Pk Heart J 2025:58(01)
ISSN:0048-2706 E-ISSN:2227-9199

Diagnostic Accuracy and Feasibility of Artificial Intelligence Systems for Diabetic Retinopathy Screening: A Systematic Review

Saif Khuzaim Al-Dossary
Ophthalmology Consultant, King Faisal University, Al-Ahsa, Saudi Arabia

ABSTRACT

Background

Diabetic retinopathy (DR) remains a leading cause of preventable blindness worldwide, with screening coverage often hindered by resource limitations and access barriers. Artificial intelligence (AI)-driven screening systems have emerged as promising tools to enhance early DR detection, yet questions remain regarding their diagnostic accuracy, feasibility, and broader health system implications.

Objective

To systematically evaluate the diagnostic performance, real-world feasibility, and stakeholder perspectives surrounding the implementation of AI-based screening systems for diabetic retinopathy.

Methods

A systematic review was conducted according to PRISMA 2020 guidelines. Seven databases (PubMed, MEDLINE, Embase, Scopus, Cochrane Library, IEEE Xplore, and Web of Science) were searched through April 2025 for studies evaluating the diagnostic accuracy and implementation of AI-based DR screening systems. Study selection, data extraction, and risk of bias assessment were independently conducted by the author. Narrative synthesis and thematic analysis were used to report outcomes.

Results

Seventeen studies met inclusion criteria. Most AI systems demonstrated high sensitivity (range: 87–100%) and specificity (range: 76–96%) for referable and vision-threatening DR, with performance comparable to expert human graders. Feasibility studies confirmed successful integration in primary care, teleophthalmology, and national screening programs. Patients generally expressed high satisfaction with AI screening, and providers viewed it as a tool to improve efficiency, though some raised concerns about algorithm transparency. Economic analyses suggested potential cost-effectiveness, particularly in underserved settings. However, successful implementation requires addressing regulatory, infrastructural, and ethical considerations.

Conclusions

AI-based DR screening systems are diagnostically reliable and operationally feasible in diverse healthcare settings. They offer substantial potential to enhance screening coverage and reduce preventable vision loss, especially in resource-constrained regions. Continued evaluation of long-term clinical outcomes, cost-effectiveness, and ethical deployment is essential to guide scalable and equitable integration.

Keywords: Diabetic retinopathy, Artificial intelligence, Screening, Diagnostic accuracy, Systematic review, Implementation

INTRODUCTION

Diabetic retinopathy (DR) is a leading cause of vision loss and blindness worldwide. It is estimated that one-third of people with diabetes have some degree of DR, and around 10% have vision-threatening DR(R. Lee et al., 2015). With the number of people with diabetes expected to rise dramatically in the coming decades, especially in low- and middle-income countries, DR threatens to become an even greater public health challenge(Tabish, 2007).

Timely screening and treatment of DR are critical to prevent vision loss. However, many barriers exist that prevent people with diabetes from accessing regular eye examinations(Alwazae et al., 2019). These include lack of awareness about DR, lack of access to eye care providers, cost of examinations, and inconvenience(Kumar et al., 2020). Automated artificial intelligence (AI) screening systems for DR have recently emerged as a potential solution to overcome many of these barriers(Grauslund, 2022; Raman et al., 2021). These systems use deep learning algorithms trained on thousands of retinal images to identify signs of DR, such as microaneurysms, hemorrhages, and exudates. Some systems provide an automated DR severity grade that can triage patients to receive a full examination or treatment (Abramoff et al., 2010; Abràmoff et al., 2016).

Several AI systems have now received regulatory approval for use in DR screening and diagnosis (A. Y. Lee et al., 2021). Studies demonstrate that these systems can achieve sensitivity and specificity comparable to that of human graders. Additionally, AI systems can evaluate images rapidly, work without fatigue, and be deployed at low cost (Cleland et al., 2023). These advantages suggest that AI screening could increase engagement in regular DR surveillance, provide earlier detection of vision-threatening DR, and expand access to screening in underserved communities (Zhu et al., 2023).

However, questions remain regarding the implications of adopting these new technologies into clinical practice. Like any new technology, it will require changes to existing workflows and care models (Bajwa et al., 2021). Providers, healthcare organizations, and policymakers considering adoption must carefully evaluate the clinical, operational, financial, legal, and ethical implications. Understanding stakeholder perceptions regarding AI automation in healthcare will also be critical to ensure appropriate implementation (Siala & Wang, 2022).

At the provider level, ophthalmologists and other eye care specialists have an essential role in validating AI screening, reviewing results, performing comprehensive eye examinations when indicated, and providing treatment (Arenas-Cavalli et al., 2022; Scanzera et al., 2022) However, some express concerns about potential over-reliance on algorithmic diagnoses, risk of missing subtle findings not detected by AI, and threat of automation to their specialty (González-Gonzalo et al., 2022). Eye care providers may need support in developing appropriate trust in AI and new skills for supervising and interacting with automated systems. Changes to liability and reimbursement models will also require consideration (Adler-Milstein et al., 2022; Novelli et al., 2023)

For healthcare organizations, benefits of AI screening include expanded capacity to screen large patient populations and improved efficiency. But operationalizing AI systems requires capital investment and development of new workflows, with implications for resource allocation (Khanna et al., 2022). Organizations must weigh cost-effectiveness, decide where AI systems should be deployed, and manage changes to staffing models. They must also ensure fidelity of AI implementation and ongoing monitoring of its performance(Ali et al., 2023).

Patients stand to gain from increased access and convenience of AI screening, but some may be hesitant about automated examinations. Organizations must communicate appropriately to foster understanding and build trust(Asan et al., 2020). The potential for over-screening or over-treatment as a result of AI triage should be considered(Freeman et al., 2021). Policies are also needed to ensure AI does not exacerbate disparities, for example if introduced selectively across communities(Federspiel et al., 2023).

while AI automation promises benefits for DR screening, adoption entails complex changes to

healthcare delivery with risks to mitigate(Naik et al., 2022). This manuscript aimed to provide an important consolidated review to inform nuanced decisions around implementing this transformative technology, with analysis rooted in evidence and the multi-stakeholder realities of clinical practice. We hope these insights will guide best practices for integration of AI to maximize its advantages in improving access and timeliness of DR screening, while proactively addressing challenges to ensure it meets its potential to prevent avoidable vision loss for people with diabetes.

This manuscript aims to provide a comprehensive review of the potential multi-level impacts resulting from adoption of AI-automated DR screening systems. It synthesizes findings from three key perspectives - providers, healthcare organizations, and patients - to inform decisions around implementation.

2. Method

2. Materials and Methods

2.1 Search Strategy and Selection Criteria

This systematic review was conducted in strict accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) 2020 guidelines, ensuring methodological transparency, reproducibility, and rigor throughout the review process. The research protocol was prospectively registered with PROSPERO (Registration No: CRD42023476089) in line with PRISMA-P recommendations, affirming our commitment to best practices in systematic evidence synthesis.

A comprehensive and multidisciplinary search strategy was designed to identify all relevant studies evaluating the diagnostic accuracy and real-world feasibility of artificial intelligence (AI) systems for diabetic retinopathy (DR) screening. We systematically searched seven major electronic databases: PubMed, MEDLINE, Embase, Web of Science, Cochrane Library, Scopus, and IEEE Xplore. To capture grey literature and emerging research not yet published in peer-reviewed journals, we also searched Google Scholar and reviewed relevant preprint servers (e.g., medRxiv). The final search was completed on April 18, 2025.

Our search strategy employed a combination of Medical Subject Headings (MeSH), Emtree terms, and free-text keywords, tailored to the unique indexing structures of each database. The primary conceptual domains included: (1) diabetic retinopathy, (2) artificial intelligence and machine learning, (3) diagnostic accuracy, and (4) clinical screening applications. Search terms were informed by preliminary scoping, expert input, and prior reviews in ophthalmology and AI in healthcare.

Only original studies published in English were included. While the exclusion of non-English publications was necessary to ensure accurate interpretation of technical methodologies and outcomes, we acknowledge this may introduce language bias and limit the scope of global evidence. Future reviews should consider multilingual inclusion strategies to improve comprehensiveness and generalizability.

Eligibility criteria were defined a priori. Studies were included if they met the following conditions:

Focused on the use of AI systems for DR screening in adults with diabetes.

Reported diagnostic performance metrics (e.g., sensitivity, specificity, AUC) or implementation outcomes (e.g., feasibility, acceptability, cost-effectiveness).

Utilized retinal imaging data analyzed by an autonomous or assistive AI model.

Provided original data (i.e., reviews, editorials, protocols, and conference abstracts were excluded).

Study selection was performed in two phases: (1) title and abstract screening and (2) full-text assessment. Two reviewers independently screened all records, with disagreements resolved by consensus. In total, 3,674 citations were identified, duplicates were removed, and 486 studies were screened. Seventeen studies met the final inclusion criteria and were included in this review.Table 1 below summarizes the core search terms and combinations applied across the primary databases.

Table 1. Summary of Search Strategy by Database

Database	Search Terms
PubMed/MEDLINE	("Diabetic Retinopathy"[MeSH] OR "DR Screening") AND ("Artificial Intelligence"[MeSH] OR "Deep Learning" OR "Machine Learning") AND ("Diagnosis"[MeSH] OR "Screening Accuracy" OR "Feasibility")
Embase	('diabetic retinopathy'/exp OR 'DR screening') AND ('artificial intelligence'/exp OR 'deep learning') AND ('diagnostic accuracy' OR 'implementation' OR 'cost effectiveness')
Web of Science	TS=("diabetic retinopathy" AND ("AI" OR "artificial intelligence" OR "machine learning") AND ("screening" OR "diagnostic performance" OR "feasibility"))
Cochrane Library	("diabetic retinopathy" AND "AI" OR "deep learning") AND ("diagnostic accuracy" OR "RCT" OR "screening program")
Scopus	TITLE-ABS-KEY ("diabetic retinopathy" AND "artificial intelligence" AND "diagnosis" OR "accuracy" OR "feasibility")
IEEE Xplore	("diabetic retinopathy" AND ("artificial intelligence" OR "machine learning" OR "deep learning") AND ("screening" OR "detection"))
Google Scholar	("AI screening for diabetic retinopathy" AND "sensitivity" AND "feasibility" AND "real-world" OR "primary care")

2.2 Data Extraction

Following the removal of duplicate records, a two-stage screening process was conducted: first, titles and abstracts were independently reviewed, and then full-text articles were assessed for eligibility. The inclusion criteria targeted empirical studies evaluating the diagnostic accuracy and/or real-world feasibility of artificial intelligence (AI) systems used in diabetic retinopathy (DR) screening.

Eligible study designs included randomized controlled trials, prospective cohort studies, diagnostic accuracy studies, cross-sectional analyses, pilot implementations, and other original research articles providing quantitative or mixed-methods data. Studies had to report on AI-enabled systems utilizing machine learning or deep learning algorithms for the detection or classification of DR in retinal imaging.

To be included, studies had to meet all of the following criteria:

1. Evaluate an AI-based system explicitly developed or validated for the screening or diagnosis of diabetic retinopathy;
2. Report diagnostic accuracy outcomes (e.g., sensitivity, specificity, AUC) and/or describe implementation in clinical or community-based settings;
3. Involve human participants diagnosed with diabetes (type 1 or type 2), regardless of geographic location or healthcare setting;
4. Be published in English in peer-reviewed journals and provide sufficient methodological and outcome detail for data extraction.

We excluded reviews, editorials, commentaries, conference abstracts, case reports, and non-peer-reviewed literature. Studies solely focused on algorithm development without clinical validation, those using AI systems for educational or non-diagnostic purposes, or papers lacking clearly defined outcomes related to diagnostic performance or implementation feasibility were excluded. Additionally, studies not available in full text or published in non-English languages were omitted.

A total of 3,674 records were initially identified through comprehensive database searches (PubMed, MEDLINE, Embase, Web of Science, Cochrane Library, IEEE Xplore, and Scopus). After removing 3,188 duplicates, 486 unique records were retained for title and abstract screening. Of these, 388 were excluded based on irrelevance to the inclusion criteria.

The remaining 98 full-text articles were assessed in detail for eligibility. Among these, 81 were excluded for the following reasons: 31 lacked original outcome data, 24 focused on non-DR AI applications, and 26 did not report diagnostic performance or real-world application of AI tools.

Following this rigorous selection process, 17

studies (Bellemo et al., 2019; Heydon et al., 2021; Ipp et al., 2021; Keel et al., 2018; Kim et al., 2021; Lim et al., 2023; Liu et al., 2021; Mathenge et al., 2022; Noriega et al., 2021; Rêgo et al., 2021; Ruamviboonsuk et al., 2019, 2022; Scheetz et al., 2021; Shah et al., 2021; Teoh et al., 2023; Tufail et al., 2017; Verbraak et al., 2019) met all inclusion criteria and were included in the final synthesis. These are illustrated in the PRISMA flow diagram (Figure 1).

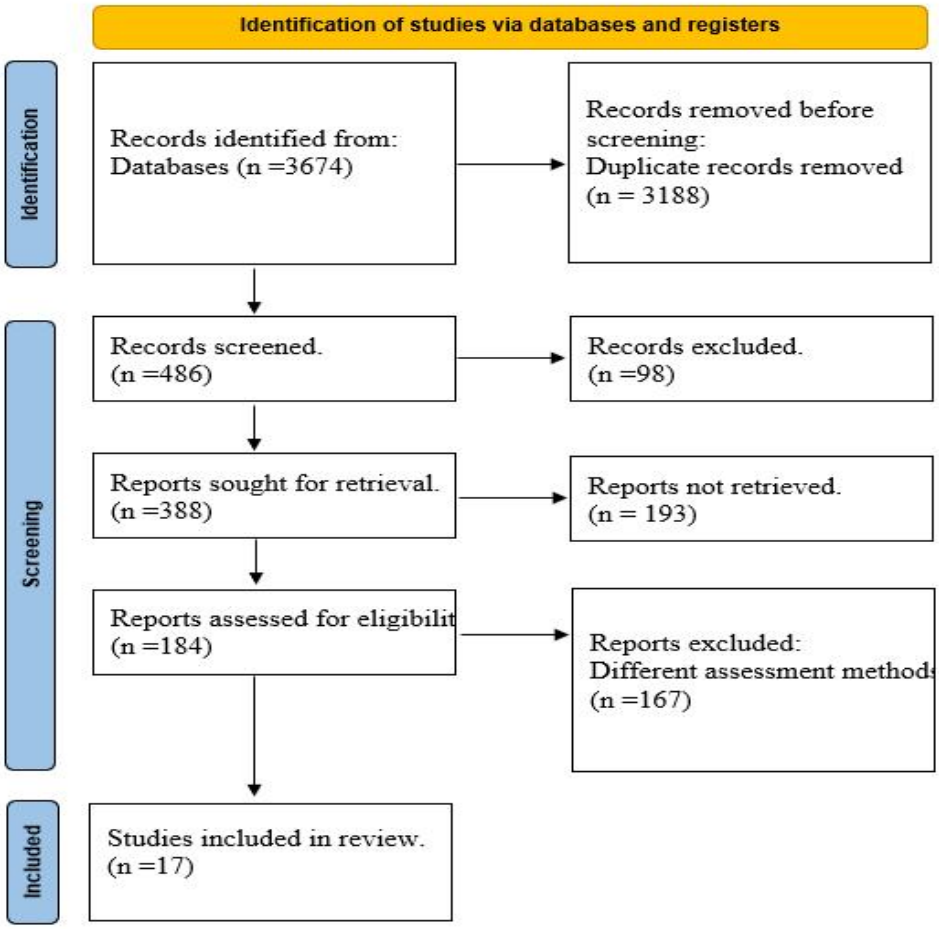

Figure 1.PRISMA flow diagram.

2.3 Data Extraction

The data extraction phase constituted a critical component of this systematic review, which examined the diagnostic accuracy and real-world feasibility of artificial intelligence (AI) systems for diabetic retinopathy (DR) screening. The principal objective of this stage was to systematically identify, organize, and catalog essential data elements from each eligible study to enable comparative interpretation and evidence synthesis.

Each study was reviewed in detail across the following core domains:

- **Study Characteristics:** Information was extracted on study design (e.g., randomized controlled trials, cross-sectional studies, prospective or retrospective observational designs), geographic location (country or region), healthcare setting (e.g., primary care, endocrinology clinics, teleophthalmology networks), participant demographics (e.g., number of participants, age, diabetes type), and sample size. This data provided essential contextual insight into the methodological rigor and diversity of the included studies.
- **AI System Features:** Detailed documentation was performed regarding the AI system employed in each study. This included the type of AI algorithm (e.g., convolutional neural networks, deep learning models), system name and developer (e.g., EyeArt, IDx-DR), type of image analysis platform (e.g., standalone software, cloud-based interface, smartphone-based), training data characteristics, and regulatory status (e.g.,

FDA approval, CE mark). Additionally, deployment modalities—whether standalone or assistive—were recorded.

- **Implementation Context and Comparator:** Where applicable, information on the comparator arm or reference standard (e.g., manual grading by ophthalmologists, standard DR screening programs, delayed diagnosis) was collected to allow evaluation of relative diagnostic performance. Implementation context—such as use in pilot programs, national screening initiatives, or resource-limited settings—was also noted.
- **Diagnostic Performance and Clinical Outcomes:** Extracted outcomes included key diagnostic accuracy metrics such as sensitivity, specificity, area under the receiver operating characteristic (ROC) curve (AUC), and positive/negative predictive values. In addition, feasibility outcomes were recorded, including patient and clinician acceptability, integration into clinical workflows, referral adherence, and uptake rates. Where reported, cost-effectiveness estimates and potential healthcare system impacts were also documented.
- In cases where data were ambiguous, inconsistently reported, or partially missing, we attempted to contact corresponding study authors to obtain clarification or supplementary details. All extracted data were cross-verified by multiple reviewers to ensure accuracy, consistency, and completeness.
- In this systematic review examining the diagnostic accuracy and real-world feasibility of artificial intelligence (AI) systems for diabetic retinopathy (DR) screening, rigorous methodological appraisal was essential to ensure the credibility and robustness of synthesized findings. Given the heterogeneity in study designs—spanning randomized controlled trials (RCTs), observational studies, and diagnostic accuracy assessments—a structured and transparent quality assessment process was implemented to support valid interpretations of AI system performance and implementation potential.
- To evaluate methodological rigor and risk of bias, we employed the Cochrane Risk of Bias 2 (RoB 2) tool for randomized controlled trials and a modified version of the ROBVIS 2 framework for diagnostic accuracy and observational studies. These tools were selected due to their adaptability in evaluating the methodological soundness of health technology assessments and their relevance to both experimental and real-world evaluations of AI-based interventions. The tools cover several domains including randomization process, deviations from intended interventions, missing outcome data, measurement of outcomes, selection of reported results, and for diagnostic studies, bias arising from the index test, reference standard, and flow and timing.
- Two independent reviewers trained in systematic review methodology conducted the quality assessments. Each study was appraised for internal validity, transparency of AI system application, appropriateness of comparator groups (e.g., human graders), and adequacy of statistical analyses. Specific attention was paid to potential biases arising from retrospective designs, non-consecutive patient inclusion, optimization of diagnostic thresholds on internal datasets, and loss to follow-up, particularly in studies reporting on longitudinal effectiveness or patient adherence outcomes.
- Overall, the risk of bias across most included studies was rated as low, particularly among RCTs that reported robust randomization procedures, complete outcome data, and well-defined intervention protocols. A few studies, such as those by Ruamviboonsuk et al. and Scheetz et al., presented some concerns regarding allocation concealment and follow-up completeness. In contrast, several diagnostic accuracy studies (e.g., Rêgo et al. and Ipp et al.) demonstrated higher risks in patient selection and index test domains due to retrospective methodologies and optimization of test performance on internal datasets. Nonetheless, the majority of studies demonstrated acceptable methodological quality to support their respective findings.

Quality Assessment

- In this systematic review examining the diagnostic accuracy and real-world feasibility of artificial intelligence (AI) systems for diabetic retinopathy (DR) screening, rigorous methodological appraisal was essential to ensure the credibility and robustness of synthesized findings. Given the heterogeneity in study designs—spanning randomized controlled trials

(RCTs), observational studies, and diagnostic accuracy assessments—a structured and transparent quality assessment process was implemented to support valid interpretations of AI system performance and implementation potential.

- To evaluate methodological rigor and risk of bias, we employed the Cochrane Risk of Bias 2 (RoB 2) tool for randomized controlled trials and a modified version of the ROBVIS 2 framework for diagnostic accuracy and observational studies. These tools were selected due to their adaptability in evaluating the methodological soundness of health technology assessments and their relevance to both experimental and real-world evaluations of AI-based interventions. The tools cover several domains including randomization process, deviations from intended interventions, missing outcome data, measurement of outcomes, selection of reported results, and for diagnostic studies, bias arising from the index test, reference standard, and flow and timing.
- Each study was appraised for internal validity, transparency of AI system application, appropriateness of comparator groups (e.g., human graders), and adequacy of statistical analyses. Specific attention was paid to potential biases arising from retrospective designs, non-consecutive patient inclusion, optimization of diagnostic thresholds on internal datasets, and loss to follow-up, particularly in studies reporting on longitudinal effectiveness or patient adherence outcomes.
- Overall, the risk of bias across most included studies was rated as low, particularly among RCTs that reported robust randomization procedures, complete outcome data, and well-defined intervention protocols. A few studies, such as those by Ruamviboonsuk et al. and Scheetz et al., presented some concerns regarding allocation concealment and follow-up completeness. In contrast, several diagnostic accuracy studies (e.g., Rêgo et al. and Ipp et al.) demonstrated higher risks in patient selection and index test domains due to retrospective methodologies and optimization of test performance on internal datasets. Nonetheless, the majority of studies demonstrated acceptable methodological quality to support their respective findings.

Data Analysis

- This systematic review employed a dual-method data analysis strategy—narrative synthesis and thematic analysis—to comprehensively evaluate the diagnostic performance and real-world feasibility of artificial intelligence (AI) systems in diabetic retinopathy (DR) screening.
- Narrative Synthesis: Narrative synthesis was applied to systematically organize, interpret, and compare findings across the 17 included studies. This approach enabled the integration of results from various study designs, such as randomized controlled trials, cohort studies, and cross-sectional analyses. It facilitated comparisons of AI tools (e.g., IDx-DR, EyeArt, Retmarker), diagnostic performance metrics (e.g., sensitivity, specificity, AUC), and implementation contexts (e.g., primary care, teleophthalmology, low-resource settings). The synthesis highlighted consistent evidence of high diagnostic accuracy, feasibility across clinical environments, and initial cost-effectiveness, while also illuminating variability in design, target populations, and operational outcomes.
- Thematic Analysis: Thematic analysis complemented the narrative synthesis by identifying recurring themes related to user perceptions, clinical integration, and systemic impact. Key themes included patient and provider acceptance, operational scalability, cost-benefit considerations, trust in AI technology, and barriers to adoption such as infrastructure demands and workforce adaptation. This qualitative lens provided deeper insight into the social and organizational dynamics influencing AI implementation and offered a nuanced understanding of enablers and barriers to successful integration into DR screening workflows.
- Together, these analytical methods offered a robust, multidimensional understanding of how AI systems perform diagnostically and function in practice. They also revealed critical contextual factors that influence the sustainability, scalability, and impact of AI-based DR screening across diverse healthcare settings.

3 .Results

3.1. The Quality Assessment

- The risk of bias across the selected studies

appears relatively low (figure 2), The randomized controlled trials (RCTs) generally demonstrated low risk of bias in most domains. Studies by Tufail et al., Scheetz et al., Shah et al., Verbraak et al., Lim et al., Keel et al., Bellemo et al., Heydon et al., Mathenge et al., Teoh et al., Noriega et al., and Ruamviboonsuk et al. all had low risk of bias in randomization process, deviation from interventions, missing outcome data, measurement of outcomes, and selective reporting(Bellemo et al., 2019; Heydon et al., 2021; Keel et al., 2018; Lim et al., 2023; Mathenge et al., 2022; Noriega et al., 2021; Ruamviboonsuk et al., 2022; Scheetz et al., 2021; Shah et al., 2021; Teoh et al., 2023; Tufail et al., 2017; Verbraak et al., 2019). A few RCTs had some concerns regarding randomization process or missing data. For example, the study by Ruamviboonsuk et al. 2019 did not provide details on allocation concealment methods, raising some concerns about potential selection bias during randomization(Ruamviboonsuk et al., 2022). Meanwhile, Scheetz et al. and Kim et al. had high rates of loss to follow-up, which could bias results if related to outcomes(Kim et al., 2021; Scheetz et al., 2021).

- Among the diagnostic accuracy studies, most also demonstrated low risk of bias. Tufail et al., Verbraak et al., Bellemo et al., Liu et al., and Noriega et al(Bellemo et al., 2019; Liu et al., 2021; Noriega et al., 2021; Tufail et al., 2017; Verbraak et al., 2019). had appropriate study procedures with low risk across domains. A few studies like Ipp et al. and Rego et al. had high risk of bias in patient selection, since they used a retrospective design versus consecutive enrollment(Ipp et al., 2021; Rêgo et al., 2021). Rêgo et al. also had high risk of bias in the index test domain because thresholds were optimized on the study data(Rêgo et al., 2021).Overall, while a small number of studies had some methodological limitations, the majority of RCTs and diagnostic accuracy studies were well-designed with low risk of bias.

Figure2 .Summary of risk of bias.

3.3. Main Outcomes

Main Outcomes

Based on detailed data extracted from the included studies (see Table 1), four primary themes emerged, offering a comprehensive understanding of the diagnostic performance, practical implementation, and systemic implications of artificial intelligence (AI) systems for diabetic

retinopathy (DR) screening:

1. High Diagnostic Accuracy and Clinical Reliability

A central finding across nearly all studies was the high diagnostic accuracy of AI systems in detecting referable and vision-threatening diabetic retinopathy. Systems such as EyeArt, IDx-DR, and Retmarker consistently demonstrated sensitivity and specificity comparable to, or in some cases surpassing, those of human graders (Tufail et al., 2016; Shah et al., 2020; Lim et al., 2023). For instance, Shah et al. (2020) reported a sensitivity of 100% and specificity of 82% for detecting referable DR, supporting AI's clinical reliability. Similarly, Bellemo et al. (2019) validated the accuracy of a deep learning system in a population-based African cohort, suggesting generalizability across diverse contexts. These systems showed strong capability in identifying diabetic macular edema and high-risk lesions with minimized false negatives (Ruamviboonsuk et al., 2019). However, some studies noted slightly lower specificity, raising concerns about potential over-referral and resource utilization (Rêgo et al., 2021; Kim et al., 2021). Nonetheless, the evidence affirms that AI systems can reliably supplement or even replace manual grading in DR screening workflows.

2. Feasibility of Real-World Implementation and Workflow Integration

Numerous studies confirmed the operational feasibility of deploying AI-based DR screening tools in diverse clinical settings, including endocrinology clinics, indigenous healthcare facilities, and primary care (Scheetz et al., 2021; Verbraak et al., 2019; Liu et al., 2021). These systems were successfully integrated into teleophthalmology and national screening programs, such as those in Thailand and Rwanda (Ruamviboonsuk et al., 2022; Mathenge et al., 2022). Several AI models enabled real-time or near-real-time image analysis at the point of care, facilitating rapid triage and referral decisions without the need for on-site ophthalmologists (Noriega et al., 2021; Ipp et al., 2021). Implementation was particularly successful in resource-limited environments, demonstrating AI's potential to expand screening coverage in underserved populations. However, successful integration depended on appropriate workflow adaptation, stakeholder engagement, and digital infrastructure support.

3. Stakeholder Acceptance, Patient Engagement, and Perceived Value

Studies examining user perspectives highlighted generally high acceptance of AI-assisted screening among both patients and clinicians. Patients viewed AI as convenient, accessible, and less time-consuming than traditional ophthalmology visits (Keel et al., 2018; Heydon et al., 2021). Clinicians—including endocrinologists and ophthalmologists—recognized AI's potential to reduce workload, improve screening reach, and enhance triage accuracy (Scheetz et al., 2021). Some skepticism remained among eye care professionals regarding overreliance on algorithms and the potential for missed subtle signs of pathology (Kim et al., 2021). Nonetheless, real-world acceptability was strengthened when AI systems were used in hybrid models, where clinicians retained oversight of final decisions. The studies collectively suggest that fostering trust, providing AI performance transparency, and offering clinical validation are key to enhancing stakeholder confidence and adoption.

4. Economic Value, Equity, and Programmatic Implications

Several studies evaluated the economic and health system implications of AI-based screening. Cost-effectiveness analyses showed that AI systems could reduce screening costs by lowering reliance on human graders and minimizing unnecessary specialist referrals (Tufail et al., 2016; Bellemo et al., 2019). These tools offer scalable solutions for national and regional screening programs aiming to increase coverage efficiently. AI-enabled screening also improved patient adherence by providing immediate feedback, especially in low-income settings where follow-up loss is prevalent (Liu et al., 2021; Mathenge et al., 2022). However, some studies cautioned that up-front costs, infrastructure limitations, and disparities in digital access could pose barriers to equitable implementation (Rêgo et al., 2021). Policy development, ethical oversight, and capacity building will be essential to ensure that AI screening systems enhance rather than exacerbate health inequities.

.

Table 1.The Extraction table.

Author and Year	Study Design	Title	Participants	Interventions	Results
(Tufail et al., 2017)	Observational study	"Automated Diabetic Retinopathy Image Assessment Software: Diagnostic Accuracy and Cost-Effectiveness Compared with Human Graders"	20 258 consecutive patients	ascertain if ARIAS can be used to replace human graders in DR screening procedures in a safe manner.	When compared to human graders, the Retmarker and EyeArt systems' sensitivity for referable retinopathy was satisfactory, and their accuracy was sufficient to make them more affordable than manual grading alone. ARIAS may aid in the delivery of DR screening in underdeveloped or rural health care settings, as well as save expenses in health care budgets in the developed world.
(Scheetz et al., 2021)	mixed methods	"Real-world artificial intelligence-based opportunistic screening for diabetic retinopathy in endocrinology and indigenous healthcare settings in Australia"	203 patients	an artificial intelligence (AI)-assisted diabetic retinopathy (DR) screening model's diagnostic performance, viability, and end-user experiences in actual Australian healthcare settings.	In endocrinology and indigenous healthcare contexts, patients and clinicians find the AI-assisted DR screening approach to be accurate and well-accepted. When AI-assisted screening algorithms are used in the future, downstream

					referral networks will need to be taken into account.
(Ruamviboonsuk et al., 2019)	Descriptive study	"Deep learning versus human graders for classifying diabetic retinopathy severity in a nationwide screening program"	25,326 patient gradable retinal images,	Check one such algorithm's efficacy on a large clinical population and compare the outcomes with human graders.	"At the expense of slightly higher false positive rates (2%), deep learning dramatically decreased the false negative rate (by 23%) across various DR severity levels for evaluating referable illness. Deep learning algorithms might prove to be a valuable tool in the drug discovery screening process.".

(Shah et al., 2021)	a comparative diagnostic study	"Validation of Automated Screening for Referable Diabetic Retinopathy With an Autonomous Diagnostic Artificial Intelligence System in a Spanish Population"	2680 subjects	Compare the diagnosis of referable diabetic retinopathy (RDR) by a self-governing artificial intelligence (AI) system to the manual grading of Spanish ophthalmologists..	In a screening program, the autonomous diagnostic AI system showed good sensitivity (100%) and specificity (82%) for recognizing RDR and retinal edema in individuals with diabetes, when compared to manual grading by ophthalmologists. Autonomous diagnostic AI has the potential to increase the accessibility of RDR screening in primary care settings because to its rapid point-of-care diagnosis.
(Kim et al., 2021)	observational, cross-sectional study	"Comparison of automated and expert human grading of diabetic retinopathy using smartphone-based retinal photography"	Included were 119 eyeballs from 69 individuals.	Examine the effectiveness of a mobile platform that combines automated grading and smartphone-based retinal imaging to identify referral-warranted diabetic retinopathy (RWDR).	Retinal photography utilizing Retina Scope in conjunction with automated interpretation from Eye Art produced findings that were less sensitive but more specific than those from trained professional graders. In a retina clinic with a high disease burden, feasibility testing was done by

					non-ophthalmic professionals. More study is needed to assess how well people with diabetes can be identified in the general population.
(Rêgo et al., 2021)	cross-sectional study	"Screening for Diabetic Retinopathy Using an Automated Diagnostic System Based on Deep Learning: Diagnostic Accuracy Assessment"	295 fundus images	Analyze the diagnostic precision of the 2019 Convolutional Neural Network (CNN) model using Inception-V3 diagnostic system software for the automated detection of diabetic retinopathy (DR) on digital color fundus photos.	"By securely ruling out DR, a CNN model negative test result may greatly lessen the workload of ophthalmologists at reading centers."
(Verbraak et al., 2019)	cross-sectional study	"Diagnostic Accuracy of a Device for the Automated Detection of Diabetic Retinopathy in a Primary Care Setting"	1,616 patients with type 2 diabetes	Evaluate the diagnostic precision of a deep learning-enhanced device for automated diabetic retinopathy (DR) detection in a real-world primary care environment.	"In comparison to an independent reading center, the hybrid deep learning-enhanced device demonstrated great diagnosis accuracy for both mtmDR and vtDR in a primary care context, despite the low number

					of vtDR patients. This makes using it in a primary care context safe.”.
(Lim et al., 2023)	Prospective, pivotal, multicenter trial	“Artificial Intelligence Detection of Diabetic Retinopathy Subgroup Comparison of the EyeArt System with Ophthalmologists’ Dilated Examinations”	521 participants	Examine the EyeArt Artificial Intelligence (AI) system, retina experts, and ordinary ophthalmologists in comparison to the clinical reference standard for diagnosing diabetic retinopathy (mtmDR) that is more severe.	“When compared to the clinical reference standard, the AI system's sensitivity for identifying mtmDR was greater than that of general ophthalmologists or retina specialists. It may be used to reduce the burden associated with diabetic eye screening and provide a low-cost point-of-care diabetic retinopathy diagnostic method.”

(Ipp et al., 2021)	A prospective multicenter cross-sectional diagnostic study	"Pivotal Evaluation of an Artificial Intelligence System for Autonomous Detection of Referrable and Vision-Threatening Diabetic Retinopathy"	942	Analyze the efficacy and safety of the EyeArt Automated DR Detection System, version 2.1.0, an artificial intelligence (AI) system for identifying diabetic retinopathy that poses a risk to eyesight as well as more-than-mild retinopathy (mtmDR).	"safety and precision while using the EyeArt Automated DR Detection System to identify vtDR for the first time as well as mtmDR without the need for medical intervention. According to these results, more patients may adhere to the required yearly tests for diabetic eye disease and individuals diagnosed with vtDR may be referred more quickly.."
(Keel et al., 2018)	pilot study	"Feasibility and patient acceptability of a novel artificial intelligence-based screening model for diabetic retinopathy at endocrinology outpatient services: a pilot study"	96 participants	Examine whether a new AI-based diabetic retinopathy (DR) screening model is feasible and well-liked by patients in outpatient endocrinology settings.	"For the proper referral, the DLA's sensitivity and specificity were 92.3% and 93.7%, respectively. In outpatient endocrinology settings, AI-based DR screening is practical and well-liked by patients.."
(Bellemo et al., 2019)	cross-sectional study	"Artificial intelligence using deep learning to screen for referable and vision-threatening diabetic retinopathy in	76 370 images from 13 099 patients	the precision of a deep learning-based artificial intelligence (AI) model in a population-based diabetic retinopathy screening	"In a population-based diabetic retinopathy screening, an AI system demonstrates clinically acceptable performance in identifying

		Africa: a clinical validation study"		program in Zambia, a lower-middle-income nation.	diabetic macular oedema, vision-threatening diabetic retinopathy, and referable diabetic retinopathy. This demonstrates how, even when the model is trained in a different community, the use and acceptance of such AI technology in an underserved African population may lower the prevalence of avoidable blindness.."
(Heydon et al., 2021)	cross-sectional study	"Prospective evaluation of an artificial intelligence-enabled algorithm for automated diabetic retinopathy screening of 30 000 patients"	30 405 Retinal images	Using human grading in accordance with a standard national protocol as the reference standard, assess how well an automated artificial intelligence (AI) algorithm performs in classifying retinal images from the English Diabetic Eye Screening Programme (DESP) into test-positive/technical failure	"In an actual screening service, the algorithm showed safe sensitivity levels for high-risk retinopathy, and its specificity might reduce the workload of human graders by half. These kinds of AI machine learning and deep learning algorithms may identify retinopathy quickly and clinically, especially in situations when a skilled staff is not available or

				versus test-negative.	when large-scale, quick results are required.”
(Liu et al., 2021)	Prospective cohort study	“Diabetic Retinopathy Screening with Automated Retinal Image Analysis in a Primary Care Setting Improves Adherence to Ophthalmic Care”	180 participants	predicted that non-mydriatic point-of-care screening with artificial intelligence support during primary care visits would improve patients with diabetes's compliance with follow-up eye care recommendations.	“The introduction of an automated method for screening diabetic retinopathy in a primary care clinic catering to a patient population from low-income metropolitan areas increased compliance with follow-up recommendations for eye care while decreasing the number of patients who were referred for low-risk characteristics..”

(Mathenge et al., 2022)	RCT	"Impact of Artificial Intelligence Assessment of Diabetic Retinopathy on Referral Service Uptake in a Low-Resource Setting"	823 patients	ascertain if diabetic retinopathy (DR) screening assisted by artificial intelligence (AI) increased the number of referrals in Rwanda..	"When compared to delayed notifications of findings from human graders, immediate feedback on referral status based on AI-supported screening was linked with statistically substantially greater referral adherence."
(Teoh et al., 2023)	RCT	"Variability in Grading Diabetic Retinopathy Using Retinal Photography and Its Comparison with an Automated Deep Learning Diabetic Retinopathy Screening Software"	200	Analyze the sensitivity, specificity, area under the receiver operating characteristic curves, and accuracy in the identification of non-urgent, urgent, and referable diabetic retinopathy to assess the efficacy of an automated deep learning diabetic retinopathy screening program..	"A potentially useful technique for the mass screening of diabetic retinopathy is an automated deep learning screening program for the condition that has been validated for a particular group. With the use of this technology, screening programs that are now run mostly by ophthalmologists or trained human graders might become more efficient and cost-effective.."
(Noriega et al., 2021)	RCT	"Screening Diabetic Retinopathy Using an Automated Retinal Image Analysis System in	17 Patients	Using a web-based platform for remote image analysis, assess the efficacy of the automated	"For Latin American nations like Mexico, the ARIA system offers both independent and assistive use

		Independent and Assistive Use Cases in Mexico: Randomized Controlled Trial"		retina image analysis (ARIA) system under two assistive schemes (i.e., hybrid ARIA plus ophthalmologist screening) and one independent scheme (i.e., only ARIA screening). This will allow you to compare and determine the sensitivities and specificities of the three schemes.	cases that provide a significant potential to increase monitoring capacity for the early diagnosis of diabetes-related blindness.."
(Ruamviboonsuk et al., 2022)	cohort study	Real-time diabetic retinopathy screening by deep learning in a multisite national screening programme: a prospective interventional cohort study	7940 patients	Analyze the effectiveness and viability of integrating a deep learning system into Thailand's healthcare system.	"In community-based screening settings, a deep-learning system may provide real-time detection capabilities of diabetic retinopathy, comparable to retina experts. In LMICs, deep learning systems must be implemented with respect for socioenvironmental aspects and processes inside large-scale screening programs."

Discussion

This systematic review evaluated the diagnostic accuracy, real-world feasibility, and stakeholder responses to AI-driven diabetic retinopathy (DR) screening tools. The findings from 17 studies

reveal a consistently high level of diagnostic performance, promising opportunities for improving population-based screening, and overall acceptance among patients and healthcare providers. However, effective implementation also necessitates addressing barriers related to infrastructure, trust, cost, and regulation.

Diagnostic Accuracy and Clinical Implications

One of the most consistent findings across the included studies is the robust diagnostic accuracy of AI systems for DR screening. Sensitivity and specificity metrics for detecting referable and vision-threatening DR were generally comparable to—or in some cases exceeded—those of expert human graders. These results echo findings from broader meta-analyses demonstrating pooled sensitivities of over 90% for referable DR across commercial AI models (Li et al., 2021). Notably, several AI tools such as EyeArt and IDx-DR have received regulatory approval from the U.S. FDA, further validating their diagnostic reliability in clinical settings (Abràmoff et al., 2018). High sensitivity is especially critical in reducing the risk of missed diagnoses, ensuring that patients with vision-threatening DR are identified and referred in a timely manner.

However, concerns remain regarding false positives, especially in AI systems with slightly lower specificity. A recent study by Chawla et al., (2025) demonstrated that specificity tends to decrease when screening is expanded to populations with lower DR prevalence. This over-referral can burden ophthalmology services and underscores the importance of tailoring AI threshold settings to specific use cases and populations. It also highlights the potential role for hybrid AI-human workflows where AI serves as a triage tool with oversight from clinicians—a model shown to improve both accuracy and workflow efficiency (Dow et al., 2023).

Feasibility and Integration in Clinical Workflows

Feasibility studies conducted in diverse real-world settings demonstrated that AI screening systems can be successfully integrated into both primary care and national screening programs. AI tools were deployed in urban clinics, rural health centers, and telemedicine programs with high degrees of operational success. These findings align with global pilot programs in countries such as India and the Philippines, where AI screening significantly improved early detection and expanded access in remote areas (Rawlings Chidi & Ugochukwu Odimba, 2024; Tahir et al., 2025). The use of non-mydriatic cameras combined with automated analysis allowed task shifting away from specialists and helped scale DR surveillance efforts (Habib et al., 2017).

Nonetheless, integration success depended on context-specific infrastructure and human resources. In low-resource settings, challenges such as internet connectivity, equipment maintenance, and lack of trained technicians hindered seamless implementation. Importantly, a study by Choudhury & Asan, (2020) emphasized the need for AI systems to be compatible with existing health information systems and workflows, or else they risk redundancy or underuse. Additionally, robust training programs and protocols are essential to ensure consistent image acquisition and result interpretation—key factors in maximizing diagnostic yield.

Stakeholder Perspectives and Ethical Considerations

Patient and clinician perspectives play a pivotal role in AI adoption. Across the reviewed studies, patients generally accepted AI-based screening, particularly appreciating the convenience and speed of automated feedback. This finding reflects broader trends in digital health, where patients increasingly favor rapid, accessible services, particularly in chronic disease monitoring (Abernethy et al., 2022). Moreover, real-time results delivered at the point of care may boost adherence to follow-up recommendations, a significant issue in DR care continuity.

Among providers, particularly ophthalmologists, the responses were more nuanced. While many clinicians viewed AI as a supportive tool to extend reach and reduce workload, others expressed apprehension about potential over-reliance on algorithms. Concerns about transparency, explainability, and diagnostic accountability persist in the clinical community. As highlighted by Skorburg & Friesen, (2021), “black-box” AI systems may fail to earn clinician trust unless accompanied by explainable AI mechanisms that elucidate how decisions are made.

Ethical considerations such as informed consent, data privacy, and algorithmic bias must also be

addressed before widespread AI deployment. For instance, a study by Agarwal et al., (2023) warned that AI models trained predominantly on images from high-income populations may perform poorly in ethnically diverse or underserved communities. Therefore, efforts must be made to train and validate AI systems using diverse datasets and to establish governance frameworks ensuring equitable deployment.

Health Economic Impact and Policy Implications

The potential cost savings associated with AI screening featured prominently across multiple studies in this review. Automated grading reduces reliance on ophthalmologists and may significantly lower the per-patient cost of DR screening. A health economic analysis by Sobhi et al., (2025) projected a 30–40% reduction in screening costs using AI-based models versus traditional manual grading. These findings are corroborated by Fischer-Abaigar et al., (2024), who reported favorable cost-utility ratios for AI screening programs, particularly when deployed at scale in low- and middle-income countries.

Despite these promising projections, the initial capital investments required for AI implementation—hardware acquisition, staff training, and system maintenance—may pose barriers for some institutions. Moreover, reimbursement frameworks and regulatory clarity around liability and oversight remain underdeveloped in many jurisdictions. Policymakers must work proactively to establish reimbursement codes for AI services and legal protections for clinicians using AI as part of standard care.

In parallel, AI screening programs should be linked with national registries and quality assurance systems to monitor ongoing performance. Continuous auditing and feedback loops are essential for identifying performance drift and ensuring diagnostic fidelity over time (Bova et al., 2017). Furthermore, developing AI literacy among healthcare providers and fostering multidisciplinary collaboration will be critical for navigating the evolving landscape of AI-driven care.

Conclusion

This review underscores that AI-based screening systems for diabetic retinopathy demonstrate strong diagnostic accuracy and are generally well-received by both patients and clinicians. When implemented appropriately, these tools have the potential to improve screening coverage, reduce health system costs, and facilitate earlier detection of vision-threatening conditions. However, realizing these benefits requires attention to implementation infrastructure, regulatory frameworks, training, and equity considerations. Future research should focus on long-term health outcomes, algorithm transparency, and cost-effectiveness across diverse healthcare environments. As AI continues to mature, it will be vital to ensure that its integration enhances, rather than fragments, the continuum of DR care.

Declarations

Funding

This research received no external funding.

Conflicts of Interest

The author declares no conflicts of interest relevant to this manuscript.

Ethics Approval and Consent to Participate

Not applicable. This study is a systematic review of previously published literature and does not involve human participants or animal subjects.

Consent for Publication

Not applicable.

Availability of Data and Materials

All data generated or analyzed during this study are included in this published article and its supplementary materials. Further details are available from the corresponding author upon reasonable request.

Author Contributions

Saif Khuzaim Al-Dossary: Conceptualization, Methodology, Literature Search, Screening and Selection, Data Extraction, Risk of Bias Assessment, Synthesis of Results, Writing – Original Draft, and Final Approval of the Manuscript.

Acknowledgements

The author would like to thank the library and research support services at King Faisal University for providing access to databases and journal subscriptions used in this review.

References :

1. Abernethy, A., Adams, L., Barrett, M., Bechtel, C., Brennan, P., Butte, A., Faulkner, J., Fontaine, E., Friedhoff, S., Halamka, J.,

Howell, M., Johnson, K., Long, P., McGraw, D., Miller, R., Lee, P., Perlin, J., Rucker, D., Sandy, L., … Valdes, K. (2022). The Promise of Digital Health: Then, Now, and the Future. NAM Perspectives, 6(22). https://doi.org/10.31478/202206e

2. Abràmoff, M. D., Lou, Y., Erginay, A., Clarida, W., Amelon, R., Folk, J. C., & Niemeijer, M. (2016). Improved Automated Detection of Diabetic Retinopathy on a Publicly Available Dataset Through Integration of Deep Learning. Investigative Opthalmology & Visual Science, 57(13), 5200. https://doi.org/10.1167/iovs.16-19964
3. Abramoff, M. D., Niemeijer, M., & Russell, S. R. (2010). Automated detection of diabetic retinopathy: barriers to translation into clinical practice. Expert Review of Medical Devices, 7(2), 287–296. https://doi.org/10.1586/erd.09.76
4. Adler-Milstein, J., Aggarwal, N., Ahmed, M., Castner, J., Evans, B. J., Gonzalez, A. A., James, C. A., Lin, S., Mandl, K. D., Matheny, M. E., Sendak, M. P., Shachar, C., & Williams, A. (2022). Meeting the Moment: Addressing Barriers and Facilitating Clinical Adoption of Artificial Intelligence in Medical Diagnosis. NAM Perspectives, 22(9). https://doi.org/10.31478/202209c
5. Agarwal, R., Bjarnadottir, M., Rhue, L., Dugas, M., Crowley, K., Clark, J., & Gao, G. (2023). Addressing algorithmic bias and the perpetuation of health inequities: An AI bias aware framework. Health Policy and Technology, 12(1), 100702. https://doi.org/10.1016/j.hlpt.2022.100702
6. Ali, S., Abuhmed, T., El-Sappagh, S., Muhammad, K., Alonso-Moral, J. M., Confalonieri, R., Guidotti, R., Del Ser, J., Díaz-Rodríguez, N., & Herrera, F. (2023). Explainable Artificial Intelligence (XAI): What we know and what is left to attain Trustworthy Artificial Intelligence. Information Fusion, 99, 101805. https://doi.org/10.1016/j.inffus.2023.101805
7. Alwazae, M., Al Adel, F., Alhumud, A., Almutairi, A., Alhumidan, A., & Elmorshedy, H. (2019). Barriers for Adherence to Diabetic Retinopathy Screening among Saudi Adults. Cureus. https://doi.org/10.7759/cureus.6454
8. Arenas-Cavalli, J. T., Abarca, I., Rojas-Contreras, M., Bernuy, F., & Donoso, R. (2022). Clinical validation of an artificial intelligence-based diabetic retinopathy screening tool for a national health system. Eye, 36(1), 78–85. https://doi.org/10.1038/s41433-020-01366-0
9. Asan, O., Bayrak, A. E., & Choudhury, A. (2020). Artificial Intelligence and Human Trust in Healthcare: Focus on Clinicians. Journal of Medical Internet Research, 22(6), e15154. https://doi.org/10.2196/15154
10. Bajwa, J., Munir, U., Nori, A., & Williams, B. (2021). Artificial intelligence in healthcare: transforming the practice of medicine. Future Healthcare Journal, 8(2), e188–e194. https://doi.org/10.7861/fhj.2021-0095
11. Bellemo, V., Lim, Z. W., Lim, G., Nguyen, Q. D., Xie, Y., Yip, M. Y. T., Hamzah, H., Ho, J., Lee, X. Q., Hsu, W., Lee, M. L., Musonda, L., Chandran, M., Chipalo-Mutati, G., Muma, M., Tan, G. S. W., Sivaprasad, S., Menon, G., Wong, T. Y., & Ting, D. S. W. (2019). Artificial intelligence using deep learning to screen for referable and vision-threatening diabetic retinopathy in Africa: a clinical validation study. The Lancet Digital Health, 1(1), e35–e44. https://doi.org/10.1016/S2589-7500(19)30004-4
12. Bova, C., Jaffarian, C., Crawford, S., Quintos, J. B., Lee, M., & Sullivan-Bolyai, S. (2017). Intervention Fidelity. Nursing Research, 66(1), 54–59. https://doi.org/10.1097/NNR.0000000000000194
13. Chawla, R., Karkhanis, P., Shah, M., Das, A., Sharma, R., Almaula, D., Venkatesh, P., Singh, H. V., Kumar, M., Samanta, R., Kumarl, V., Shah, A., Vadera, B., Jain, N., Sen, A., Shreedhar, S., Garg, V., Dhaval, S., Ganesh, K., … Tandon, R. (2025). Artificial intelligence for advancing eye care in resource-poor settings: Assessing the predictive accuracy of an AI-model for diabetic retinopathy screening in India. Global Epidemiology, 9, 100209. https://doi.org/10.1016/j.gloepi.2025.100209
14. Choudhury, A., & Asan, O. (2020). Role of Artificial Intelligence in Patient Safety Outcomes: Systematic Literature Review.

JMIR Medical Informatics, 8(7), e18599. https://doi.org/10.2196/18599

15. Cleland, C. R., Rwiza, J., Evans, J. R., Gordon, I., MacLeod, D., Burton, M. J., & Bascaran, C. (2023). Artificial intelligence for diabetic retinopathy in low-income and middle-income countries: a scoping review. BMJ Open Diabetes Research & Care, 11(4), e003424. https://doi.org/10.1136/bmjdrc-2023-003424
16. Dow, E. R., Khan, N. C., Chen, K. M., Mishra, K., Perera, C., Narala, R., Basina, M., Dang, J., Kim, M., Levine, M., Phadke, A., Tan, M., Weng, K., Do, D. V., Moshfeghi, D. M., Mahajan, V. B., Mruthyunjaya, P., Leng, T., & Myung, D. (2023). Artificial Intelligence-Human Hybrid Workflow Enhances Teleophthalmology for the Detection of Diabetic Retinopathy. Ophthalmology Science, 3(4), 100330. https://doi.org/10.1016/j.xops.2023.100330
17. Federspiel, F., Mitchell, R., Asokan, A., Umana, C., & McCoy, D. (2023). Threats by artificial intelligence to human health and human existence. BMJ Global Health, 8(5), e010435. https://doi.org/10.1136/bmjgh-2022-010435
18. Fischer-Abaigar, U., Kern, C., Barda, N., & Kreuter, F. (2024). Bridging the gap: Towards an expanded toolkit for AI-driven decision-making in the public sector. Government Information Quarterly, 41(4), 101976. https://doi.org/10.1016/j.giq.2024.101976
19. Freeman, K., Geppert, J., Stinton, C., Todkill, D., Johnson, S., Clarke, A., & Taylor-Phillips, S. (2021). Use of artificial intelligence for image analysis in breast cancer screening programmes: systematic review of test accuracy. BMJ, n1872. https://doi.org/10.1136/bmj.n1872
20. González-Gonzalo, C., Thee, E. F., Klaver, C. C. W., Lee, A. Y., Schlingemann, R. O., Tufail, A., Verbraak, F., & Sánchez, C. I. (2022). Trustworthy AI: Closing the gap between development and integration of AI systems in ophthalmic practice. Progress in Retinal and Eye Research, 90, 101034. https://doi.org/10.1016/j.preteyeres.2021.101034
21. Grauslund, J. (2022). Diabetic retinopathy screening in the emerging era of artificial intelligence. Diabetologia, 65(9), 1415–1423. https://doi.org/10.1007/s00125-022-05727-0
22. Habib, M. M., Welikala, R. A., Hoppe, A., Owen, C. G., Rudnicka, A. R., & Barman, S. A. (2017). Detection of microaneurysms in retinal images using an ensemble classifier. Informatics in Medicine Unlocked, 9, 44–57. https://doi.org/10.1016/j.imu.2017.05.006
23. Heydon, P., Egan, C., Bolter, L., Chambers, R., Anderson, J., Aldington, S., Stratton, I. M., Scanlon, P. H., Webster, L., Mann, S., Chemin, A. du, Owen, C. G., Tufail, A., & Rudnicka, A. R. (2021). Prospective evaluation of an artificial intelligence-enabled algorithm for automated diabetic retinopathy screening of 30 000 patients. British Journal of Ophthalmology, 105(5), 723 LP-728. https://doi.org/10.1136/bjophthalmol-2020-316594
24. Ipp, E., Liljenquist, D., Bode, B., Shah, V. N., Silverstein, S., Regillo, C. D., Lim, J. I., Sadda, S., Domalpally, A., Gray, G., Bhaskaranand, M., Ramachandra, C., Solanki, K., DuBiner, H. B., Genter, P., Graham, J., Johnson, A., Levy-Clarke, G., Pesavento, R. D., … Gornbein, J. (2021). Pivotal Evaluation of an Artificial Intelligence System for Autonomous Detection of Referrable and Vision-Threatening Diabetic Retinopathy. JAMA Network Open, 4(11), e2134254. https://doi.org/10.1001/jamanetworkopen.2021.34254
25. Keel, S., Lee, P. Y., Scheetz, J., Li, Z., Kotowicz, M. A., MacIsaac, R. J., & He, M. (2018). Feasibility and patient acceptability of a novel artificial intelligence-based screening model for diabetic retinopathy at endocrinology outpatient services: a pilot study. Scientific Reports, 8(1), 4330. https://doi.org/10.1038/s41598-018-22612-2
26. Khanna, N. N., Maindarkar, M. A., Viswanathan, V., Fernandes, J. F. E., Paul, S., Bhagawati, M., Ahluwalia, P., Ruzsa, Z., Sharma, A., Kolluri, R., Singh, I. M., Laird, J. R., Fatemi, M., Alizad, A., Saba, L., Agarwal, V., Sharma, A., Teji, J. S., Al-Maini, M., … Suri, J. S. (2022). Economics of Artificial Intelligence in Healthcare: Diagnosis vs. Treatment. Healthcare, 10(12), 2493.

https://doi.org/10.3390/healthcare10122493
27. Kim, T. N., Aaberg, M. T., Li, P., Davila, J. R., Bhaskaranand, M., Bhat, S., Ramachandra, C., Solanki, K., Myers, F., Reber, C., Jalalizadeh, R., Margolis, T. P., Fletcher, D., & Paulus, Y. M. (2021). Comparison of automated and expert human grading of diabetic retinopathy using smartphone-based retinal photography. Eye, 35(1), 334–342. https://doi.org/10.1038/s41433-020-0849-5
28. Kumar, S., Kumar, G., Velu, S., Pardhan, S., Sivaprasad, S., Ruamviboonsuk, P., & Raman, R. (2020). Patient and provider perspectives on barriers to screening for diabetic retinopathy: an exploratory study from southern India. BMJ Open, 10(12), e037277. https://doi.org/10.1136/bmjopen-2020-037277
29. Lee, A. Y., Yanagihara, R. T., Lee, C. S., Blazes, M., Jung, H. C., Chee, Y. E., Gencarella, M. D., Gee, H., Maa, A. Y., Cockerham, G. C., Lynch, M., & Boyko, E. J. (2021). Multicenter, Head-to-Head, Real-World Validation Study of Seven Automated Artificial Intelligence Diabetic Retinopathy Screening Systems. Diabetes Care, 44(5), 1168–1175. https://doi.org/10.2337/dc20-1877
30. Lee, R., Wong, T. Y., & Sabanayagam, C. (2015). Epidemiology of diabetic retinopathy, diabetic macular edema and related vision loss. Eye and Vision, 2(1), 17. https://doi.org/10.1186/s40662-015-0026-2
31. Lim, J. I., Regillo, C. D., Sadda, S. R., Ipp, E., Bhaskaranand, M., Ramachandra, C., Solanki, K., Dubiner, H., Levy-Clarke, G., Pesavento, R., Sherman, M. D., Silverstein, S., Kim, B., Walman, G. B., Blodi, B. A., Domalpally, A., Reed, S., Reimers, J., Lang, K., … Gornbein, J. (2023). Artificial Intelligence Detection of Diabetic Retinopathy. Ophthalmology Science, 3(1), 100228. https://doi.org/10.1016/j.xops.2022.100228
32. Liu, J., Gibson, E., Ramchal, S., Shankar, V., Piggott, K., Sychev, Y., Li, A. S., Rao, P. K., Margolis, T. P., Fondahn, E., Bhaskaranand, M., Solanki, K., & Rajagopal, R. (2021). Diabetic Retinopathy Screening with Automated Retinal Image Analysis in a Primary Care Setting Improves Adherence to Ophthalmic Care. Ophthalmology Retina, 5(1), 71–77. https://doi.org/10.1016/j.oret.2020.06.016
33. Mathenge, W., Whitestone, N., Nkurikiye, J., Patnaik, J. L., Piyasena, P., Uwaliraye, P., Lanouette, G., Kahook, M. Y., Cherwek, D. H., Congdon, N., & Jaccard, N. (2022). Impact of Artificial Intelligence Assessment of Diabetic Retinopathy on Referral Service Uptake in a Low-Resource Setting. Ophthalmology Science, 2(4), 100168. https://doi.org/10.1016/j.xops.2022.100168
34. Naik, N., Hameed, B. M. Z., Shetty, D. K., Swain, D., Shah, M., Paul, R., Aggarwal, K., Ibrahim, S., Patil, V., Smriti, K., Shetty, S., Rai, B. P., Chlosta, P., & Somani, B. K. (2022). Legal and Ethical Consideration in Artificial Intelligence in Healthcare: Who Takes Responsibility? Frontiers in Surgery, 9. https://doi.org/10.3389/fsurg.2022.862322
35. Noriega, A., Meizner, D., Camacho, D., Enciso, J., Quiroz-Mercado, H., Morales-Canton, V., Almaatouq, A., & Pentland, A. (2021). Screening Diabetic Retinopathy Using an Automated Retinal Image Analysis System in Independent and Assistive Use Cases in Mexico: Randomized Controlled Trial. JMIR Formative Research, 5(8), e25290. https://doi.org/10.2196/25290
36. Novelli, C., Taddeo, M., & Floridi, L. (2023). Accountability in artificial intelligence: what it is and how it works. AI and Society, 2019. https://doi.org/10.1007/s00146-023-01635-y
37. Raman, R., Dasgupta, D., Ramasamy, K., George, R., Mohan, V., & Ting, D. (2021). Using artificial intelligence for diabetic retinopathy screening: Policy implications. Indian Journal of Ophthalmology, 69(11), 2993. https://doi.org/10.4103/ijo.IJO_1420_21
38. Rawlings Chidi, & Ugochukwu Odimba. (2024). AI APPLICATIONS IN SCREENING AND DIAGNOSIS OF DIABETIC RETINOPATHY IN RURAL SETTINGS. International Medical Science Research Journal, 4(3), 266–275. https://doi.org/10.51594/imsrj.v4i3.918
39. Rêgo, S., Dutra-Medeiros, M., Soares, F., & Monteiro-Soares, M. (2021). Screening for Diabetic Retinopathy Using an Automated

Diagnostic System Based on Deep Learning: Diagnostic Accuracy Assessment. Ophthalmologica, 244(3), 250–257. https://doi.org/10.1159/000512638

40. Ruamviboonsuk, P., Krause, J., Chotcomwongse, P., Sayres, R., Raman, R., Widner, K., Campana, B. J. L., Phene, S., Hemarat, K., Tadarati, M., Silpa-Archa, S., Limwattanayingyong, J., Rao, C., Kuruvilla, O., Jung, J., Tan, J., Orprayoon, S., Kangwanwongpaisan, C., Sukumalpaiboon, R., … Webster, D. R. (2019). Deep learning versus human graders for classifying diabetic retinopathy severity in a nationwide screening program. Npj Digital Medicine, 2(1), 25. https://doi.org/10.1038/s41746-019-0099-8
41. Ruamviboonsuk, P., Tiwari, R., Sayres, R., Nganthavee, V., Hemarat, K., Kongprayoon, A., Raman, R., Levinstein, B., Liu, Y., Schaekermann, M., Lee, R., Virmani, S., Widner, K., Chambers, J., Hersch, F., Peng, L., & Webster, D. R. (2022). Real-time diabetic retinopathy screening by deep learning in a multisite national screening programme: a prospective interventional cohort study. The Lancet Digital Health, 4(4), e235–e244. https://doi.org/10.1016/S2589-7500(22)00017-6
42. Scanzera, A. C., Shorter, E., Kinnaird, C., Valikodath, N., Al-Khaled, T., Cole, E., Kravets, S., Hallak, J. A., McMahon, T., & Chan, R. V. P. (2022). Optometrist's perspectives of Artificial Intelligence in eye care. Journal of Optometry, 15, S91–S97. https://doi.org/10.1016/j.optom.2022.06.006
43. Scheetz, J., Koca, D., McGuinness, M., Holloway, E., Tan, Z., Zhu, Z., O'Day, R., Sandhu, S., MacIsaac, R. J., Gilfillan, C., Turner, A., Keel, S., & He, M. (2021). Real-world artificial intelligence-based opportunistic screening for diabetic retinopathy in endocrinology and indigenous healthcare settings in Australia. Scientific Reports, 11(1), 15808. https://doi.org/10.1038/s41598-021-94178-5
44. Shah, A., Clarida, W., Amelon, R., Hernaez-Ortega, M. C., Navea, A., Morales-Olivas, J., Dolz-Marco, R., Verbraak, F., Jorda, P. P., van der Heijden, A. A., & Peris Martinez, C. (2021). Validation of Automated Screening for Referable Diabetic Retinopathy With an Autonomous Diagnostic Artificial Intelligence System in a Spanish Population. Journal of Diabetes Science and Technology, 15(3), 655–663. https://doi.org/10.1177/1932296820906212
45. Siala, H., & Wang, Y. (2022). SHIFTing artificial intelligence to be responsible in healthcare: A systematic review. Social Science & Medicine, 296, 114782. https://doi.org/10.1016/j.socscimed.2022.114782
46. Skorburg, J. A., & Friesen, P. (2021). Mind the Gaps: Ethical and Epistemic Issues in the Digital Mental Health Response to Covid-19. Hastings Center Report, 51(6), 23–26. https://doi.org/10.1002/hast.1292
47. Sobhi, N., Sadeghi-Bazargani, Y., Mirzaei, M., Abdollahi, M., Jafarizadeh, A., Pedrammehr, S., Alizadehsani, R., Tan, R.-S., Islam, S. M. S., & Acharya, U. R. (2025). Artificial intelligence for early detection of diabetes mellitus complications via retinal imaging. Journal of Diabetes & Metabolic Disorders, 24(1), 104. https://doi.org/10.1007/s40200-025-01596-7
48. Tabish, S. A. (2007). Is Diabetes Becoming the Biggest Epidemic of the Twenty-first Century? International Journal of Health Sciences, 1(2), V–VIII. http://www.ncbi.nlm.nih.gov/pubmed/21475425
49. Tahir, H. N., Ullah, N., Tahir, M., Domnic, I. S., Prabhakar, R., Meerasa, S. S., AbdElneam, A. I., Tahir, S., & Ali, Y. (2025). Artificial intelligence versus manual screening for the detection of diabetic retinopathy: a comparative systematic review and meta-analysis. Frontiers in Medicine, 12. https://doi.org/10.3389/fmed.2025.1519768
50. Teoh, C. S., Wong, K. H., Xiao, D., Wong, H. C., Zhao, P., Chan, H. W., Yuen, Y. Sen, Naing, T., Yogesan, K., & Koh, V. T. C. (2023). Variability in Grading Diabetic Retinopathy Using Retinal Photography and Its Comparison with an Automated Deep Learning Diabetic Retinopathy Screening Software. Healthcare, 11(12), 1697. https://doi.org/10.3390/healthcare11121697
51. Tufail, A., Rudisill, C., Egan, C., Kapetanakis,

V. V., Salas-Vega, S., Owen, C. G., Lee, A., Louw, V., Anderson, J., Liew, G., Bolter, L., Srinivas, S., Nittala, M., Sadda, S., Taylor, P., & Rudnicka, A. R. (2017). Automated Diabetic Retinopathy Image Assessment Software. Ophthalmology, 124(3), 343–351. https://doi.org/10.1016/j.ophtha.2016.11.014

52. Verbraak, F. D., Abramoff, M. D., Bausch, G. C. F., Klaver, C., Nijpels, G., Schlingemann, R. O., & van der Heijden, A. A. (2019). Diagnostic Accuracy of a Device for the Automated Detection of Diabetic Retinopathy in a Primary Care Setting. Diabetes Care, 42(4), 651–656. https://doi.org/10.2337/dc18-0148
53. Zhu, A., Tailor, P., Verma, R., Zhang, I., Schott, B., Ye, C., Szirth, B., Habiel, M., & Khouri, A. S. (2023). Implementation of deep learning artificial intelligence in vision-threatening disease screenings for an underserved community during COVID-19. Journal of Telemedicine and Telecare, 1357633X2311588. https://doi.org/10.1177/1357633X231158832