IJRETINA
International Journal of Retina

EVALUATING THE ACCURACY OF ARTIFICIAL INTELLIGENCE (AI)-INTEGRATED, SMARTPHONE-BASED SCREENING FOR DIABETIC RETINOPATHY: SYSTEMATIC REVIEW

Shofia Medina Samara, Ariyoga Kun Laksono
Resident of Ophthalmology Department, General Hospital of Saiful Anwar Malang, Indonesia
Faculty of Medicine, Universitas Brawijaya, Malang, Indonesia

Abstract

Background: *Diabetic retinopathy (DR is the most common microvascular complication of diabetes that can cause vision problems and blindness that poses a significant health risk and financial burden, increasing the needs to effectively screen and manage diabetic eye disease. The current method of screening for diabetic eye disease relies on human experts to analyze the results. Alternatively, recent advancements in artificial intelligence (AI) especially deep learning (DL) and retinal imaging using smartphones offer a promising solution for both patients and ophthalmologists, potentially improving patient compliance and making telemedicine more efficient for DR screening.*

Purpose : *To represent on accuracy of AI-integrated process in smartphone-based DR screening and to compare the various study methods and settings used to achieve this accuracy.*

Method: *Literature search on current DR screening programs was conducted following the Preferred Reporting Items for Systematic Reviews and Meta-Analyzes (PRISMA) framework on Google Scholar, Scopus, Web of Science, PubMed, Medline, and Embase with most recent search was updated on June 1st, 2024. Key information was extracted from the studies included author names, journal, year of publication, country, sensitivity, specificity, positive and negative predictive values (if available), study methods, and settings.*

Result: *The study identification process resulting in 9 selected studies. The performance metrics reported included intergrader/intramodality agreement, sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV). The sensitivity of AI in detecting DR ranged from 77-100%, while specificity ranged from 61.4 - 95.5%. PPV and NPV were reported less frequently, with ranges of 48.1 - 92.92% and 91.3 - 99.46%, respectively. Intergrader agreement was within range κ= 0.45 – 0.91.*

Conclusion: *The studies reviewed in this paper collectively represents the potential of smartphone based integrated with AI in revolutionizing DR screening. The high sensitivity and specificity achieved by various AI algorithms, often exceeding the standards set by regulatory bodies like the FDA and ETDRS, highlight their accuracy in detecting DR and its severity levels. The accessibility and user-friendliness of smartphone-based retinal imaging further enhance the coverage of DR screening, particularly in underserved areas with limited resources and internet connectivity.*

Correspondence to:
Shofia Medina Samara,
General Hospital of Saiful
Anwar Malang, Indonesia,
shofiamedina@gmail.com

INTRODUCTION

Diabetic eye disease, a common issue for people with Type 1 or Type 2 diabetes, is a major reason for vision loss in working-age adults. In 2017, an estimated 425 million adults worldwide had diabetes, double the number in 1980, and this figure is expected to reach 629 million by 2045. Diabetic retinopathy (DR) roughly affected 30% to 45% people with diabetes and is the most common microvascular complication where 10% of these cases are vision-threatening, meaning they could significantly impair vision, even blindness.[1]

The increasing prevalence of diabetic eye disease is a worldwide problem that poses a significant health risk and financial burden for both individuals and societies, particularly in developing countries where it is becoming more common.[2] This surge in diabetes cases puts increasing pressure on healthcare providers to effectively screen and manage diabetic eye disease.[3] A fully implemented national diabetic eye screening program (DESP) by UK successfully screened a vast majority of diabetic patients resulting in decrease of DR as the primary cause of blindness in working-age adults in the UK, highlighting the effectiveness of such programs in preventing and treating DR-related vision loss.[4,5]

Despite the importance of DR screening, compliance with recommended guidelines is low because of a lack of understanding about the disease and its potential complications,[6] difficulty accessing medical resources and inadequate insurance coverage for these exams.[7] One study from Indonesia reflects that despite being knowledgeable and having a positive attitude towards DR screening, general practitioners (GP)s did not consistently implement it in practice because of limited experience, lack of confidence in diagnosing fundus abnormalities, and lack of equipment in primary care.[8]

Telemedicine offers a way to make fundus screening more accessible by allowing patients to be screened at convenient times and locations, eliminating the need to travel to far-off hospitals for an eye examination with an ophthalmologist. While telemedicine makes DR screening more accessible, it still requires human experts to analyze the images.[9] However, recent progress in deep learning (DL)-based artificial intelligence (AI) presents a potential solution that benefits both patients and ophthalmologists, as a way of detecting retinal images which may be sight-threatening,[10] potentially improving patient compliance and making telemedicine more efficient for DR screening.[9]

DL in AI software can automatically analyze the retinal images and provide recommendations for follow-up care or referrals, reducing the workload of ophthalmologists. This increased convenience and efficiency can significantly improve patient participation in DR screenings and detections.[1,9] While some research has shown that AI can accurately screen for diabetic retinopathy in developed country like UK,[3,11] it often relies on expensive, desktop-based fundus cameras that may not be accessible in rural areas.[12]

Retinal imaging using smartphones has proven to be an effective and affordable method for diabetic retinopathy screening.[13,14] Similarly, using AI to analyze images from inexpensive smartphone devices has also been shown to be valid for screening in community settings.[15] This study aims to represent on accuracy of AI-integrated process in smartphone-based DR screening and to compare the various study methods and settings used to achieve this accuracy

METHOD

The author evaluated the current state of smartphone-based DR screening programs by searching Google scholar and PubMed via Medline for open-access studies published in English using

these keywords: "diabetic retinopathy, artificial intelligence, DR screening, smartphone". In addition, we sought reference lists and publicly accessible websites for information on established, commercially available diabetic retinopathy (DR) screening algorithms.

A comprehensive study was done to compare the accuracy and methods on the pre-existing research on smartphone-based, AI-integrated of diabetic retinopathy screening. The PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyzes) framework guided the review process by establishing a set of predefined criteria and guidelines. The PRISMA diagram (Figure 1) outlines the search strategy used, including the criteria for including and excluding studies. The most recent search was conducted on June 1st, 2024. Key information extracted from the studies included author names, journal, year of publication, country, study type, sensitivity, specificity, positive and negative predictive values (if available), and study methods and settings.

Eligibility criteria

We sought to include studies that had outcome measures of accuracy for sensitivity and specificity (and predictive values, if available) for smartphone-based, AI-integrated of diabetic retinopathy screening. Only studies in English were included. Studies which not stated the accuracy on DR screening are excluded. Researches on data set without the patient are excluded. Conference abstracts, review articles, letters to the editor, editorials, and correspondence notes were excluded.

Risk of Bias

We used the QUADAS-2 (Quality Assessment of Diagnostic Accuracy Studies) tool to assess the risk of bias and applicability of all included studies. Two reviewers (SMS and AKL) performed assessment of selected studies independently. Any disagreements between the review authors were resolved by discussion until a consensus was reached.

Data Extraction and Analysis

The data extraction process involved systematically gathering and combining relevant data from the selected studies. Data extraction included meticulously reviewing each study to identify information such as the study's methods, setting, sample size, the AI software or tools used, intergrader/intermodality agreement (kappa values) and the study's accuracy as indicated by sensitivity, specificity, and predictive values.

In our analysis, we divided the results into two categories to assess the accuracy of AI-integrated, smartphone-based devices in diagnosing any level of DR and the more severe, referable DR (RDR) defined as moderate non-proliferative DR with diabetic macular edema (DME), severe NPDR, or worse, regardless of DME presence. This assessment was based on clinical grading using retinal images without any OCT examination.[16] Sensitivity and specificity data from studies reporting detailed test outcomes such as true positive (TP), false positive (FP), true negative (TN), and false negative (FN) were visually summarized in a forest plot using RevMan 5.4. Studies without this detailed information were excluded from the forest plot. This approach allowed for a clear and concise presentation of the accuracy of smartphone-based devices in diagnosing any DR and RDR.

RESULTS

The study identification process included searching databases, resulting in 24 records. These reports were screened for eligibility, leading to the inclusion of 9 studies in the review [Figure 1]. Twelve reports were excluded for reasons such as focusing on data sets rather than patients, not using smartphone, not stating sensitivity and/or specificity, and studies using smartphone but not integrated with AI. Ultimately, 9 studies were selected for
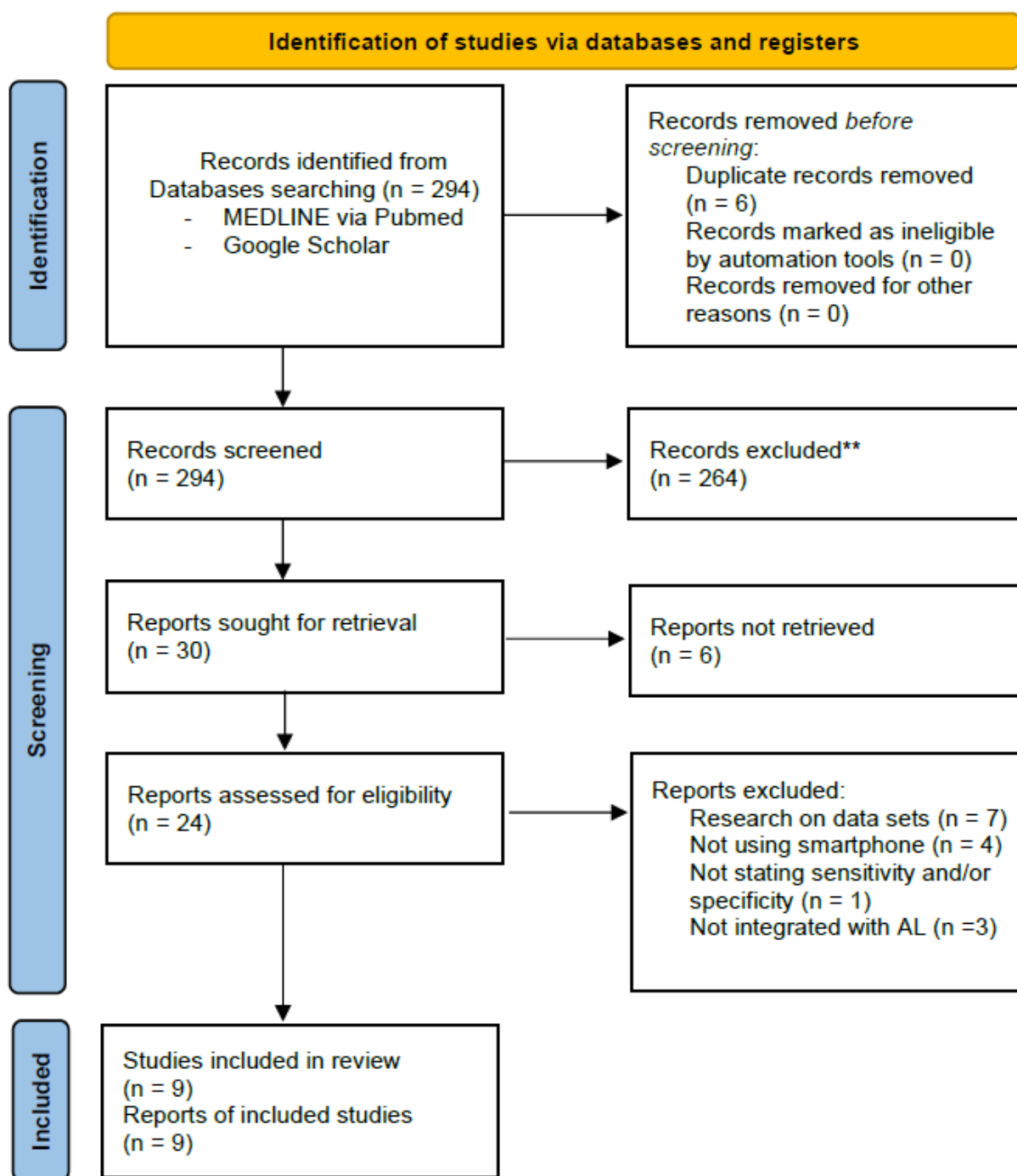
Figure 1 Data selection steps using the Preferred Reporting Items for Systematic Review and Meta-Analyzes.

detailed analysis and synthesis in the review. Forest plot contains the summary of sensitivity and specificity from five studies in detecting any DR [Figure 2] and five studies in detecting RDR [Figure 3]. The study by Malerbi (2022), which categorized outcomes as "more than mild DR (mtmDR)," was incorporated into the RDR forest plot due to the equal definition of RDR.

Any DR

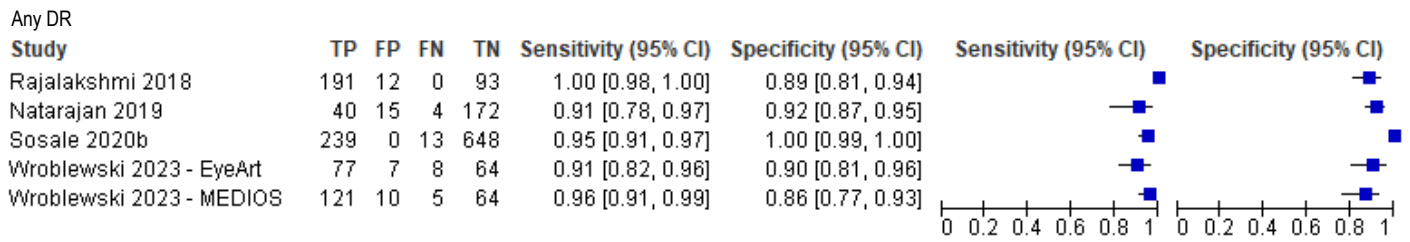| Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Rajalakshmi 2018	191	12	0	93	1.00 [0.98, 1.00]	0.89 [0.81, 0.94]
Natarajan 2019	40	15	4	172	0.91 [0.78, 0.97]	0.92 [0.87, 0.95]
Sosale 2020b	239	0	13	648	0.95 [0.91, 0.97]	1.00 [0.99, 1.00]
Wroblewski 2023 - EyeArt	77	7	8	64	0.91 [0.82, 0.96]	0.90 [0.81, 0.96]
Wroblewski 2023 - MEDIOS	121	10	5	64	0.96 [0.91, 0.99]	0.86 [0.77, 0.93]

Figure 2 Forest plot showing sensitivity and specificity range of AI software in detecting any severity of diabetic retinopathy (DR) from smartphone-based images

RDR

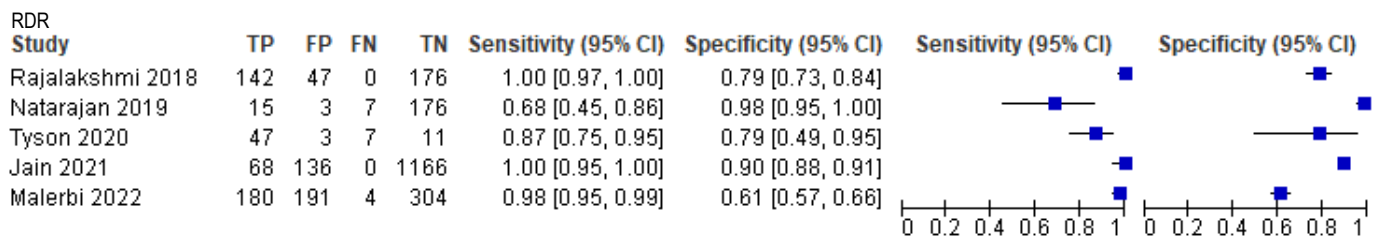| Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Rajalakshmi 2018	142	47	0	176	1.00 [0.97, 1.00]	0.79 [0.73, 0.84]
Natarajan 2019	15	3	7	176	0.68 [0.45, 0.86]	0.98 [0.95, 1.00]
Tyson 2020	47	3	7	11	0.87 [0.75, 0.95]	0.79 [0.49, 0.95]
Jain 2021	68	136	0	1166	1.00 [0.95, 1.00]	0.90 [0.88, 0.91]
Malerbi 2022	180	191	4	304	0.98 [0.95, 0.99]	0.61 [0.57, 0.66]

Figure 3 Forest plot showing sensitivity and specificity range of AI software in detecting referable diabetic retinopathy (RDR) from smartphone-based images

Table 1 illustrates the risk of bias assessment of included studies using the QUADAS-2 tool. Most studies employed low risk of bias on reference standard except from Wroblewski, et al. 2023[17] because this study did not state exact reference used to grade DR. Some studies were deemed high risk of bias in index test because the images used were labelled by a single grader. Risk of bias in patient selection from couple of studies were scored high because the study did not present demographic of the patient. Overall, the applicability concern of the studies is low.

Table 2 summarized studies evaluating the performance of various smartphone-based, AI integrated screening process in detecting DR. The studies were arranged chronologically, starting with the research published in 2018 progressing towards more recent publications. The studies employed different DR classification systems, with the International Clinical Diabetic Retinopathy (ICDR) scale being the most common. Intergrader and intramodality agreement varied, suggesting inconsistencies in image interpretation.

Study	RISK OF BIAS				APPLICABILITY CONCERNS		
	PATIENT SELECTION	INDEX TEST	REFERENCE STANDARD	FLOW AND TIMING	PATIENT SELECTION	INDEX TEST	REFERENCE STANDARD
Sengupta, 2018[14]	☹ High	?	☺ Low	☺ Low	☺ Low	?	?
Rajalakshmi, 2018[18]	☺ Low	?	☺ Low	☺ Low	☺ Low	☺ Low	☺ Low
Natarajan, 2019[13]	☹ High	☺ Low	☺ Low	☹ High	☺ Low	☺ Low	☺ Low
Sosale, 2020a[19]	☹ High	☺ Low	☺ Low	☹ High	☺ Low	☺ Low	☺ Low
Tyson, 2020[20]	☹ High	?	☺ Low	☹ High	☹ High	☺ Low	?
Sosale, 2020b[21]	☺ Low	☺ Low	☺ Low	☺ Low	☺ Low	☺ Low	☺ Low
Jain, 2021[12]	☺ Low	☺ Low	☺ Low	☺ Low	☺ Low	☺ Low	☺ Low
Malerbi, 2022[22]	☹ High	☹ High	☺ Low	☺ Low	☹ High	☺ Low	?
Wroblewski, 2023[17]	☺ Low	☹ High	☹ High	☺ Low	☺ Low	☺ Low	☺ Low

☺ Low Risk   ☹ High Risk   ? Unclear Risk

Table 1 Risk of Bias Assessment Using Quadas-2 Tools

However, the accuracy results were promising, with sensitivity ranging from 83.3% to 98.84% and specificity from 61.4% to 95.5 PPV and NPV were also generally high, indicating the reliability of smartphone-based screening in identifying individuals with and without DR.

Table 3 provides a summary of different study settings DR screening. data from nine studies conducted in India, the USA, Brazil, and Mexico, which explored the use of smartphone-based AI software for diabetic retinopathy screening. The studies used various AI software (Medios AI, EyeArt, PhelcomNet), smartphones (HTC One, iPhone 6), and image capture methods (with and without pupil dilation). The number of fundus images taken and the fields of view varied across the studies. Some studies specified the healthcare workers involved in the screening process (trained technicians, medical students, health workers), while others did not.

Table 2 Summary of selected study showing study samples and accuracy on DR screening. CI: Confidence Interval; DR: Diabetic Retinopathy; NPDR: Non-Proliferative Diabetic Retinopathy; ICDR: International Clinical Diabetic Retinopathy; STDR: Sight Threatening Diabetic Retinopathy; DME: Diabetic Macular Edema; RDR; Referable Diabetic Retinopathy; PDR: Proliferative Diabetic Retinopathy; ETDRS: Early Treatment Diabetic Retinopathy Study; DESP: Diabetic Eye Screening Programme; vtDR: Vision-threatening Diabetic Retinopathy; mtmDR: more than mild Diabetic Retinopathy

Author	Sample Size	DR classification	Intergrader Agreement	Intramodality Agreement	Accuracy results, 95% CI			
					Sensitivity	Specificity	PPV	NPV
Sengupta, et al. 2018[14]	135 individuals 233 eyes	ICDR[18], DESP[19,20]	any DR κ= 0.55 VTDR κ= 0.76	-	Any DR 93.1 (88.3-96.4)	Any DR 89.1% 89.1 (68.2-92.2)	-	-
Rajalakshmi, et al. 2018[21]	296 individuals 2408 images	ICDR,[18] Referable DR (RDR) defined as moderate NPDR and above	-	any DR κ=0.78 STDR κ= 0.75 RDR κ= 0.67	Any DR 95.8% (92.9-98.7) DME 97% (91.5-99.4) PDR 78.1% (63.8-83.3) STDR 99.1% (95.1-99.9) RDR 99.3% (96.1-99.9)	Any DR 80.2% (72.6-87.8) DME 75.8% (69.7-81.8) PDR 89.8% (86.1-93.4) STDR 80.4% (73.9-85.9) RDR 68.8% (61.5-76.2)	Any DR 89.7% DME 67.4% PDR 48.1% STDR 75.3% RDR 74.6%	Any DR 91.4% DME 98% PDR 97.1% STDR 99.3% RDR 99.1%
Natarajan, et al. 2019[13]	231 individuals	ICDR[18]	Any DR κ= 0.85	-	Any DR 85.2% (66.3-95.8) RDR 100% (78.2-100)	Any DR 92% (91.1-95.4) RDR 88.4% (83.1-92.5)	-	-
Sosale, et al. 2020a[22]	297 individuals	ICDR, DME[18]	Any DR κ= 0.89 DME κ= 0.9	-	Any DR 86.78% (82.9–90.6) RDR 98.84% (97.6–100)	Any DR 95.45% (93–97.8) RDR 86.73% (82.8–90.5)	Any DR 92.9% RDR 75.2%	Any DR 91.3% RDR 99.4%
Tyson, et al. 2020[23]	69 individuals 119 eyes	modified Airlie House classification system used in ETDRS[24,25] and ICDR[18]	κ= 0.45 ± 0.33.	-	Per patient/eye RDR 87.0% (74.5-94.2) / 77.8% (67.3-85.7)	Per patient/eye RDR 78.6% (44.8-94.3) / 71.5% (48.7-86.9)	-	-
Sosale, et al. 2020b[26]	922 patients	ICDR, DME[18]	κ= 0.79 - 0.91	κ= 0.8	Any DR 83.3% (80.9-85.7) RDR 93% (91.3-94.7) STDR 95.2%	Any DR 95.5% (94.1-96.8) RDR 92.5% (90.8-94.2)	Any DR 87.8% RDR 78.2%	Any DR 93.6% RDR 97.8%
Jain, et al. 2021[12]	1378 individuals	ICDR[18]	κ= 0.89	-	Per Patient/Eye Any DR 89.1% (82.7-93.7)/88.6% (83.5-92.5) RDR 100% (94.7-100.0) /100% (96.6-100.00	Per Patient/Eye Any DR 94.4% (91.8-94.7)/94.8% (93.8-95.6) RDR 89.5% (87.7-91.1)/91.8% (90.7-92.9)	-	-
Malerbi, et al. 2022[27]	824 individuals 3255 images	ICDR[18]	-	.	mtmDR 97.8% (96.7-98.9)	mtmDR 61.4% (57.7-65.1)	mtmDR 48.5%	mtmDR 98.7%
Wroblewski, et al. 2023[17]	248 patients 2130 images	Any grade DR by a retinologist	-	-	Any DR Medios 94% (88-97) EyeArt 94% (86-98)	Any DR Medios 94% (88-98) EyeArt 86% (77-93)	Any DR Medios 95% EyeArt 89%	Any DR Medios 93% EyeArt 93%

Table 3 Summary of selected studies showing study settings, methods, study samples, tools, and images used in DR screening. AI: Artificial Intelligence.

Author	Country	Screener	AI Software Used	Smartphone Used	Pupil Dilation	Fundus Images
Sengupta, *et al.* 2018[14]	India	Not specified	Medios AI[28]	HTC One with Remidio FOP (Fundus on Phone)[29]	Yes	45° Field of View (FOV) 3 fields (posterior pole [macula-centered], nasal field, and superotemporal field)
Rajalakshmi, *et al.* 2018[21]	India	Not specified	EyeArt™ (v2.1.0)[30]	Not specified	Yes	45° FOV 4 fields (macula centred, disc centred, superior-temporal and inferior-temporal)
Natarajan, *et al.* 2019[13]	India	Health worker	Medios AI[28]	Not specified	Yes	Anterior segment 3 fields (posterior pole [including disc and macula], nasal and temporal field)
Sosale, *et al.* 2020[22]	India	Trained technician[14]	Medios AI[28]	Iphone 6 Remidio Non Mydriatic (NM) FOP[29]	No	3 fields (posterior pole [macula centred], nasal and supero-temporal field)
Tyson, *et al.* 2020[23]	USA	Medical student and medical intern	EyeArt® (v2.0) system[30]	Iphone with Retinascape[31]	Yes	5 sequential images (central, inferior, superior, nasal, and temporal)
Sosale, *et al.* 2020[26]	India	Trained technician[14]	Medios AI[28]	Iphone 6 with Remidio NM FOP[29]	No	2 images, disc and macula centred
Jain, *et al.* 2021[12] (2)	India	Healthcare workers	Medios AI[28]	Smartphone with Remidio NM FOP[29]	No	3 fields (posterior pole, nasal and temporal fields)
Malerbi, *et al.* 2022[27]	Brazil	9 examiners including med students	PhelcomNet, Modified Xception[32]	Smartphone with Eyer, Phelcom Technologies[33]	Yes	45° FOV, 2 images of posterior pole (macula and disc centred)
Wroblewski, et al. 2023[17]	Mexico	3 graduate students	Medios AI,[28] EyeArt™ (v2.1.0)[30]	Smartphone with Remidio FOP[29]	Yes	3 fundus fields (posterior pole [disc and macula], nasal, temporal)

DISCUSSION

In this review, we reported the sensitivity, specificity, predictive values, intergrader/intermodality agreement and compared the different study settings from nine studies. These studies showed from good sensitivity of 67.3 - 100% and specificity ranged from 61.4 - 95.5%. There is a wide range of PPV with 48.1 - 92.92%, but display astounding NPV range of 91.3 - 99.46%. This founding suggests that even though using smartphone-based screening integrated with AI for DR might need some work to increase the detection rate, it already shows promising results in differentiating those with no disease. The U.S. Food and Drug Administration (FDA) requires superiority cut-offs for AI algorithms used in DR screening to demonstrate a minimum sensitivity of 85% and specificity of 82.5%.[34] Most studies reviewed shown that AI algorithms can meet or exceed these accuracy requirements.

Within the study of Rajalakshmi et al. (2018),[21] the EyeArt software demonstrated a high sensitivity exceeding 95% for detecting DR, STDR, and RDR in retinal images captured using the FOP smartphone device. The accuracy achieved in this study is comparable to that of the Google AI algorithm, which demonstrated high sensitivity and specificity for detecting RDR in both the EYEPACS-1 (97.5% sensitivity, 93.4% specificity) and Messidor-2 (96.1% sensitivity, 93.9% specificity) datasets.[35]

Lower specificity is found in study of Rajalakshmi et al. 2018 in detecting RDR (68.8%) because AI algorithm tends to overestimate the presence of moderate NPDR in retinal images. This misdiagnosis frequently occurs because the AI system mistakenly identifies certain retinal features not associated with DR, such as drusen, RPE atrophy, RPE hypertrophy, telangiectatic vessels, and retinal vein occlusion, as indicators of DR.[21] Specificity of not mtmDR in Malerbi et al. (2022) is also somewhat lower (61.4%)[27] than previous report,[34,36] and from further

examination of the false positives showed that the algorithm often mistook normal variations in fundus pigmentation or image imperfections for signs of disease.[27]

In a study by Jain et al. (2021), the AI system demonstrated 100% sensitivity for detecting RDR, but only 89.55% specificity. This resulted from 136 false positives (10.5%), including 55 misdiagnoses of mild NPDR and other non-DR conditions such as glaucoma, retinitis pigmentosa, age-related macular degeneration, gliosis, macular scars and asteroid hyalosis. This high false-positive rate (and the low PPV of 33.33%) might be attributed to the combination of the AI's extreme sensitivity and potential image quality issues due to less-experienced operators. While a high rate of false positives may result in more patients being referred to specialists and potentially cause them unnecessary anxiety, it can also be beneficial by identifying other eye conditions other than DR that require medical attention.[12]

Overall, the intergrader agreement (quadratic weighted kappa) reported in selected studies is moderate to high (κ= 0.45-0.91). Sengupta, et al. 2018 compared photographic modality (Remidio FOP vs. Topcon desktop fundus camera) and found intergrader agreement using Remidio FOP is moderate in detecting any DR (κ= 0.55; CI = 0.50 - 0.57), not too far off from using Topcon (κ= 0.68; CI = 0.67 -0.73). Both imaging methods showed substantial agreement in diagnosing VTDR with Remidio FOP (κ= 0.76; CI = 0.68 -0.85) and Topcon (κ= 0.81; CI = 0.73 -0.89). However, the slightly higher agreement for Topcon may be partially attributed to the misclassification of three cases of PDR as no DR (R0) by one grader using the Remidio FOP. This error occurred despite adequate image quality, likely because the characteristic neovascularization of PDR was not captured within the limited field of view of the images.[14]

In this review, the Medios AI software was primarily used for retinal image analysis, working seamlessly with both the Remidio FOP and Remidio NM-FOP applications installed on the smartphone used for capturing retinal images. By utilizing the smartphone's powerful CoreML and OpenGL capabilities, image processing occurred directly on the device's graphics processing unit (GPU), eliminating the need for an internet connection to a server. The AI algorithm was run offline by the technician on the smartphone itself after image capture. The technician was trained to retake images if the AI indicated poor quality. The AI initially assessed image quality, then provided a binary output such as DR present or No DR.[17] Furthermore, AL software enhances the captured images with visual map highlighting potential lesions on the retinal images, assisting healthcare providers in their assessments and educating patients about potential issues of their eyes.[13] Even though Remidio FOP only utilized 4 FOV compared to previous study with 7 FOV using Digital Fundus Camera (Zeiss FF450 Plus)[37], it produced high sensitivity of 92.7% (87.8–96.1) in grading Any DR, 87.9% (83.2–92.9) for stDR with high specificity of 98.4% (94.3–99.8) for any DR and 94.9% (89.7–98.2) for stDR.

EyeArt's high accuracy was also reflected in the UK's National DESP program, where it screened 30,000 patients across three regions. EyeArt showed a 95.7% (94.8 - 96.5%) sensitivity with 95% CI for detecting sight-threatening retinopathy, though its specificity was lower of 68% (67-69%) for no retinopathy, and 54.0% (53.4-54.5%) when combined with non-referable retinopathy. However, combined with its high sensitivity, EyeArt still provided significant cost savings for the NHS.[3] The large number of people screened also marked the reproducibility of EyeArt in greater population. While the results are promising, it is worth noting that these studies were carried out under ideal conditions where experienced professionals used desktop fundus cameras and had stable internet access. This may not be feasible in countries with limited resources, particularly in outreach settings where access to trained personnel, reliable internet, and expensive equipment may be limited.[12,21,26]

Tyson, et al. 2020[23] used smartphone with Retinascape[31] conjugated with EyeArt® (v2.0) system[30] has relatively lower score of intergrader agreement in detecting any DR (κ= 0.45 ± 0.33) compared to study by Rajalakshmi, et al. 2018[21] that used the same AI software (κ= 0.78; 0.7 - 0.86). The lower specificity reported by Tyson, et al. 2020[23] for detecting RDR of 78.6% (44.8 - 94.3%) per patient and 71.5% (48.7 - 86.9%) per eye is because the grading is based on gold-standard dilated examination by a retina specialist. Another reason is because EyeArt AI system was trained from a dataset of conventional retinal pictures, which may limit its ability to identify certain pathologies in smartphone images. Incorporating smartphone images into the training data could improve the algorithm's performance. Tyson, et al. 2020[23] emphasizes that studies that validate new screening modalities by comparing clinician grading of mobile device images to traditional images may have erroneously high sensitivity due to the assumption that both methods are equivalent while recent research has shown that there is a notable difference between the two methods.[38–40] Tyson also states that it is critical for researchers to rely dilated examination as the gold standard to make clinical diagnosis when validating the sensitivity and specificity of new diagnostic tools, particularly when using them in combination.

Rajalakshmi, et al. 2018[21] stated that retinal photography, evaluated and interpreted by eye doctors specializing in the retina or other trained professionals, is a widely acknowledged and established method for diabetic retinopathy screening.[37,41,42] However, there is shortage of trained professionals to evaluate retinal images in countries like India. If even available, busy schedules

have resulted in delays in providing DR grading and recommendations. These delays can result in miscommunication, loss of follow-up, and ultimately hinder the timely management of sight-threatening DR.[43] That is along the current trend of tele-ophthalmology and telemedicine,[44] use of AI software to evaluate retinal images for automated DR grading as it has potential to reduce the workload and costs for healthcare providers in screening the increasing number of individuals with diabetes,[45] resulting only those who have stDR and RDR would require appointment with the ophthalmologist or retina specialist.

Along the trend of using artificial intelligence (AI)[35] that accurately detect and grade DR in digital fundus images[13,21,34,46] combined with the feasibility of DR detection using smartphone-based fundus photography, Wroblewski, et al. 2023[17] from Mexico made a noninterventional, retrospective analysis to compare the diagnostic accuracy of the offline Medios AI software and online EyeArt AI software in detecting DR. The analysis used a single set of patient images taken with the Remidio-FOP camera in a field setting. Medios AI have a sensitivity of 94% and specificity of 94% for detecting any DR when including poor quality images, and 99% and 88% when excluding poor quality images, respectively. These results are comparable with previous study from India.[13,14] For EyeArt AI analysis, the sensitivities and specificities are 94% and 86% with all gradable images and 95% and 88% after excluding poor quality images. These results are similar to. those from previous studies detecting for any DR and stDR conducted in India,[21] English,[3] and America.[23] Surprisingly, the two AIs achieved similar levels of sensitivity and specificity, despite being trained using different methods. [15,47] International DR data set where Medios were trained from patients from Indian lineage[22,26,46] while American and Northern Mexican lineage were used for EyeArt.[15,48] It is imperative to note that from the same image sets using Remidio FOP, Medios AI was able to evaluate

all 248 patients despite having poor quality images, whereas EyeArt is lower with 63% (156/248) of patients. One of the reasons for the high ungradable rate by EyeArt is that the camera operators were not trained and unfamiliar with the specific image and patient criteria required by EyeArt.[21]

In 2022, Malerbi et al. made a study that used the Eyer camera from Phelcom Technologies to capture fundus images. These images were then analyzed remotely using the EyerCloud platform with a deep learning-enhanced method called PhelcomNet. This method assigned a prediction score (x) between 0 and 1, indicating the likelihood of diabetic retinopathy (DR). Notably, the device grouped "no DR" and "mild DR" together, while all other DR severities were considered "more than mild DR" in comparison to the reference standard.[27,33] Despite having high sensitivity of 97.8% and NPV 98.7% for mtmDR, it produced lower specificity of 61.4% and PPV of 48.5%. The team for obtaining images consist of a mix of trained professionals and inexperienced volunteers, captured images from over 900 individuals in six hours. Since image quality from portable devices depends heavily on the operator skill,[49] this combination of factors may have affected the quality of the images obtained in the study.[27]

The ease of use and portability of smartphone-based retinal cameras have been demonstrated by the successful acquisition of gradable images by a variety of operators, including healthcare workers [12,13] trained technicians,[14,22,26] medical interns and students,[17,23,27] with minimal to no experience in ophthalmic examination with reliable result of sensitivity and specificity. This highlights the accessibility and user-friendliness of smartphone-based retinal imaging compared to traditional tabletop fundus cameras, which are expensive and require specialized training.[50]

Dilated fundus examination by an ophthalmologist, the main screening method for DR screening, time-consuming as the patients need to

wait till their pupil is dilated, and accessibility to pharmacologic agents and ophthalmologist are not readily available. The gold standard method defined by the ETDRS group to screen photographically for DR is by stereoscopic color fundus photographs in seven standard fields, and this presents challenges in terms of time, accessibility, and cost. It requires a skilled photographer and special costly equipment, and is also time-consuming to the patient, limiting expanded use, especially in resource-constrained settings.[51,52]

While most study use pupil dilation for part of the screening process, Sosale *et al* (2020)[22,26] and Jain *et al* (2021)[12] opted for non-dilated using Remidio NM FOP. The sensitivity and specificity using Remidio NM FOP is considerably high, even exceeding mandated superiority cut-offs by FDA. These results are comparable with a study of Alfejri (2020)[52] that compare non mydriatic fundus camera with optical coherence tomography (OCT) in screening for DR. This study shows 75.2% (69.3-80.5) sensitivity, 96.0% (94.8-96.9) specificity, 75.8% (69.9-81.1) of PPV and 95.8% (94.7-96.8) NPV from a total 2406 patients in Riyadh, Saudi Arabia. This confirmed that using non mydriatic DR screening, both from smartphone or fundus camera can produce high sensitivity and specificity, allowing for more time-effective screening process and eliminating pharmacologic availability barrier in performing examination.

Indonesia, the country where the author came from, is the world's largest archipelagic country with a population of over 280 million, making it the fourth most populous country in the world.[53] A population-based study in Indonesia revealed that 43.1% of adults with type 2 diabetes in both urban and rural areas had DR, with 26.3% having the more severe VTDR.[45] However, Indonesian healthcare services still heavily focus on curative efforts rather than promotive and preventive services,[54] resulting limited study to focus on preventive strategy, such as DR screening.

In 2023, Lestari and colleagues conducted a study to evaluate the knowledge, attitudes, and practices of general practitioners (GP) in Jakarta regarding the screening of diabetic retinopathy in primary care settings and revealed a discrepancy between GPs' theoretical understanding and positive attitudes towards DR screening and their actual implementation of the practice, which was found to be poor. Most GPs referred patients for ophthalmic without attempting it themselves, believing DR screening was not their responsibility. Limited experience, lack of confidence in diagnosing fundus abnormalities, and lack of equipment in primary care also hinder GPs from conducting DR screening.[8] The author hoped with emergence of AI-integrated screening process combined with the convenience of smartphone to capture retinal images will expand healthcare coverage, particularly for preventive services.

The strength of this study is that we used comprehensive search methodology that was specified beforehand in our study's design and gave comparative details about different settings and methods used in selected studies that provide options if it were to be applied to a health care system. However, limitations are noted such as difference reference methods, difference definition of DR grading such as RDR, vtDR, and mtmDR. Secondly, we did not specify whether the AI software used in smartphone retinal imaging was compared to actual retinal examination by ophthalmologist or assessing the digital retinal image obtained from tabletop fundus camera. However, in DR screening, it is important use a system that can be implemented existing workflow especially in primary healthcare settings that facilitated expanded screening and early referral for increasing population of underserved diabetic patients to get intervention.[55]

CONCLUSION

The studies reviewed in this paper collectively represents the potential of smartphone-based integrated with AI in revolutionizing DR screening. The high sensitivity and specificity achieved by various AI algorithms, often exceeding the standards set by regulatory bodies like the FDA and ETDRS, highlight their accuracy in detecting DR and its severity levels. The accessibility and user-friendliness of smartphone-based retinal imaging further enhance DR screening coverage, particularly in underserved areas with limited resources and internet connectivity.

Funding and endorsement

None.

Conflicts of interest

This review had no conflicts of interest.

REFERENCES

1. Yau JW, Rogers SL & Kawasaki R, et al. Global prevalence and major risk factors of diabetic retinopathy. *Diabetes Care* **Vols. 35**, 556–564 (2012).

2. Chan, J., Malik, V. & Jia W, et al. Diabetes in Asia: epidemiology, risk factors, *JAMA* **Vols. 301**, 2129–2140 (2009).

3. Heydon, P. *et al.* Prospective evaluation of an artificial intelligence enabled algorithm for automated diabetic retinopathy screening of 30 000 patients. *Br J Ophthalmol* **Vols. 105**, 723–728 (2021).

4. Public Health England. Diabetic Eye Screening: Programme Overview. [Online] 2014. *https://www.gov.uk/guidance/diabetic-eye-screening-programme-overview* (2024).

5. Scanlon PH. The English national screening programme for diabetic retinopathy 2003-2016. *Acta Diabetol* **Vols. 54**, 515–525 (2017).

6. Alwazae M, Al Adel F, Alhumud A, Almutairi A & Alhumidan A. Barriers for adherence to diabetic retinopathy screening among Saudi adults. *Cureus* **Vol. 11:6454**, (2019).

7. Eppley SE, Mansberger SL, Ramanathan S & Lowry EA. Characteristics associated with adherence to annual dilated eye examinations among US patients with diagnosed diabetes. *Ophthalmology,* **Vols. 126**, 1492–1499 (2019).

8. Yeni Dwi Lestari, Gitalisa Andayani Adriono, Rizka Ratmilia, Christy Magdalena & Ratna Sitompul. Knowledge, attitude, and practice pattern towards diabetic retinopathy screening among general practitioners in primary health centres in Jakarta, the capital of Indonesia. *BMC Prim Care* **24**, 114 (2023).

9. Sicong Li, Ruiwei Zhao & Haidong Zou. Artificial intelligence for diabetic retinopathy. *Chin Med J (Engl)* **Vols. 135:**, 253–260 (2022).

10. Dankwa Mullan I *et al.* Transforming diabetes care through artificial intelligence: The future is here. *Popul Health Manag* **Vols. 22**, 229–242 (2019).

11. Meredith, S., Grinsven, M. & Engelberts, J. Performance of an artificial intelligence automated system for diabetic eye screening in a large English population. *Diabetic Medicine* **Vol. 40:e15055**, (2023).

12. Jain A, Krishnan R, Rogye A & Natarajan S. Use of offline artificial intelligence in a smartphone based fundus camera for community screening of diabetic retinopathy. . *Indian J Ophthalmol* **Vols.69:3150-4.**, (2021).

13. Natarajan, S., Jain, A., Krishnan, R. & Rogye, A. Diagnostic Accuracy of Community-Based Diabetic Retinopathy Screening with an Offline Artificial Intelligence System on a Smartphone. *JAMA Ophthalmol* **Vols.137(10)**, 1182–1188 (2019).

14. Sengupta S, Sindal MD, Baskaran P, Pan U & Venkatesh R. Sensitivity and specificity of smartphone based retinal imaging for diabetic retinopathy: A comparative study. *Ophthalmol Retina* **Vols. 019;3**, 146–53 (2019).

15. Bhaskaranand M, Ramachandra C & Bhat S, et al. Automated diabetic retinopathy screening and monitoring using retinal fundus image analysis. *J Diabetes SciTechnol* **Vol. 2016.**, (2016).

16. Wong TY, Sun J & Kawasaki R, et al. Guidelines on diabetic eye care: the international council of ophthalmology recommendations for screening, follow-up, referral, and treatment based on resource settings. *Ophthalmology* **2018;125(10)**, 1608–1622 (2018).

17. Wroblewski, J. *et al.* Diabetic Retinopathy Screening Using Smartphone-Based Fundus Photography and Deep-Learning Artificial Intelligence in the Yucatan Peninsula: A Field Study. *J Diabetes Sci Technol* (2023).

18. Wilkinson CP *et al.* Proposed International Clinical Diabetic Retinopathy and diabetic macular edema disease severity scales. *Ophthalmology* **Vols. 110**, 1677–82 (2003).

19. Core NDESP team. Guidance on standard feature-based grading forms to be used in the NHS diabetic eye screening programme. in *Diabetic Eye Screening Feature Based Grading* (2013).

20. Public Health England. NHS Diabetic Eye Screening Programme grading definitions for referable disease. Preprint at (2017).

21. Rajalakshmi, R., Subashini, R., Anjana, R. & Mohan, V. Automated diabetic retinopathy detection in smartphone-based fundus photography using artificial intelligence. *Springer Nature* **Vols. 32:**, 1138–1144 (2018).

22. Sosale, B., Sosale, A., Murthy, H., Sengupta, S. & Naveenam, M. Medios– An offline, smartphone based artificial intelligence algorithm for the diagnosis of diabetic retinopathy. *Indian J Ophthalmol* **68**, (2020).

23. Tyson, K. *et al.* Comparison of automated and expert human grading of diabetic retinopathy using smartphone-based retinal photography. *Springer Nature* **Vols. 35**, 334–342 (2020).

24. Early Treatment Diabetic Retinopathy Study Research Group. Grading diabetic retinopathy from stereoscopic color fundus photographs--an extension of the modified Airlie House classification. *Ophthalmology* **Vols:98**, 786–806 (1991).

25. Yang, Z., Tan, T., Shao, Y., Wong, T. & Li, X. Classification of diabetic retinopathy: Past, present and future. *Front Endocrinol* **Vol. 13: 1079217.**, (2022).

26. Sosale, B. *et al.* Simple, Mobile-based Intelligence Algorithm in the detection of Diabetic Retinopathy (SMART) study. *BMJ Open Diab Res Care* **Vol. 8**, (2020).

27. Malerbi, F. *et al.* Diabetic Retinopathy Screening Using Artificial Intelligence and Handheld Smartphone-Based Retinal Camera. *J Diabetes Sci Technol* **Vol. 16(3)**, 716–723 (2022).

28. Remidio. Medios DR. [Online]. https://www.eyenuk.com/en/products/eyeart/.

29. Remidio. Hand Held Fundus Camera | FOP NM-10. [Online]. https://www.remidio.com/products/fop.

30. EYENUK. EyeArt® AI Eye Screening System. [Online]. https://www.eyenuk.com/en/products/eyeart/.

31. Maamari RN, Keenan JD, Fletcher DA & Margolis TP. A mobile phone-based retinal camera for portable wide field imaging. *Br J Ophthalmol* **Vols. 2014;98**, 438–41 (2014).

32. Chollet. F. Xception: deep learning with depthwise separable convolutions. In: Proceedings of the IEEE conference on computer vision and pattern recognition. [Online]. https://openaccess.thecvf.com/content_cvpr_2017/papers/Chollet_Xception_Deep_Learning_CVPR_2017_paper.pdf (2017).

33. Phelcom. Eyer. [Online]. https://phelcom.com/products/eyer/.

34. Abràmoff MD, Lavin PT & Birch M. Pivotal trial of an autonomous AI-based diagnostic system for detection of diabetic retinopathy in primary care offices. *NPJ Digit Med* **Vol. 2018;1:39.**, (2018).

35. Gulshan V *et al.* Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. *JAMA* **Vols. 316**, 2402–10 (2016).

36. Vujosevic S, Aldington SJ & Silva P. Screening for diabetic retinopathy: new perspectives and challenges. *Lancet Diabetes Endocrinol* **Vols. 8**, 337–347 (2020).

37. Rajalakshmi R *et al.* Validation of smartphone based retinal photography for diabetic retinopathy screening. *PLoS One* **Vol. 0:e0138285**, (2015).

38. Toy BC *et al.* Smartphone-based dilated fundus photography and near visual acuity testing as inexpensive screening tools to detect referral warranted diabetic eye disease. *Retin Philos* **Vols. 36**, 1000–8 (2016).

39. Russo A, Morescalchi F, Costagliola C, Delcassi L & Semeraro F. Comparison of smartphone ophthalmoscopy with slit-lamp biomicroscopy for grading diabetic retinopathy. *Am J Ophthalmol* **Vols. 159**, 360–364 (2015).

40. Ryan ME *et al.* Comparison Among Methods of Retinopathy Assessment (CAMRA) study: smartphone, nonmydriatic, and mydriatic photography. *Ophthalmology* **Vols. 122**, 2038–43 (2015).

41. Bhargava M *et al.* Accuracy of diabetic retinopathy screening by trained non-physician graders using non-mydriatic fundus camera. *Singapore Med J* **Vols. 53**, 715–19 (2012).

42. Raman R, Padmaja RK & Sharma T. The sensitivity and specificity of non-mydriatic digital stereoscopic retinal imaging in detecting diabetic retinopathy. *Diabetes Care* **Vol. 30:1**.

43. Rachapelle S *et al.* The cost-utility of telemedicine to screen for diabetic retinopathy in India. *Ophthalmology* **Vols. 120**, 566–73 (2013).

44. Mohan V, Prathiba V & Pradeepa R. Tele-diabetology to screen for diabetes and associated complications in rural India: The Chunampet Rural Diabetes Prevention Project Model. *J Diabetes SciTechnol* **Vols. 2017;181**, 79–87 (2014).

45. Sasongko MB *et al.* Prevalence of diabetic retinopathy and blindness in indonesian adults with type 2 diabetes. *Am J Ophthalmol* (2017).

46. Gulshan V, Rajan RP & Widner K, et al. Performance of a deeplearning algorithm vs manual grading for detecting diabetic retinopathy in India. *JAMA Ophthalmol* **Vols. 137(9)**, 987–993 (2019).

47. Rogan WJ & Gladen B. Estimating prevalence from the results of a screening test. *Am J Epidemiol* **Vols. 107(1)**, 71–76 (1978).

48. Cuadros J & Bresnick G. EyePACS: an adaptable telemedicine system for diabetic retinopathy screening. *J Diabetes SciTechnol* **Vols. (3)**, 509–516 (2009).

49. Davila JR, Sengupta SS & Niziol LM, et al. Predictors of photographic quality with a handheld nonmydriatic fundus camera used for screening of vision-threatening diabetic retinopathy. *Ophthalmologica* **Vols. 238**, 89–99 (2017).

50. Graham-Rowe E, Lorencatto F & Lawrenson JG, et al. Barriers to and enablers of diabetic retinopathy screening attendance: a systematic review of published and grey literature. *Diabet Med* **Vols. 35**, 1308–19 (2018).

51. arly treatment diabetic retinopathy study research group. Early photocoagulation for diabetic retinopathy. ETDRS report number 9. *Ophthalmology* **Vols. 98**, 766–85 (1991).

52. Aljefri, S. & Adel, F. The validity of diabetic retinopathy screening using nonmydriatic fundus camera and optical coherence tomography in comparison to clinical examination. *Saudi J Ophthalmol* **Vols. 34**, 266–72 (2020).

53. Badan Pusat Statistik. Jumlah Penduduk Pertengahan Tahun (Ribu Jiwa), 2022-2024. [Online] . https://www.bps.go.id/id/statistics-table/2/MTk3NSMy/jumlah-penduduk-pertengahan-tahun--ribu-jiwa-.html. (2024).

54. Darmawansyah. Upaya Kesehatan di Indonesia: Tantangan dan Harapan. *Jurnal AKK* **Vol 2 No 2, Mei,** (2013).

55. Jéssica Deponti Gobbi *et al.* Efficacy of smartphone based retinal photography by undergraduate students in screening and early diagnosing diabetic retinopathy. *Int J Retina Vitreous* **Vol. (2022) 8:35**, (2022).