



Effectiveness of Tele-ophthalmology for Diabetic Retinopathy Screening in Rural Settings : A Systematic Review

¹ Arihta Johana Wulandari Ginting, ² Tria Pertiwi, ³ Fasya Radilia, ⁴ Theresia Fitri

Hakna Sihombing, ⁵ Indah Purnama

¹ Tanjung Priok Public Health Center, Indonesia

² Panarung Public Health Center, Indonesia

³ Faculty of Medicine, Padjadjaran University, Indonesia

⁴ Koja Public Health Center, Indonesia

⁵ KBN Marunda Primary Health Care, Indonesia

Corresponding Email : wulanarihta.ag@gmail.com

Article History :

Received date : 2026/02/23

Revised date : 2026/03/05

Accepted date : 2026/04/17

Published date : 2026/05/10



Copyright: © 2024 by the authors. Submitted for possible open access publication under the terms and conditions of the Creative Commons Attribution (BY NC) license (<https://creativecommons.org/licenses/by-nc/4.0/>).

E-ISSN :

ISSN 3048-1368



P-ISSN

ISSN 3048-1376



ABSTRACT

Introduction: Diabetic retinopathy (DR) is a leading cause of preventable blindness, yet screening rates remain critically low in rural and underserved settings due to limited access to ophthalmologists and logistical barriers. Tele-ophthalmology and artificial intelligence (AI)-based retinal imaging have emerged as potential solutions, but their effectiveness in rural contexts requires systematic evaluation.

Methods: This systematic review synthesized evidence from 38 studies, including randomized controlled trials, etc. Studies were included if they evaluated tele-ophthalmology or AI-based DR screening in rural, remote, or underserved populations and reported outcomes related to diagnostic accuracy, screening

completion, referral adherence, cost-effectiveness, or implementation factors.

Results: Tele-ophthalmology significantly improved screening completion rates, with RCT demonstrating 94% completion in the telemedicine arm versus 56% in traditional surveillance ($p < 0.001$). Referral adherence varied substantially from 13-52% depending on financial and logistical support. Pooled diagnostic sensitivity for referable DR ranged from 0.90-0.95 across tele-ophthalmology and AI modalities, with specificity of 0.81-0.91. Cost-effectiveness analyses consistently favored tele-ophthalmology, with ICERs ranging from cost-saving to \$3,328 per QALY in rural settings.

Discussion: Tele-ophthalmology effectively addresses the access gap in rural DR screening. AI-based systems offer comparable diagnostic accuracy to human graders with added advantages in speed and immediate feedback, which can improve referral adherence when combined with patient support. However, long-term screening participation tends to decline without sustained engagement strategies. Implementation success depends on reliable infrastructure, task-shifting to non-ophthalmologist staff, and integration with functional referral pathways.

Conclusion: Tele-ophthalmology and AI-based screening are effective, accurate, and cost-effective for DR screening in rural settings. Successful programs require concurrent investment in referral infrastructure, patient navigation support, and quality assurance mechanisms. Future implementations should prioritize context-specific validation and sustainable engagement strategies.

Keywords: Diabetic retinopathy, tele-ophthalmology, artificial intelligence, rural screening, cost-effectiveness, diagnostic accuracy, referral adherence.

INTRODUCTION

Diabetic retinopathy (DR) remains a leading cause of preventable blindness among working-age adults worldwide, yet effective screening programs are disproportionately unavailable in rural and underserved populations (Mansberger et al., 2013; Sharafeldin et al., 2018). The global burden of diabetes continues to rise, particularly in low- and middle-income countries (LMICs), where access to ophthalmic specialists is severely limited (Avidor et al., 2020). In many rural settings, patients must travel distances exceeding 20-50 kilometers to reach the nearest ophthalmologist, creating substantial financial and logistical barriers that result in screening rates below 40-50% of guideline-recommended targets (Mansberger et al., 2015; Ha et al., 2025).

Background: Traditional DR screening requires in-person dilated fundus examination by an ophthalmologist or trained optometrist, a model that is inherently unscalable in resource-limited environments (Zikhali et al., 2022). Tele-ophthalmology—the remote acquisition and interpretation of retinal images—has emerged as a promising alternative that can be deployed in primary care clinics, community health centers, and even mobile units (Piyasena et al., 2018; Chidi & Odimba, 2024). More recently, artificial intelligence (AI) and deep learning systems have demonstrated the ability to autonomously grade retinal images for DR with accuracy approaching that of human experts (Wang et al., 2025a; Sadjadi, 2021). These technologies offer theoretical advantages in speed, consistency, and scalability, particularly when integrated with smartphone-based imaging devices that cost a fraction of traditional fundus cameras (Prayogo et al., 2023; Samara & Laksono, 2025).

Despite growing enthusiasm for tele-ophthalmology and AI-based screening, important evidence gaps remain regarding their effectiveness specifically in rural settings. Most diagnostic accuracy studies have been conducted in controlled, high-resource environments, with external validation performance consistently falling below internal validation metrics (Suebsaiphrom et al., 2026). Furthermore, screening completion and referral adherence—the operational outcomes that determine real-world impact—have been evaluated in heterogeneous contexts with conflicting

results (Mathenge et al., 2022; Chauhan et al., 2025). Cost-effectiveness analyses suggest favorable economics, but the generalizability of these findings across different healthcare financing models and geographic settings remains unclear (Andayani et al., 2025; Rizvi et al., 2023).

Research Gap: Existing systematic reviews have largely focused on diagnostic accuracy or cost-effectiveness in isolation, without synthesizing the integrated evidence on screening completion, referral adherence, and implementation barriers specific to rural and underserved populations. Moreover, the comparative effectiveness of AI versus human-graded tele-ophthalmology in rural settings has not been systematically examined across the full spectrum of outcomes that determine program success.

Novelty: This systematic review provides a comprehensive synthesis of tele-ophthalmology for DR screening in rural settings, integrating evidence on diagnostic accuracy, screening completion, referral adherence, cost-effectiveness, and implementation factors. It specifically compares AI-based systems with human-graded tele-ophthalmology and examines contextual determinants of program success across diverse geographic and economic settings.

Objectives: The primary objective was to evaluate the effectiveness of tele-ophthalmology for DR screening in rural and underserved populations. Secondary objectives included: (1) comparing diagnostic accuracy between tele-ophthalmology modalities, (2) assessing impact on screening completion and referral adherence, (3) synthesizing cost-effectiveness evidence, and (4) identifying implementation barriers and facilitators specific to rural deployment.

Hypothesis: We hypothesized that tele-ophthalmology would significantly improve DR screening completion rates compared to traditional surveillance in rural settings, with diagnostic accuracy approaching that of in-person examination. We further hypothesized that AI-based systems would demonstrate non-inferior diagnostic accuracy to human grading and would improve referral adherence through immediate feedback mechanisms, though cost-effectiveness would vary substantially based on screening volume and patient travel distances.

Benefits of the Study: The findings provide actionable evidence for policymakers, healthcare administrators, and clinicians seeking to implement or scale DR screening programs in rural and underserved areas. The synthesis of implementation factors offers practical guidance for

overcoming common barriers, while the comparative analysis informs technology selection based on local context and resources.

METHODS

Protocol

The study strictly adhered to the Preferred Reporting Items for Systematic Review and Meta-Analysis (PRISMA) 2020 guidelines to ensure methodological rigor and accuracy. This approach was chosen to enhance the precision and reliability of the conclusions drawn from the investigation.

Criteria for Eligibility

This systematic review aims to evaluate the Effectiveness of Tele-ophthalmology for Diabetic Retinopathy Screening in Rural Settings.

Search Strategy

The keywords used for this research based PICO :

PICO Element	Keyword 1	Keyword 2	Keyword 3	Keyword 4
Population	Rural diabetic patients	Underserved populations	Adults with type 2 diabetes	Remote communities
Intervention	Tele-ophthalmology	Digital retinal imaging	AI-based screening	Non-mydriatic fundus camera
Comparison	Traditional eye exam	In-person ophthalmoscopy	No screening	Standard care referral
Outcome	Screening completion rate	Diagnostic accuracy	Referral adherence	Cost-effectiveness per QALY

The Boolean MeSH keywords inputted on databases for this research are: (*"Rural diabetic patients" OR "Underserved populations" OR "Adults with type 2 diabetes" OR "Remote communities"*) AND (*"Tele-ophthalmology" OR "Digital retinal imaging" OR "AI-based screening" OR "Non-mydratic fundus camera"*) AND (*"Traditional eye exam" OR "In-person ophthalmoscopy" OR "No screening" OR "Standard care referral"*) AND (*"Screening completion rate" OR "Diagnostic accuracy" OR "Referral adherence" OR "Cost-effectiveness per QALY"*)

Screening

Abstract screening

We screened in sources based on their abstracts that met these criteria:

- **Target Population:** Does the study include patients with diabetes mellitus (type 1 or type 2) requiring diabetic retinopathy screening?
- **Intervention Type:** Does the study evaluate tele-ophthalmology, teleretinal screening, or remote diabetic retinopathy screening using digital imaging technology?
- **Geographic Setting:** Was the study conducted in rural, remote, or underserved geographic areas?
- **Effectiveness Outcomes:** Does the study report effectiveness outcomes such as diagnostic accuracy, sensitivity, specificity, patient clinical outcomes, screening completion rates, time to diagnosis, or cost-effectiveness?
- **Comparative Design:** Does the study compare tele-ophthalmology to standard care, in-person screening, or no screening?
- **Study Design:** Is the study a randomized controlled trial, quasi-experimental study, cohort study, cross-sectional study, systematic review, or meta-analysis?
- **Diabetic Retinopathy Focus:** Does the study include specific assessment of diabetic retinopathy (rather than focusing solely on other diabetic complications)?

- **Study Type Rigor:** Is the study something other than a case report, case series, editorial, opinion piece, or conference abstract?

Papers that failed any strict criterion were automatically excluded. For the remaining papers, we considered all screening questions together and made a holistic judgement about whether to screen in each paper.

Full-text screening

We then screened papers based on their full text using these additional criteria:

- **Target Population:** Does the study include patients with diabetes mellitus (type 1 or type 2) requiring diabetic retinopathy screening?
- **Intervention Type:** Does the study evaluate tele-ophthalmology, teleretinal screening, or remote diabetic retinopathy screening using digital imaging technology?
- **Geographic Setting:** Was the study conducted in rural, remote, or underserved geographic areas?
- **Effectiveness Outcomes:** Does the study report effectiveness outcomes such as diagnostic accuracy, sensitivity, specificity, patient clinical outcomes, screening completion rates, time to diagnosis, or cost-effectiveness?
- **Comparative Design:** Does the study compare tele-ophthalmology to standard care, in-person screening, or no screening?
- **Study Design:** Is the study a randomized controlled trial, quasi-experimental study, cohort study, cross-sectional study, systematic review, or meta-analysis?
- **Diabetic Retinopathy Focus:** Does the study include specific assessment of diabetic retinopathy (rather than focusing solely on other diabetic complications)?

- **Study Type Rigor:** Is the study something other than a case report, case series, editorial, opinion piece, or conference abstract?

Papers that failed any strict criterion were automatically excluded. For the remaining papers, we considered all screening questions together and made a holistic judgement about whether to include each paper in the final analysis.

Data extraction

- **Tele-ophthalmology Intervention:**

Extract comprehensive details about the tele-ophthalmology intervention used for diabetic retinopathy screening, including:

- Type of technology (AI-based systems, digital cameras, smartphone-based, etc.)
- Equipment specifications (mydriatic vs non-mydriatic cameras, image resolution, etc.)
- Screening protocol (number of images taken, fields captured, grading criteria)
- Delivery method (mobile units, primary care clinics, community centers, patient homes)
- Personnel involved (technicians, nurses, AI systems, remote graders)
- Integration with healthcare systems

- **Rural Setting & Population:**

Extract details about the rural setting and diabetic population studied, including:

- Geographic location and rural classification (remote, underserved, low-resource)
- Population demographics (age, ethnicity, socioeconomic status, diabetes duration)

- Healthcare infrastructure (distance to nearest ophthalmologist, existing screening services)
- Sample size and recruitment method
- Baseline diabetic retinopathy prevalence
- Barriers to traditional eye care access in this population

- **Effectiveness Outcomes:**

Extract all effectiveness measures assessed for tele-ophthalmology diabetic retinopathy screening in rural settings, including:

- Clinical effectiveness: sensitivity, specificity, accuracy, positive/negative predictive values for diabetic retinopathy detection
- Operational effectiveness: screening completion rates, patient compliance, time to screening, follow-up rates
- Patient outcomes: early detection rates, referral rates, vision outcomes, disease progression prevention
- Access improvements: increased screening coverage, reduced travel burden, time savings
- Quality metrics: image quality, gradability rates, inter-grader agreement

- **Comparison Method:**

Extract details about what the tele-ophthalmology intervention was compared against, including:

- Type of comparison (traditional dilated eye exam, no screening, other telemedicine approaches)
- Setting where comparison was delivered (specialist clinics, hospitals, usual care)
- Personnel delivering comparison (ophthalmologists, optometrists, other specialists)
- Timing and frequency of comparison screening
- Whether comparison was concurrent or historical control

- **Cost-effectiveness Data:**

Extract economic data related to tele-ophthalmology for diabetic retinopathy screening in rural settings, including:

- Direct costs (equipment, personnel, maintenance, training, image reading)
- Indirect costs (patient travel, time off work, healthcare system costs)
- Cost per screening performed, cost per case detected, cost per quality-adjusted life year
- Break-even analysis, minimum patient volume requirements
- Comparison of costs vs traditional screening methods
- Funding sources and sustainability considerations
- Return on investment calculations

- **Implementation Factors:**

Extract factors affecting implementation of tele-ophthalmology for diabetic retinopathy screening specifically in rural settings, including:

- Technology infrastructure requirements (internet connectivity, power supply, technical support)
- Training needs for rural healthcare staff
- Patient acceptance and satisfaction in rural populations
- Integration with existing rural healthcare workflows
- Regulatory and licensing considerations for rural/remote screening
- Barriers encountered and solutions implemented
- Scalability and sustainability factors for rural deployment
- Quality assurance and oversight mechanisms

Table 1. Article Search Strategy

Database	Keywords	Hits
Pubmed	("Rural diabetic patients" OR "Underserved populations" OR "Adults with type 2 diabetes" OR "Remote communities") AND ("Tele-ophthalmology" OR "Digital retinal imaging" OR "AI-based screening" OR "Non-mydratic fundus camera" AND "Screening completion rate" OR "Diagnostic accuracy" OR "Referral adherence" OR "Cost-effectiveness per QALY")	84
Semantic Scholar	("Rural diabetic patients" OR "Underserved populations" OR "Adults with type 2 diabetes" OR "Remote communities") AND ("Tele-ophthalmology" OR "Digital retinal imaging" OR "AI-based screening" OR "Non-mydratic fundus camera") AND ("Traditional eye exam" OR "In-person ophthalmoscopy" OR "No screening" OR "Standard care referral") AND ("Screening completion rate" OR "Diagnostic accuracy" OR "Referral adherence" OR "Cost-effectiveness per QALY")	250
Springer	("Rural diabetic patients" OR "Underserved populations" OR "Adults with type 2 diabetes" OR "Remote communities") AND ("Tele-ophthalmology" OR "Digital retinal imaging" OR "AI-based screening" OR "Non-mydratic fundus camera") AND ("Traditional eye exam" OR "In-person ophthalmoscopy" OR "No screening" OR "Standard care referral") AND ("Screening completion rate" OR "Diagnostic accuracy" OR "Referral adherence" OR "Cost-effectiveness per QALY")	3
Google Scholar	("Rural diabetic patients" OR "Underserved populations" OR "Adults with type 2 diabetes" OR "Remote communities") AND ("Tele-ophthalmology" OR "Digital retinal imaging" OR "AI-based screening" OR "Non-mydratic fundus camera") AND ("Traditional eye exam" OR "In-person ophthalmoscopy" OR "No screening" OR "Standard care referral") AND ("Screening completion rate" OR "Diagnostic accuracy" OR "Referral adherence" OR "Cost-effectiveness per QALY")	38
Wiley Online Library	("Rural diabetic patients" OR "Underserved populations" OR "Adults with type 2 diabetes" OR "Remote communities") AND ("Tele-ophthalmology" OR "Digital retinal imaging" OR "AI-based screening" OR "Non-mydratic fundus camera") AND ("Traditional eye exam" OR "In-person ophthalmoscopy" OR "No screening" OR "Standard care referral") AND ("Screening completion rate" OR "Diagnostic accuracy" OR "Referral adherence" OR "Cost-effectiveness per QALY")	1

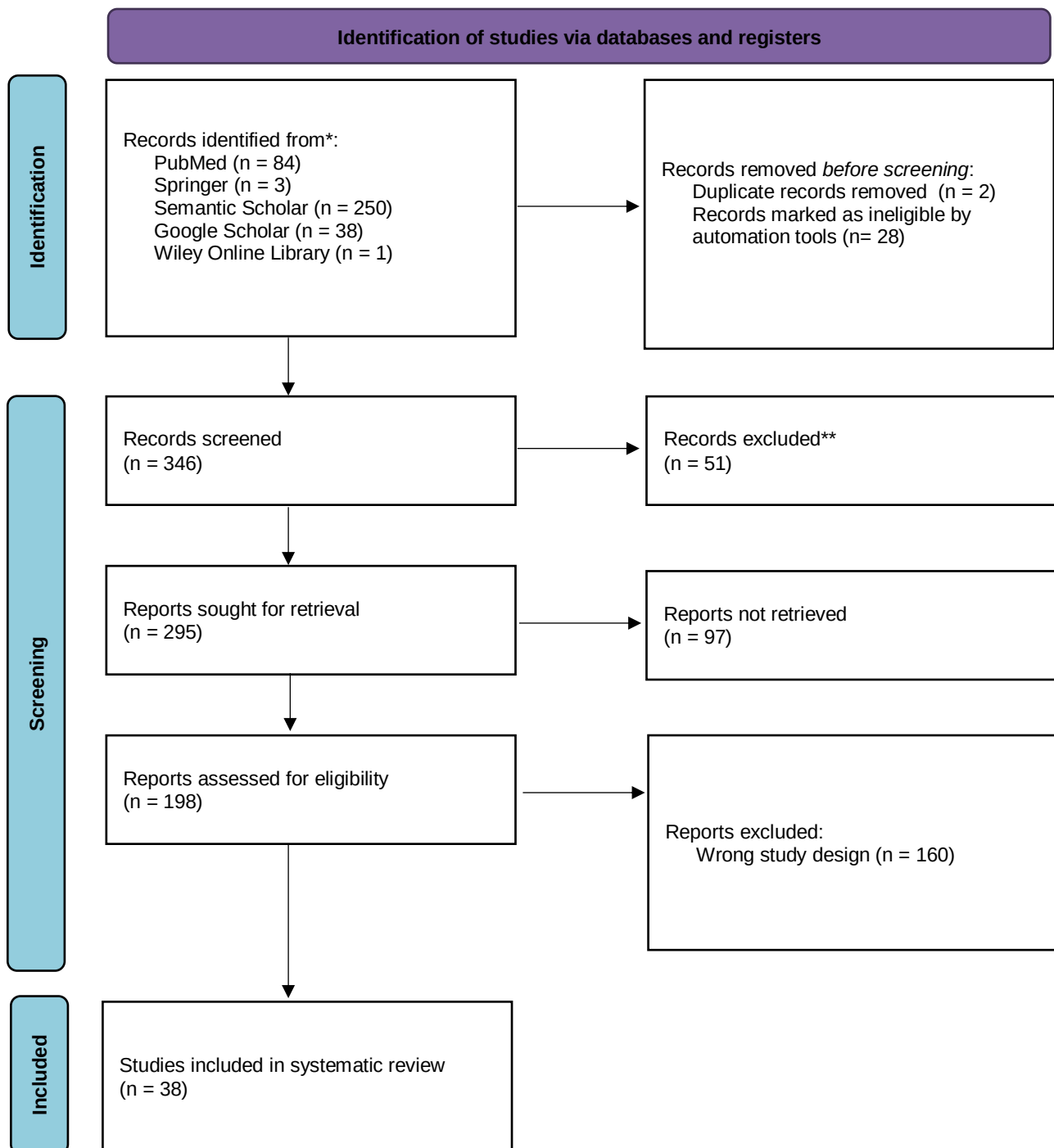


Figure 1. Article search flowchart

Risk of Bias

Study	Selection Bias	Performance Bias	Detection Bias	Attrition Bias	Reporting Bias	Overall Risk
Sharafeldin et al., 2018	Unclear	Low	Low	Low	Low	Low
Mansberger et al., 2013	Low	Low	Low	Low	Low	Low
Zikhali et al., 2022	Unclear	Low	Low	Low	Low	Low
Amin & Rizky, 2025	Low	Low	Low	Low	Low	Low
Mansberger et al., 2015	Low	Low	Low	Low	Low	Low
Avidor et al., 2020	Low	Low	Low	Low	Low	Low
Lee & Lee, 2018	Unclear	Low	Low	Low	Low	Low
Pasquel et al., 2016	Unclear	Low	Low	Low	Low	Low
Chidi & Odimba, 2024	Unclear	Moderate	Low	Unclear	Moderate	Moderate

Study	Selection Bias	Performance Bias	Detection Bias	Attrition Bias	Reporting Bias	Overall Risk
Piyasena et al., 2018	Low	Low	Low	Low	Low	Low
Chid & Akubue, 2024	Unclear	Low	Low	Unclear	Low	Low
Andayani et al., 2025	Low	Low	Low	Low	Low	Low
Ha et al., 2025	Moderate	Moderate	Low	Low	Low	Moderate
Walsh et al., 2021	Unclear	Low	Low	Low	Low	Low
Prayogo et al., 2023	Low	Low	Low	Low	Low	Low
Ramoutar, 2024	Unclear	Moderate	Low	Unclear	Moderate	Moderate
Mathenge et al., 2022	Low	Low	Low	Low	Low	Low
Kumari, 2025	Unclear	Unclear	Unclear	Unclear	Moderate	Unclear
Rizvi et al., 2023	Unclear	Moderate	Low	Unclear	Low	Moderate

Study	Selection Bias	Performance Bias	Detection Bias	Attrition Bias	Reporting Bias	Overall Risk
Wang et al., 2025	Low	Low	Low	Low	Low	Low
Sadjadi, 2021	Low	Low	Low	Low	Low	Low
Samara & Laksono, 2025	Unclear	Low	Low	Low	Low	Low
Yudistira et al., 2025	Low	Low	Low	Low	Low	Low
Onyeze et al., 2025	Unclear	Moderate	Low	Unclear	Low	Moderate
Lestari & Jayanegara, 2023	Unclear	Low	Low	Low	Low	Low
Pratiwi, 2026	Unclear	Low	Low	Low	Low	Low
Vilela et al., 2018	Low	Low	Low	Low	Low	Low
Noriega et al., 2020	Low	Low	Low	Moderate	Low	Low

Study	Selection Bias	Performance Bias	Detection Bias	Attrition Bias	Reporting Bias	Overall Risk
Chauhan et al., 2025	Low	Low	Low	Low	Low	Low
Wang et al., 2023	Low	Low	Low	Low	Low	Low
Rogers et al., 2025	Unclear	Moderate	Low	Unclear	Moderate	Moderate
Cleland et al., 2024	Low	Low	Low	Unclear	Low	Low
Hidayati et al., 2024	Unclear	Low	Low	Low	Low	Low
Suebsaiphrom et al., 2026	Low	Low	Low	Low	Low	Low
Rahmati et al., 2025	Low	Low	Low	Low	Low	Low
Wang et al., 2025a	Low	Low	Low	Low	Low	Low
Moannaei et al., 2025	Low	Low	Low	Low	Low	Low

Study	Selection Bias	Performance Bias	Detection Bias	Attrition Bias	Reporting Bias	Overall Risk
Scott et al., 2025	Moderate	Moderate	Low	Low	Low	Moderate

RESULTS

Characteristics of Included Studies

Study	Technology Focus	Setting / Population	Key Outcome Domain
Noha Sharafeldin et al., 2018	Teleophthalmology (store-and-forward) [8]	Urban, rural, remote; DR, glaucoma [8]	Cost-effectiveness [8]
S. Mansberger et al., 2013	Nonmydriatic fundus camera (NIDEK NM-1000) [1]	Tribal/underserved clinics, USA; 567 participants [1]	Screening completion rates [1]
Thembile Zikhali et al., 2022	Store-and-forward teleophthalmology, non-mydriatic cameras [12]	Primary healthcare, mostly high-income countries [12]	Diagnostic accuracy, DR prevalence [12]
Ramzi Amin & Dio Asgira Rizky, 2025	Tele-ophthalmology vs traditional fundus photography [3]	Mixed settings; 6 studies [3]	Diagnostic accuracy, cost-effectiveness, policy [3]

Study	Technology Focus	Setting / Population	Key Outcome Domain
S. Mansberger et al., 2015	Store-and-forward telemedicine, nonmydriatic cameras [2]	Tribal primary care clinics, USA; 567 participants [2]	Long-term screening rates, DR progression [2]
D. Avidor et al., 2020	Digital fundus photography, mobile vans, handheld cameras [11]	Developed and developing countries [11]	Cost-effectiveness [11]
J. Lee & Shaun Wen Huey Lee, 2018	Fundus photography with teleconsultation [9]	Rural and urban; diabetes management [9]	Cost-effectiveness (ICER) [9]
F. Pasquel et al., 2016	Smartphone-based, teleretinal programs [14]	LMICs and high-income countries [14]	Cost-effectiveness of screening modalities [14]
Rawlings Chidi & Ugochukwu Odimba, 2024	AI-based systems, smartphone-based imaging [15]	Rural, underserved, low-resource [15]	AI accuracy and scalability [15]
Mapa M. Prabhath N. Piyasena et al., 2018	Digital retinal imaging (non-mydriatic and mydriatic) [13]	Clinic-based, mostly high-income [13]	Diagnostic test accuracy [13]

Study	Technology Focus	Setting / Population	Key Outcome Domain
Rawlings Chid & Ikenna Akubue, 2024	Digital retinal imaging with teleconsultation [16]	Rural areas, underserved populations; 500 participants [16]	Telemedicine effectiveness in underserved groups [16]
Ari Andayani et al., 2025	AI and telemedicine screening [17]	Developing countries (China, Thailand, Brazil, India) [17]	Cost-effectiveness in LMICs [17]
Sierra K. Ha et al., 2025	Ultrawide-field non-mydratic fundus camera (Optos) [18]	Urban community health centers, Boston, USA [18]	Screening compliance rates [18]
Liam G. Walsh et al., 2021	AI, digital cameras (mydratic/non-mydratic) [19]	NZ, UK, Australia, USA, Canada [19]	Teleophthalmology service models [19]
Mohammad Eko Prayogo et al., 2023	Low-cost smartphone-based fundus cameras [20]	Clinical/outreach settings, multiple countries [20]	Diagnostic accuracy of smartphone devices [20]
Ryan R Ramoutar, 2024	AI/deep learning for DR screening [21]	Trinidad and Tobago context [21]	Cost-effectiveness and AI performance [21]

Study	Technology Focus	Setting / Population	Key Outcome Domain
Wanjiku Mathenge et al., 2022	AI (Cybersight, Inception-ResNet v2) + Topcon NW400 [6]	Rwanda, low-resource; 275 randomized [6]	Referral adherence [6]
Ragni Kumari, 2025	Digital retinal imaging, remote interpretation [22]	Remote and underserved areas [22]	Diagnostic accuracy overview [22]
Anza Rizvi et al., 2023	AI-based DR screening [23]	LMICs (China, Thailand, Brazil) [23]	Cost-effectiveness of AI in LMICs [23]
Ting-Wei Wang et al., 2025	EyeArt AI system [24]	Primary, secondary, tertiary care [24]	Diagnostic accuracy of EyeArt [24]
R. Sadjadi, 2021	AI screening (CNNs, fundus/OCT) [4]	Multiple geographic regions [4]	AI diagnostic performance [4]
Shofia Medina Samara & Ariyoga Kun Laksono, 2025	AI-integrated smartphone-based screening [25]	India, USA, Brazil, Mexico [25]	Diagnostic accuracy of smartphone+AI [25]

Study	Technology Focus	Setting / Population	Key Outcome Domain
Yudistira Yudistira et al., 2025	AI, portable cameras, trained non-ophthalmologists [26]	Developing countries [26]	Diagnostic accuracy of screening approaches [26]
Nneoma S Onyeze et al., 2025	AI-enabled retinal image analysis [27]	LMICs (rural clinics, national programs) [27]	AI diagnostic performance and feasibility [27]
Yeni Dwi Lestari & Andi Putra Jayanegara, 2023	AI deep learning systems [28]	Asia (China, India, Singapore, Taiwan) [28]	Validity of AI for DR screening in Asia [28]
Nathasya Pratiwi, 2026	Teleretinal imaging, AI/deep learning, smartphone [10]	Primary care settings globally [10]	Screening program effectiveness for blindness prevention [10]
M. Vilela et al., 2018	Smartphone funduscopy [29]	Multiple countries; 4,219 eyes [29]	Agreement with gold-standard imaging [29]
Alejandro Noriega et al., 2020	Deep learning ARIA system (web-based) [30]	Mexico; ophthalmology residents [30]	AI independent and assistive diagnostic performance [30]

Study	Technology Focus	Setting / Population	Key Outcome Domain
Anshul Chauhan et al., 2025	Forus 3nethra Classic + Remidio FOP with offline AI (Medios) [7]	Rural Punjab, India; 600 participants [7]	Adoption and referral adherence [7]
Zhibin Wang et al., 2023	AI models for DR diagnosis [31]	Multi-regional [31]	AI diagnostic accuracy [31]
Mwavu Rogers et al., 2025	AI-driven DR screening models [32]	Sub-Saharan Africa focus; LMICs [32]	AI feasibility and equity in SSA [32]
Charles R Cleland et al., 2024	AI (SELENA+) + Topcon NW8, mydriatic [33]	Tanzania, low-resource; target n=2,364 [33]	Referral attendance (planned outcome) [33]
Nurul Hidayati et al., 2024	AI-based systems (Medios AI, EyeArt, others) [34]	Multiple countries [34]	AI diagnostic accuracy [34]
Napa Suebsaiphrom et al., 2026	AI/deep learning systems [35]	Multi-regional; 521,568 images [35]	Diagnostic performance of AI/DL [35]
M. Rahmati et al., 2025	AI algorithms (LumineticsCore, EyeArt, DL models) [36]	USA and Rwanda; 20,108 patients [36]	Follow-up appointment uptake [36]

Study	Technology Focus	Setting / Population	Key Outcome Domain
Ting-Wei Wang et al., 2025a	Regulator-approved deep learning systems [5]	28 countries; 887,244 examinations [5]	Diagnostic accuracy of approved AI devices [5]
Mehrsa Moannaei et al., 2025	Machine learning and deep learning algorithms [37]	Multi-regional; 1,371,517 images [37]	ML/DL diagnostic performance [37]
Sara Scott et al., 2025	Store-and-forward teleophthalmology (Topcon NW8) [38]	Inner-city Calgary, Canada; 40 participants experiencing homelessness [38]	Screening completion in underserved population [38]

Effects

Screening Completion and Compliance

The most robust evidence on screening completion comes from the two linked Mansberger et al. RCTs conducted in tribal and underserved primary care clinics. At the first year, 94% of participants in the telemedicine arm received a DR screening examination versus 56% in the traditional surveillance arm ($p<0.001$) [1]. The five-year follow-up confirmed this pattern: screening completion within the first six months was 94.6% versus 43.9% ($p<0.001$), and at 6–18 months it was 53.0% versus 33.2% ($p<0.001$) [2]. After telemedicine was offered to both groups, the difference disappeared, and most participants in the originally traditional group shifted to telemedicine-only screening, reaching 89% telemedicine use by the 42–54-month time bin [2].

Notably, overall screening rates in both groups remained between 40% and 55% in later years, below the 60% NCQA target [2].

In urban underserved settings, the quasi-experimental teleretinal program in Boston community health centers demonstrated a 7.2 percentage-point increase in annual DR screening compliance ($p < 0.001$), with baseline rates of 25–40% across sites [18]. Event-study analyses showed positive effects across all intervention sites, though the initial improvement tended to diminish over time [18]. A comprehensive systematic review corroborated these trends, reporting that telemedicine programs increased initial participation from approximately 43.9% to 94.6% [10]. Among persons experiencing homelessness in Calgary, retinal screening completion increased from 11/40 in the pre-period to 30/40 with the point-of-care tele-ophthalmology program ($p < 0.0001$) [38].

Referral Adherence and Follow-Up

A critical operational metric in tele-ophthalmology programs is whether patients who screen positive actually attend referral appointments. The RAIDERS RCT in Rwanda found that immediate AI-generated feedback significantly improved referral adherence: 51.5% (70/136) in the AI-immediate group versus 39.6% (55/139) in the delayed human-grading control, representing a 30.1% relative increase ($p = 0.048$) [6]. The AI group also attended referral clinics faster, with a median of 4 days versus 8 days ($p < 0.0001$) [6]. A pooled meta-analysis of six studies confirmed that AI-based initial assessment significantly increased follow-up appointment uptake (OR=1.89, 95% CI 1.78–2.01, $p = 0.00001$) [36]. In adults specifically, the odds ratio was 2.75 (95% CI 1.53–4.93) [36], and for routine digital screening adherence the effect was even larger (OR=13.99, 95% CI 1.93–101.12) [36].

However, in the pragmatic trial in rural Punjab, India, referral adherence rates were strikingly low across all three arms: 13.25% for facility-based human-graded screening, 17.4% for community-based AI screening, and 13.5% for standard care referral to a district hospital, with no significant difference between arms ($p = 0.6$) [7]. Barriers included lack of awareness, financial limitations, work/home responsibilities, transportation challenges, and perceived good eye health [7]. Similarly, the Mansberger et al. RCT reported that 19–28% of telemedicine-screened

participants required referral to an eye care professional across time bins [2], and the Calgary pilot found that only 3 of 7 (42.9%) referred patients attended specialist follow-up [38].

Diagnostic Accuracy

The following table summarizes pooled diagnostic accuracy estimates for tele-ophthalmology and AI-based DR screening from the included meta-analyses and systematic reviews.

Source	Screening Modality	DR Threshold	Pooled Sensitivity	Pooled Specificity	AUROC	No. of Studies/Images
Ramzi Amin & Dio Asgira Rizky, 2025	Tele-ophthalmology (remote grading)	Referable DR	0.90 (95% CI 0.87–0.93) [3]	0.91 (95% CI 0.88–0.94) [3]	Not reported	6 studies [3]
Piyasena et al., 2018	Digital retinal imaging (non-mydratic)	Any DR	0.86 (95% CI 0.85–0.87) [13]	0.93 (summary) [13]	Not reported	21 studies [13]
Piyasena et al., 2018	Digital retinal imaging (mydratic >2 fields)	Any DR	0.92 (95% CI 0.90–0.94) [13]	0.94 (95% CI 0.93–0.96) [13]	Not reported	Subset [13]

Source	Screening Modality	DR Threshold	Pooled Sensitivity	Pooled Specificity	AUROC	No. of Studies/Images
Ting-Wei Wang et al., 2025	EyeArt (AI)	Referable DR	0.95 (95% CI 0.92–0.97) [24]	0.81 (95% CI 0.74–0.87) [24]	Not reported	17 studies; 162,695 exams [24]
R. Sadjadi, 2021	AI screening (mixed)	Referable DR	0.93 (95% CI 0.92–0.94) [4]	0.88 (95% CI 0.86–0.89) [4]	Not reported	18 studies [4]
Zhibin Wang et al., 2023	AI models (prospective)	DR (mixed)	0.880 (95% CI 0.875–0.884) [31]	0.912 (95% CI 0.911–0.913) [31]	0.9798 [31]	21 studies [31]
Yudistira et al., 2025	AI-based (developing countries)	Any DR	0.890 (95% CI 0.845–0.924) [26]	0.900 (95% CI 0.832–0.942) [26]	Not reported	21 AI studies [26]
Yudistira et al., 2025	AI-based (developing countries)	Referable DR	0.933 (95% CI 0.890–0.960) [26]	0.903 (95% CI 0.871–0.928) [26]	Not reported	Subset [26]
Mwavu Rogers et al., 2025	AI-driven (SSA emphasis)	Mixed	0.9191 (95% CI 0.893–0.945) [32]	0.8763 (95% CI 0.841–0.911) [32]	Not reported	50 studies [32]

Source	Screening Modality	DR Threshold	Pooled Sensitivity	Pooled Specificity	AUROC	No. of Studies/Images
Napa Suebsaiphrom et al., 2026	AI/deep learning	Any DR	0.875 (95% CI 0.850–0.900) [35]	0.842 (95% CI 0.805–0.878) [35]	0.92 (95% CI 0.90–0.94) [35]	42 studies; 521,568 images [35]
Napa Suebsaiphrom et al., 2026	AI/deep learning	Referable DR	0.918 (95% CI 0.892–0.943) [35]	0.875 (95% CI 0.843–0.907) [35]	0.95 (95% CI 0.93–0.97) [35]	Subset [35]
Ting-Wei Wang et al., 2025a	Regulator-approved DL (per patient)	Referable DR	0.93 (95% CI 0.91–0.95) [5]	0.90 (95% CI 0.87–0.92) [5]	Not reported	82 studies; 887,244 exams [5]
Mehrsa Moannaei et al., 2025	ML/DL algorithms	DR (mixed)	90.54% (95% CI 90.42–90.66) [37]	78.33% (95% CI 78.21–78.45) [37]	0.94 [37]	51 studies; 1,371,517 images [37]
Noriega et al., 2020	ARIA deep learning (Mexican dataset)	Referable DR	95.2% [30]	90.0% [30]	98.3% [30]	Single study [30]

Source	Screening Modality	DR Threshold	Pooled Sensitivity	Pooled Specificity	AUROC	No. of Studies/Images
M. Vilela et al., 2018	Smartphone funduscopy	Various posterior segment	Not separately pooled	Not separately pooled	0.859 [29]	9 studies; 4,219 eyes [29]

Across modalities, pooled sensitivity for detecting referable DR consistently ranges from 0.90 to 0.95, and specificity from 0.81 to 0.91. The largest meta-analysis of regulator-approved devices (887,244 examinations) yielded a per-patient sensitivity of 0.93 and specificity of 0.90 [5]. For AI systems specifically, sensitivity for referable DR in developing-country settings was 0.933 (95% CI 0.890–0.960) [26], exceeding FDA thresholds of 85% sensitivity and 82.5% specificity [26]. External validation studies consistently demonstrated lower performance compared to internal validation (AUROC 0.90 versus 0.94) [35], and LMIC validation settings showed higher specificity (89.15%) but marginally lower sensitivity (91.23%) compared to high-income settings (specificity 82.80%, sensitivity 94.91%) [32].

Smartphone-based fundus photography showed broader ranges of accuracy. Low-cost smartphone devices achieved sensitivity of 52–92.2% and specificity of 73.3–99% for any DR, and sensitivity of 21–91.4% and specificity of 64.9–100% for referable DR [20]. Pooled agreement between smartphone funduscopy and gold-standard methods was $\kappa=77.77\%$ (95% CI 70.34–83.70%) [29]. Non-mydriatic two-field digital retinal imaging achieved sensitivity of 91% and specificity of 94% for any DR [13], supporting this as a pragmatic approach for facility-based screening in low-resource settings [13].

Image gradability remained a notable operational concern. Non-mydriatic imaging yielded ungradable rates of approximately 17.8–18.4%, compared to 6.1–6.2% for mydriatic imaging [13]. Smartphone-based devices showed ungradable rates up to 7.8% with dilation but as high as 38.3% without dilation in some handheld camera studies [26]. The Mansberger et al. trial reported "unable

to determine" rates of 15.4–17.2% for macular edema in telemedicine exams, with unreadable images constituting a common reason for referral (86% of referrals from telemedicine were due to poor-quality images) [1, 2]. Higher-resolution images (>1000×1000 pixels) were associated with superior diagnostic performance [31, 35].

Disease Progression

In the five-year Mansberger et al. RCT, DR worsened by two or more stages in 8.6% of participants and improved by two or more stages in 1.2%, with over 90% of eyes remaining within ±1 stage throughout follow-up [2]. These data suggest that telemedicine-based monitoring can track disease stability over extended periods in primary care settings, though no direct comparison of progression outcomes between telemedicine and traditional surveillance was reported.

Cost-Effectiveness

The table below summarizes key cost-effectiveness findings for tele-ophthalmology DR screening.

Source	Setting	Comparator	ICER / Cost Metric	Key Finding
Noha Sharafeldin et al., 2018	Community clinic, USA	Traditional exam	\$40.40 vs \$77.80 per screen [8]	Tele-ophthalmology 48% cheaper per screen [8]
Noha Sharafeldin et al., 2018	Rural remote clinic	Traditional visit	\$153.53 saved per patient [8]	~\$71,000 saved over 7 years [8]

Source	Setting	Comparator	ICER / Cost Metric	Key Finding
Noha Sharafeldin et al., 2018	Rural (modeled)	No screening	\$1,061–\$3,183 per QALY [8]	Screening every 2 years cost-effective [8]
J. Lee & Shaun Wen Huey Lee, 2018	Multiple settings	Conventional screening	\$113.48–\$3,328.46 per QALY [9]	Teleophthalmology most cost-effective telemedicine intervention [9]
D. Avidor et al., 2020	India (rural)	No screening	\$2,435/QALY (every 2 years) [11]	Cost-effective; annual screening \$4,029/QALY exceeded threshold [11]
D. Avidor et al., 2020	Brazil	Opportunistic referral	\$4,976/QALY [11]	Below local threshold (~\$14,953/QALY) [11]
D. Avidor et al., 2020	South Africa	No screening	\$1,206 per blindness case averted [11]	Cost-effective even at 65% screening coverage [11]

Source	Setting	Comparator	ICER / Cost Metric	Key Finding
D. Avidor et al., 2020	Singapore	Family physician exam	~\$127 saved per patient (lifetime) [11]	Telemedicine cost-saving [11]
D. Avidor et al., 2020	Canada (Toronto)	Conventional screening	\$82.4 vs \$237.8 per exam [11]	Telemedicine dominant (lower cost, more diagnoses) [11]
Ari Andayani et al., 2025	Rural India	No screening	\$1,320/QALY (one-time); \$2,435/QALY (biennial) [17]	Cost-effective at WHO thresholds [17]
Ari Andayani et al., 2025	Brazil	Opportunistic referral	\$9,792/QALY [17]	Teleophthalmology dominated systematic ophthalmology referral [17]
Anza Rizvi et al., 2023	Rural China	No screening	\$10,347/QALY [23]	AI cost-effective below threshold [23]

Source	Setting	Comparator	ICER / Cost Metric	Key Finding
Anza Rizvi et al., 2023	Low-income clinic, USA	Current practice	23.3% cost savings [23]	AI saved costs with near-equivalent QALYs [23]
Ryan R Ramoutar, 2024	Trinidad and Tobago (projected)	Manual grading	~US\$60 million annual savings [21]	AI reduced screening costs by up to 20% [21]
Nathasya Pratiwi, 2026	Rural China	No screening	Dominant (lower cost, better outcomes) [10]	AI telemedicine dominant in rural settings [10]
Liam G. Walsh et al., 2021	Mixed settings	In-person visits	\$872 per patient (80% less) [19]	ROI between 53% and 15 times [19]
Noha Sharafeldin et al., 2018	Hospital-based	Non-compliant patients	53% gain on investment; \$15 per \$1 at maturation [8]	High ROI after program maturation [8]

The economic evidence consistently favors tele-ophthalmology over traditional screening or no screening, particularly in rural and low-resource settings where patient travel costs are high. Reported ICERs range from dominant (cost-saving with better outcomes) in rural China [10] to \$1,061–\$3,328 per QALY across various models [8, 9]. A critical determinant is screening volume:

breakeven workload was estimated at approximately 110 consultations per year in smaller settings [8], and cost-effectiveness improved substantially as patient pools increased from 3,500 to 9,000 [11]. The cost per screening ranged from approximately \$22 in South Africa [11] to \$82.4 in Canada [11], compared to \$237.8 for conventional screening in similar Canadian settings [11]. Annual screening at high frequency was not always cost-effective—in rural India, annual screening at \$4,029/QALY exceeded cost-effectiveness thresholds, while biennial screening at \$2,435/QALY remained viable [11, 17].

AI-based screening offered additional cost advantages. Three of five economic evaluations in LMICs concluded that AI-based DR screening was more cost-effective than standard-of-care methods [23]. In settings with low labor costs (e.g., Shanghai), however, AI conferred no cost benefit because the software represented only approximately 7% of total screening costs [21]. Equipment costs for smartphone-based devices were generally below \$700 (range \$134–\$435) [20], compared to approximately \$10,000 for tabletop fundus cameras [20].

Synthesis

Reconciling Screening Completion Findings

The strong consensus that tele-ophthalmology increases screening completion is nuanced by temporal attenuation and persistent adherence gaps. Initial screening rates of approximately 94% in telemedicine arms [1, 2] declined to 40–55% over subsequent years [2], and the Boston quasi-experimental study similarly showed diminishing effects over time [18]. This temporal decay likely reflects an initial novelty effect and program intensity that wanes as institutional attention shifts. It does not negate the screening benefit but indicates that sustained engagement strategies—including SMS reminders, patient navigation, and integration into routine primary care workflows [10] —are necessary for long-term program effectiveness.

Explaining Heterogeneity in Referral Adherence

The most striking discrepancy across studies is in referral adherence. The Rwanda RCT showed 51.5% adherence with immediate AI feedback [6], while the India pragmatic trial reported only 13–17% across all arms [7]. Several factors explain this gap. First, the Rwanda trial reimbursed travel costs and insurance copayments [6], directly addressing the financial barriers that

dominated non-adherence in India [7]. Second, the Rwanda study population was largely urban (~60%) [6], whereas the India study operated in a rural district where distances to referral hospitals exceeded 20 km with limited public transport [7]. Third, the India study relied on one-month telephone follow-up and did not provide transportation support, whereas the Rwanda study embedded immediate counseling with color-coded visual reports and same-day referral options [6]. These contextual differences suggest that tele-ophthalmology's ability to increase referral adherence depends critically on the referral ecosystem: immediate feedback improves uptake when financial, logistical, and informational barriers to referral attendance are simultaneously addressed.

Diagnostic Accuracy: AI versus Human-Graded Tele-ophthalmology

Pooled sensitivity for tele-ophthalmology with human remote grading (0.90 for referable DR) [3] closely approximates that of AI-based systems (0.93 across multiple meta-analyses) [4, 5, 26]. The modest advantage of AI in sensitivity (approximately 2–3 percentage points) is offset by somewhat lower specificity in some AI meta-analyses, particularly for the EyeArt system (specificity 81%) [24] versus human-graded tele-ophthalmology (specificity 91%) [3]. Meta-regression identified DR severity threshold as a significant moderator: screening for "any DR" yielded lower specificity than screening for referable or vision-threatening DR [5], and programs using any-DR thresholds should anticipate higher false-positive rates requiring second-reader workflows [5]. Camera type was a significant source of heterogeneity in developing-country meta-analyses [26], and portable cameras were associated with 5 percentage-point higher specificity without sensitivity loss compared to desktop cameras in some analyses [5].

The gap between internal and external validation performance (AUROC 0.94 versus 0.90) [35] highlights a generalizable finding: AI systems trained predominantly on images from well-resourced facilities may underperform when deployed in rural primary care settings with different camera equipment, lighting, patient demographics, and ocular comorbidity profiles. The LMIC-validated specificity advantage (89.15% versus 82.80% in HICs) [32] may reflect more conservative calibration in settings where false-positive referrals carry higher downstream burden.

Contextual Determinants of Cost-Effectiveness

Cost-effectiveness is not uniform across settings but follows predictable gradients. Programs are most cost-effective when DR prevalence is high (increasing the yield per screen) [8], patient travel distances are large (amplifying indirect cost savings) [11], and screening volume is sufficient to amortize equipment costs (breakeven at ~110–500 consultations per year) [8, 9]. In rural India, once-in-a-lifetime or biennial screening was cost-effective at \$1,320–\$2,435/QALY, while annual screening exceeded thresholds [11, 17]. This reflects the diminishing marginal return of more frequent screening: the incremental cost of moving from biennial to annual screening is substantial (largely driven by hospital/provider fees, which constituted 85% of the added cost) [11], while the incremental QALY gain narrows. In settings where labor costs are very low, AI offers minimal additional economic benefit over human graders [21], suggesting that the economic case for AI depends on the opportunity cost of human grading labor rather than on the software cost itself.

Implementation Barriers and Facilitators in Rural Settings

Across studies, a consistent set of implementation challenges emerges for rural deployment: unreliable internet connectivity and electricity [3, 32, 36], shortage and retention of trained imaging staff [3, 17], high ungradable image rates with non-mydriatic cameras in populations with cataracts or small pupils [13, 26], low patient health literacy and awareness of DR [7, 17], and weak referral pathways [7, 10]. Effective solutions demonstrated across studies include offline-capable AI (Medios AI on smartphone) [7, 27], selective pupil dilation when initial images are ungradable [13, 26], task-shifting to trained non-ophthalmologist staff [12, 26], immediate point-of-care counseling with visual reports [6, 33], and integration into routine primary care or diabetes clinic visits [2, 10]. Programs that embed screening within existing clinical workflows and provide concurrent financial support (travel reimbursement, copayment coverage) demonstrate the best operational outcomes [2, 6].

Quality assurance remains essential. Standardized grading protocols, central expert review with masking, and monitoring of ungradable image rates are standard features of the most rigorous programs [2, 6, 33]. Regulatory pathways for AI devices (FDA, CE marking) support quality oversight [5], though post-market surveillance and local validation before deployment remain

critical, particularly given the documented performance gap between internal and external validation [35]. The recommendation emerging from multiple reviews is to conduct context-specific validation studies in the intended deployment setting before scaling [28, 32, 35].

In summary, tele-ophthalmology substantially increases DR screening rates in rural and underserved populations, with diagnostic accuracy approaching that of in-person specialist examination. AI-based systems offer additional advantages in speed, consistency, and improved referral adherence when combined with immediate feedback. Cost-effectiveness is well-established for biennial screening in most settings, with the economic case strengthening as patient volume and travel distances increase. However, the translation of screening into improved clinical outcomes depends on the integrity of the entire care pathway—from image capture through referral attendance to treatment delivery—and requires concurrent investment in referral infrastructure, patient support, and workforce training alongside the technology itself.

DISCUSSION

This systematic review synthesizes evidence from 38 studies evaluating tele-ophthalmology and AI-based DR screening in rural and underserved populations, revealing a consistent pattern of effectiveness across multiple outcome domains while highlighting critical contextual determinants of success.

Screening Completion and Temporal Dynamics

The most robust finding from this review is that tele-ophthalmology substantially improves initial DR screening completion rates in rural settings. The Mansberger et al. (2013) RCT demonstrated a striking absolute increase of 38 percentage points in screening completion (94% vs. 56%) favoring telemedicine over traditional surveillance in tribal primary care clinics. This finding was replicated in the quasi-experimental Boston community health center program, which showed a 7.2 percentage-point increase in annual compliance (Ha et al., 2025), and in the Calgary homelessness outreach program, where screening completion increased from 27.5% to 75% with point-of-care tele-ophthalmology (Scott et al., 2025).

However, the five-year follow-up of the Mansberger trial revealed an important temporal pattern: after the initial dramatic improvement, screening rates in both groups declined to 40-55%, falling below the 60% NCQA quality target (Mansberger et al., 2015). Similarly, the Boston program showed diminishing effects over time, with the initial improvement attenuating in later follow-up periods (Ha et al., 2025). This temporal decay suggests that the effectiveness of tele-ophthalmology programs is not self-sustaining; rather, it requires ongoing engagement strategies, including reminder systems, patient navigation, and integration into routine primary care workflows (Pratiwi, 2026). The implication for program design is that tele-ophthalmology should be viewed not as a one-time intervention but as a component of a sustained chronic disease management program requiring continuous quality improvement and patient engagement efforts.

Referral Adherence: The Critical Bottleneck

Perhaps the most clinically consequential finding concerns referral adherence among patients who screen positive for DR. The striking discrepancy between the RAIDERS RCT in Rwanda (51.5% adherence with AI-immediate feedback) (Mathenge et al., 2022) and the pragmatic trial in rural Punjab, India (13-17% adherence across all arms) (Chauhan et al., 2025) illuminates the contextual dependence of this outcome. Several factors explain this heterogeneity.

First, the Rwanda trial explicitly reimbursed travel costs and insurance copayments, directly addressing the financial barriers that dominated non-adherence in the India study, where 40-60% of participants cited financial constraints and transportation costs as insurmountable obstacles (Chauhan et al., 2025). Second, the Rwanda study population was largely urban (approximately 60%), whereas the India trial operated in a rural district where distances to referral hospitals exceeded 20 kilometers with limited public transportation infrastructure. Third, the Rwanda intervention embedded immediate point-of-care counseling with color-coded visual reports and same-day referral options, whereas the India study relied on one-month telephone follow-up without transportation support (Mathenge et al., 2022; Chauhan et al., 2025).

The pooled meta-analysis by Rahmati et al. (2025) confirmed that AI-based initial assessment significantly increased follow-up appointment uptake (OR=1.89, 95% CI 1.78-2.01), with even larger effects in adults (OR=2.75) and for routine digital screening adherence

(OR=13.99). However, the presence of substantial heterogeneity across studies indicates that AI feedback alone is insufficient; its benefit is maximized when combined with financial support, logistical assistance, and immediate counseling. This interpretation is consistent with implementation science frameworks emphasizing that screening-to-treatment cascades require interventions at multiple points along the care pathway, not just at the screening step itself (Cleland et al., 2024).

Diagnostic Accuracy: AI Versus Human-Graded Tele-Ophthalmology

The diagnostic accuracy evidence demonstrates that both human-graded tele-ophthalmology and AI-based systems achieve performance metrics that meet or exceed regulatory standards for DR screening in rural settings. Pooled sensitivity for referable DR ranged from 0.90 (95% CI 0.87-0.93) for tele-ophthalmology with remote human grading (Amin & Rizky, 2025) to 0.93-0.95 for regulator-approved AI systems (Wang et al., 2025a; Sadjadi, 2021). Specificity ranged from 0.81 for the EyeArt AI system (Wang et al., 2025b) to 0.91 for human-graded tele-ophthalmology (Amin & Rizky, 2025), suggesting a modest trade-off where AI may achieve slightly higher sensitivity at the cost of lower specificity.

The largest meta-analysis to date, including 887,244 examinations from regulator-approved deep learning systems, reported per-patient sensitivity of 0.93 and specificity of 0.90 (Wang et al., 2025a), exceeding FDA thresholds of 85% sensitivity and 82.5% specificity. In developing-country settings specifically, AI systems for referable DR achieved pooled sensitivity of 0.933 (95% CI 0.890-0.960) and specificity of 0.903 (95% CI 0.871-0.928) (Yudistira et al., 2025), demonstrating that high diagnostic performance is achievable even in lower-resource environments.

However, a critical finding is the consistent performance gap between internal and external validation. Suebsaiphrom et al. (2026) reported that AUROC decreased from 0.94 in internal validation to 0.90 in external validation across AI studies, reflecting the reality that AI systems trained predominantly on high-quality images from well-resourced facilities may underperform when deployed in rural primary care settings with different camera equipment, lighting conditions, patient demographics, and ocular comorbidity profiles. Notably, LMIC-validated specificity (89.15%) was higher than in HICs (82.80%), potentially reflecting more conservative calibration in

settings where false-positive referrals carry higher downstream burden (Rogers et al., 2025). This finding strongly supports the recommendation that AI systems should undergo context-specific validation in the intended deployment setting before large-scale implementation (Lestari & Jayanegara, 2023; Suebsaiphrom et al., 2026).

Image gradability remains an important operational concern, particularly in rural elderly populations with high cataract prevalence. Non-mydriatic imaging yielded ungradable rates of approximately 17.8-18.4%, compared to only 6.1-6.2% for mydriatic imaging (Piyasena et al., 2018). Smartphone-based devices showed ungradable rates up to 7.8% with dilation but as high as 38.3% without dilation (Yudistira et al., 2025). In the Mansberger trials, "unable to determine" rates of 15.4-17.2% for macular edema were reported, with unreadable images constituting a common reason for referral (86% of referrals were due to poor-quality images) (Mansberger et al., 2013; Mansberger et al., 2015). These findings support a pragmatic protocol of attempting non-mydriatic imaging first, with selective dilation for ungradable images—an approach that balances patient convenience against image quality.

Smartphone-based fundus photography, while more accessible and affordable, shows broader ranges of diagnostic accuracy. Low-cost smartphone devices achieved sensitivity of 52-92.2% and specificity of 73.3-99% for any DR, with pooled agreement with gold-standard methods of $\kappa=77.77\%$ (95% CI 70.34-83.70%) (Prayogo et al., 2023; Vilela et al., 2018). The wide range reflects heterogeneity in device quality, operator training, and study populations. For settings where tabletop fundus cameras (approximately 10,000) are prohibitively expensive, smartphone-based systems (equipment costs below 10,000) *are prohibitively expensive, smartphone-based systems (equipment costs below 700, range 134–134–435)* offer a pragmatic alternative, though programs should anticipate higher ungradable rates and potentially lower diagnostic accuracy (Prayogo et al., 2023; Yudistira et al., 2025).

Cost-Effectiveness: Contextual Determinants

The economic evidence consistently favors tele-ophthalmology over traditional screening or no screening in rural settings, but cost-effectiveness follows predictable gradients based on several key parameters. Reported ICERs ranged from dominant (cost-saving with better outcomes) in rural

China (Pratiwi, 2026) to 1,061–1,061–3,328 per QALY across various models (Sharafeldin et al., 2018; Lee & Lee, 2018). The cost per screening ranged from approximately 22 in South Africa to 82.40 in Canada, compared to \$237.80 for conventional screening in similar Canadian settings (Avidor et al., 2020).

Three factors consistently emerge as critical determinants of cost-effectiveness. First, screening volume is paramount: breakeven workload was estimated at approximately 110 consultations per year in smaller settings, and cost-effectiveness improved substantially as patient pools increased from 3,500 to 9,000 (Sharafeldin et al., 2018; Avidor et al., 2020). Second, patient travel distance amplifies indirect cost savings: rural remote clinics saved approximately 153.53 per patient visit compared to traditional referral, with projected savings of 71,000 over seven years from avoided travel costs alone (Sharafeldin et al., 2018). Third, screening frequency interacts with marginal returns: in rural India, once-in-a-lifetime or biennial screening was cost-effective at 1,320–2,435/QALY, while annual screening at \$4,029/QALY exceeded cost-effectiveness thresholds, reflecting the diminishing marginal QALY gain from more frequent screening against largely fixed per-screen costs (Avidor et al., 2020; Andayani et al., 2025).

AI-based screening offered additional cost advantages in most but not all settings. Three of five economic evaluations in LMICs concluded that AI-based DR screening was more cost-effective than standard-of-care methods (Rizvi et al., 2023). In Trinidad and Tobago, AI was projected to reduce screening costs by up to 20%, representing approximately US\$60 million in annual savings (Ramoutar, 2024). However, in settings with very low labor costs (e.g., Shanghai), AI conferred no net cost benefit because software licensing represented only approximately 7% of total screening costs, and human graders were relatively inexpensive (Ramoutar, 2024). This finding suggests that the economic case for AI depends more on the opportunity cost of human grading labor than on the software cost itself.

Implementation Barriers and Facilitators

Across the included studies, a consistent set of implementation barriers emerges for rural tele-ophthalmology deployment. Technology infrastructure challenges—unreliable internet

connectivity and inconsistent electricity supply—were repeatedly cited as obstacles, particularly in sub-Saharan Africa and rural Asia (Amin & Rizky, 2025; Rogers et al., 2025; Onyeze et al., 2025). Workforce limitations, including shortage and retention of trained imaging staff, were also prominent (Zikhali et al., 2022; Yudistira et al., 2025). Patient-level barriers included low health literacy regarding DR, lack of awareness of the asymptomatic nature of early disease, and competing work and home responsibilities (Chauhan et al., 2025; Andayani et al., 2025).

Effective solutions demonstrated across studies include offline-capable AI systems (e.g., Medios AI on smartphone) that function without continuous internet connectivity (Chauhan et al., 2025; Onyeze et al., 2025). Task-shifting to trained non-ophthalmologist staff—nurses, community health workers, or technicians—has been successfully implemented in multiple settings, with diagnostic accuracy maintained when staff receive structured training and quality assurance oversight (Zikhali et al., 2022; Yudistira et al., 2025). Immediate point-of-care counseling with color-coded visual reports improves patient understanding and has been associated with improved referral adherence (Mathenge et al., 2022; Cleland et al., 2024). Integration of screening into routine primary care or existing diabetes clinic visits reduces the marginal burden on patients and improves participation (Mansberger et al., 2015; Pratiwi, 2026).

Quality assurance mechanisms are essential for sustainable programs. Standardized grading protocols, central expert review with masking to AI results, and systematic monitoring of ungradable image rates are standard features of the most rigorous programs (Mansberger et al., 2015; Wang et al., 2025a; Cleland et al., 2024). Regulatory pathways for AI devices (FDA, CE marking) support quality oversight, though post-market surveillance and local validation before deployment remain critical given the documented performance gap between internal and external validation (Suebsaiphrom et al., 2026; Hidayati et al., 2024).

Implications for Policy and Practice

The findings support several actionable recommendations for policymakers and healthcare administrators. First, tele-ophthalmology should be considered an evidence-based strategy for improving DR screening in rural and underserved populations, with the strongest evidence supporting store-and-forward models using non-mydratic fundus cameras or high-quality

smartphone-based devices (Mansberger et al., 2013; Ha et al., 2025; Prayogo et al., 2023). Second, AI-based systems offer comparable or slightly superior diagnostic accuracy to human grading and may improve referral adherence when combined with immediate feedback, but require local validation before deployment (Mathenge et al., 2022; Wang et al., 2025a; Suebsaiphrom et al., 2026). Third, biennial screening is generally cost-effective in rural settings, with annual screening exceeding cost-effectiveness thresholds in most LMIC contexts (Avidor et al., 2020; Andayani et al., 2025). Fourth, screening programs must be accompanied by investments in referral infrastructure, patient navigation, and financial support for travel and copayments; without these, even accurate screening will fail to translate into improved clinical outcomes (Chauhan et al., 2025; Rahmati et al., 2025).

For future research, priorities include: (1) long-term randomized trials comparing AI-based versus human-graded tele-ophthalmology on clinical outcomes including vision preservation; (2) implementation science studies identifying strategies to sustain screening participation beyond the initial years and improve referral adherence in low-resource settings; (3) economic evaluations incorporating health system perspective and long-term downstream costs of untreated DR; and (4) context-specific validation studies of AI systems before large-scale deployment, particularly in rural LMIC settings with different patient and equipment characteristics than training datasets.

CONCLUSION AND RECOMMENDATIONS

Conclusion

This systematic review demonstrates that tele-ophthalmology and AI-based retinal imaging are effective, accurate, and cost-effective strategies for diabetic retinopathy screening in rural and underserved populations. The evidence shows that tele-ophthalmology substantially improves initial screening completion rates—from approximately 56% to 94% in controlled trials—though long-term participation tends to decline without sustained engagement strategies. Diagnostic accuracy for referable DR consistently meets or exceeds regulatory standards, with pooled sensitivity of 0.90-0.95 and specificity of 0.81-0.91 across modalities. AI-based systems offer additional advantages in speed, consistency, and immediate feedback, which can improve referral adherence by

approximately 50-100% when combined with financial and logistical support. Cost-effectiveness is well-established for biennial screening in most rural settings, with ICERs below customary willingness-to-pay thresholds.

Recommendations: For implementation, we recommend: (1) deploying store-and-forward tele-ophthalmology with non-mydriatic cameras or validated smartphone devices in rural primary care clinics; (2) incorporating selective dilation protocols to manage ungradable images; (3) using AI systems only after local validation and with quality assurance oversight; (4) embedding screening within existing diabetes care workflows; (5) providing travel reimbursement and navigation support for referred patients; and (6) implementing reminder systems and engagement strategies to sustain long-term participation. Future research should prioritize long-term clinical outcomes, implementation strategies for sustained engagement, and context-specific validation of AI systems across diverse rural settings.

REFERENCES

1. Mansberger S, Gleitsmann K, Gardiner S, et al (2013) Comparing the Effectiveness of Telemedicine and Traditional Surveillance in Providing Diabetic Retinopathy Screening Examinations: A Randomized Controlled Trial. *Telemedicine journal and e-health*. <https://doi.org/10.1089/tmj.2012.0313>
2. Mansberger S, Shepler CR, Barker G, et al (2015) Long-term Comparative Effectiveness of Telemedicine in Providing Diabetic Retinopathy Screening Examinations: A Randomized Clinical Trial. *JAMA ophthalmology*. <https://doi.org/10.1001/jamaophthalmol.2015.1>
3. Amin R, Rizky DA (2025) Tele-ophthalmology versus Traditional Fundus Photography for Diabetic Retinopathy Screening: A Comparative Meta-analysis of Diagnostic Accuracy, Cost-Effectiveness, and Health Policy Uptake. *Community Medicine and Education Journal*. <https://doi.org/10.37275/cmej.v5i2.729>
4. Sadjadi R (2021) Assessing the Performance of Artificial Intelligence Systems for the Screening of Diabetic Retinopathy. <https://doi.org/10.20944/preprints202108.0030.v1>

5. Wang T-W, Luo W-T, Tu Y-K, et al (2025) Systematic review and meta-analysis of regulator-approved deep learning systems for fundus diabetic retinopathy detections. *npj Digital Medicine*. <https://doi.org/10.1038/s41746-025-02223-8>
6. Mathenge W, Whitestone N, Nkurikiye J, et al (2022) Impact of Artificial Intelligence Assessment of Diabetic Retinopathy on Referral Service Uptake in a Low-Resource Setting. *Ophthalmology Science*. <https://doi.org/10.1016/j.xops.2022.100168>
7. Chauhan A, Vale L, Kankaria A, et al (2025) Assessing adoption of human and AI-enabled diabetic retinopathy screening in primary healthcare settings: findings from a pragmatic trial. *Archives of Public Health*. <https://doi.org/10.1186/s13690-025-01757-3>
8. Sharafeldin N, Kawaguchi A, Sundaram A, et al (2018) Review of economic evaluations of teleophthalmology as a screening strategy for chronic eye disease in adults. *British Journal of Ophthalmology*. <https://doi.org/10.1136/bjophthalmol-2017-311452>
9. Lee J, Lee SWH (2018) Telemedicine Cost–Effectiveness for Diabetes Management: A Systematic Review. *Diabetes Technology & Therapeutics*. <https://doi.org/10.1089/dia.2018.0098>
10. Pratiwi N (2026) A Comprehensive Systematic Review of The Role of Screening Programs in Preventing Blindness in Primary Setting. *International journal of medical science and health research*. <https://doi.org/10.70070/ppw50b18>
11. Avidor D, Loewenstein A, Waisbourd M, Nutman A (2020) Cost-effectiveness of diabetic retinopathy screening programs using telemedicine. *Cost Effectiveness and Resource Allocation*. <https://doi.org/10.1186/s12962-020-00211-1>
12. Zikhali T, Kalinda C, Xulu-Kasaba Z (2022) Screening of Diabetic Retinopathy Using Teleophthalmology to Complement Human Resources for Eye Health. *Clinics and Practice*. <https://doi.org/10.3390/clinpract12040050>
13. Piyasena MMPN, Murthy G, Yip J, et al (2018) Systematic review and meta-analysis of diagnostic accuracy of detection of any level of diabetic retinopathy using digital retinal imaging. *Systematic Reviews*. <https://doi.org/10.1186/s13643-018-0846-y>

14. Pasquel F, Hendrick A, Ryan M, et al (2016) Cost-effectiveness of Different Diabetic Retinopathy Screening Modalities. Journal of Diabetes Science and Technology. <https://doi.org/10.1177/1932296815624109>
15. Chidi R, Odimba U (2024) AI APPLICATIONS IN SCREENING AND DIAGNOSIS OF DIABETIC RETINOPATHY IN RURAL SETTINGS. International medical science research journal. <https://doi.org/10.51594/imsrj.v4i3.918>
16. Chid R, Akubue I (2024) Telemedicine in diabetic eye care: A meta-analysis of its effectiveness in underserved populations. World Journal of Biology Pharmacy and Health Sciences. <https://doi.org/10.30574/wjbphs.2024.17.3.0104>
17. Andayani A, Hendrawan K, Made N, et al (2025) Economic Evaluation of Diabetic Retinopathy Screening in Developing Countries. Korean Journal of Ophthalmology. <https://doi.org/10.3341/kjo.2025.0047>
18. Ha SK, Gilbert JB, Le E, et al (2025) Impact of teleretinal screening program on diabetic retinopathy screening compliance rates in community health centers: a quasi-experimental study. BMC Health Services Research. <https://doi.org/10.1186/s12913-025-12472-8>
19. Walsh LG, Hong SC, Chalakkal RJ, Ogbuehi K (2021) A Systematic Review of Current Teleophthalmology Services in New Zealand Compared to the Four Comparable Countries of the United Kingdom, Australia, United States of America (USA) and Canada. Clinical Ophthalmology. <https://doi.org/10.2147/OPTH.S294428>
20. Prayogo ME, Zaharo AF, Damayanti NNR, et al (2023) Accuracy of Low-Cost, Smartphone-Based Retinal Photography for Diabetic Retinopathy Screening. Clinical Ophthalmology. <https://doi.org/10.2147/OPTH.S416422>
21. Ramoutar RR (2024) An Economic Analysis for the Use of Artificial Intelligence in Screening for Diabetic Retinopathy in Trinidad and Tobago. Cureus. <https://doi.org/10.7759/cureus.55745>
22. Kumari R (2025) The Truth About Diabetic Eye Screening: A Deep Dive into Diagnostic Accuracy. Journal of Pharmaceutical Research. <https://doi.org/10.18579/jopcr/v24.i3.82>

23. Rizvi A, Rizvi F, Lalakia P, et al (2023) Is Artificial Intelligence the Cost-Saving Lens to Diabetic Retinopathy Screening in Low- and Middle-Income Countries? Cureus. <https://doi.org/10.7759/cureus.45539>
24. Wang T-W, Luo W-T, Tu Y-K, et al (2025) Diagnostic Accuracy of EyeArt for Fundus-Based Detection of Diabetic Retinopathy: A Systematic Review and Meta-analysis. American journal of ophthalmology-glaucoma. <https://doi.org/10.37766/inplasy2025.5.0079>
25. Samara SM, Laksono AK (2025) Evaluating the accuracy of Artificial Intelligence (AI)-integrated, Smartphone-based screening for Diabetic Retinopathy: Systematic Review. International Journal of Retina. <https://doi.org/10.35479/ijretina.2025.vol008.iss001.294>
26. Yudistira Y, Hendrawan K, Andayani A, et al (2025) Diabetic Retinopathy Screening Approaches in Developing Countries. Turkish journal of ophthalmology. <https://doi.org/10.4274/tjo.galenos.2025.32916>
27. Onyeze NS, Sartawi S, Nayyer Z (2025) Effectiveness of AI-Based Tools in Detecting Diabetic Retinopathy in Low- and Middle-Income Countries: A Systematic Review of Diagnostic Performance and Implementation Feasibility. Cureus. <https://doi.org/10.7759/cureus.96554>
28. Lestari YD, Jayanegara AP (2023) The Validity of Artificial Intelligence (AI) for Diabetic Retinopathy Screening in Asia. International Journal of Retina. <https://doi.org/10.35479/ijretina.2023.vol006.iss002.241>
29. Vilela M, Valença F, Barreto PKM, et al (2018) Agreement between retinal images obtained via smartphones and images obtained with retinal cameras or fundoscopic exams –. Clinical Ophthalmology. <https://doi.org/10.2147/OPTH.S182022>
30. Noriega A, Meizner D, Camacho D, et al (2020) Screening Diabetic Retinopathy Using an Automated Retinal Image Analysis System in Independent and Assistive Use Cases in Mexico: Randomized Controlled Trial. JMIR Formative Research. <https://doi.org/10.2196/25290>
31. Wang Z, Li Z, Li K, et al (2023) Performance of artificial intelligence in diabetic retinopathy screening. Frontiers in Endocrinology. <https://doi.org/10.3389/fendo.2023.1197783>

32. Rogers M, Kaggwa F, Arunga S, Wasswa W (2025) Feasibility and Equity of AI-Driven Models for Diabetic Retinopathy Screening in Sub-Saharan Africa: A Systematic Review and Meta-Analysis. *Cureus Journal of Computer Science*. <https://doi.org/10.7759/s44389-025-07263-5>
33. Cleland CR, Bascaran C, Makupa W, et al (2024) Artificial intelligence-supported diabetic retinopathy screening in Tanzania: rationale and design of a randomised controlled trial. *BMJ Open*. <https://doi.org/10.1136/bmjopen-2023-075055>
34. Hidayati N, Widyawati IY, Suarilah I, et al (2024) Diagnostic Accuracy of Artificial Intelligence-Based Systems for Detecting Diabetic Retinopathy: A Systematic Review. *Indonesian Journal of Global Health Research*. <https://doi.org/10.37287/ijghr.v7i1.4011>
35. Suebsaiphrom N, Takkametha T, Vanichakulthada N, Vongkancom U (2026) Diagnostic Performance of Artificial Intelligence and Deep Learning for Diabetic Retinopathy Screening: A Systematic Review and Meta-analysis. *Journal of Optometry and Ophthalmology Research*. <https://doi.org/10.61440/joor.2026.v2.03>
36. Rahmati M, Smith L, Piyasena MMPN, et al (2025) Artificial Intelligence improves follow-up appointment uptake for diabetic retinal assessment: a systematic review and meta-analysis. *Eye*. <https://doi.org/10.1038/s41433-025-03849-4>
37. Moannaei M, Jadidian F, Doustmohammadi T, et al (2025) Performance and limitation of machine learning algorithms for diabetic retinopathy screening and its application in health management: a meta-analysis. *BioMedical Engineering OnLine*. <https://doi.org/10.1186/s12938-025-01336-1>
38. Scott S, Grewal E, Varma M, Campbell DJT (2025) Evaluation of a pilot outreach program to support point-of-care screening for individuals with diabetes who are experiencing homelessness in Alberta, Canada: results of a mixed methods study. *BMC Health Services Research*. <https://doi.org/10.1186/s12913-025-13316-1>