dataset	97.3	94.6	0.968
Shah et al. 2020 [27]	Not specified	1533	MESSIDOR dataset	99.7	98.5	0.991

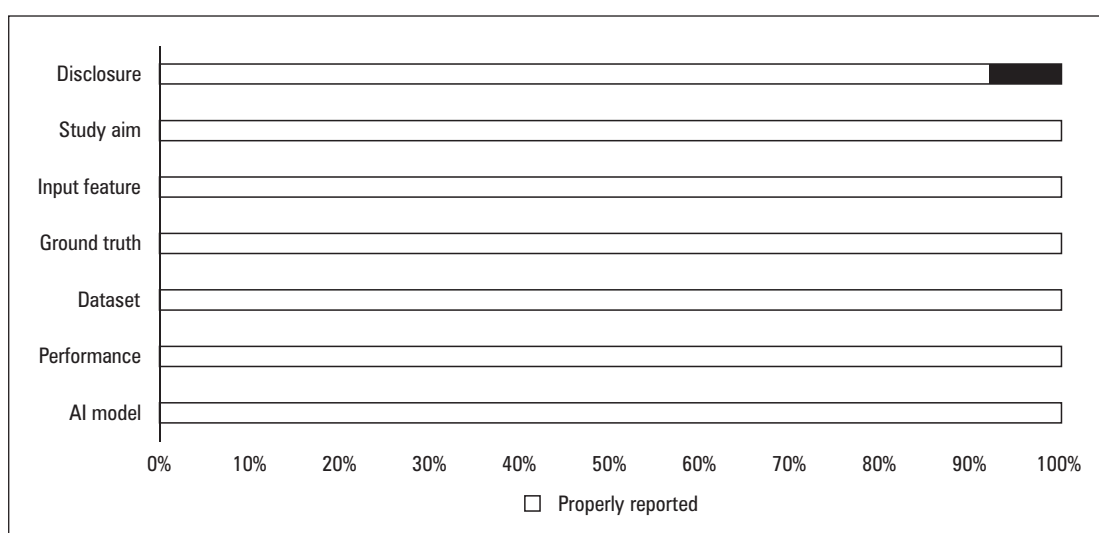AUROC — area under the receiver operating characteristic curve

FIGURE 2. Quality assessment using modified MINORS checklist. AI — artificial intelligence

Quality assessment

An overview of quality assessment is provided in Figure 2. All the evaluated studies had a clearly defined study aim. Articles using locally gathered datasets described patients' exclusion and inclusion criteria in detail, while studies using publicly available datasets had a comprehensive description of the visual image selection procedure. Each study had explicitly stated that the index test was compared with ophthalmologist diagnosis as a ground truth. All the studies described how their AI-based algorithm was trained, validated, and tested using specific datasets. 100% of studies explained how the diagnostic performance of their AI model was evaluated and provided outcome measures of Se, Sp, and AUROC. Six articles didn't specify the market name of the AI system they used. However, every study explained in depth how their AI model works.

DISCUSSION

Screening and diagnosing DR is one of the most widely explored fields of AI research in ophthalmology, with more and more studies reporting new techniques being published every year [29]. It is important to systematically review and evaluate the emerging studies to identify the direction of work needed for full integration into our healthcare systems. The UK National Institute for Clinical Excellence (NICE) guidelines states that a DR screening test should have sensitivity and specificity of at least 80% and 95%, respectively, while the specificity of 0.5 was defined by another commission of specialists and has been shown to be cost-effective [30,31]. Out of all studies included in this review, He et al., Gargeya et al., Ming et al., and Shah et al. exceeded the requirements. At the same time, Baget-Bernaldiz et al. and Li et al. came close with sensitivity and specificity, reaching 97.3%/94.6% and 96.9%/93.5%, respectively. Furthermore, all of the studies included in this review proved to be cost-effective.

To reduce preventable vision loss globally, early detection and timely treatment of DR are needed. An efficient DR screening system must be established [32]. Reviewed studies show that AI-based, fully automated systems can operate with high accuracy while assessing retinal fundus images and potentially reduce the manual workload needed. Moreover, collected outcome measures display that these systems have made a significant advance from the early days and perform on par with trained professionals. It is expected that as technology advances AI will be able to autonomously identify patients with preventable vision loss and refer them to clinicians. This is particularly important in the developing world, where healthcare systems are chronically underfunded, and trained professionals are difficult to reach [33]. Some places have already started to integrate AI into the healthcare system. For example, in 2018 US Food and Drug Administration (FDA) approved the AI-based algorithm IDX-DR for DR detection, making it the world's first AI medical equipment for disease diagnosis [32].

Due to the lack of standardized testing procedures, comparing the results of different AI-based algorithms can pose various challenges. Most of the studies included in this review use different algorithms to analyze different datasets, which makes direct comparisons less accurate. Moreover, the sample sizes of the studies vary drastically. For example, Ming et al. used retinal fundus images collected from local hospitals with a total sample size of 321 images, while Gargeya et al. used multiple publicly available datasets with a total sample size of 75137 retinal fundus images. Such a big difference in sample size might distort the results and negatively impact the reliability of the results. Additionally, none of the publications disclosed whether their algorithms were trained on DR images only. AI systems benefit from constant exposure to a variety of images, while algorithms exposed only to DR samples tend to have lower sensitivity and specificity over time [34].

Even though AI has already achieved significant results in DR detection, there is still room for improvement. AI-based systems still struggle to correctly identify diseases based solely on fundus photos, such as macular edema [35]. It is believed that the ability to evaluate multiple types of images could boost the diagnostic performance of AI [36]. Some AI-based algorithms currently in development use additional data besides images to make a diagnosis. EyeWisdom system used by Pei et al. and Ming et al. can screen for nearly 20 different eye diseases, such as DR, glaucoma, and age-related macular degeneration, based on the fundus photographs and disease history [15, 16]. These examples show that adding additional data inputs could increase the scope and efficacy of AI technology.

An additional topic of discussion is the potential collaboration of AI and human professionals. There are some growing worries that advancing

AI could make current-age physicians redundant. However, studies show that a combination of human specialists and AI technology provide the best quality of care compared to either acting individually. AI-based algorithms can meticulously analyze provided data and make purely objective decisions, while human specialists can see the bigger picture that allows them to make nuanced decisions. This combination of different strengths ensures the best possible care is being provided [37].

CONCLUSION

Our results show that AI-based algorithms can accurately detect DR in retinal fundus images. AI technology can quicken the screening process and reduce the cost of care for a constantly growing population of patients. Some studies included in the review already meet the required criteria to be successfully used in clinical practice. However, due to various AI algorithms and datasets, further research and a more standardized approach are needed to accurately compare the efficacy of AI technology.

Conflict of interests

The authors report no competing interests.

REFERENCES

1. Shukla U, Tripathy K. Diabetic Retinopathy. StatPearls, Treasure Island 2022.
2. Bourne RR, Steinmetz JD, Saylan M, 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., doi: 10.1016/S2214-109X(20)30489-7, indexed in Pubmed: 33275949.
3. Ting DS, Pasquale LR, Peng L, et al. Artificial intelligence and deep learning in ophthalmology. Br J Ophthalmol. 2019; 103(2): 167–175, doi: 10.1136/bjophthalmol-2018-313173, indexed in Pubmed: 30361278.
4. WHO Global Report on Diabetes. Global Report on Diabetes.2016;978:6–86. https://sci-hub.si/https://apps.who.int/iris/handle/10665/204874%0Ahttps://apps.who.int/iris/bit-stream/handle/10665/204874/WHO_NMH_NVI_16.3_eng.pdf?sequence=1%0Ahttp://www.who.int/about/licensing/copyright_form/index.html%0Ahttp://www.who.int/about/licens.
5. Yau JWY, Rogers SL, Kawasaki R, et al. Meta-Analysis for Eye Disease (META-EYE) Study Group. Global prevalence and major risk factors of diabetic retinopathy. Diabetes Care. 2012; 35(3): 556–564, doi: 10.2337/dc11-1909, indexed in Pubmed: 22301125.
6. Kanasi E, Ayilavarapu S, Jones J. The aging population: demographics and the biology of aging. Periodontol 2000. 2016; 72(1): 13–18, doi: 10.1111/prd.12126, indexed in Pubmed: 27501488.
7. Ramesh AN, Kambhampati C, Monson JRT, et al. Artificial intelligence in medicine. Ann R Coll Surg Engl. 2004; 86(5): 334–338, doi: 10.1308/147870804290, indexed in Pubmed: 15333167.
8. Keskinbora K, Güven F. Artificial Intelligence and Ophthalmology. Turk J Ophthalmol. 2020; 50(1): 37–43, doi: 10.4274/tjo.galenos.2020.78989, indexed in Pubmed: 32167262.
9. Page MJ, McKenzie JE, Bossuyt PM, et al. The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. PLoS Med. 2021; 18(3): e1003583–189, doi: 10.1371/journal.pmed.1003583, indexed in Pubmed: 33780438.
10. Eriksen MB, Frandsen TF. The impact of patient, intervention, comparison, outcome (PICO) as a search strategy tool on literature search quality: a systematic review. J Med Libr Assoc. 2018; 106(4): 420–431, doi: 10.5195/jmla.2018.345, indexed in Pubmed: 30271283.
11. Park SHo, Goo JMo, Jo CH. Receiver operating characteristic (ROC) curve: practical review for radiologists. Korean J Radiol. 2004; 5(1): 11–18, doi: 10.3348/kjr.2004.5.1.11, indexed in Pubmed: 15064554.
12. Obuchowski NA. Receiver operating characteristic curves and their use in radiology. Radiology. 2003; 229(1): 3–8, doi: 10.1148/radiol.2291010898, indexed in Pubmed: 14519861.
13. Janssens AC, Martens FK. Reflection on modern methods: Revisiting the area under the ROC Curve. Int J Epidemiol. 2020; 49(4): 1397–1403, doi: 10.1093/ije/dyz274, indexed in Pubmed: 31967640.
14. Langerhuizen DWG, Janssen SJ, Mallee WH, et al. What Are the Applications and Limitations of Artificial Intelligence for Fracture Detection and Classification in Orthopaedic Trauma Imaging? A Systematic Review. Clin Orthop Relat Res. 2019; 477(11): 2482–2491, doi: 10.1097/CORR.0000000000000848, indexed in Pubmed: 31283727.
15. Pei X, Yao Xi, Yang Y, et al. Efficacy of artificial intelligence-based screening for diabetic retinopathy in type 2 diabetes mellitus patients. Diabetes Res Clin Pract. 2022; 184: 109190, doi: 10.1016/j.diabres.2022.109190, indexed in Pubmed: 35031348.
16. Ming S, Xie K, Lei X, et al. Evaluation of a novel artificial intelligence-based screening system for diabetic retinopathy in community of China: a real-world study. Int Ophthalmol. 2021; 41(4): 1291–1299, doi: 10.1007/s10792-020-01685-x, indexed in Pubmed: 33389425.
17. Zhang X, Li F, Wei Q, et al. Automated detection of severe diabetic retinopathy using deep learning method. Graefes Arch Clin Exp Ophthalmol. 2022; 260(3): 849–856, doi: 10.1007/s00417-021-05402-x, indexed in Pubmed: 34591173.
18. Li F, Liu Z, Chen H, et al. Automatic Detection of Diabetic Retinopathy in Retinal Fundus Photographs Based on Deep Learning Algorithm. Transl Vis Sci Technol. 2019; 8(6): 4, doi: 10.1167/tvst.8.6.4, indexed in Pubmed: 31737428.
19. Arenas-Cavalli JT, Abarca I, Rojas-Contreras M, et al. Clinical validation of an artificial intelligence-based diabetic retinopathy screening tool for a national health system. Eye (Lond). 2022; 36(1): 78–85, doi: 10.1038/s41433-020-01366-0, indexed in Pubmed: 33432168.
20. Roychowdhury S, Koozekanani DD, Parhi KK. DREAM: diabetic retinopathy analysis using machine learning. IEEE J Biomed Health Inform. 2014; 18(5): 1717–1728, doi: 10.1109/JBHI.2013.2294635, indexed in Pubmed: 25192577.
21. Wang XN, Dai L, Li ST, et al. Automatic Grading System for Diabetic Retinopathy Diagnosis Using Deep Learning Artificial Intelligence Software. Curr Eye Res. 2020; 45(12): 1550–1555, doi: 10.1080/02713683.2020.1764975, indexed in Pubmed: 32410471.
22. Wang Y, Yu M, Hu B, et al. Deep learning-based detection and stage grading for optimising diagnosis of diabetic retinopathy. Diabetes Metab Res Rev. 2021; 37(4): e3445, doi: 10.1002/dmrr.3445, indexed in Pubmed: 33713564.
23. He J, Cao T, Xu F, et al. Artificial intelligence-based screening for diabetic retinopathy at community hospital. Eye (Lond). 2020; 34(3): 572–576, doi: 10.1038/s41433-019-0562-4, indexed in Pubmed: 31455902.
24. Gargeya R, Leng T. Automated Identification of Diabetic Retinopathy Using Deep Learning. Ophthalmology. 2017; 124(7): 962–969, doi: 10.1016/j.ophtha.2017.02.008, indexed in Pubmed: 28359545.
25. Saxena G, Verma D, Paraye A, et al. Improved and robust deep learning agent for preliminary detection of diabetic retinopathy using public datasets. Intelligence-Based Med. 2020; 3-4: 100022, doi: 10.1016/j.ibmed.2020.100022.
26. Baget-Bernaldiz M, Pedro RA, Santos-Blanco E, et al. Testing a Deep Learning Algorithm for Detection of Diabetic Retinopathy in

a Spanish Diabetic Population and with MESSIDOR Database. Diagnostics (Basel). 2021; 11(8), doi: 10.3390/diagnostics11081385, indexed in Pubmed: 34441319.

27. Shah P, Mishra DK, Shanmugam MP, et al. Validation of Deep Convolutional Neural Network-based algorithm for detection of diabetic retinopathy - Artificial intelligence versus clinician for screening. Indian J Ophthalmol. 2020; 68(2): 398–405, doi: 10.4103/ijo.IJO_966_19, indexed in Pubmed: 31957737.

28. Cao P, Ren F, Wan C, et al. Efficient multi-kernel multi-instance learning using weakly supervised and imbalanced data for diabetic retinopathy diagnosis. Comput Med Imaging Graph. 2018; 69: 112–124, doi: 10.1016/j.compmedimag.2018.08.008, indexed in Pubmed: 30237145.

29. Zhang X, Li F, Li D, et al. Automated detection of severe diabetic retinopathy using deep learning method. Graefes Arch Clin Exp Ophthalmol. 2022; 260(3): 849–856, doi: 10.1007/s00417-021-05402-x, indexed in Pubmed: 34591173.

30. 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, doi: 10.1016/j.ophtha.2016.11.014, indexed in Pubmed: 28024825.

31. Wong TY, Bressler NM. Artificial Intelligence With Deep Learning Technology Looks Into Diabetic Retinopathy Screening. JAMA. 2016; 316(22): 2366–2367, doi: 10.1001/jama.2016.17563, indexed in Pubmed: 27898977.

32. Goh JK, Cheung CY, Sim SS, et al. Retinal Imaging Techniques for Diabetic Retinopathy Screening. J Diabetes Sci Technol. 2016; 10(2): 282–294, doi: 10.1177/1932296816629491, indexed in Pubmed: 26830491.

33. Sommer A, Taylor HR, Ravilla TD, et al. Council of the American Ophthalmological Society. Challenges of ophthalmic care in the developing world. JAMA Ophthalmol. 2014; 132(5): 640–644, doi: 10.1001/jamaophthalmol.2014.84, indexed in Pubmed: 24604415.

34. Ting DS, Pasquale LR, Peng L, et al. Artificial intelligence and deep learning in ophthalmology. Br J Ophthalmol. 2019; 103(2): 167–175, doi: 10.1136/bjophthalmol-2018-313173, indexed in Pubmed: 30361278.

35. He J, Cao T, Xu F, et al. Artificial intelligence-based screening for diabetic retinopathy at community hospital. Eye (Lond). 2020; 34(3): 572–576, doi: 10.1038/s41433-019-0562-4, indexed in Pubmed: 31455902.

36. Hassan B, Hassan T, Li Bo, et al. Deep Ensemble Learning Based Objective Grading of Macular Edema by Extracting Clinically Significant Findings from Fused Retinal Imaging Modalities. Sensors (Basel). 2019; 19(13), doi: 10.3390/s19132970, indexed in Pubmed: 31284442.

37. Bellemo V, Lim G, Rim TH, et al. Artificial Intelligence Screening for Diabetic Retinopathy: the Real-World Emerging Application. Curr Diab Rep. 2019; 19(9): 72, doi: 10.1007/s11892-019-1189-3, indexed in Pubmed: 31367962.