Indonesian Journal of Global Health Research
Volume 7 Number 1, February 2025
e-ISSN 2715-1972; p-ISSN 2714-9749
http://jurnal.globalhealthsciencegroup.com/index.php/IJGHR

DIAGNOSTIC ACCURACY OF ARTIFICIAL INTELLIGENCE-BASED SYSTEMS FOR DETECTING DIABETIC RETINOPATHY: A SYSTEMATIC REVIEW

Nurul Hidayati[1]*, Ika Yuni Widyawati[1], Ira Suarilah[1], Trihaningsih Puji Astuti[1], Fauzi Tsanifiandi[1], Andi Safutra Suraya[1,2]
[1]Faculty of Nursing, Universitas Airlangga, Mulyorejo, Surabaya, East Java 60115, Indonesia
[2]Research Centre of Advancing Community Healthcare- REACH, Mulyorejo, Surabaya, East Java 60115, Indonesia
*hnurul597@gmail.com

ABSTRACT
Diabetic retinopathy (DR) represents a leading cause of blindness worldwide, early detection is critical to prevent vision loss. However, traditional screening methods, which rely on human experts, prove to be costly and time-consuming. The systematic review aims to assess the validity of artificial intelligence (AI) as a screening tool for detecting DR among diabetic patients. A systematic literature search was performed of the following databases: PubMed, Scopus, CINAHL, and Web of Science. The last date of our search was January 31, 2024. We included all observational studies, including cohort, case–control and cross-sectional studies and evaluated their quality using the Joanna Briggs Institute tool. We included diagnostic test accuracy studies evaluating the use of AI algorithms for DR screening in patients with diabetes. Studies were excluded if they exclusively assessed diagnostic accuracy for DR that did not use AI algorithms as a diagnostic tool and studies with incomplete or inaccessible data. Thirteen studies with sample sizes ranging from 69 to 1378 participants, reported good sensitivity of AI for detecting visually threatening DR (VTDR). The lowest sensitivity was 89.2%, and the highest was 100%. In terms of specificity, Any DR exhibited higher specificity compared to RDR and VTDR, ranging from 80.2% to 100%. The sensitivity and specificity of the Artificial Intelligence (AI)-based tools available for DR screening was considered acceptable, especially in detecting VTDR and Any DR, was regarded as good. These results implied the potential usefulness of these tools for DR screening in settings with limited resources. However, further high-quality comparative studies were deemed necessary to evaluate their effectiveness in real-world clinical settings.

Keywords: artificial intelligence; diabetic retinopathy; screening

How to cite (in APA style)
Hidayati, N., Widyawati, I. Y., Suarilah, I., Astuti, T. P., Tsanifiandi, F., & Suraya, A. S. (2024). Diagnostic Accuracy of Artificial Intelligence-Based Systems for Detecting Diabetic Retinopathy: A Systematic Review. Indonesian Journal of Global Health Research, 7(1), 21-36. https://doi.org/10.37287/ijghr.v7i1.4011.

INTRODUCTION

Globally, the leading cause of blindness in middle-aged and elderly individuals is diabetic retinopathy (DR) (Tan & Wong, 2023). DR is a prevalent microvascular complication of diabetic retinal disease, affecting approximately one in every three individuals with diabetes. If left untreated, this condition has the potential to cause irreversible blindness (Kim et al., 2021; Sasongko et al., 2017). The prevalence of DR is expected to rise with the increasing incidence of non-communicable diseasese. Despite its notable impact, a considerable number of DR cases go undiagnosed due to restricted access to healthcare facilities staffed with adequately trained personnel (Fenner et al., 2018).

Worldwide, the estimated prevalence of diabetic retinopathy among individuals with diabetes stands at 27.0%, contributing to approximately 0.4 million cases of blindness globally (Thomas et al., 2019). And the prevalence of diabetic retinopathy among type two diabetes patients was 36.3% (Zegeye et al., 2023). The World Health Organization (WHO) estimates

that by 2030, approximately 400 million people worldwide will have diabetes. Among them, an estimated 10% will be at risk of developing diabetic retinopathy, diabetic macular edema, cataracts, and glaucoma (Bourne et al., 2013; CDC, 2020). Vision-threatening diseases, unfortunately, often manifest without symptoms, underscoring the vital importance of regular retina screenings. Early diagnosis enables early treatment, which can minimize the likelihood of vision loss.

Regular screening and prompt treatment are essential for effectively managing DR and preventing vision loss (Royle et al., 2015). Early identification enables timely intervention, potentially reducing the risk of DR progression and visual loss by up to 50% within a year (Prayogo et al., 2023), emphasizes the importance of early DR examination in diabetes patients even if there are no symptoms. Unfortunately, numerous individuals, particularly those from lower socioeconomic backgrounds, do not pursue medical evaluation and intervention (Murchison et al., 2017; Tao et al., 2016). Additionally, obstacles to accessing screening may emerge due to limited transportation choices, a lack of awareness regarding the disease, and retinal photography for DR community screening remains insufficiently available in many countries, especially those with limited health financial resources (Fenner et al., 2018; Vujosevic et al., 2020). Despite the fact that the American Diabetes Association (ADA) have recommended that every diabetic patient undergo annual visual acuity and retinal examinations to prevent the occurrence of DR (Care & Suppl, 2021). Thus, limited diabetic retinopathy (DR) screening coverage results in more individuals with diabetes mellitus (DM) missing out on treatment opportunities (Burgess et al., 2013; Jayanegara & Lestari, 2023). New screening approaches that offer enhanced efficiency and accessibility are imperative to tackle the escalating prevalence of DR.

In recent times, there has been a growing interest in the use of Artificial Intelligence (AI) within ophthalmology, largely driven by Deep Learning (DL), a new breed of AI capable of handling vast datasets autonomously (Jamaaluddin & Indah, 2021), without human involvement (Nagendran et al., 2020). The integration of DL into diabetic retinopathy (DR) screening methods relies on an image-based approach. AI systems are trained to recognize a vast array of retinal images and can be programmed to assess the severity of diabetic retinopathy (DR) (Sheikh et al., 2021). Although the potential expense associated with developing Deep Learning (DL) technology, employing DL for diabetic retinopathy (DR) screening has demonstrated significant societal cost savings and enhanced health outcomes (Srisubat et al., 2023). Additionally, it is imperative to validate AI algorithms using datasets specific before real-world implementation. Several literature reviews (Lupidi et al., 2023a; Malerbi et al., 2022; Natarajan et al., 2019a) have discussed Artificial Intelligence for screening diabetic retinopathy. The systematic review aims to assess the validity of artificial intelligence (AI) as a screening tool for detecting DR among diabetic patients.

METHOD

We conducted this systematic review by following the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-analysis) guidelines (Moher et al., 2009). This study is a systematic review aimed at elucidating the validity of artificial intelligence (AI) as a screening tool for detecting diabetic retinopathy (DR) in diabetic patients based on the latest and best scientific evidence. The literature search was conducted in February 2024. The data used in this study are secondary data obtained from previous research conducted by earlier researchers and not from direct observation results. The sources of secondary data obtained are reputable international journal articles according to the predetermined theme. The literature used was obtained from the Scopus, PubMed, Web of Science, and CINAHL

databases. Specific keywords used to search for articles are guided by Medical Subject Heading (MeSH) terms using the following terms: “Diabetic Retinopathy” OR “Diabetic Retinopathies” OR “Retinopathies, Diabetic” OR “Retinopathy, Diabetic” AND “Artificial intelligence” OR “AI” OR “machine learning” OR “deep learning” OR “neural network” AND “smartphone” OR “smartphone-based” OR “mobile-based” OR “handheld” OR “iPhone” OR “mobile camera” OR “mobile” AND “diagnosis” OR “screen” OR “screening” OR “classification” OR “discriminate”. The articles included in the search were those published in English up to January 31, 2024. During the search process, we employed Boolean operators and wildcard characters precisely to focus our search and detect singular or plural forms of the same terms across all databases utilized.

Study selection and eligibility criteria

The inclusion criteria for article selection were based on the PICOS method (Population, Issue of interest, Comparison, Outcome, and Study design) (Liberati et al., 2009): 1. Patients with type 1 or type 2 diabetes mellitus aged over 18 years; 2. Studies utilizing AI algorithms as diagnostic tests; 3. Reporting sensitivity and specificity; and 4. All types of observational studies, including cohort studies, published in English. Letter to editor, commentaries, qualitative studies, abstract only, case series, case reports, reviews, discussion papers, meta-analyses, conference papers, oral presentations and articles available in abstract form only are excluded. Studies were excluded if they exclusively assessed diagnostic accuracy for DR that did not use AI algorithms as a diagnostic tool and studies with incomplete or inaccessible data.

The data collection process in this systematic review study begins with the article search. We conduct separate screening of articles, commencing with the selection of titles deemed relevant (Figure 1). Subsequently, abstract reading is conducted to filter articles deemed eligible. Finally, full-text reading is performed to identify articles that meet the predetermined inclusion and exclusion criteria within the scope of the review. We select studies independently based on the inclusion and exclusion criteria mentioned above. In cases of disagreement among the reviewers, a joint consultation with a third reviewer is conducted before making any decisions. In terms of data extraction, we collectively decide which information may be relevant based on the review's focus, clinical experience, and previously published reviews. Here, the population, intervention, comparison/exposure, outcome, and study design (PICOS) are integral to our research. Population: Our study population consists of patients with type 1 or type 2 diabetes with diabetic retinopathy (DR); Intervention: Patients with DR diagnosed using AI technology; Comparison/Exposure: The diagnostic measure was AI technology, and DR was diagnosed from fundus color images; Outcome: The validity of artificial intelligence (AI) as a screening tool for detecting DR in diabetics patients; Study design: The design of this study is observational study. Therefore, we create a table to highlight relevant data, then input data from each article into the table and discuss the possibility of inconsistencies. Finally, we re-read the full text that we will use to ensure the accuracy of relevant data.

Quality assessment

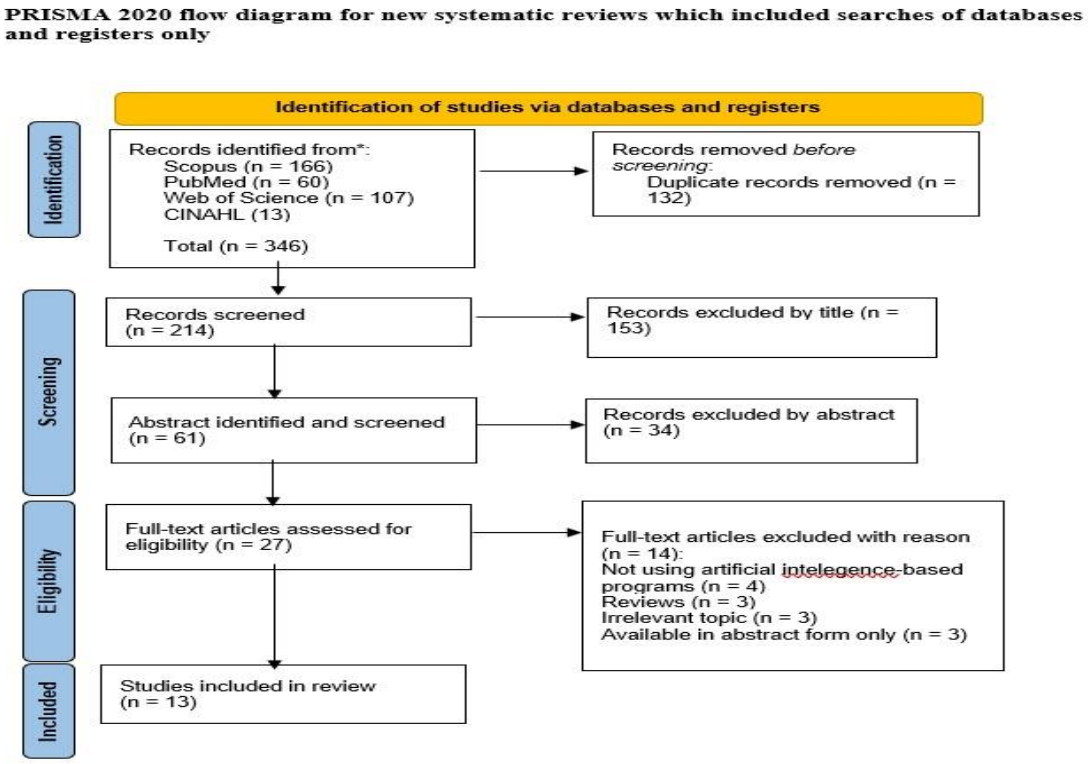
In this study, we employed the Joanna Briggs Institute tool comprising 8 questions for cross-sectional studies and 10 questions for cohort studies to evaluate both the level and quality of each article (Aromataris E, 2020; JBI, 2020). We utilized the Joanna Briggs Institute assessment tool for both cross-sectional and cohort studies to gauge the level of evidence in each reviewed source. To assess the overall methodological quality, we utilized the 11-item JBI Critical Appraisal Checklist for cohort studies and the 8-item JBI Critical Appraisal

Checklist for cross-sectional studies, updated and released in 2020. These tools classified overall quality into categories of high, moderate, and low. Each item was assigned a score of either 0 (indicating high risk of bias) or 1 (indicating low risk of bias). Quality scores were then categorized as high quality (100 points), moderate quality (74 points), or low quality (49 points).

RESULTS

Study selection

After conducting the literature review, a total of 346 articles that were potentially relevant were identified for screening. Using Mendeley software, 132 studies were removed due to duplication. Based on title and abstract screening, we excluded 187 studies. After stage two (full text) screening of the remaining 27 articles. Out of these, 14 were considered ineligible as they did not meet the specified eligibility criteria, as follows: the study was not using artificial intelegence-based program (n = 4); the study was not an original article (n = 3); the study was not relevant topic (n = 3); and the study was available in abstract form only (n = 3). Finally, 13 sources were included in our final analysis. The process employed to select these study sources is presented in Fig. 1 as a PRISMA flow diagram.



Studies characteristics

All the included articles were either cross-sectional or cohort studies, primarily focusing on diagnosing Diabetic Retinopathy (DR) in patients with or without diabetes. Studies included in this review were conducted in the United States of America (USA) (Kim et al., 2021), Italy (Lupidi et al., 2023a; Midena et al., 2022), Dominica (Kemp et al., 2023), Brazil (Malerbi et al., 2022), Croatia (Tomić et al., 2023), China (Ruan et al., 2022) and India (Krishnan et al., 2021; Natarajan et al., 2019c; Rajalakshmi et al., 2018a; Rao et al., 2022; Sosale, Aravind, et al., 2020; Sosale, Sosale, et al., 2020) with a sample size ranging from 69 to 1378 participants. Participants' mean age ranged from 53.1 to 65.5 years. Table 1 provides a summary of the methodological characteristics of the included studies. All seven papers used the Remidio NM Fundus on Phone (FOP) 10 camera (Remidio Innovative Solutions Pvt. Ltd,

Bangalore, India) (Kemp et al., 2023; Krishnan et al., 2021; Natarajan et al., 2019c; Rajalakshmi et al., 2018a; Rao et al., 2022; Sosale, Aravind, et al., 2020; Sosale, Sosale, et al., 2020) to image the retina of patients. Six out of thirteen studies used Medios AI (Kemp et al., 2023; Krishnan et al., 2021; Natarajan et al., 2019b; Rao et al., 2022; Sosale, Aravind, et al., 2020; Sosale, Sosale, et al., 2020), two used EyeArt software (Kim et al., 2021; Rajalakshmi et al., 2018b), one used PhelcomNet (Malerbi et al., 2022), one used AI algorithm Selena (Lupidi et al., 2023b), one used Phoebus algorithm (Ruan et al., 2022), one used AI-based automated software (DeepDR) (Tomić et al., 2024) and one article does not clearly specify (Midena et al., 2022). Image retina capture was performed by a trained technician and health worker included ophthalmologists, a medical student and a medical intern rather than ophthalmologists or ophthalmic photographers. The reference test or gold standard utilized was the grading of images by a retinal specialist, vitreoretinal surgeons, and ophthalmologists. Overall, ten studies utilized the International Clinical DR Disease Severity Scale to grade DR. Additionally, one study employed the Airlie House or modified ETDRS criteria (Kim et al., 2021), while another study utilized the English Grading System for DR (Kemp et al., 2023) and one study did not specify the DR measurement instrument used (Malerbi et al., 2022)

Efficacy outcomes

In this review, we assessed the sensitivity, specificity, and area under the curve (AUC) as the outcomes to evaluate the validity of AI screening for Diabetic Retinopathy (DR), as presented in Table 2. The categories for Diabetic Retinopathy (DR) are divided into three gradings: referable DR (RDR), Any DR, and visual-threatening DR (VTDR). Our findings indicated that the sensitivity for detecting any DR varied between 83,2%-97,6%, with the study by Tomic et al. (Tomić et al., 2024) reporting the lowest sensitivity value and the study by Rao et al. reporting the highest sensitivity value (Rao et al., 2022). For specificity, the lowest was 80,2% (Rajalakshmi et al., 2018b) and the highest was 100% (Tomić et al., 2024). The AUC values within this group ranged from the lowest, reported by Ruan et al. (Ruan et al., 2022) with 0,88, to the highest, reported by Tomic et al. (Tomić et al., 2024), with 0,92. The AI software's ability to detect RDR exhibited a sensitivity range of 77,5% to 100%, while its specificity in diagnosing RDR ranged from 61,4% to 92,5%. The AUC value varied among studies, with Kemp et al. (Kemp et al., 2023) reporting the lowest at 0.84, while Sosale et al. (Sosale, Sosale, et al., 2020) reported the highest at 0.92. In identifying vision-threatening diabetic retinopathy (VTDR), only four studies within this group conducted comparisons (Kemp et al., 2023; Rajalakshmi et al., 2018b; Rao et al., 2022; Sosale, Sosale, et al., 2020). In this group, the lowest sensitivity was 89,2%, while the highest sensitivity was 100%. As for specificity, only one study discusses the specificity of AI software in detecting VTDR (Rajalakshmi et al., 2018b) and out of the 13 articles, none of the studies in this group compared AUC in their research. Furthermore, there are two articles that did not mention RDR, Any DR, or STDR. These articles only assessed the sensitivity and specificity of the presence or absence of DR in patients (Kim et al., 2021; Lupidi et al., 2023b).

Table 1.
Characteristics of the reviewed study

Author/Year	Country	Study design	Sample size (patients)	Age (years), mean SD or range	AI	Retinal camera	Gold standard	Diabetic retinopathy severity scale
Kemp et al, 2023 (Kemp et al., 2023)	Dominica	Cross-sectional	587	Range 26–94 years	Medios DR AI software	A Non-Mydriatic Fundus on Phone Camera, Model FOPNM-10 (Remidio)	The senior Dominican screener–grader. The grading by the field grader was compared with remote grading by senior graders in the English National Screening Programme, and the interobserver reliability kappa coefficient was calculated.	English Grading System for DR
Kim et al., 2021 (Kim et al., 2021)	USA	Cohort	69	57.0 (15,7)	EyeArt®	A smartphone-based camera (RetinaScope) by a medical student and a medical intern rather than ophthalmologists or ophthalmic photographers.	A retina specialist and a comprehensive ophthalmologist	The severity of the DR and the modified Airlie House classification system used in the Early Treatment Diabetic Retinopathy Study (ETDRS) severity classification criteria
Krishnan et al, 2021 (Krishnan et al., 2021)	India	Cross-sectional	1378	54.9 (10.43)	Medios AI	Remidio Non-Mydriatic Fundus on Phone by health workers	Two vitreoretinal surgeons	ICDR severity scale; no DR, mild nonproliferative DR (mild NPDR), moderate nonproliferative DR (moderate NPDR), severe nonproliferative DR (severe NPDR), and proliferate DR (PDR)
Lupidi et al, 2023 (Lupidi et al., 2023a)	Italy	Cross-sectional	251	60.1 (16.2)	AI algorithm Selena	Handheld non-mydriatic digital fundus camera (Optomed Aurora IQ)	An experienced retina specialist	ICDR severity scale; as pathological and non-pathological
Marlebi et al, 2022 (Malerbi et al., 2022)	Brazil	Retrospective study	824	60.8 (11.4)	PhelcomNet	a portable retinal camera (Phelcom Eyer) by a team of nine examiners, including medical students, with variable degrees of previous experience in this kind of procedure.	A single retinal specialist	N/A for the DR severity scale used but referable DR defined as no or mild non-proliferative DR (NPDR) versus moderate NPDR, severe NPDR, proliferative DR, or apparently present macular oedema (more than mild DR, mtmDR) in at least one eye. Vision-threatening DR (vtDR) was considered

Author/Year	Country	Study design	Sample size (patients)	Age (years), mean SD or range	AI	Retinal camera	Gold standard	Diabetic retinopathy severity scale
								as the presence of severe NPDR, proliferative DR, or apparently present macular oedema in at least one eye.
Midena et al, 2022 (Midena et al., 2022)	Italy	Cross-sectional	213	62.6 (12.6)	N/A	- Portable non-mydriatic digital fundus camera (the Aurora® handheld fundus camera) by a trained technician Table-top non-mydriatic digital fundus camera (AFC-230, Nidek, Gamagori, Japan)	Retinal specialist	International Clinical Diabetic Retinopathy and Diabetic Macular Edema Severity Scale: absent, mild, moderate, severe non-proliferative (NP) and proliferative (PDR) DR, and absent, mild, moderate and severe DM.
Natarajan et al., 2019 (Natarajan et al., 2019c)	India	Cross-sectional	231	53.1 (10.3)	Medios AI	Remidio nonmydriatic retinal imaging system by a health care worker	a vitreoretinal resident and a vitreoretinal surgeon (A.J.) at the Aditya Jyot Eye Hospital and Aditya	ICDR severity scale
Rajalakshmi et al., 2018 (Rajalakshmi et al., 2018a)	India	Cross-sectional	301	N/A	EyeArt™	Remidio fundus on phone (FOP) images graded by ophthalmologists	Two ophthalmologists (retina specialists)	ICDR severity scale: STDR defined as severe NPDR, PDR, or DME, RDR defined as moderate NPDR or worse, or DME
Rao et al, 2022 (Rao et al., 2022)	India	Retrospective study	135	54.1 (8.3)	Medios AI	Kowa VX-10α Mydriatic and Remidio Non Mydriatic FOP NM10	A single retina specialist	ICDR severity scale: No DR, Mild NPDR, Moderate NPDR, Severe NPDR, PDR and DME. Referable DR (RDR) was defined as moderate NPDR or worse disease and/or the presence of DME
Ruan et al., 2022 (Ruan et al., 2022)	China	Observational study	315	65.5 (11.1)	Phoebus	Handheld fundus camera optomed aurora	Three masked and experienced ophthalmologists.	ICDR severity scale
Sosale et al, 2020 (Sosale, Sosale, et al., 2020)	India	Cross-sectional	297	55 (11)	Medios AI	Remidio Non Mydriatic Fundus on Phone Camera (NM FOP 10) by a trained technician	Retinal surgeons	ICDR severity scale; no DR, mild nonproliferative DR (mild NPDR), moderate nonproliferative DR (moderate NPDR), severe nonproliferative DR (severe NPDR), and proliferate DR (PDR)

Author/Year	Country	Study design	Sample size (patients)	Age (years), mean SD or range	AI	Retinal camera	Gold standard	Diabetic retinopathy severity scale
Sosale et al, 2020 (Sosale, Aravind, et al., 2020)	India	Cross-sectional	900	N/A	Medios AI	The smartphone based 'Remidio FOP camera' by a trained technician	Five retina specialists, that is, three fellowshiptrained vitreoretinal surgeons and two medical retina specialists	ICDR severity scale; no DR, mild non-proliferative DR (mild NPDR), moderate non-proliferate DR (moderate NPDR), severe non-proliferate DR (severe NPDR) and proliferate DR (PDR). The images with DR were then evaluated for DME.
Tomic et al., 2023 (Tomić et al., 2023)	Croatia	Cross-sectional	160	Range 45–83 years	AI-based automated software (DeepDR)	Handheld fundus camera TANG	Medical retina specialists	ICDR and diabetic macular edema disease severity ratings

Table 2.
Outcome assessed in each study

Author/Year	Any DR			RDR			VTDR		
	Sensitivity	Specificity	AUC	Sensitivity	Specificity	AUC	Sensitivity	Specificity	AUC
Kemp et al, 2023 (Kemp et al., 2023)	-	-	-	77,5%	91,5%	0.84	89,2%	-	-
Kim et al., 2021 (Kim et al., 2021)	87,0% Not mentioned RDR or Any DR or STDR, but RWDR	78,6% Not mentioned RDR or Any DR or STDR, but RWDR	-	-	-	-	-	-	-
Krishnan et al, 2021 (Krishnan et al., 2021)	89,13%	94,43%	-	100%	89,55%	-	-	-	-
Lupidi et al, 2023 (Lupidi et al., 2023a)	96,8% Not mentioned RDR or Any DR or STDR	96,8% Not mentioned RDR or Any DR or STDR	-	-	-	-	-	-	-
Marlebi et al, 2022 (Malerbi et al., 2022)	-	-	-	97,8 %	61,4%	0,89	-	-	-
Midena et al, 2022 (Midena et al., 2022)	96,9% Any degree of DR	94,8% Any degree of DR	-	-	-	-	-	-	-
Natarajan et al., 2019 (Natarajan et al., 2019c)	85,2%	92%	-	100%	88,4%	-	-	-	-
Rajalakshmi et al., 2018 (Rajalakshmi et al., 2018a)	95,8 %	80,2 %	-	99,3 %	68,8 %	-	99,1 %	80,4 %	-

Author/Year	Any DR			RDR			VTDR		
	Sensitivity	Specificity	AUC	Sensitivity	Specificity	AUC	Sensitivity	Specificity	AUC
Rao et al, 2022 (Rao et al., 2022)	97,6%	90,9%	-	98,3%	83,7%	-	99,0%	N/A	-
Ruan et al., 2022 (Ruan et al., 2022)	-	-	-	82,1%	97,4%	0,88	-	-	-
Sosale et al, 2020 (Sosale, Sosale, et al., 2020)	86,78 %	95,45 %	0,91	98,84 %	86,73 %	0,92	100%	-	-
Sosale et al, 2020 (Sosale, Aravind, et al., 2020)	83,3 %	95,5 %	0,90	93 %	92,5 %	0,88	-	-	-
Tomic et al., 2023 (Tomić et al., 2023)	83,2%	100%	0,921	-	-	-	-	-	-

DISCUSSION

This review aimed to determine the validity of artificial intelligence (AI) as a screening tool for detecting diabetic retinopathy (DR) in diabetic patients. This device exhibited quite good sensitivity (83.2%–97.6%) and specificity (ranging from 80.2% to 100%) in detecting the presence of any DR. However, a broader range of sensitivity was reported in detecting VTDR (ranging from 89.2% to 100%), in contrast to the sensitivity and specificity reported for any DR and RDR. The accessibility of ophthalmologists and retinal consultants has emerged as a significant barrier to sufficient DR screening (Huemer et al., 2020). This indicates that the utilization of artificial intelligence (AI)-based tools for DR screening in the community could have mitigated such obstacles and was feasible, especially in countries or regions with limited facilities, encountering challenging geographical conditions, or grappling with healthcare financing constraints (Prayogo et al., 2023).

The advancement of computing technologies has led to the development of deep learning (DL). In recent years, deep learning (DL) has emerged as the dominant computational approach in the field of machine learning (ML) (Taye, 2023). DL, as a branch of AI, specialized in algorithms learning to accomplish tasks without explicit instructions (Esteva et al., 2019). Particularly, artificial intelligence (AI) has the capability to undergo training in order to distinguish diagnostic results by utilizing various inputs, including a retinal image. Every image undergoes scrutiny and comparison with the comprehensive data stored in the database. AI has the ability to identify distinctive features on the retina associated with diabetic retinopathy (DR), such as exudates, microaneurysms, and hemorrhages. Subsequently, the AI categorizes the identified feature as either normal or abnormal, thereby producing the ultimate output (Cheung et al., 2019). The number of previously published studies for comparison was very limited. A recent systematic review by Jayanegara & Lestari (2023) also analyzed twelve different AI-based devices from twelve articles, but only four of these devices met our inclusion criteria.

Sensitivity ranged from 79.2% as the lowest to 100% as the highest. Regarding specificity, eleven studies indicated reasonable values, while only one reported a lower figure, with just 68.8% for detecting RDR (Jayanegara & Lestari, 2023). The accuracy of this artificial intelligence-based device was nearly comparable to the recommended retinal examination for DR screening using conventional methods (Wong et al., 2018). According to the guidelines from the UK National Institute for Clinical Excellence (NICE), DR screening tests were required to achieve a sensitivity and specificity of at least 80% and 95%, respectively, with a technical failure rate of less than 5% (National Institute for Clinical Excellence, 2002). In a study conducted by Ahsan et al. (2014) using conventional methods, photos taken with a fundal camera (Canon CR-1) were compared with digital retinal photography of a single macular center spanning 45 degrees, equipped with a disc center plane, against the gold standard of seven standard 30° stereoscopic plane photos as proposed by ETDRS. The study yielded sensitivity and specificity results for detecting any DR of 55.67% and 71.27%, NSTDR of 37.36% and 76.78%, and STDR of 68.25% and 90% (Ahsan et al., 2014).

The rapid progression of technology in retinal imaging has elevated the precision and reduced the time required for detecting DR (Gulshan et al., 2019). Furthermore, the introduction of artificial intelligence (AI) has led to a reduction in the manpower or resources required for conducting screenings. Significant studies have shown that machine learning systems can identify referable DR from retinal photographs with a sensitivity and specificity of over 90%, as compared to expert assessments (Prayogo et al., 2023; Ting et al., 2017). For example, a notable advancement in AI for DR detection occurred through the development of an active deep learning (ADL) technique utilizing an artificial bee colony (ABC) algorithm. This method exhibited enhanced efficacy in identifying five levels of DR severity, surpassing an earlier AI approach which could only discern two levels of DR: referable and non-referable DR (Özbay, 2023). These findings suggested the potential for improving efficiency in DR screening coverage within clinical settings.

Despite numerous publications showcasing the reliability and accuracy of these DL systems in detecting DR. Its implementation faced several barriers, such as: In real-time screening, capturing images with artificial intellegence-based retinal imaging was often challenging due to small pupils, cataracts, or uncooperative patients. Despite the portability of handheld camera or smartphone-based imaging, using a portable table and chin rest helped stabilize image acquisition, simplifying the screening process (Balasopoulou et al., 2017). In addition, implementing AI for DR screening faced challenges with older participants having small pupils and cataracts, leading to poor image quality and hindering AI assessment. AI couldn't overcome physical limitations like small pupils or cataracts, making screening elderly patients challenging (Sosale, Sosale, et al., 2020).

Limitations of our systematic review included the absence of studies utilizing an RCT design. Randomized controlled trials for assessing diagnostic test accuracy posed unique challenges for researchers. It proved challenging to randomize diagnostic tests within a clinical workflow while ensuring comprehensive coverage for clinicians. Moreover, certain studies lacked detailed descriptions regarding the sensitivity and specificity of AI in screening for various levels of DR (Any DR, RDR, STDR). Another limitation pertained to the heterogeneity among studies in terms of baseline characteristics, particularly regarding camera usage and the AI algorithms employed for image acquisition and classification. AI algorithms utilized across different studies might have exhibited varied diagnostic accuracy depending on the training set used for the algorithm.

CONCLUSION

The review suggests that current AI-based devices exhibit commendable accuracy in detecting DR, with satisfactory levels of sensitivity and specificity. Integrating AI with conventional or portable retinal camera devices can significantly improve DR screening, enhancing the efficiency and accessibility of screening programs. This would help reduce vision loss and blindness in DR populations worldwide, especially in countries with limited DR screening infrastructure. Additionally, it would benefit regions with constrained healthcare resources. Furthermore, the review highlights opportunities for researchers to create more affordable yet precise DR screening kits, thus broadening the availability of affordable screening options for DR.

REFERENCES

Ahsan, S., Basit, A., Ahmed, K. R., Ali, L., Shaheen, F., Ulhaque, M. S., & Fawwad, A. (2014). Diagnostic accuracy of direct ophthalmoscopy for detection of diabetic retinopathy using fundus photographs as a reference standard. Diabetes and Metabolic Syndrome: Clinical Research and Reviews, 8(2), 96–101. https://doi.org/10.1016/j.dsx.2014.04.015

Aromataris E, M. Z. (Editors). (2020). CHECKLIST FOR COHORT STUDIES. Critical Appraisal tools for use in JBI Systematic Reviews. JBI Manual for Evidence Synthesis.

Balasopoulou, A., Kokkinos, P., Pagoulatos, D., Plotas, P., Makri, O. E., Georgakopoulos, C. D., Vantarakis, A., Li, Y., Liu, J. J., Qi, P., Rapoport, Y., Wayman, L. L., Chomsky, A. S., Joshi, R. S., Press, D., Rung, L., Ademola-popoola, D., Africa, S., Article, O., … Loukovaara, S. (2017). Symposium Recent advances and challenges in the management of retinoblastoma Globe - saving Treatments. BMC Ophthalmology, 17(1), 1. https://doi.org/10.4103/ijo.IJO

Bourne, R. R. A., Stevens, G. A., White, R. A., Smith, J. L., Flaxman, S. R., Price, H., Jonas, J. B., Keeffe, J., Leasher, J., Naidoo, K., Pesudovs, K., Resnikoff, S., & Taylor, H. R. (2013). Causes of vision loss worldwide, 1990-2010: A systematic analysis. The Lancet Global Health, 1(6), 339–349. https://doi.org/10.1016/S2214-109X(13)70113-X

Burgess, P. I., MacCormick, I. J. C., Harding, S. P., Bastawrous, A., Beare, N. A. V., & Garner, P. (2013). Epidemiology of diabetic retinopathy and maculopathy in Africa: A systematic review. Diabetic Medicine, 30(4), 399–412. https://doi.org/10.1111/j.1464-5491.2012.03756.x

Care, D., & Suppl, S. S. (2021). Microvascular complications and foot care: Standards of medical care in diabetes—2021. Diabetes Care, 44(January), S151–S167. https://doi.org/10.2337/dc21-S011

Cheung, C. Y., Tang, F., Wei Ting, D. S., Wei Tan, G. S., & Wong, T. Y. (2019). Artificial intelligence in diabetic eye disease screening. Asia-Pacific Journal of Ophthalmology, 8(2), 158–164. https://doi.org/10.22608/APO.201976

Esteva, A., Robicquet, A., Ramsundar, B., Kuleshov, V., DePristo, M., Chou, K., Cui, C., Corrado, G., Thrun, S., & Dean, J. (2019). A guide to deep learning in healthcare. Nature Medicine, 25(1), 24–29. https://doi.org/10.1038/s41591-018-0316-z

Fenner, B. J., Wong, R. L. M. M., Lam, W.-C. W. C., Tan, G. S. W. W., & Cheung, G. C. M. M. (2018). Advances in Retinal Imaging and Applications in Diabetic Retinopathy

Screening: A Review. Ophthalmology and Therapy, 7(2), 333–346. https://doi.org/10.1007/s40123-018-0153-7

Gulshan, V., Rajan, R. P., Widner, K., Wu, D., Wubbels, P., Rhodes, T., Whitehouse, K., Coram, M., Corrado, G., Ramasamy, K., Raman, R., Peng, L., & Webster, D. R. (2019). Performance of a Deep-Learning Algorithm vs Manual Grading for Detecting Diabetic Retinopathy in India. JAMA OPHTHALMOLOGY, 137(9), 987–993. https://doi.org/10.1001/jamaophthalmol.2019.2004 WE - Science Citation Index Expanded (SCI-EXPANDED)

Huemer, J., Wagner, S. K., & Sim, D. A. (2020). The evolution of diabetic retinopathy screening programmes: A chronology of retinal photography from 35 mm slides to artificial intelligence. Clinical Ophthalmology, 14, 2021–2035. https://doi.org/10.2147/OPTH.S261629

Jamaaluddin, & Indah, S. (2021). Buku Ajar Kecerdasan Buatan. Umsida Press, 121.

Jayanegara, A. P., & Lestari, Y. D. (2023). The Validity Of Artificial Intelligence (Ai) For Diabetic Retinopathy Screening In Asia : 006(2), 118–129.

JBI. (2020). Checklist for Analytical Cross Sectional Studies - Critical Appraisal tools for use in JBI Systematic Reviews. The Joanna Briggs Institute Critical Appraisal Tools for Use in JBI Systematic Reviews, 1–6.

Kemp, O., Bascaran, C., Cartwright, E., McQuillan, L., Matthew, N., Shillingford-Ricketts, H., Zondervan, M., Foster, A., & Burton, M. (2023). Real-world evaluation of smartphone-based artificial intelligence to screen for diabetic retinopathy in Dominica: a clinical validation study. BMJ OPEN OPHTHALMOLOGY, 8(1). https://doi.org/10.1136/bmjophth-2023-001491 WE - Emerging Sources Citation Index (ESCI)

Kim, T. N., Aaberg, M. T., Li, P. R., Davila, J. R., Bhaskaranand, M., Bhat, S., Ramachandra, C., Solanki, K., Myers, F., Reber, C., Jalalizadeh, R., Margolis, T. P., Fletcher, D., & Paulus, Y. M. (2021). Comparison of automated and expert human grading of diabetic retinopathy using smartphone-based retinal photography. EYE, 35(1), 334–342. https://doi.org/10.1038/s41433-020-0849-5

Krishnan, R., Jain, A., Rogye, A., & Natarajan, S. (2021). Use of offline artificial intelligence in a smartphone-based fundus camera for community screening of diabetic retinopathy. INDIAN JOURNAL OF OPHTHALMOLOGY, 69(11), 3150-+. https://doi.org/10.4103/ijo.IJO_3808_20 WE - Science Citation Index Expanded (SCI-EXPANDED)

Liberati, A., Altman, D. G., Tetzlaff, J., Mulrow, C., Gøtzsche, P. C., Ioannidis, J. P. A., Clarke, M., Devereaux, P. J., Kleijnen, J., & Moher, D. (2009). The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: Explanation and elaboration. PLoS Medicine, 6(7). https://doi.org/10.1371/journal.pmed.1000100

Lupidi, M., Danieli, L., Fruttini, D., Nicolai, M., Lassandro, N., Chhablani, J., & Mariotti, C. (2023a). Artificial intelligence in diabetic retinopathy screening: clinical assessment using handheld fundus camera in a real-life setting. ACTA DIABETOLOGICA, 60(8), 1083–1088. https://doi.org/10.1007/s00592-023-02104-0

Lupidi, M., Danieli, L., Fruttini, D., Nicolai, M., Lassandro, N., Chhablani, J., & Mariotti, C. (2023b). Artificial intelligence in diabetic retinopathy screening: clinical assessment using handheld fundus camera in a real-life setting. Acta Diabetologica, 60(8), 1083–1088. https://doi.org/10.1007/s00592-023-02104-0

Malerbi, F. K., Andrade, R. E., Morales, P. H., Stuchi, J. A., Lencione, D., de Paulo, J. V, Carvalho, M. P., Nunes, F. S., Rocha, R. M., Ferraz, D. A., & Belfort, R. (2022). Diabetic Retinopathy Screening Using Artificial Intelligence and Handheld Smartphone-Based Retinal Camera. JOURNAL OF DIABETES SCIENCE AND TECHNOLOGY, 16(3), 716–723. https://doi.org/10.1177/1932296820985567 WE - Emerging Sources Citation Index (ESCI)

Midena, E., Zennaro, L., Lapo, C., Torresin, T., Midena, G., Pilotto, E., & Frizziero, L. (2022). Handheld Fundus Camera for Diabetic Retinopathy Screening: A Comparison Study with Table-Top Fundus Camera in Real-Life Setting. JOURNAL OF CLINICAL MEDICINE, 11(9). https://doi.org/10.3390/jcm11092352 WE - Science Citation Index Expanded (SCI-EXPANDED)

Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., Antes, G., Atkins, D., Barbour, V., Barrowman, N., Berlin, J. A., Clark, J., Clarke, M., Cook, D., D'Amico, R., Deeks, J. J., Devereaux, P. J., Dickersin, K., Egger, M., Ernst, E., Gøtzsche, P. C., … Tugwell, P. (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. PLoS Medicine, 6(7). https://doi.org/10.1371/journal.pmed.1000097

Murchison, A. P., Hark, L., Pizzi, L. T., Dai, Y., Mayro, E. L., Storey, P. P., Leiby, B. E., & Haller, J. A. (2017). Non-adherence to eye care in people with diabetes. BMJ Open Diabetes Research and Care, 5(1), 1–10. https://doi.org/10.1136/bmjdrc-2016-000333

Nagendran, M., Chen, Y., Lovejoy, C. A., Gordon, A. C., Komorowski, M., Harvey, H., Topol, E. J., Ioannidis, J. P. A., Collins, G. S., & Maruthappu, M. (2020). Artificial intelligence versus clinicians: Systematic review of design, reporting standards, and claims of deep learning studies in medical imaging. The BMJ, 368, 1–12. https://doi.org/10.1136/bmj.m689

Natarajan, S., Jain, A., Krishnan, R., Rogye, A., & Sivaprasad, S. (2019a). Diagnostic Accuracy of Community-Based Diabetic Retinopathy Screening with an Offline Artificial Intelligence System on a Smartphone. JAMA Ophthalmology, 137(10), 1182–1188. https://doi.org/10.1001/jamaophthalmol.2019.2923

Natarajan, S., Jain, A., Krishnan, R., Rogye, A., & Sivaprasad, S. (2019b). Diagnostic Accuracy of Community-Based Diabetic Retinopathy Screening with an Offline Artificial Intelligence System on a Smartphone. JAMA Ophthalmology, 137(10), 1182–1188. https://doi.org/10.1001/jamaophthalmol.2019.2923

Natarajan, S., Jain, A., Krishnan, R., Rogye, A., & Sivaprasad, S. (2019c). Diagnostic Accuracy of Community-Based Diabetic Retinopathy Screening With an Offline Artificial Intelligence System on a Smartphone. JAMA OPHTHALMOLOGY, 137(10), 1182–1188. https://doi.org/10.1001/jamaophthalmol.2019.2923 WE - Science Citation Index Expanded (SCI-EXPANDED)

National Institute for Clinical Excellence. (2002). Management of Type 2 Diabetes: Retinopathy Screening and Early Management. Inherited Clinical Guideline.

Özbay, E. (2023). An active deep learning method for diabetic retinopathy detection in segmented fundus images using artificial bee colony algorithm. Artif Intell Rev, 56, 3291–3318. https://doi.org/https://doi.org/10.1007/s10462-022-10231-3

Prayogo, M. E., Zaharo, A. F., Damayanti, N. N. R., Widyaputri, F., Thobari, J. A., Susanti, V. Y., & Sasongko, M. B. (2023). Accuracy of Low-Cost, Smartphone-Based Retinal Photography for Diabetic Retinopathy Screening: A Systematic Review. Clinical Ophthalmology, 17(July), 2459–2470. https://doi.org/10.2147/OPTH.S416422

Rajalakshmi, R., Subashini, R., Anjana, R. M., & Mohan, V. (2018a). Automated diabetic retinopathy detection in smartphone-based fundus photography using artificial intelligence. EYE, 32(6), 1138–1144. https://doi.org/10.1038/s41433-018-0064-9 WE - Science Citation Index Expanded (SCI-EXPANDED)

Rajalakshmi, R., Subashini, R., Anjana, R. M., & Mohan, V. (2018b). Automated diabetic retinopathy detection in smartphone-based fundus photography using artificial intelligence. Eye (Basingstoke), 32(6), 1138–1144. https://doi.org/10.1038/s41433-018-0064-9

Rao, D. P., Sindal, M., Sengupta, S., Baskaran, P., Venkatesh, R., Sivaraman, A., & Savoy, F. M. (2022). Towards a Device Agnostic AI for Diabetic Retinopathy Screening: An External Validation Study. CLINICAL OPHTHALMOLOGY, 16, 2659–2667. https://doi.org/10.2147/OPTH.S369675 WE - Emerging Sources Citation Index (ESCI)

Royle, P., Mistry, H., Auguste, P., Shyangdan, D., Freeman, K., Lois, N., & Waugh, N. (2015). Pan-retinal photocoagulation and other forms of laser treatment and drug therapies for non-proliferative diabetic retinopathy: Systematic review and economic evaluation. Health Technology Assessment, 19(51), 1–248. https://doi.org/10.3310/hta19510

Ruan, S., Liu, Y., Hu, W. T., Jia, H. X., Wang, S. S., Song, M. L., Shen, M. X., Luo, D. W., Ye, T., & Wang, F. H. (2022). A new handheld fundus camera combined with visual artificial intelligence facilitates diabetic retinopathy screening. INTERNATIONAL JOURNAL OF OPHTHALMOLOGY, 15(4), 620–627. https://doi.org/10.18240/ijo.2022.04.16 WE - Science Citation Index Expanded (SCI-EXPANDED)

Sasongko, M. B., Widyaputri, F., Agni, A. N., Wardhana, F. S., Kotha, S., Gupta, P., Widayanti, T. W., Haryanto, S., Widyaningrum, R., Wong, T. Y., Kawasaki, R., & Wang, J. J. (2017). Prevalence of Diabetic Retinopathy and Blindness in Indonesian Adults With Type 2 Diabetes. American Journal of Ophthalmology, 181, 79–87. https://doi.org/10.1016/j.ajo.2017.06.019

Sheikh, A., Bhatti, A., Adeyemi, O., Raja, M., & Sheikh, I. (2021). The Utility of Smartphone-Based Artificial Intelligence Approaches for Diabetic Retinopathy: A Literature Review and Meta-Analysis. Journal of Current Ophthalmology, 33(3), 219–226. https://doi.org/10.4103/2452-2325.329064

Sosale, B., Aravind, S. R., Murthy, H., Narayana, S., Sharma, U., Gowda, S. G. V, & Naveenam, M. (2020). Simple, Mobile-based Artificial Intelligence Algorithm in the detection of Diabetic Retinopathy (SMART) study. BMJ OPEN DIABETES RESEARCH & CARE, 8(1). https://doi.org/10.1136/bmjdrc-2019-000892 WE - Science Citation Index Expanded (SCI-EXPANDED)

Sosale, B., Sosale, A. R., Murthy, H., Sengupta, S., & Naveenam, M. (2020). Medios- An offline, smartphone-based artificial intelligence algorithm for the diagnosis of diabetic retinopathy. INDIAN JOURNAL OF OPHTHALMOLOGY, 68(2), 391–395. https://doi.org/10.4103/ijo.IJO_1203_19 WE - Science Citation Index Expanded (SCI-EXPANDED)

Srisubat, A., Kittrongsiri, K., Sangroongruangsri, S., Khemvaranan, C., Shreibati, J. B., Ching, J., Hernandez, J., Tiwari, R., Hersch, F., Liu, Y., Hanutsaha, P., Ruamviboonsuk, V., Turongkaravee, S., Raman, R., & Ruamviboonsuk, P. (2023). Cost-Utility Analysis of Deep Learning and Trained Human Graders for Diabetic Retinopathy Screening in a Nationwide Program. OPHTHALMOLOGY AND THERAPY, 12(2), 1339–1357. https://doi.org/10.1007/s40123-023-00688-y

Tan, T. E., & Wong, T. Y. (2023). Diabetic retinopathy: Looking forward to 2030. Frontiers in Endocrinology, 13(January), 1–8. https://doi.org/10.3389/fendo.2022.1077669

Tao, X., Li, J., Zhu, X., Zhao, B., Sun, J., Ji, L., Hu, D., Pan, C., Huang, Y., Jiang, S., Feng, Q., & Jiang, C. (2016). Association between socioeconomic status and metabolic control and diabetes complications: A cross-sectional nationwide study in Chinese adults with type 2 diabetes mellitus. Cardiovascular Diabetology, 15(1), 1–10. https://doi.org/10.1186/s12933-016-0376-7

Taye, M. M. (2023). Understanding of Machine Learning with Deep Learning: Architectures, Workflow, Applications and Future Directions. Computers, 12(5). https://doi.org/10.3390/computers12050091

Thomas, R. L., Halim, S., Gurudas, S., Sivaprasad, S., & Owens, D. R. (2019). IDF Diabetes Atlas: A review of studies utilising retinal photography on the global prevalence of diabetes related retinopathy between 2015 and 2018. Diabetes Research and Clinical Practice, 157, 107840. https://doi.org/10.1016/j.diabres.2019.107840

Ting, D. S. W., Cheung, C. Y. L., Lim, G., Tan, G. S. W., Quang, N. D., Gan, A., Hamzah, H., Garcia-Franco, R., Yeo, I. Y. S., Lee, S. Y., Wong, E. Y. M., Sabanayagam, C., Baskaran, M., Ibrahim, F., Tan, N. C., Finkelstein, E. A., Lamoureux, E. L., Wong, I. Y., Bressler, N. M., … Wong, T. Y. (2017). Development and validation of a deep learning system for diabetic retinopathy and related eye diseases using retinal images from multiethnic populations with diabetes. JAMA - Journal of the American Medical Association, 318(22), 2211–2223. https://doi.org/10.1001/jama.2017.18152

Tomić, M., Vrabec, R., Hendelja, Đ., Kolarić, V., Bulum, T., & Rahelić, D. (2023). Diagnostic Accuracy of Hand-Held Fundus Camera and Artificial Intelligence in Diabetic Retinopathy Screening. Biomedicines, 12(1). https://doi.org/10.3390/biomedicines12010034

Tomić, M., Vrabec, R., Hendelja, Đ., Kolarić, V., Bulum, T., & Rahelić, D. (2024). Diagnostic Accuracy of Hand-Held Fundus Camera and Artificial Intelligence in Diabetic Retinopathy Screening. Biomedicines, 12(1). https://doi.org/10.3390/biomedicines12010034

US Department of Health and Human Services. (2020). National Diabetes Statistics Report, 2020. National Diabetes Statistics Report, 2.

Vujosevic, S., Aldington, S. J., Silva, P., Hernández, C., Scanlon, P., Peto, T., & Simó, R. (2020). Screening for diabetic retinopathy: new perspectives and challenges. The Lancet. Diabetes & Endocrinology, 8(4), 337–347. https://doi.org/10.1016/S2213-8587(19)30411-5

Wong, T. Y., Sun, J., Kawasaki, R., Ruamviboonsuk, P., Gupta, N., Lansingh, V. C., Maia, M., Mathenge, W., Moreker, S., Muqit, M. M. K., Resnikoff, S., Verdaguer, J., Zhao, P., Ferris, F., Aiello, L. P., & Taylor, H. R. (2018). Guidelines on Diabetic Eye Care: The International Council of Ophthalmology Recommendations for Screening, Follow-up, Referral, and Treatment Based on Resource Settings. Ophthalmology, 125(10), 1608–1622. https://doi.org/10.1016/j.ophtha.2018.04.007

Zegeye, A. F., Temachu, Y. Z., & Mekonnen, C. K. (2023). Prevalence and factors associated with Diabetes retinopathy among type 2 diabetic patients at Northwest Amhara Comprehensive Specialized Hospitals, Northwest Ethiopia 2021. BMC Ophthalmology, 23(1), 1–8. https://doi.org/10.1186/s12886-022-02746-8.