











BMJ Open Evaluation of non-ophthalmologist-led and offline AI-assisted models for diabetic retinopathy screening in India: a pragmatic diagnostic accuracy study

Anshul Chauhan ¹, Luke Vale ², Ankita Kankaria ³, Basavaraj Tigari ¹, Sonam Kumar ¹, Mitesh Yadav ¹, Lakshay Kumar ¹, Gagandeep Kaur ¹, Sanjay Kumar Bhadada ⁴, Vishali Gupta,¹ Mona Duggal ⁵

To cite: Chauhan A, Vale L, Kankaria A, *et al.* Evaluation of non-ophthalmologist-led and offline AI-assisted models for diabetic retinopathy screening in India: a pragmatic diagnostic accuracy study. *BMJ Open* 2026;**16**:e106397. doi:10.1136/bmjopen-2025-106397

► Prepublication history and additional supplemental material for this paper are available online. To view these files, please visit the journal online (<https://doi.org/10.1136/bmjopen-2025-106397>).

Received 11 June 2025
Accepted 24 February 2026



© Author(s) (or their employer(s)) 2026. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ Group.

For numbered affiliations see end of article.

Correspondence to

Dr Mona Duggal;
monaduggal2@gmail.com

ABSTRACT

Objectives To assess and compare the diagnostic accuracy of non-ophthalmologist-led diabetic retinopathy screening (DRS) at health and wellness centres (HWCs) and offline artificial intelligence (AI)-assisted community-based screening, using specialist grading as the reference standard in India.

Design, settings and participants Pragmatic diagnostic accuracy study in primary healthcare settings. The settings included HWCs and community-based screening sites in rural Block Boothgarh, Mohali District, Punjab, India. A total of 600 people with diabetes aged ≥ 30 years were enrolled across three screening models: (1) non-ophthalmologist-led DRS at the HWC, (2) AI-assisted smartphone-based DRS in the community and (3) standard referral-based care. Retinal images were captured using non-mydriatic fundus cameras and independently graded by two masked human graders; a senior retina specialist resolved any disagreements. The AI was assessed for its ability to detect diabetic retinopathy (DR) and referable diabetic retinopathy (RDR). Diagnostic performance metrics were reported.

Results The non-ophthalmologist-led model demonstrated 86.4% sensitivity (95% CI 65.1% to 97.1%) and 94.3% specificity (95% CI 88.5% to 97.7%) for DR detection, with an ungradability rate of 8%. For RDR, sensitivity reached 95.8% (95% CI 78.9% to 99.9%) and specificity was 93.1% (95% CI 88.0% to 96.5%). The offline AI-assisted model achieved 93.3% sensitivity (95% CI 68.1% to 99.8%) and 85.1% specificity (95% CI 76.9% to 91.2%) for RDR, but with a higher ungradability rate (38%), mainly due to cataracts and poor image quality. Both approaches effectively identified referable cases; however, the non-ophthalmologist-led model demonstrated greater accuracy and operational feasibility.

Conclusions This study demonstrates that non-ophthalmologist-led DRS at HWCs can enhance access to primary care. Offline AI-enabled screening demonstrates potential for community use but is currently limited by image quality and binary classification outputs. Integrating both approaches may strengthen DRS coverage in resource-limited settings.

Clinical trials registry of India CTRI/2022/10/046283.

STRENGTHS AND LIMITATIONS OF THIS STUDY

- ⇒ Conducted in real-world primary and community healthcare settings in rural Punjab, enhancing the pragmatic relevance of the findings.
- ⇒ A non-ophthalmologist was formally trained and certified at a tertiary retina centre to perform retinal imaging, demonstrating the feasibility of task shifting in resource-limited contexts.
- ⇒ Both benchtop and handheld non-mydriatic fundus cameras were evaluated, allowing comparison of screening performance across facility-based and community-based delivery models.
- ⇒ The offline artificial intelligence provided only a binary referable diabetic retinopathy (RDR)/non-RDR output, without grading severity or diabetic macular oedema, and excluded ungradable images, thereby limiting its applicability in populations with a high cataract burden or small pupils.

INTRODUCTION

Diabetic retinopathy (DR) is a major microvascular complication of diabetes and a rapidly growing global public health concern.¹ In India, the prevalence of DR is estimated at 18% among people with diabetes, significantly higher than the global average of 9.3%.²⁻³ The global diabetes burden is projected to reach 700 million by 2045⁴ and requires early detection of DR through scalable and efficient screening. This is critical to prevent vision loss and reduce long-term healthcare costs.^{5,6}

Despite its importance, diabetic retinopathy screening (DRS) remains insufficient in many low-resource settings.¹ Conventional models, which rely on trained graders, are costly, time-consuming and challenging to scale.¹⁻⁶ In contrast, artificial intelligence (AI)-based DRS offers automated, rapid and cost-effective screening with the potential to reach underserved populations more

efficiently.⁷⁻⁹ Community-based smartphone AI retinal imaging offers a potential pathway to reach individuals in their homes.^{10 11} Offline AI systems may offer particular advantages in remote settings, where limited connectivity and non-mydriatic (without pupil dilation) image quality can compromise diagnostic accuracy.^{7 12} While AI algorithms have demonstrated over 90% sensitivity and specificity in tertiary settings, their application in primary and community care remains underexplored.^{13 14} Real-world evaluation is therefore essential to assess effectiveness across clinical settings with varying disease prevalence, image quality and patient characteristics that may differ from those in the training dataset.^{15 16} These challenges can lead to misclassification and incorrect referrals, limiting the effectiveness of AI in primary and community care.¹⁷

This diagnostic accuracy study compares DRS performance across primary healthcare settings: (1) health and wellness centre (HWC)¹⁸ (2) community (home) and (3) standard referral-based screening. It assesses the sensitivity and specificity of DR and referable diabetic retinopathy (RDR) detection and evaluates the feasibility

of deploying AI-integrated smartphone fundus cameras operated by non-ophthalmologists.

METHODS

This study adhered to the Standards for Reporting Diagnostic Accuracy Studies guidelines to ensure methodological rigour and transparent reporting of diagnostic accuracy findings.¹⁹ An overview of the study design is presented in figure 1.

Study design and participants

This diagnostic accuracy assessment is embedded within a larger pragmatic trial guided by the Reach, Effectiveness, Adoption, Implementation, Maintenance implementation framework, which systematically evaluates the real-world implementation, adoption and effectiveness of DRS strategies in primary care settings.²⁰ While the pragmatic trial explores a comprehensive range of outcomes, including screening coverage, follow-up compliance and integration within primary health systems, this paper focuses exclusively on the diagnostic accuracy

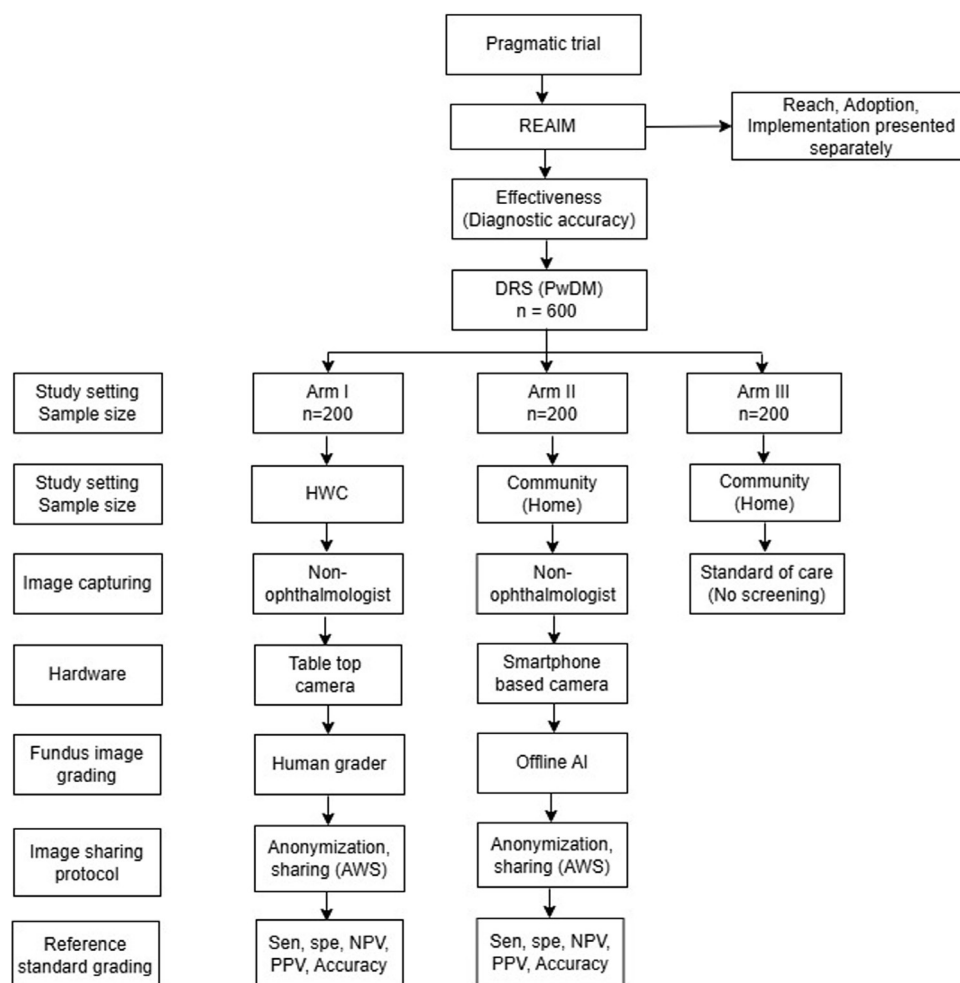


Figure 1 Study flow chart. AI, artificial intelligence; AWS, Amazon Web Services; DRS, diabetic retinopathy screening; HWC, health and wellness centre; NPV, negative predictive value; PPV, positive predictive value; PwDM, people with diabetes mellitus; REAIM, Reach, Effectiveness, Adoption, Implementation, Maintenance; Sen, sensitivity; Spe, specificity.

of AI-assisted and non-ophthalmologist-led screening approaches (online supplemental figure 1).

This study was conducted in primary healthcare settings at Block Boothgarh, Mohali, Punjab, India, between February 2023 and January 2024. It enrolled people with diabetes mellitus (PwDM) aged ≥ 30 , excluding those with conjunctivitis, red eyes or active ocular inflammation, precluding image quality and patient cooperation.²⁰ Participants (n=600) were allocated to each arm: Arm I—non-ophthalmologist-led screening at HWC Khijrabad, Arm II—AI-assisted screening in the community (home) and Arm III—standard care (referral counselling without active screening) in a community setting. Following screening, participants diagnosed with RDR were referred to the district hospital in Arms I and II. The referral was based on the reference standard grading. In contrast, all participants in Arm III were referred irrespective of their screening results. This study will solely present the results of the diagnostic accuracy test. The reach, implementation and adoption, as well as screening outcomes and referral follow-up, will be reported separately.

Patient and public involvement

Patients or the public were not involved in this research's design, conduct, reporting or dissemination.

Sample size

Sample size estimation was based on a binary outcome (DR: yes/no), assuming equal group allocation, $\alpha=0.05$ (1.96), $\text{power}=0.80$ ($\beta=0.20$),²¹ with expected DR prevalence: $p_1=0.16$, $q_1=0.84$ (human grading),²² and $p_2=0.28$, $q_2=0.75$ (AI screening).²³ Based on prior studies

evaluating DRS interventions in similar settings, the dropout rate was estimated at 15%. Therefore, the final required sample size was 600 PwDM to account for potential participant attrition and ungradable images.

Image acquisition and training

A non-ophthalmologist with a bachelor's degree in optometry acquired fundus images after a structured 15-day training programme at Postgraduate Institute of Medical Education and Research's Advanced Eye Centre retina clinic. The training covered image acquisition using two non-mydratic fundus cameras: Forus 3 Nethra Classic (benchtop),^{24 25} and Remidio NM FOP 10 (smartphone-based, integrated with offline Medios AI).^{26 27} The Forus camera was used at HWCs, while the portable fundus on phone (FOP) was preferred in community settings due to its mobility and battery backup²⁶ (figure 1).

The non-ophthalmologist captured 320 retinal images (80 participants) in a retina clinic for proficiency assessment during the training. Their performance was evaluated based on predefined quality metrics,²⁸ ensuring that at least 90% of captured images met the gradability criteria. Participants were instructed to retake images if the initial captures were inadequate. Additional manufacturer-led training supplemented this learning. On-site training was conducted at HWC, where 30 test images were captured before recruitment and were excluded from the final sample.

At HWC Khijrabad, darkroom conditions were optimised by covering the windows, using black chart paper and installing thick curtains. In community settings,

Table 1 Sociodemographic characteristics of the study participants (arm-wise and overall)

Characteristics	Overall (N=600)	Arm I (HWC, N=200)	Arm II (community, N=200)	Arm III (community, N=200)
Age (years), mean \pm SD	58.22 (11.52)	57.9 (11.41)	58.22 (11.45)	58.55 (11.74)
Age (years), median (IQR)	59 (50–65)	56.5 (50–65)	60 (50–67)	60 (50–65)
Age categories (years), n (%)				
≤40	46 (7.65)	10 (5)	19 (9.5)	17 (8.5)
41–50	119 (19.80)	53 (26.5)	32 (16)	34 (17)
51–60	181 (30.12)	56 (28)	61 (30.5)	64 (32)
61–70	177 (29.45)	56 (28)	64 (32)	57 (28.5)
>70	78 (12.98)	25 (12.5)	24 (12)	28 (14)
Gender, n (%)				
Male	245 (40.77)	84 (42)	73 (36.5)	88 (44)
Female	355 (59.23)	116 (58)	127 (63.5)	112 (56)
Diabetes duration (years), n (%)				
≤5	286 (48.31)	96 (50)	92 (46.23)	98 (49)
5.1–10	170 (28.72)	53 (27.6)	62 (31.16)	54 (27)
10.1–15	89 (15.03)	29 (15.1)	28 (14.07)	32 (16)
≥15.1	47 (7.94)	14 (7.29)	17 (8.54)	16 (8)

HWC, health and wellness centre.



smartphone-based imaging was performed in dimly lit areas of participants' households. Basic demographic data, including age, sex and diabetes duration, were collected before imaging.

Reference standard grading

Data confidentiality was ensured by anonymising all patient identifiers before image storage and analysis. Deidentified images were securely uploaded and stored on a cloud-based platform (Amazon Web Services (AWS)). Fundus images were deidentified and uploaded to a secure cloud-based platform, AWS, for remote grading. Three masked human graders (HGs) independently graded the image: a certified optometrist (HG1), an ophthalmologist (HG2) and a senior retina specialist (HG3). HG3 adjudication identified discrepancies between HG1 and HG2, establishing a consensus agreement for finalising DR grading decisions (figure 1).

The graders affiliated with different institutions were blinded to each other's and AI diagnoses to prevent bias. DR lesions were classified as either DR or no DR, with further grading into mild non-proliferative diabetic retinopathy (NPDR), moderate NPDR, severe NPDR or proliferative DR, following the International Classification of Diabetic Retinopathy (ICDR).²⁹ Diabetic macular oedema (DME) was defined as retinal thickening within one disc diameter of the foveal centre.²⁵ RDR included moderate NPDR and above or any DME. Images were considered gradable if at least 80% of the retina was visible and the third vascular branch could be assessed.²⁸ Participants with ungradable images were also included for referral to a higher centre for ophthalmological opinion. Both eye- and patient-level analyses were conducted.

AI grading workflow

The Medios offline deep-learning AI software was used to automate the detection of DR. The AI algorithm was embedded within the smartphone fundus camera control app, enabling real-time analysis.^{23 26} The non-ophthalmologist ran the AI algorithm offline postimage acquisition, generating two outputs: 'no signs of DR detected' (non-RDR) or 'signs of DR detected' (RDR). AI-based grading preceded HG assessment. Misclassified cases were identified post hoc, and their characteristics were analysed to assess potential failure points in AI-based detection.

Statistical analysis

Data were collected electronically using Research Electronic Data Capture³⁰ and exported for analysis in STATA V.15 SE.³¹ Continuous variables were summarised as means and SD. The diagnostic accuracy of AI and human grading was assessed via 2×2 contingency tables comparing each method to the reference standard. Sensitivity, specificity, positive predictive value (PPV) and negative predictive value (NPV) were computed with 95% CIs. Inter-rater agreement between AI and HGs was evaluated

Table 2 DR screening outputs at HWC and community settings (eye-wise analysis)

Screening output	Arm I (HWC, N=400)	Arm II* (community, N=362)
Image gradability, n (%)		
Gradable	367 (91.75)	224 (62)
Ungradable	33 (8.25)	138 (38)
DR, n (%)		
Yes	110 (28.72)	40 (17.86)
No	257 (70.03)	184 (82.14)
DR grades, n (%)		
Mild NPDR	70 (63.64)	–
Moderate NPDR	31 (28.18)	–
Severe NPDR	6 (5.45)	–
PDR	3 (2.73)	–
DME, n (%)		
Yes	21 (5.72)	–
No	346 (94.28)	–
Referral pattern, n (%)		
Referral	40 (10.9)	40 (17.86)
Non-referral	327 (89.1)	184 (82.14)
*Arm II screening was conducted by the Medios AI algorithm, which does not provide DR grades. DME, diabetic macular oedema; DR, diabetic retinopathy; HWC, health and wellness centre; NPDR, non-proliferative diabetic retinopathy; PDR, proliferative diabetic retinopathy.		

using Cohen's kappa. Statistical significance was set at $p < 0.05$.

RESULTS

Table 1 presents the sociodemographic characteristics of the study population. Of the 600 participants, 355 (59%) were female. The mean age was 58.2 ± 11.5 years, with a median of 59 years (IQR 50–65). Most participants (60%) were aged 51–70 years, and nearly half (48%) had diabetes for ≤ 5 years. Arm I had a notably higher proportion of participants aged 41–50 (26.5%) than Arms II and III (16% and 16.9%, respectively).

DRS outputs

Table 2 presents the DRS outputs across HWC and community settings. Arm I had a lower proportion of ungradable images (8.3%) than Arm II (38%). AI in Arms II misclassified 38 ungradable images as DR, primarily due to poor image quality, necessitating their exclusion from the final analysis ($n=362$). Among the eyes screened, 40 were identified for referral in both Arms I and II. The AI system did not provide DME grading; in Arm I, DME was detected in 5.7% of eyes.

Table 3 Diagnostic accuracy of human grader-based DR screening at the health and wellness centre (per patient analysis)

Diagnostic parameter	Sensitivity (95% CI)	Specificity (95% CI)	PPV (95% CI)	NPV (95% CI)	Accuracy (95% CI)	Kappa (p value)
DR	86.36% (65.09% to 97.09%)	94.26% (88.54% to 97.66%)	86.27% (75.06% to 92.92%)	86.47% (81.20% to 90.43%)	86.41% (80.60% to 91.01%)	0.68 (<0.001)
RDR	95.83% (78.88% to 99.89%)	93.12% (88.03% to 96.52%)	67.65% (54.02% to 78.8%2)	99.33% (95.63% to 99.90%)	93.48% (88.89% to 96.59%)	0.75 (<0.001)

DR, diabetic retinopathy; NPV, negative predictive value; PPV, positive predictive value; RDR, referable diabetic retinopathy.

Diagnostic performance of DR and RDR at HWC

Table 3 presents the diagnostic performance of HG-based DRS at the HWC. The sensitivity for DR detection was 86.4% (95% CI 65.1% to 97.1%), with a specificity of 94.3% (95% CI 88.5% to 97.7%). The PPV and NPV were 86.3% (95% CI 75.1% to 92.9%) and 86.5% (95% CI 81.2% to 90.4%), respectively. The inter-rater agreement (kappa) for DR diagnosis was moderate ($\kappa=0.68$, $p<0.001$), while the agreement for RDR was substantial ($\kappa=0.75$, $p<0.001$).

For RDR detection, sensitivity was 95.8% (95% CI 78.9% to 99.9%), and specificity was 93.1% (95% CI 88.0% to 96.5%). The high specificity suggests minimal false positives (FPs), supporting the reliability of HG-based screening for RDR identification in primary care settings. The high NPV (99.3%) further indicates the system's strength in ruling out RDR among screened individuals.

Diagnostic performance of DR and RDR in community settings

In Arm II, the sensitivity for DR detection decreased from 80.95% (95% CI 58.01% to 94.5%) to 73.33% (95% CI 44.90% to 92.21%) when misclassified images were excluded. Similarly, for RDR detection, sensitivity decreased from 93.33% (95% CI 68.05% to 99.83%)

to 88.89% (95% CI 51.75% to 99.72%) after excluding misclassified images. Notably, specificity increased for DR and RDR, indicating a lower FP rate in AI-based detection after misclassified images were excluded. The AI algorithm exhibited weak inter-rater agreement (kappa) for diagnosing DR and RDR (table 4).

DISCUSSION

This study is the first to evaluate the diagnostic accuracy of a DRS programme across primary healthcare centres, HWC and community settings in a real-world context. Implementing DRS at HWCs aligns with India's broader healthcare strategy of integrating eye care services into primary care.^{32 33} Nearly 70% of India's population resides in rural areas, and the ophthalmologist-to-population ratio remains critically low at approximately 1 per 100 000 individuals.²³ However, establishing and sustaining such programmes requires significant investments in infrastructure, resource allocation and workforce training to capture high-quality images and attain high proficiency levels.³⁴ This study demonstrates that a non-ophthalmologist trained in a structured programme

Table 4 Diagnostic accuracy of artificial intelligence-based DR screening in community settings (per patient analysis, with and without misclassified images)

Diagnostic parameter	DR		RDR	
	Including misclassified images	Excluding misclassified images	Including misclassified images	Excluding misclassified images
Sensitivity (95% CI)	80.95% (58% to 94.5%)	73.33% (44.90% to 92.21%)	93.33% (68.05% to 99.83%)	88.89% (51.75% to 99.72%)
Specificity (95% CI)	87.13% (79% to 92.96%)	93.62% (86.62% to 97.62%)	85.05% (76.86% to 91.20%)	91% (83.60% to 95.80%)
PPV (95% CI)	56.67% (43.05% to 69.35%)	64.71% (44.37% to 80.82%)	46.67% (35.32 to 58.37)	47.06% (31.38% to 63.34%)
NPV (95% CI)	95.65% (90% to 98.12%)	95.65% (90.47% to 98.08%)	98.91% (93.19 to 99.84)	98.91% (93.47% to 99.83%)
Accuracy (95% CI)	86.07% (78.63% to 91.67%)	90.83% (83.77% to 95.51%)	86.07% (78.63% to 91.67%)	90.83% (83.77% to 95.51%)
Kappa (p value)	0.58 (<0.001)	0.63 (<0.001)	0.55 (<0.001)	0.57 (<0.001)

DR, diabetic retinopathy; NPV, negative predictive value; PPV, positive predictive value; RDR, referable diabetic retinopathy.



can achieve high diagnostic accuracy in detecting DR, supporting the feasibility of task shifting in primary care settings.³⁵

Unlike previous studies that used specialised optometrists, reporting sensitivities of 73%,³⁶ and specificities of 97%,³⁷ our study employed a non-ophthalmologist trained in retinal imaging at a tertiary centre and an HWC. This approach achieved a sensitivity of 86.36% (95% CI 65.09% to 97.09%) and a specificity of 94.26% (95% CI 88.54% to 97.66%). As reported in previous studies, image quality significantly influences screening performance.^{38–39} Poor-quality images remain a significant challenge in non-mydriatic imaging. Operator training influences them, and uncontrolled environmental conditions in community-based settings exacerbate the problem.^{38–40}

Ungradable images increase false negatives (FNs), reduce sensitivity and may hinder the effectiveness of DRS programmes.⁴⁰ The situation becomes more problematic when non-mydriatic systems are used in regions with a high prevalence of cataracts.³⁴ In this study, ungradability rates were 8% in a darkroom-controlled HWC setting but substantially higher (38%) in home-based screenings, exceeding the 18–30% range reported in other low- and middle-income countries (LMICs) studies.^{26–34} Factors such as patient age, diabetes duration and ocular opacities like cataracts contributed to image gradability issues.⁴¹ While pupillary dilation is known to enhance image quality in DRS,⁴² this study opted for non-mydriatic imaging due to logistical constraints, the absence of ophthalmologists and concerns over patient discomfort.^{26–34} Although staged mydriasis with 1% tropicamide may be considered to enhance image quality for ungradable images, it is not staged mydriasis.⁴³ Online supplemental table 1 presents a cross-tabulation of cataract status and ungradable images across three study arms.

Online supplemental table 2 compares DRS methodologies across different healthcare settings. This study's HWC-based screening demonstrated high diagnostic

accuracy, with a sensitivity of 86.36% and specificity of 94.26%. By contrast, slit-lamp examinations in tertiary settings exhibit variable performance, with sensitivities ranging from 67% to 73% and specificities ranging from 77% to 90%.^{36–37–44–45} These findings highlight the potential of non-mydriatic imaging as a scalable approach for DRS in primary care.

Additionally, this study evaluated an offline AI-based DRS algorithm deployed via a smartphone fundus camera in community settings. This approach offers a potential solution to the shortage of trained ophthalmologists, particularly in rural and underserved regions.⁴⁶ The Medios AI offline algorithm generated screening results without internet connectivity, a crucial advantage in resource-limited settings. Prior studies in India^{23–46–47} have assessed the accuracy of offline AI-based DRS, though methodological differences limit direct comparisons. Table 5 presents the diagnostic performance of the Medios AI algorithm across different studies. Sensitivity is a critical parameter for patient safety, as AI-based screening must effectively identify PwDM requiring referral.⁴⁸

Our findings indicate that offline AI screening in primary care settings met and exceeded Food and Drug Administration (FDA) regulatory thresholds, achieving an RDR sensitivity of 93.33% (>85%) and specificity of 85% (>82.5%). A study by Natarajan *et al*^{23–46} reported 100% sensitivity for RDR detection when ungradable images were included, though specificity dropped from 88.4% to 81.9%. Some studies have reported higher sensitivity for mydriatic imaging compared with non-mydriatic approaches for DR detection,^{3–10–17} while others have found no significant difference, with sensitivity remaining at 86% (95% CI 85% to 87%) even after dilation.³⁴

While the effectiveness of automated AI for detecting DR is best evaluated in diverse clinical or population settings,⁴⁶ which was beyond the scope of our study. Despite its promise, AI-based DRS must account for misclassification risks, particularly FPs. In this study, while RDR detection exceeded FDA thresholds, the AI

Table 5 Comparison of offline AI performance for referable diabetic retinopathy detection in different healthcare settings

	AI software	Settings	Pupil status	Image capturing	Sensitivity (%)	Specificity (%)
FDA cut-off					85	82.5
Present study: including misclassified images	Medios AI	Primary (community)	Non-dilated	Community optometrist	93.33	85.05
Present study: excluding misclassified images	Medios AI	Primary (community)	Non-dilated	Community optometrist	88.89	91
Sosale <i>et al</i> ²⁶	Medios AI	Tertiary	Dilated	Trained technician	98.8	86.7
Natarajan <i>et al</i> ²³	Medios AI	Primary (dispensary)	Dilated	Healthcare worker	100	88.4
Jain <i>et al</i> ⁴⁶	Medios AI	Primary (dispensary)	Dilated	Healthcare worker	100	89.55

AI, artificial intelligence; FDA, Food and Drug Administration.

algorithm misclassified 19/200 cases (9.5%) as RDR. In our research, image ungradability was primarily attributed to cataract (online supplemental table 1), meiosis, older age, macular scars and the inability to control daylight conditions at home (table 2). FP cases can burden healthcare systems by increasing unnecessary referrals and patient anxiety.⁴⁹ Conversely, FNs pose a greater risk because they result in missed treatment referrals, increasing the likelihood of preventable vision loss.^{46 50} In some cases, images labelled as ungradable were still identified by the Medios platform as showing 'signs of retinopathy detected' (online supplemental figures 2 and 3). The AI system in this study demonstrated a high NPV of 99%, minimising FN-related screening failures and aligning with previous Indian screening studies.^{23 46} While AI has strong diagnostic potential, its integration into clinical workflows must be carefully planned to strike a balance between sensitivity and specificity, while minimising disruption to healthcare services.^{51 52}

A key strength of this study is its demonstration of the feasibility of DRS at HWCs, supporting India's broader goal of integrating eye care into primary healthcare. AI-based screening may provide a scalable solution to standardise DR detection and reduce intergrader variability.⁵³ AI's ability to process large volumes of images rapidly enhances efficiency. In LMICs such as India, where internet connectivity and power supply are inconsistent, offline AI algorithms may ensure uninterrupted screening programmes.^{26 49} Future research should explore AI's integration into public health systems, including its potential to supplement or replace HCs. However, AI misclassification of ungradable images as DR could lead to unnecessary referrals and patient anxiety, particularly in populations with high rates of cataracts, small pupils and advanced age.⁴⁰

This study has limitations, including a relatively small sample size for community-based screening. While the sample size may limit generalisability, it provides valuable preliminary data on AI performance in primary care settings, which future large-scale studies can build on. This study did not explore the integration of DRS into HWCs as outlined in India's comprehensive eye care guidelines.³² Although the overall sample size was adequate, the small number of DR/RDR cases in some arms, particularly Arm II, produced wide CIs and reduced the precision of the sensitivity estimate. Prior research has reported high ungradability rates (38%) for non-mydratic imaging in Indian populations, mainly due to cataract comorbidities and smaller mesopic pupil sizes, which may necessitate pupillary dilation for accurate RDR detection.⁵⁴ Furthermore, the offline AI version used in this study does not grade DR severity based on the ICDR scale. This limits its clinical applicability, as management and referral decisions differ across stages of NPDR, highlighting the need for algorithms capable of multigrade classification. Additionally, repeated non-mydratic imaging was challenging during community screenings

due to pupillary constriction caused by the camera flash, which degraded image quality.

CONCLUSION

The study employs a comprehensive DRS strategy in primary healthcare facilities, demonstrating the diagnostic accuracy of non-ophthalmologist-led screening at HWCs using non-mydratic fundus cameras. These findings highlight the need for further research on training and involving Community Health Officers⁵⁵ in DRS. Our study also underscores the effectiveness of AI in screening for PwDM in primary care, achieving high sensitivity and specificity for DR and RDR detection. Handheld fundus cameras in these settings meet US FDA diagnostic thresholds, supporting their feasibility for broader implementation. From a patient perspective, expanding DRS will enable earlier diagnosis and timely intervention and prevent the debilitating effects of sight loss. Further algorithm training on ungradable images is recommended to optimise AI performance and reduce FP referrals, rather than relying solely on AI-based decisions. While non-ophthalmologist grading and AI-assisted screening demonstrate strong diagnostic accuracy, a cost-effectiveness analysis is needed to inform evidence-based policy decisions on scaling DRS interventions.

Author affiliations

¹Advanced Eye Centre, Post Graduate Institute of Medical Education and Research, Chandigarh, India

²Global Health Economics Centre, London School of Hygiene & Tropical Medicine, London, UK

³Department of Community Medicine and School of Public Health, Post Graduate Institute of Medical Education and Research, Chandigarh, India

⁴Department of Endocrinology, Post Graduate Institute of Medical Education and Research, Chandigarh, India

⁵National Institute for Research in Digital Health and Data Science, Indian Council of Medical Research, New Delhi, Delhi, India

Acknowledgements We thank the Punjab Health System for facilitating interviews with patients and healthcare providers. We also sincerely thank the ASHA workers and the Senior Medical Officer for their assistance in contacting the PwDM during the study.

Contributors AC, MD and LV contributed to the study's conception and design. GK, LK, SK and MY assisted with data collection, transcription and translation. AC prepared the initial draft, while MD, LV, AK, VG, BT and SKB reviewed it and provided significant revisions. MD is the guarantor.

Funding This research was funded by the Indian Council of Medical Research (ICMR) under approval number 5/4/6/13/OPH/2020-NCD-II. The funders have no role in our research's study design, conduct, reporting or dissemination plans.

Competing interests None declared.

Patient and public involvement Patients and/or the public were not involved in the design, conduct, reporting or dissemination plans of this research.

Patient consent for publication Consent obtained directly from patient(s).

Ethics approval This study involves human participants and adhered to the Declaration of Helsinki and received approval from the Institutional Ethics Committee of Postgraduate Institute of Medical Education and Research (PGIMER), Chandigarh (PGI/IEC/2020/000741). Before enrolment, written informed consent was obtained from all participants. The protocol was prospectively registered with the Clinical Trials Registry India (CTRI/2022/10/046283).

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement All data relevant to the study are included in the article or uploaded as supplementary information. The datasets generated and/or analysed during the current study are not publicly available due to ethical considerations and confidentiality agreements related to state health systems data, but are available from the corresponding author at a reasonable request.

Supplemental material This content has been supplied by the author(s). It has not been vetted by BMJ Publishing Group Limited (BMJ) and may not have been peer-reviewed. Any opinions or recommendations discussed are solely those of the author(s) and are not endorsed by BMJ. BMJ disclaims all liability and responsibility arising from any reliance placed on the content. Where the content includes any translated material, BMJ does not warrant the accuracy and reliability of the translations (including but not limited to local regulations, clinical guidelines, terminology, drug names and drug dosages), and is not responsible for any error and/or omissions arising from translation and adaptation or otherwise.

Open access This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: <https://creativecommons.org/licenses/by-nc/4.0/>.

ORCID iDs

Anshul Chauhan <https://orcid.org/0000-0002-2434-935X>

Luke Vale <https://orcid.org/0000-0001-8574-8429>

Ankita Kankaria <https://orcid.org/0000-0002-4641-7333>

Basavaraj Tigari <https://orcid.org/0000-0002-5486-263X>

Sonam Kumar <https://orcid.org/0009-0008-7257-9644>

Mitesh Yadav <https://orcid.org/0009-0003-3293-2152>

Lakshay Kumar <https://orcid.org/0009-0007-4026-9935>

Gagandeep Kaur <https://orcid.org/0009-0003-0403-8997>

Sanjay Kumar Bhadada <https://orcid.org/0000-0002-0410-8778>

Mona Duggal <https://orcid.org/0000-0002-9404-2871>

REFERENCES

- Grzybowski A, Brona P, Lim G, *et al*. Artificial intelligence for diabetic retinopathy screening: a review. *Eye (Lond)* 2020;34:451–60.
- Gulshan V, Peng L, Coram M, *et al*. Development and Validation of a Deep Learning Algorithm for Detection of Diabetic Retinopathy in Retinal Fundus Photographs. *JAMA* 2016;316:2402.
- Ting DS, Cheung GCM, Wong TY. Diabetic retinopathy: global prevalence, major risk factors, screening practices and public health challenges: a review. *Clin Exp Ophthalmol* 2016;44:260–77.
- Saeedi P, Petersohn I, Salpea P, *et al*. Global and regional diabetes prevalence estimates for 2019 and projections for 2030 and 2045: Results from the International Diabetes Federation Diabetes Atlas, 9th edition. *Diabetes Res Clin Pract* 2019;157:107843.
- Flaxman SR, Bourne RRA, Resnikoff S, *et al*. Global causes of blindness and distance vision impairment 1990–2020: a systematic review and meta-analysis. *Lancet Glob Health* 2017;5:e1221–34.
- Padhy S, Takkar B, Chawla R, *et al*. Artificial intelligence in diabetic retinopathy: A natural step to the future. *Indian J Ophthalmol* 2019;67:1004.
- Beede E, Baylor E, Hersch F, *et al*. A human-centered evaluation of a deep learning system deployed in clinics for the detection of diabetic retinopathy. CHI '20: Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems; Honolulu HI USA, 2020 10.1145/3313831.3376718 Available: <https://dl.acm.org/doi/proceedings/10.1145/3313831>
- Chassagnon G, Vakalopoulou M, Paragios N, *et al*. Artificial intelligence applications for thoracic imaging. *Eur J Radiol* 2020;123:108774.
- Shah P, Kendall F, Khozin S, *et al*. Artificial intelligence and machine learning in clinical development: a translational perspective. *NPJ Digit Med* 2019;2:69.
- Raman R, Ramasamy K, Rajalakshmi R, *et al*. Diabetic retinopathy screening guidelines in India: All India Ophthalmological Society diabetic retinopathy task force and Vitreoretinal Society of India Consensus Statement. *Indian J Ophthalmol* 2021;69:678–88.
- Sengupta S, Sindal MD, Baskaran P, *et al*. Sensitivity and Specificity of Smartphone-Based Retinal Imaging for Diabetic Retinopathy: A Comparative Study. *Ophthalmol Retina* 2019;3:146–53.
- Klonoff DC, Gutierrez A, Fleming A, *et al*. Real-World Evidence Should Be Used in Regulatory Decisions About New Pharmaceutical and Medical Device Products for Diabetes. *J Diabetes Sci Technol* 2019;13:995–1000.
- Grassmann F, Mengelkamp J, Brandl C, *et al*. A Deep Learning Algorithm for Prediction of Age-Related Eye Disease Study Severity Scale for Age-Related Macular Degeneration from Color Fundus Photography. *Ophthalmology* 2018;125:1410–20.
- Lin SY, Mahoney MR, Sinsky CA. Ten Ways Artificial Intelligence Will Transform Primary Care. *J Gen Intern Med* 2019;34:1626–30.
- Kohne EM, *et al*. United Kingdom Prospective Diabetes Study, 30: Diabetic Retinopathy at Diagnosis of Non-Insulin-Dependent Diabetes Mellitus and Associated Risk Factors. *Arch Ophthalmol* 1998;116:297–303.
- Topol EJ. High-performance medicine: the convergence of human and artificial intelligence. *Nat Med* 2019;25:44–56.
- Banaee T, Ansari-Astaneh M-R, Pourreza H, *et al*. Utility of 1% Tropicamide in Improving the Quality of Images for Tele-Screening of Diabetic Retinopathy in Patients with Dark Irides. *Ophthalmic Epidemiol* 2017;24:217–21.
- Ministry of Health and Family Welfare. Ayushman bharat comprehensive primary health care through health and wellness centres operational guidelines. Available: <https://nhsrcindia.org/node/736> [Accessed 13 Oct 2023].
- Cohen JF, Korevaar DA, Altman DG, *et al*. STARD 2015 guidelines for reporting diagnostic accuracy studies: explanation and elaboration. *BMJ Open* 2016;6:e012799.
- Duggal M, Chauhan A, Gupta V, *et al*. Comparing the Implementation of Different Diabetic Retinopathy Screening Models in Primary Health Care Settings in Northern India: Pragmatic Three-Arm Observational Study. *Series Endo Diab Met* 2024;6:1–15.
- Noordzij M, Tripepi G, Dekker FW, *et al*. Sample size calculations: basic principles and common pitfalls. *Nephrol Dial Transplant* 2010;25:1388–93.
- Prasad R. Key findings for national diabetes and diabetic retinopathy survey among population aged ≥50 years.
- Natarajan S, Jain A, Krishnan R, *et al*. Diagnostic Accuracy of Community-Based Diabetic Retinopathy Screening With an Offline Artificial Intelligence System on a Smartphone. *JAMA Ophthalmol* 2019;137:1182–8.
- Shah P, Mishra D, Shanmugam M, *et al*. Validation of Deep Convolutional Neural Network-based algorithm for detection of diabetic retinopathy – Artificial intelligence versus clinician for screening. *Indian J Ophthalmol* 2020;68:398.
- Darwish DY, Patel SN, Gao Y, *et al*. Diagnostic accuracy and reliability of retinal pathology using the Forus 3nethra fundus camera compared to ultra wide-field imaging. *Eye (Lond)* 2019;33:856–7.
- Sosale B, Sosale AR, Murthy H, *et al*. Medios- An offline, smartphone-based artificial intelligence algorithm for the diagnosis of diabetic retinopathy. *Indian J Ophthalmol* 2020;68:391–5.
- Rajalakshmi R, Subashini R, Anjana RM, *et al*. Automated diabetic retinopathy detection in smartphone-based fundus photography using artificial intelligence. *Eye (Lond)* 2018;32:1138–44.
- Shi C, Lee J, Wang G, *et al*. Assessment of image quality on color fundus retinal images using the automatic retinal image analysis. *Sci Rep* 2022;12:10455.
- 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;125:1608–22.
- Harris PA, Taylor R, Thielke R, *et al*. Research electronic data capture (REDCap)—a metadata-driven methodology and workflow process for providing translational research informatics support. *J Biomed Inform* 2009;42:377–81.
- StataCorp. Stata statistical software: release 17. College Station, TX StataCorp LLC; 2021.
- Ministry of Health and Family Welfare, Government of India. Operational guidelines, Eye care at health and wellness centres (Part of comprehensive primary health care).
- Murthy GVS, Sundar G, Gilbert C, *et al*. Operational guidelines for diabetic retinopathy in India: Summary. *Indian J Ophthalmol* 2020;68:S59–62.
- Piyasena MMPN, Murthy GVS, Yip JLY, *et al*. Systematic review and meta-analysis of diagnostic accuracy of detection of any level of diabetic retinopathy using digital retinal imaging. *Syst Rev* 2018;7:1.
- Bascaran C, Mwangi N, D'Esposito F, *et al*. Effectiveness of task-shifting for the detection of diabetic retinopathy in low- and middle-income countries: a rapid review protocol. *Syst Rev* 2021;10:1.
- Olson JA, Strachan FM, Hipwell JH, *et al*. A comparative evaluation of digital imaging, retinal photography and optometrist examination in screening for diabetic retinopathy. *Diabet Med* 2003;20:528–34.

- 37 Prasad S, Kamath GG, Jones K, *et al.* Effectiveness of optometrist screening for diabetic retinopathy using slit-lamp biomicroscopy. *Eye (Lond)* 2001;15:595–601.
- 38 Li HK, Horton M, Bursell S-E, *et al.* Telehealth practice recommendations for diabetic retinopathy, second edition. *Telemed J E Health* 2011;17:814–37.
- 39 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* 2017;238:89–99.
- 40 Gupta V, Bansal R, Gupta A, *et al.* Sensitivity and specificity of nonmydriatic digital imaging in screening diabetic retinopathy in Indian eyes. *Indian J Ophthalmol* 2014;62:851–6.
- 41 Sosale AR. Screening for diabetic retinopathy—is the use of artificial intelligence and cost-effective fundus imaging the answer? *Int J Diabetes Dev Ctries* 2019;39:1–3.
- 42 Mead A, Burnett S, Davey C. Diabetic retinal screening in the UK. *J R Soc Med* 2001;94:127–9.
- 43 Murgatroyd H, Cox A, Ellingford A, *et al.* Can we predict which patients are at risk of having an ungradeable digital image for screening for diabetic retinopathy? *Eye (Lond)* 2008;22:344–8.
- 44 Sundling V, Platou CGP, Jansson RW, *et al.* Retinopathy and visual impairment in diabetes, impaired glucose tolerance and normal glucose tolerance: the Nord-Trøndelag Health Study (the HUNT study). *Acta Ophthalmol* 2012;90:237–43.
- 45 Hulme SA, Tin-U A, Hardy KJ, *et al.* Evaluation of a district-wide screening programme for diabetic retinopathy utilizing trained optometrists using slit-lamp and Volk lenses. *Diabet Med* 2002;19:741–5.
- 46 Jain A, Krishnan R, Rogye A, *et al.* Use of offline artificial intelligence in a smartphone-based fundus camera for community screening of diabetic retinopathy. *Indian J Ophthalmol* 2021;69:3150–4.
- 47 Sosale B, Aravind SR, Murthy H, *et al.* Simple, Mobile-based Artificial Intelligence Algo r ithm in the detection of Diabetic Retinopathy (SMART) study . *BMJ Open Diab Res Care* 2020;8:e000892.
- 48 Abràmoff MD, Lavin PT, Birch M, *et al.* Pivotal trial of an autonomous AI-based diagnostic system for detection of diabetic retinopathy in primary care offices. *NPJ Digit Med* 2018;1:39.
- 49 Chauhan A, Bascaran C, Duggal M. Effects of the Quality of Power Supply on Diabetic Retinopathy Screenings: An Experience From India. *J Vis Impair Blind* 2022;116:850–2.
- 50 Limwattanayingyong J, Nganthavee V, Seresirikachorn K, *et al.* Longitudinal Screening for Diabetic Retinopathy in a Nationwide Screening Program: Comparing Deep Learning and Human Graders. *J Diabetes Res* 2020;2020:1–8.
- 51 Dawoodbhoj FM, Delaney J, Cecula P, *et al.* AI in patient flow: applications of artificial intelligence to improve patient flow in NHS acute mental health inpatient units. *Heliyon* 2021;7:e06993.
- 52 Maleki Varnosfaderani S, Forouzanfar M. The Role of AI in Hospitals and Clinics: Transforming Healthcare in the 21st Century. *Bioengineering (Base)* 2024;11:337.
- 53 Uy H, Fielding C, Hohlfeld A, *et al.* Diagnostic test accuracy of artificial intelligence in screening for referable diabetic retinopathy in real-world settings: A systematic review and meta-analysis. *PLoS Glob Public Health* 2023;3:e0002160.
- 54 Vujosevic S, Benetti E, Massignan F, *et al.* Screening for Diabetic Retinopathy: 1 and 3 Nonmydriatic 45-degree Digital Fundus Photographs vs 7 Standard Early Treatment Diabetic Retinopathy Study Fields. *Am J Ophthalmol* 2009;148:111–8.
- 55 Ministry of Health and Family Welfare, Government of India. Ayushman bharat. Health and Wellness Centre. Available: <https://abhwc.nhp.gov.in/#about> [Accessed 10 Oct 2023].