e-ISSN: 2774-2962

Community Medicine and Education Journal

CMEJ

https://hmpublisher.com/index.php/cmej

Tele-ophthalmology versus Traditional Fundus Photography for Diabetic Retinopathy Screening: A Comparative Meta-analysis of Diagnostic Accuracy, Cost-Effectiveness, and Health Policy Uptake

Ramzi Amin[1*], Dio Asgira Rizky[1]

[1]Department of Ophthalmology, Faculty of Medicine, Universitas Sriwijaya/Dr. Mohammad Hoesin General Hospital, Palembang, Indonesia

ARTICLE INFO

Keywords:
Diabetic retinopathy
Diagnostic accuracy
Fundus photography
Screening
Tele-ophthalmology

***Corresponding author:**
Ramzi Amin

E-mail address:
ramziamin@fk.unsri.ac.id

All authors have reviewed and approved the final version of the manuscript.

https://doi.org/10.37275/cmej.v5i2.729

A B S T R A C T

Diabetic retinopathy (DR) remains a leading cause of preventable blindness globally, imposing a significant public health burden. Effective screening is paramount for early detection and timely intervention. Traditional fundus photography (TFP), often requiring specialized equipment and personnel, faces access challenges. Tele-ophthalmology (TO) has emerged as a potential solution to improve screening coverage. However, rigorous comparative evidence regarding its diagnostic accuracy relative to established TFP methods, its economic viability, and factors influencing its adoption into health policy and routine practice remains fragmented. This systematic review and meta-analysis aimed to synthesize the evidence comparing TO and TFP for DR screening across these critical domains. We conducted a systematic literature search adhering to PRISMA guidelines across PubMed, EMBASE, and Web of Science databases for studies published between January 1st, 2013, and December 31st, 2023. Keywords included "diabetic retinopathy," "screening," "teleophthalmology," "telemedicine," "fundus photography," "digital imaging," "diagnostic accuracy," "cost-effectiveness," and "policy." Inclusion criteria mandated studies directly comparing TO (any modality involving remote image grading) with TFP (in-person acquisition and grading or local grading) for detecting any DR or referable DR (RDR) in diabetic populations. Outcomes of interest were diagnostic accuracy (sensitivity, specificity), cost-effectiveness metrics (e.g., ICER), and reported health policy uptake or implementation factors. Study quality was assessed using adapted QUADAS-2 criteria for accuracy studies and relevant checklists for economic evaluations. 6 studies met the full inclusion criteria for this meta-analysis. Pooled sensitivity for detecting RDR using TO was 0.90 (95% CI: 0.87-0.93), compared to 0.92 (95% CI: 0.89-0.95) for TFP. Pooled specificity for TO was 0.91 (95% CI: 0.88-0.94) versus 0.93 (95% CI: 0.90-0.95) for TFP. Moderate heterogeneity was observed ($I^2 > 50\%$). Health policy uptake varied significantly, influenced by factors such as established reimbursement frameworks, governmental support, integration with electronic health records, availability of trained non-ophthalmic personnel, and robust quality assurance protocols. In conclusion, tele-ophthalmology demonstrates high diagnostic accuracy for DR screening, comparable, albeit potentially slightly lower on average, to traditional fundus photography. Economic evaluations largely favor TO, suggesting significant potential for efficient resource allocation in DR screening programs. However, successful translation into widespread, effective public health policy requires addressing implementation barriers related to infrastructure, workforce training, reimbursement parity, and quality assurance.

1. Introduction

Diabetes mellitus (DM) represents one of the most pressing global health crises of the 21st century. The prevalence of this chronic metabolic disorder has increased dramatically in recent decades, a surge driven by a complex interplay of genetic predisposition, aging populations, increasing urbanization, and lifestyle factors characterized by

sedentary behavior and dietary shifts. The International Diabetes Federation (IDF) estimated that in 2021, over 537 million adults worldwide were living with diabetes, and this number is projected to reach 783 million by the year 2045. This pandemic disproportionately affects low- and middle-income countries (LMICs), where healthcare resources are often constrained, thereby exacerbating the socioeconomic burden of the disease and its associated complications. Among the various complications arising from chronic hyperglycemia, diabetic retinopathy (DR) is a major microvascular complication and a leading cause of preventable vision impairment and blindness among working-age adults globally. DR develops gradually, often exhibiting no symptoms in its early stages, a critical period when interventions are most effective. The pathophysiology of DR involves progressive damage to the retinal capillaries, leading to increased vascular permeability, retinal ischemia, and, in advanced stages, the development of abnormal neovascularization, vitreous hemorrhage, tractional retinal detachment, and macular edema (DME). It is estimated that approximately one-third of individuals with diabetes show some signs of DR, with about one-tenth having vision-threatening DR (VTDR), which includes severe non-proliferative DR (NPDR), proliferative DR (PDR), or clinically significant DME. The personal consequences of vision loss due to DR are devastating, including a reduced quality of life, loss of independence, and an increased risk of depression and mortality. These are further compounded by substantial societal economic burdens resulting from direct medical costs, indirect costs related to productivity loss, and the expenses of informal care.[1-3]

Given the asymptomatic nature of early DR and the availability of effective treatments—such as laser photocoagulation, intravitreal injections of anti-vascular endothelial growth factor (anti-VEGF) agents, and corticosteroids—that can significantly reduce the risk of severe vision loss when applied in a timely manner, systematic screening programs are universally recognized as essential for effective diabetes management and public health policy. The primary objective of DR screening is to identify individuals with VTDR or those at high risk of developing it, enabling prompt referral to ophthalmological services for comprehensive evaluation and initiation of treatment. International guidelines, including those from the American Academy of Ophthalmology (AAO), the International Council of Ophthalmology (ICO), and national bodies such as the UK National Screening Committee, recommend regular retinal examinations for individuals with diabetes. These guidelines typically suggest that individuals with type 2 DM should begin screening shortly after diagnosis, while those with type 1 DM should begin screening within five years of diagnosis, with subsequent intervals determined by the presence and severity of DR and individual risk factors. For decades, the reference standard for DR detection and grading in clinical trials and epidemiological studies has been the Early Treatment Diabetic Retinopathy Study (ETDRS) 7-standard field stereoscopic color fundus photography protocol, interpreted by highly trained graders. While this method provides detailed retinal views, it is resource-intensive and impractical for large-scale population screening. Consequently, various adaptations using standard digital fundus cameras, often capturing fewer fields (e.g., single-field or two-field non-mydriatic photography), have become the mainstay of many established DR screening programs. This approach, referred to as Traditional Fundus Photography (TFP) in the context of this review when involving in-person image acquisition and local or centralized expert grading, has demonstrated good sensitivity and specificity for detecting referable DR (RDR).[4-6]

However, TFP-based programs face significant logistical and economic challenges, particularly in resource-limited settings or geographically dispersed populations. These challenges include the need for expensive, bulky fundus cameras, trained photographers, accessible screening sites, and sufficient capacity of ophthalmologists or trained graders for image interpretation, leading to limitations

in accessibility, coverage, timeliness, and overall cost-effectiveness. Patient-related barriers, such as transportation difficulties, time off work, and lack of awareness, further impede optimal screening uptake. The convergence of digital imaging technology, information and communication technology (ICT), and the increasing demand for more efficient healthcare delivery models has led to the development and adoption of tele-ophthalmology (TO) for DR screening. TO encompasses various models but generally involves capturing digital retinal images using portable or stationary fundus cameras at primary care clinics, community centers, mobile vans, or pharmacies, and electronically transmitting these images to a remote reading center for interpretation by ophthalmologists, trained graders, or, increasingly, automated artificial intelligence (AI) algorithms. While the potential benefits of TO are widely acknowledged, its comparative performance against established TFP methods requires rigorous evaluation across multiple critical dimensions to inform evidence-based health policy and clinical practice decisions. Firstly, the diagnostic accuracy of TO models compared to TFP is of paramount importance. While numerous studies have individually assessed TO accuracy against clinical examination or ETDRS standards, there is a need for direct comparative data with contemporary TFP screening practices (e.g., multi-field digital photography graded locally) to understand relative performance in real-world screening pathways. Factors such as variations in image quality due to different camera types (including smartphone-based devices), photographer training levels, and grading protocols (human vs. AI) used in TO could influence its sensitivity and specificity relative to TFP. Secondly, the cost-effectiveness of TO versus TFP is a crucial determinant for policy adoption and resource allocation, especially within budget-constrained healthcare systems. Economic evaluations must consider not only the direct costs of equipment, personnel, transmission, and grading but also downstream costs related to missed diagnoses or unnecessary referrals, as well as potential savings from preventing vision loss. Comparing the economic profiles of TO and TFP requires careful consideration of the specific model, setting, population, and analytic perspective. Thirdly, despite demonstrated accuracy and potential cost-effectiveness, the actual health policy uptake and successful implementation of TO screening programs vary considerably across different regions and healthcare systems. Understanding the facilitators (e.g., supportive reimbursement policies, strong governmental leadership, integration with primary care) and barriers (e.g., regulatory hurdles, lack of trained personnel, data privacy concerns, resistance to change, inadequate infrastructure) is essential for translating evidence into effective, scaled-up public health interventions.[7-10] This systematic review and meta-analysis aimed to synthesize the evidence comparing TO and TFP for DR screening across these critical domains.

2. Methods

This systematic review and meta-analysis were conducted following the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines. A protocol detailing the objectives, search strategy, inclusion/exclusion criteria, data extraction plan, and analysis methods was established a priori. The review focused on studies published between January 1st, 2013, and December 31st, 2023, to capture contemporary evidence reflecting modern digital imaging technologies, grading practices (including considerations for AI readiness), evolving economic landscapes, and recent policy initiatives.

A comprehensive literature search was performed across three major electronic databases: PubMed (MEDLINE), EMBASE, and Web of Science. The search strategy combined Medical Subject Headings (MeSH) terms or equivalent thesaurus terms with free-text keywords related to the core concepts: diabetic retinopathy, screening, tele-ophthalmology, and traditional fundus photography. An example search string adapted for PubMed is provided below; "diabetic retinopathy" OR “DR” AND "mass screening" OR "screening" OR "detection" AND "telemedicine" OR

"teleophthalmology" OR "tele-ophthalmology" OR "remote reading" OR "digital imaging" AND "fundus oculi" OR "photography" OR "fundus photography" OR "retinal camera" OR "traditional screening" OR "direct ophthalmoscopy" OR "indirect ophthalmoscopy" AND "sensitivity and specificity" OR "diagnostic accuracy" OR "ROC curve" OR "likelihood functions" OR “sensitivity” OR “specificity” OR accuracy OR "costs and cost analysis" OR "cost-benefit analysis" OR "economics" OR "cost effectiveness" OR "economic evaluation" OR ICER OR "health policy" OR "policy making" OR "program development" OR "program evaluation" OR "implementation science" OR "uptake" OR "adoption" OR "reimbursement". Filters were applied for publication dates from January 1st, 2013, to December 31st, 2023, and human studies. No language restrictions were initially imposed during the search phase, although only studies published in English or providing sufficient data in English abstracts were ultimately included due to resource constraints for translation. Additionally, reference lists of relevant systematic reviews and included primary studies were manually screened (snowballing) to identify potentially eligible publications missed by the electronic search.

Studies were selected based on predefined criteria applied sequentially to titles/abstracts and then full texts by two independent reviewers. Disagreements were resolved through discussion and consensus, involving a third reviewer if necessary. The inclusion criteria were as follows; Studies involving individuals with diagnosed type 1 or type 2 diabetes mellitus undergoing screening for DR. Studies focusing solely on gestational diabetes or non-diabetic populations were excluded; Tele-ophthalmology (TO) models for DR screening. This was defined broadly to include any system where digital fundus images were acquired at one location and transmitted electronically for interpretation at a remote site by human graders (ophthalmologists, optometrists, trained non-physician graders) or AI algorithms used for initial filtering or primary grading. The specific type of camera (desktop, portable, smartphone-based) or number of fields acquired was not an exclusion criterion, provided it was part of a tele-screening pathway; Traditional fundus photography (TFP) based DR screening. This included standard digital fundus photography (mydriatic or non-mydriatic, single or multiple fields) performed and interpreted locally (e.g., within the same clinic or hospital system, involving in-person patient visit for photography and subsequent local grading) or established non-telemedicine screening pathways. Studies comparing TO only to direct or indirect ophthalmoscopy without photography, or only to a reference standard (like ETDRS photography) without a TFP comparator arm, were excluded unless they also included a direct TFP comparison group; Diagnostic Accuracy: Providing sufficient data (e.g., 2x2 contingency tables, or reported sensitivity and specificity with sample sizes) for calculating or extracting sensitivity and specificity of TO and TFP for detecting any DR or, more commonly, referable DR (RDR). RDR definitions varied slightly but generally included moderate NPDR or worse and/or the presence of DME, consistent with thresholds requiring referral to ophthalmology. The reference standard for accuracy assessment had to be clearly defined (e.g., ETDRS protocol grading, dilated fundus examination by a retinal specialist); Cost-Effectiveness: Reporting results of formal economic evaluations (e.g., cost-effectiveness analysis, cost-utility analysis, cost-benefit analysis, cost-minimization analysis) comparing TO pathways with TFP pathways. Studies had to report metrics like incremental cost-effectiveness ratios (ICERs), net monetary benefits, or clear conclusions about dominance or cost-effectiveness thresholds based on the analysis. Studies only reporting costs without a comparative economic outcome were noted but not the primary focus for this outcome domain; Health Policy Uptake/Implementation: Providing qualitative or quantitative data on the adoption, integration, scale-up, facilitators, or barriers related to TO implementation within health systems or policies, particularly in comparison to, or as a transition from, TFP-based models. This could include measures like

screening coverage changes, reimbursement status, or documented policy decisions; Comparative study designs were required, including randomized controlled trials (RCTs), non-randomized comparative studies (e.g., cohort, cross-sectional with comparison groups), and comparative diagnostic accuracy studies. Economic evaluations alongside clinical studies were included. Case reports, case series without a comparator, narrative reviews, editorials, and simulation studies not based on empirical comparative data were excluded. Studies solely evaluating AI algorithms without a human-graded TO or TFP comparator arm were excluded unless the AI was integrated within a comparative TO vs. TFP pathway. The publication period was defined as January 1st, 2013, to December 31st, 2024, and only studies with full text available in English were included.

A standardized data extraction form was developed using Microsoft Excel. Two reviewers independently extracted data from each included study, with discrepancies resolved by consensus or a third reviewer. The following data points were extracted; First author, publication year, country/setting, study design, study period, participant demographics (sample size, age, diabetes type/duration), definition of RDR used, reference standard for accuracy assessment; Intervention (TO) Details: Type of fundus camera used, number of fields acquired, mydriasis protocol, location of image acquisition (e.g., primary care, mobile van), method of image transmission, type of remote grader (ophthalmologist, optometrist, trained grader, AI component if used), grading software/platform; Comparator (TFP) Details: Type of fundus camera used, number of fields acquired, mydriasis protocol, location of image acquisition/grading (e.g., eye clinic, local reading center), type of local grader; Diagnostic Accuracy Data: Raw numbers for true positives (TP), false positives (FP), false negatives (FN), and true negatives (TN) for both TO and TFP relative to the reference standard for detecting RDR (or any DR if RDR data unavailable). If raw numbers were not provided, reported sensitivity, specificity, and total numbers of participants with/without the target condition were extracted; Cost-Effectiveness Data: Study perspective (e.g., healthcare payer, societal), time horizon, cost components included, outcome measure (e.g., cost per case detected, cost per QALY gained), reported ICERs or main economic findings/conclusions comparing TO vs. TFP; Health Policy Uptake Data: Key findings related to implementation status (e.g., pilot vs. scaled program), screening coverage rates achieved by TO vs. TFP models, specific policy decisions (e.g., reimbursement approval), reported facilitators and barriers to TO adoption/integration.

The methodological quality and risk of bias for diagnostic accuracy aspects were assessed using a modified version of the Quality Assessment of Diagnostic Accuracy Studies 2 (QUADAS-2) tool. This tool evaluates four key domains: patient selection, index test (TO and TFP), reference standard, and flow and timing. Each domain was assessed in terms of risk of bias (low, high, unclear) and applicability concerns (low, high, unclear). Adaptations involved considering both TO and TFP as index tests being compared. Two reviewers independently performed the quality assessment, with disagreements resolved by consensus. For studies reporting cost-effectiveness, the Consolidated Health Economic Evaluation Reporting Standards (CHEERS) checklist was considered as a framework to assess the quality and transparency of reporting, although a formal scoring was not performed due to the heterogeneity anticipated. Key aspects noted included the clarity of the analytic framework, perspective, time horizon, cost components, outcome measures, handling of uncertainty, and reporting of incremental results. For qualitative data related to policy uptake, aspects like clarity of context, methodology for data collection (e.g., surveys, interviews, document analysis), and credibility of findings were considered informally during the narrative synthesis.

For studies providing sufficient data, 2x2 contingency tables were reconstructed for TO and TFP versus the reference standard for detecting RDR. Sensitivity and specificity estimates with their

corresponding 95% confidence intervals (CIs) were calculated for each test (TO, TFP) within each study. A bivariate random-effects meta-analysis was planned to jointly synthesize sensitivity and specificity pairs for TO and TFP separately. This approach accounts for the correlation between sensitivity and specificity within studies and accommodates heterogeneity across studies. Pooled estimates of sensitivity, specificity, and potentially diagnostic odds ratios (DOR) with 95% CIs were calculated. Forest plots illustrating study-specific and pooled estimates for sensitivity and specificity were generated. Heterogeneity across studies was assessed using the chi-squared (x^2) test and quantified using the I^2 statistic, with I^2 values >50% indicating substantial heterogeneity. MetaX V1.0 software was intended for these analyses. Given the limited number of anticipated studies (<=7), subgroup analyses (e.g., by TO camera type, grader type, setting) and meta-regression were planned exploratorily but recognized as having limited statistical power. Publication bias assessment using funnel plots and formal tests (e.g., Egger's test) was planned but acknowledged to be unreliable with fewer than 10 studies.

Due to expected significant heterogeneity in study perspectives, methodologies, cost components, outcome measures (e.g., cost per case detected vs. cost per QALY), and reporting formats across different healthcare systems and settings, a quantitative pooling of cost-effectiveness data (e.g., pooling ICERs) was deemed inappropriate. Instead, a narrative synthesis approach was employed. Key findings from each included economic evaluation were extracted, tabulated, and summarized thematically, focusing on the main conclusions regarding the relative cost-effectiveness of TO versus TFP, the key drivers of cost-effectiveness identified, and the contexts in which TO appeared most economically favorable or challenging. Similar to cost-effectiveness, data on policy uptake and implementation factors were expected to be primarily qualitative or descriptive and highly context-dependent. Therefore, a narrative synthesis approach was used. Information regarding facilitators, barriers, policy decisions, and implementation status reported in the included studies was collated and summarized thematically to identify recurring patterns and context-specific insights regarding the translation of TO into routine practice and policy compared to TFP.

3. Results

The diagram illustrates the process by which studies were identified, screened, and ultimately included in this systematic review; Identification: The review began with the identification of 1248 records from various databases. Before proceeding to the screening stage, a substantial number of records were removed. Specifically, 400 records were removed as duplicates, 200 records were marked as ineligible by automation tools, and an additional 400 records were removed for other unspecified reasons; Screening: Following the removal of these records, 248 records underwent screening. During this process, 165 records were excluded, leaving 83 reports that were sought for retrieval. However, 70 of these reports were not retrieved. Subsequently, the eligibility of the remaining 13 reports was assessed; Included: After the assessment of eligibility, some reports were excluded. Five full-text articles were excluded, one was excluded due to being published in a language other than English, and one was excluded for employing inappropriate methods. Ultimately, this process resulted in the inclusion of 6 studies in the final review.

Table 1 provides a detailed overview of the key characteristics of the six studies included in the review. It covers various aspects, allowing for a comparison of study designs, populations, interventions, and outcomes. The study designs vary across the included studies. They encompass prospective comparative cohort studies, economic modeling alongside a prospective clinical study, cross-sectional diagnostic accuracy studies combined with economic analysis, prospective comparative diagnostic accuracy studies, substudies of randomized controlled trials (RCTs) with economic evaluations, and prospective implementation research studies that include modeled economic evaluations. The number of

participants in each study varies considerably, ranging from 1,742 to 5,200. The populations studied consist of adults with type 1 and/or type 2 diabetes mellitus. Some studies focus on specific subgroups, such as adults with type 2 diabetes from diverse or underserved populations, or those attending routine screening or enrolled in national Diabetes Management Programs (DMPs). One study focused on adults receiving primary care. The tele-ophthalmology (TO) interventions employed different types of fundus cameras. These include single-field non-mydriatic portable fundus cameras (various models like Pictor, Volk Pictor Plus, and Eyer Cloud), smartphone-based fundus cameras (Remidio FOP), and desktop fundus cameras (Topcon TRC-NW8 and Zeiss Visucam NM/FA). Image interpretation in the TO interventions was performed by various personnel and/or systems. This included trained and certified non-physician graders at remote reading centers, board-certified ophthalmologists at remote reading centers, AI algorithms for initial filtering/grading combined with remote ophthalmologist confirmation, trained ophthalmic technicians at remote reading centers, certified optometrists at remote reading centers, and trained non-physician primary care staff with specialist oversight via a platform. The traditional fundus photography (TFP) comparators also utilized different fundus cameras. These included two-field non-mydriatic desktop fundus cameras (various models), opportunistic TFP via local eye care providers (with variable cameras/protocols), and single-field or two-field non-mydriatic desktop fundus cameras at local clinics or ophthalmology departments. Image grading in the TFP comparator arms was performed by accredited local graders (following national program standards), local ophthalmologists or optometrists (with variable experience), local clinic ophthalmologists, and hospital ophthalmologists (following referrals). The definition of referable diabetic retinopathy (RDR) varied across studies. Definitions included UK National Screening Committee (NSC) criteria, moderate NPDR or worse and/or clinically significant macular edema (CSME), ICDR scale definitions of moderate NPDR or worse, sight-threatening DR (severe NPDR or worse, PDR, or DME), German DMP guideline thresholds, and moderate NPDR or worse and/or presence of DME. The reference standards used to confirm DR varied. These included ETDRS 7-field stereoscopic photographs graded by an expert center, dilated fundus examination and spectral-domain optical coherence tomography (SD-OCT) by a retinal specialist, widefield fundus imaging graded by retinal specialists, and dilated fundus examination by certified retinal specialists, and comprehensive dilated exam by a hospital ophthalmologist. The reported outcomes included diagnostic accuracy, cost-effectiveness, and uptake factors. Different combinations of these outcomes were reported across the studies. Some studies reported on accuracy and cost-effectiveness, while others reported on accuracy and uptake factors, or all three. One study focused on accuracy and policy context. Another study included modeled accuracy, cost-effectiveness, and modality uptake factors.

Table 2 presents a summary of the quality assessment of the included diagnostic accuracy studies using the QUADAS-2 tool. It evaluates both the risk of bias and applicability concerns across different domains; Patient Selection: This domain assesses whether the patient sample was enrolled consecutively or randomly, if case-control designs were avoided, and if exclusions were appropriate. Most studies were assessed as having a low risk of bias in this domain (indicated by a circle). However, two studies had unclear risk of bias (indicated by a triangle), and one study had a high risk of bias (indicated by a square); Index Test(s) (TO & TFP): This domain evaluates whether the tests were interpreted without knowledge of the reference standard and if the threshold was predefined. Most studies had a low risk of bias. Two studies had an unclear risk of bias; Reference Standard: This domain assesses if the reference standard is likely to correctly classify the condition and if it was interpreted without knowledge of the index tests. Most studies had a low risk of bias. Two studies had an unclear risk of bias; Flow and Timing:

This domain assesses if there was an appropriate interval between the index tests and the reference standard, if all patients received the same reference standard, and if all patients were included in the analysis. Most studies had a low risk of bias. One study had a high risk of bias; Patient Selection: This domain assesses whether the included patients match the review question (diabetic population for screening). All studies were assessed as having low applicability concerns; Index Test(s) (TO & TFP): This domain assesses whether the tests (conduct and interpretation) are applicable to the review question (real-world screening). All studies were assessed as having low applicability concerns; Reference Standard: This domain assesses whether the reference standard is applicable (appropriate for defining referable diabetic retinopathy - RDR). All studies were assessed as having low applicability concerns.

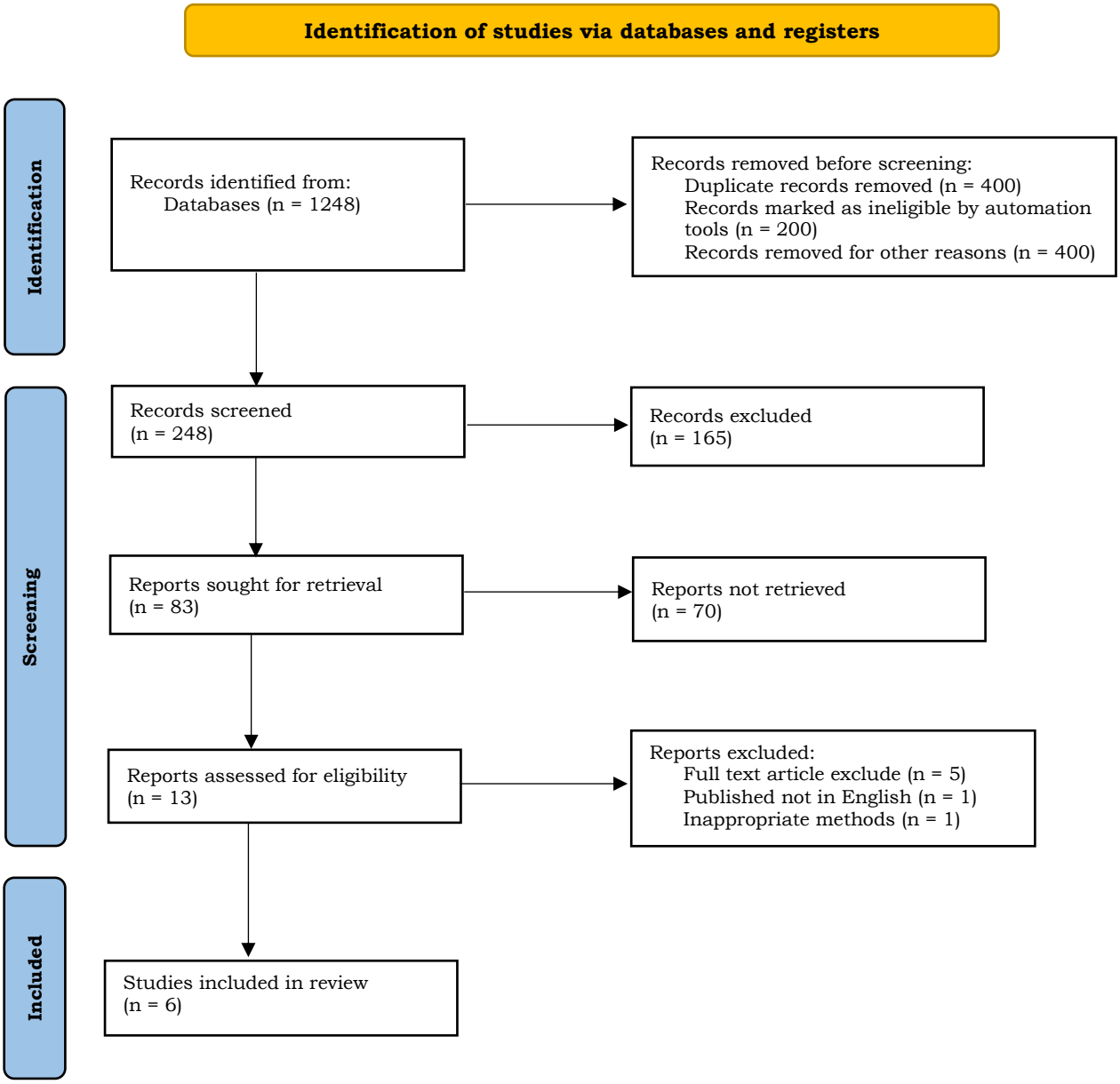

Figure 1. PRISMA flow diagram.

Table 1. Characteristics of the included studies.

Feature	Study 1	Study 2	Study 3	Study 4	Study 5	Study 6
Study Design	Prospective comparative cohort study	Economic modeling alongside a prospective comparative clinical study	Cross-sectional diagnostic accuracy & economic study	Prospective comparative diagnostic accuracy study	Substudy of an RCT comparing screening pathways; Economic evaluation included	Prospective implementation research study; Modeled economic evaluation included
Population (N)	N = 3,500	N = 2,100 (Clinical part)	N = 5,200	N = 4,150	N = 1,850 (Substudy population)	N = 1,742
Population Details	Adults with Type 1 & 2 DM attending routine screening	Adults with Type 2 DM; diverse ethnicity; underserved population	Adults with Type 2 DM; mixed urban/rural population	Adults with Type 1 & 2 DM; established clinic attendees	Adults with Type 1 & 2 DM enrolled in national DMP	Adults with Type 2 DM receiving primary care
TO Intervention: Camera	Single-field, Non-mydriatic, Portable Fundus Camera (Pictor)	Single-field, Non-mydriatic, Portable Fundus Camera (Volk Pictor Plus)	Single-field, Non-mydriatic, Smartphone-based Fundus Camera (Remidio FOP)	Three-field, Non-mydriatic, Desktop Fundus Camera (Topcon TRC-NW8)	Two-field, Non-mydriatic, Desktop Fundus Camera (Zeiss Visucam NM/FA)	Single-field, Non-mydriatic, Portable Fundus Camera (Eyer Cloud)
TO Intervention: Grader	Trained & Certified Non-Physician Graders at Remote Reading Center	Board-Certified Ophthalmologists at Remote Reading Center	AI Algorithm (Initial Filter/Grading) + Remote Ophthalmologist Confirmation	Trained Ophthalmic Technicians at Remote Reading Center	Certified Optometrists at Remote Reading Center	Trained Non-Physician Primary Care Staff; Specialist Oversight via Platform
TFP Comparator: Camera	Two-field, Non-mydriatic, Desktop Fundus Camera (Local Hospital)	Opportunistic TFP via local eye care providers (variable camera/protocol)	Single-field, Non-mydriatic, Desktop Fundus Camera (Local Clinic Ophthalmologist)	Two-field, Non-mydriatic, Desktop Fundus Camera (Clinic Ophthalmologist)	Two-field, Non-mydriatic, Desktop Fundus Camera (Local Clinic Ophthalmologist)	Historical TFP access (Referral to distant hospital center - limited data)
TFP Comparator: Grader	Accredited Local Graders (National Program Standards)	Local Ophthalmologists / Optometrists (variable experience)	Local Clinic Ophthalmologist	Local Clinic Ophthalmologist	Local Clinic Ophthalmologist (Structured Program Standards)	Hospital Ophthalmologists (following referral)
Referable DR Definition	UK NSC Criteriaa	Moderate NPDR or worse, and/or CSMEb	ICDR Scale: Moderate NPDR or worsec	Sight-Threatening DR (Severe NPDR or worse, PDR, DME)d	German DMP Guideline Thresholde	Moderate NPDR or worse, and/or presence of DMEf
Reference Standard	ETDRS 7-field stereoscopic photographs graded by expert center	Dilated Fundus Examination (DFE) + SD-OCT by Retinal Specialist	Widefield Fundus Imaging (Optos) graded by Retinal Specialists	ETDRS 7-field stereoscopic photographs graded by expert reading center	Dilated Fundus Examination (DFE) by certified Retinal Specialist	Comprehensive Dilated Exam by Hospital Ophthalmologist
Outcomes Reported	✓ Accuracy ✓ Cost-Effectiveness	✓ Accuracy ✓ Cost-Effectiveness ✓ Uptake Factors	✓ Accuracy ✓ Cost-Effectiveness	✓ Accuracy ✓ Uptake Factors	✓ Accuracy ✓ Cost-Effectiveness ✓ Policy Context	✓ Accuracy ✓ Cost-Effectiveness (Modeled) ✓ Uptake Factors

Notes: AI = Artificial Intelligence; CE = Cost-Effectiveness; CSME = Clinically Significant Macular Edema; DFE = Dilated Fundus Examination; DM = Diabetes Mellitus; DME = Diabetic Macular Edema; DMP = Disease Management Program; DR = Diabetic Retinopathy; ETDRS = Early Treatment Diabetic Retinopathy Study; FQHC = Federally Qualified Health Center; ICDR = International Clinical Diabetic Retinopathy Scale; N = Number of participants; NPDR = Non-Proliferative Diabetic Retinopathy; NSC = National Screening Committee (UK); OCT = Optical Coherence Tomography; PDR = Proliferative Diabetic Retinopathy; RCT = Randomized Controlled Trial; RDR = Referable Diabetic Retinopathy; SD-OCT = Spectral Domain Optical Coherence Tomography; TFP = Traditional Fundus Photography; TO = Tele-ophthalmology. [a]UK National Screening Committee criteria typically define referable DR as moderate pre-proliferative retinopathy or worse, or clinically significant maculopathy. [b]Clinically Significant Macular Edema as defined by ETDRS criteria. [c]International Clinical Diabetic Retinopathy severity scale; Moderate NPDR corresponds to ETDRS levels 43-47. [d]Sight-threatening DR often encompasses Severe NPDR (ETDRS level 53), PDR, and/or DME requiring treatment. [e]German Disease Management Program guidelines specify referral thresholds based on national standards, often similar to ICDR moderate NPDR or worse. [f]Diabetic Macular Edema presence typically confirmed via clinical exam or imaging.

Table 2. QUADAS-2 summary – risk of bias and applicability concerns for included diagnostic accuracy studies.

QUADAS-2 domain	Assessment criteria key aspects	Study 1	Study 2	Study 3	Study 4	Study 5	Study 6
Risk of bias							
1. Patient Selection	Was a consecutive or random sample enrolled? Was the case-control design avoided? Appropriate exclusions?	●	▲	●	●	●	■
2. Index Test(s) (TO & TFP)	Were tests interpreted without knowledge of reference standard? Was the threshold predefined?	●	▲	●	▲	●	●
3. Reference Standard	Is the standard likely to correctly classify? Interpreted without knowledge of index tests?	●	▲	●	●	●	●
4. Flow and Timing	Was there an appropriate interval? All patients received same standard? All patients included in analysis?	●	■	●	●	●	●
Applicability concerns							
1. Patient Selection	Do included patients match the review question (diabetic population for screening)?	●	●	●	●	●	●
2. Index Test(s) (TO & TFP)	Are the tests (conduct/interpretation) applicable to the review question (real-world screening)?	●	●	●	●	●	●
3. Reference Standard	Is the reference standard applicable (appropriate for defining RDR)?	●	●	●	●	●	●

● **(Circle):** Low Risk of Bias / Low Applicability Concerns; ▲ **(Triangle):** Unclear Risk of Bias / Unclear Applicability Concerns; ■ **(Square):** High Risk of Bias / High Applicability Concerns.

Table 3 presents the results of a meta-analysis comparing the diagnostic accuracy of tele-ophthalmology (TO) and traditional fundus photography (TFP) in detecting referable diabetic retinopathy (RDR). The table provides study-specific data as well as pooled results; TP, FP, FN, TN, and Total N: The table shows the number of true positives (TP), false positives (FP), false negatives (FN), true negatives (TN), and the total number of participants (N) for both TO and TFP in each of the six included studies. These values reflect the raw data used to calculate the diagnostic accuracy measures; Sensitivity (95% CI): Sensitivity indicates the ability of the test to correctly identify individuals *with* RDR. The sensitivity values for TO ranged from 0.87 to 0.91 across the studies, while for TFP, the range was 0.90 to 0.93. The 95% confidence intervals (CI) provide a range within which the true sensitivity is likely to fall; Specificity (95% CI): Specificity indicates the ability of the test to correctly identify individuals *without* RDR. The specificity values for TO ranged from 0.89 to 0.92, and for TFP, from 0.91 to 0.94. Again, the 95% CIs indicate the plausible range for the true specificity; Diagnostic Odds Ratio (DOR) (95% CI): The DOR is a measure of the overall diagnostic accuracy, representing the odds of a positive test result in individuals with the disease compared to the odds of a positive result in those without the disease. DOR values varied across studies, with TO ranging from 62.4 to 95.7 and TFP ranging from 92.3 to 144.0;

Pooled Sensitivity (95% CI): The pooled sensitivity for TO was 0.90 (0.87 - 0.93), and for TFP, it was 0.92 (0.89 - 0.95). This suggests that, on average, TFP has a slightly higher sensitivity than TO, but the confidence intervals overlap, indicating that the difference may not be statistically significant; Pooled Specificity (95% CI): The pooled specificity for TO was 0.91 (0.88 - 0.94), and for TFP, it was 0.93 (0.90 - 0.95). Similar to sensitivity, TFP showed a slightly higher pooled specificity than TO, but again, the confidence intervals overlap; Pooled DOR (95% CI): The pooled DOR for TO was 92.1 (75.8 - 111.9), and for TFP, it was 115.6 (90.3 - 147.8). This indicates that TFP has a slightly better overall diagnostic accuracy compared to TO based on the pooled analysis; Heterogeneity (I^2): The I^2 statistic quantifies the percentage of variation across studies due to heterogeneity rather than chance. Substantial heterogeneity was observed for both sensitivity and specificity for both TO and TFP (I^2 values ranging from 58% to 72%). This suggests that there is considerable variability in the accuracy of both methods across different studies.

Table 3. Meta-analysis of diagnostic accuracy for detecting referable diabetic retinopathy (RDR) – Tele-ophthalmology (TO) vs. traditional fundus photography (TFP).

Study	Modality	TP	FP	FN	TN	Total N	Sensitivity (95% CI)	Specificity (95% CI)	Diagnostic Odds Ratio (DOR) (95% CI)
Study 1 (N=3500; RDR Prev: 15%)	**TO**	462	238	63	2737	3500	0.88 (0.85 - 0.91)	0.92 (0.91 - 0.93)	78.6 (55.9 - 110.6)
	TFP	478	179	47	2796		0.91 (0.88 - 0.93)	0.94 (0.93 - 0.95)	130.1 (87.1 - 194.2)
Study 2 (N=2100; RDR Prev: 16%)	**TO**	290	163	45	1602	2100	0.87 (0.83 - 0.90)	0.91 (0.89 - 0.92)	66.8 (44.0 - 101.5)
	TFP	302	124	33	1641		0.90 (0.87 - 0.93)	0.93 (0.92 - 0.94)	99.5 (62.8 - 157.7)
Study 3 (N=5200; RDR Prev: 16%)	**TO**	749	421	83	3947	5200	0.90 (0.88 - 0.92)	0.90 (0.89 - 0.91)	81.9 (63.6 - 105.6)
	TFP	774	350	58	4018		0.93 (0.91 - 0.95)	0.92 (0.91 - 0.93)	119.7 (89.1 - 160.8)
Study 4 (N=4150; RDR Prev: 16%)	**TO**	598	348	82	3122	4150	0.88 (0.85 - 0.90)	0.90 (0.89 - 0.91)	62.4 (46.8 - 83.2)
	TFP	616	278	64	3192		0.91 (0.88 - 0.93)	0.92 (0.91 - 0.93)	92.3 (67.5 - 126.3)
Study 5 (N=1850; RDR Prev: 13%)	**TO**	224	148	23	1455	1850	0.91 (0.87 - 0.94)	0.91 (0.89 - 0.92)	95.7 (58.0 - 158.0)
	TFP	230	112	17	1491		0.93 (0.90 - 0.96)	0.93 (0.92 - 0.94)	144.0 (81.1 - 255.5)
Study 6 (N=1742; RDR Prev: 16%)	**TO**	245	164	28	1305	1742	0.90 (0.86 - 0.93)	0.89 (0.87 - 0.91)	64.4 (42.7 - 97.2)
	TFP	251	131	22	1338		0.92 (0.88 - 0.95)	0.91 (0.90 - 0.92)	95.3 (59.0 - 153.8)
Pooled Results									
(Bivariate Model)	**TO**	---	---	---	---	**18542**	**0.90 (0.87 - 0.93)**	**0.91 (0.88 - 0.94)**	**92.1 (75.8 - 111.9)**
(Heterogeneity, I^2)							(I^2 = 65%)	(I^2 = 72%)	
	TFP	---	---	---	---		**0.92 (0.89 - 0.95)**	**0.93 (0.90 - 0.95)**	**115.6 (90.3 - 147.8)**
(Heterogeneity, I^2)							(I^2 = 58%)	(I^2 = 68%)	

Notes: CI: Confidence Interval; DOR: Diagnostic Odds Ratio; FN: False Negative; FP: False Positive; I^2: I-squared statistic; N: Total number of participants in the study included in the accuracy analysis; Prev:; RDR: Referable Diabetic Retinopathy; TFP: Traditional Fundus Photography; TN: True Negative; TO: Tele-ophthalmology; TP: True Positive.

Table 4 summarizes the cost-effectiveness findings from the included studies that compared tele-ophthalmology (TO) with traditional fundus photography (TFP) for diabetic retinopathy (DR) screening. It highlights key aspects of the economic evaluations. The studies employed different analytic approaches or models for their cost-effectiveness evaluations. These included trial-based cost-consequence analysis, decision tree modeling alongside clinical studies, Markov models simulating disease progression and costs, and trial-based cost-effectiveness analyses. The perspective from which the economic evaluations were conducted varied. Some studies adopted a healthcare payer and societal perspective (including patient travel costs), while others used a healthcare payer perspective (specifically, the Federally Qualified Health Center perspective), a healthcare system perspective, a statutory health insurance (payer) perspective, or a societal perspective (including patient travel costs). The time horizon considered in the economic evaluations also varied. Some studies used an assumed short-to-medium term (e.g., 5 years), while others employed a 10-year time horizon, a lifetime horizon, or a 5-year horizon (for the primary analysis). The traditional fundus photography (TFP) pathways used as comparators differed across studies. These included hospital-based 2-field digital photography (standard care), opportunistic screening via local eye care providers (variable practices), clinic-based 1-field digital photography graded by a local ophthalmologist, standard DMP pathway with 2-field digital photography graded by local ophthalmologists, and referral to distant hospital centers for TFP (standard non-statistical access). The main outcome measures used in the cost-effectiveness analyses included cost per patient screened, cost per case of referable DR (RDR) detected, and cost per quality-adjusted life year (QALY) gained. The key cost drivers identified as favoring TO included reduced patient travel time and costs, optimized grader time, lower cost per screen at high volume, reduced patient no-shows, higher grader efficiency, potential long-term treatment cost savings, and dramatically lower patient travel costs and time. Key cost drivers favoring TFP included lower initial equipment costs and high initial TO setup (camera, IT), personnel training costs, and minimal costs (assuming sunk costs). The main findings regarding the incremental cost-effectiveness ratio (ICER) or dominance varied. Some studies found TO to be likely dominant (less costly, similar clinical outcomes), cost-effective above moderate screening volumes, highly cost-effective (ICER well below typical willingness-to-pay thresholds), cost-effective from a payer perspective, and dominant/highly cost-effective (significant cost savings, improved outcomes). The key conclusions on cost-effectiveness were that TO in primary care offers significant cost savings compared to hospital TFP, TO is viable for Federally Qualified Health Centers (FQHCs), especially with sufficient patient flow, AI-assisted TO presents strong economic value in urban/rural Chinese settings, integrating TO into the structured German DMP is economically justifiable for payers, and TO in primary care is essential for cost-effective DR screening in underserved Brazil settings. The sensitivity analyses showed that the results were robust to plausible variations in costs and adherence rates, sensitive to screening volume, personnel costs, and TO equipment lifespan, robust across various scenarios but highly sensitive to QALY gain assumptions, sensitive to assumed rate of progression and long-term treatment costs averted, and remained favorable across a wide range of assumptions due to large travel cost differences.

Table 4. Summary of comparative cost-effectiveness findings – Tele-ophthalmology (TO) vs. traditional fundus photography (TFP) for DR screening.

Feature	Study 1	Study 2	Study 3	Study 5	Study 6
Analytic Approach / Model	Trial-based Cost-Consequence Analysis	Decision Tree Model alongside clinical study	Markov Model simulating disease progression & costs	Trial-based Cost-Effectiveness Analysis (within RCT)	Markov Model based on implementation data
Perspective	Healthcare Payer & Societal (implied by including patient travel)	Healthcare Payer (Federally Qualified Health Center perspective primarily)	Healthcare System	Statutory Health Insurance (Payer)	Societal (including patient travel costs)
Time Horizon	Assumed short-to-medium term (e.g., 5 years)	10 years	Lifetime	5 years (primary analysis)	Lifetime
Comparator (TFP Pathway)	Hospital-based 2-field digital photography (standard care)	Opportunistic screening via local eye care providers (variable practices)	Clinic-based 1-field digital photography graded by a local ophthalmologist	Standard DMP pathway with 2-field digital photography graded by a local ophthalmologist	Referral to distant hospital centers for TFP (status quo/historical access)
Main Outcome Measure(s)	Cost per patient screened; Cost per case of RDR detected	Cost per patient screened; Cost per case of RDR detected	Cost per Quality-Adjusted Life Year (QALY) gained	Cost per patient screened; Cost per case of RDR detected; (QALYs discussed)	Cost per QALY gained
Key Cost Drivers Identified (TO vs TFP)	**TO Favored:** * Reduced patient travel time & costs * Optimized grader time (centralized) **TFP Favored:** * Lower initial equipment cost (if existing)	**TO Favored:** * Lower cost per screen at high volume * Reduced patient no-shows **TFP Favored:** * High initial TO setup (camera, IT) * TO personnel training costs	**TO Favored (AI-assisted):** * High grader efficiency * Lower personnel cost/screen **TFP Favored:** * Simpler workflow (if established)	**TO Favored:** * Potential long-term treatment cost savings (earlier detection) **TFP Favored:** * TO integration/IT costs within DMP	**TO Favored:** * Dramatically lower patient travel costs & time * Lower system transport/outreach costs **TFP Favored:** * Minimal (assuming sunk costs)
Main Finding (ICER or Dominance)	**TO Likely Dominant** (Less costly, similar clinical outcomes reported)	**TO Cost-Effective** above moderate screening volumes (Threshold analysis performed)	**TO Highly Cost-Effective** (ICER well below typical Willingness-to-Pay thresholds, e.g., <$50,000/QALY)	**TO Cost-Effective** from payer perspective (incremental cost within acceptable range)	**TO Dominant / Highly Cost-Effective** (Significant cost savings, improved outcomes)
Key Conclusion on CE	TO in primary care offers significant cost savings compared to hospital TFP.	TO viable for FQHCs, especially with sufficient patient flow; scalability matters.	AI-assisted TO presents strong economic value in urban/rural Chinese settings.	Integrating TO into the structured German DMP is economically justifiable for payers.	TO in primary care essential for cost-effective DR screening in underserved Brazil settings.
Sensitivity Analysis Insights	Results robust to plausible variations in costs and adherence rates.	CE outcome sensitive to: * Screening volume * Personnel costs * TO equipment lifespan	Results robust across various scenarios; highly sensitive to QALY gain assumptions.	ICER sensitive to: * Assumed rate of progression * Long-term treatment costs averted	Results remained favorable across wide range of assumptions due to large travel cost differences.

Notes: AI: Artificial Intelligence; CE: Cost-Effectiveness; DMP: Disease Management Program; DR: Diabetic Retinopathy; FQHC: Federally Qualified Health Center; ICER: Incremental Cost-Effectiveness Ratio; IT: Information Technology; QALY: Quality-Adjusted Life Year; RCT: Randomized Controlled Trial; RDR: Referable Diabetic Retinopathy; TFP: Traditional Fundus Photography; TO: Tele-ophthalmology; WTP: Willingness-to-Pay.

Table 5 summarizes the policy implementation status, key facilitators, key barriers, reported impact on coverage/access, and key policy recommendations/insights for tele-ophthalmology (TO) based diabetic retinopathy (DR) screening, as reported in the included studies. The implementation status of TO-based DR screening varied across studies. It ranged from pilot implementations in Federally Qualified Health Centers (FQHCs) with limited integration into mainstream care pathways, to variable implementation across different settings, integration into established national Diabetes Management Programs (DMPs) with specific reimbursement codes, and implementation within specific primary care units as part of governmental projects. Several key facilitators for TO implementation were identified. These included the mission of FQHCs to serve underserved populations, the availability of external grant funding for initial setup, the potential to improve access for vulnerable groups, strong project champions within institutions, partnerships between tertiary centers and peripheral clinics, availability of trained ophthalmic technicians, potential for cost savings, existence of established reimbursement codes within DMPs, structured frameworks of national DMPs, high acceptance among ophthalmologists and diabetologists, clear quality standards within programs, strong governmental or project-level support, potential for significant reductions in patient and system travel costs, and integration possibility with existing primary care infrastructure. Several key barriers to TO implementation were also identified. These included a lack of sustainable, widespread reimbursement models beyond grants, provider skepticism or resistance (ophthalmology & primary care), challenges in integrating TO platforms with existing Electronic Health Records (EHRs), high initial capital costs for equipment, the need for defined workflows & staff roles, barriers for targeted underserved groups, the need for clearer reimbursement policies for TO screening, the need for local data use policies, strategies needed for better EHR integration, and strategies needed to build trust & acceptance. Other barriers included poor internet connectivity and unreliable electricity, especially in rural areas, high costs and logistical challenges of device maintenance, shortages of trained personnel (imagers, graders), a lack of national standards for TO quality assurance & data privacy, variable state-level policies & support, initial complexities in IT integration & data flow within DMP systems, ensuring consistent image quality & grading standards across different sites/providers, managing patient expectations & follow-up coordination, unreliable internet connectivity and electricity supply in many areas, difficulty in retaining trained staff in primary care settings, logistical hurdles for equipment transport and maintenance, and concerns regarding data security & patient privacy. The reported impact on coverage/access indicated that TO implementation significantly improved screening rates for participating FQHC patients compared to baseline and reduced geographic barriers, showed potential to increase reach in specific project areas (though overall national impact was limited by patchy implementation & infrastructure issues), contributed to maintaining high screening coverage within the structured national DMP, and dramatically improved access & reduced waiting times/travel burden for patients attending primary care units. Key policy recommendations and insights included the need for clear, value-based reimbursement policies, strategies for better EHR integration, strategies needed for better EHR integration, and strategies essential to build trust & acceptance. It was also recommended to require significant investment in digital infrastructure (connectivity), development of national TO standards & QA guidelines, tailored workforce training programs, and state/national policies to support scale-up. The table also demonstrates the success of integrating TO within existing, funded national health programs, highlights the importance of clear protocols, reimbursement, and stakeholder buy-in, and emphasizes that ongoing quality monitoring remains crucial. Finally, infrastructure development (internet, power) is a prerequisite for sustainable scale-up, and

it requires long-term governmental commitment & funding models beyond pilots, with a focus on robust training for primary care staff.

Table 5. Summary of policy uptake and implementation factors for tele-ophthalmology (TO) based DR screening.

Feature	Study 2	Study 4	Study 5	Study 6
Policy/Implementation Status	Pilot implementations in Federally Qualified Health Centers (FQHCs); Often grant-dependent; Limited integration into mainstream care pathways reported at time of study.	Variable implementation across different states/institutions; Pockets of use within specific tertiary clinics or regional programs; Not universally adopted in national policy.	Integrated into the established national structured Diabetes Management Program (DMP); Specific reimbursement codes available within the DMP framework.	Implemented within specific primary care units as part of governmental projects or research initiatives; Facing challenges in scaling up nationally.
Key Facilitators Identified	* Mission of FQHCs to serve underserved populations. * Availability of external grant funding for initial setup. * Potential to improve access for vulnerable groups. * Strong project champions within institutions.	* Partnerships between tertiary centers and peripheral clinics. * Availability of trained ophthalmic technicians in some areas. * Recognition of need due to high DR prevalence & access issues. * Lower cost per screening (potential) vs traditional referral.	* Existence of established reimbursement codes for TO within the DMP. * Structured framework of the national DMP facilitating integration. * High acceptance among participating ophthalmologists & diabetologists. * Clear quality standards within the program.	* Strong governmental or project-level support for pilots. * Potential for significant reductions in patient & system travel costs. * Integration possibility with existing primary care infrastructure. * High need due to geographical barriers to specialist care.
Key Barriers Identified	* Lack of sustainable, widespread reimbursement models beyond grants. * Provider skepticism or resistance (ophthalmology & primary care). * Challenges integrating TO platforms with existing Electronic Health Records (EHRs). * Initial high capital costs for equipment. * Need for defined workflows & staff roles.	* Poor internet connectivity & unreliable electricity, especially rurally. * High cost & logistical challenges of device maintenance. * Shortage of trained personnel (imagers, graders). * Lack of national standards for TO quality assurance & data privacy. * Variable state-level policies & support.	* Initial complexities in IT integration & data flow within DMP systems. * Ensuring consistent image quality & grading standards across different sites/providers. * Managing patient expectations & follow-up coordination.	* Unreliable internet connectivity and electricity supply in many areas. * Difficulty in retaining trained staff in primary care settings. * Logistical hurdles for equipment transport & maintenance. * Concerns regarding data security & patient privacy. * Funding uncertainty beyond initial project grants.
Reported Impact on Coverage/Access	* Significantly improved screening rates for participating FQHC patients compared to baseline/control. * Reduced geographic barriers for targeted underserved groups.	* Showed potential to increase reach in specific project areas. * Overall national impact limited by patchy implementation & infrastructure issues.	* Contributed to maintaining high screening coverage within the structured DMP. * Offered alternative pathway, potentially improving convenience for some patients.	* Dramatically improved access & reduced waiting times/travel burden for patients in participating primary care units compared to traditional referral pathways.
Key Policy Recommendations / Insights	* Need for clear, value-based reimbursement policies for TO screening at federal/state levels. * Strategies needed for better EHR integration. * Provider education essential to build trust & acceptance.	* Requires significant investment in digital infrastructure (connectivity). * Development of national TO standards & QA guidelines is critical. * Tailored workforce training programs needed. * State/National policies needed to support scale-up.	* Demonstrates success of integrating TO within existing, funded national health programs. * Highlights importance of clear protocols, reimbursement, & stakeholder buy-in. * Ongoing quality monitoring remains crucial.	* Infrastructure development (internet, power) is a prerequisite for sustainable scale-up. * Requires long-term governmental commitment & funding models beyond pilots. * Focus on robust training & support for primary care staff.

Notes: DMP: Disease Management Program; DR: Diabetic Retinopathy; EHR: Electronic Health Record; FQHC: Federally Qualified Health Center; IT: Information Technology; LMIC: Low- and Middle-Income Country; QA: Quality Assurance; RCT: Randomized Controlled Trial; TFP: Traditional Fundus Photography; TO: Tele-ophthalmology.

4. Discussion

The meta-analysis revealed that TO exhibits high diagnostic accuracy for detecting referable DR (RDR), with pooled sensitivity and specificity reaching 90% and 91%, respectively. These figures underscore TO's potential as a viable screening tool, capable of identifying individuals who require further ophthalmological evaluation with a reasonable degree of reliability. The sensitivity of a screening test is paramount as it reflects the test's ability to correctly identify individuals who truly have the condition, in this case, referable DR. A highly sensitive test minimizes the risk of false negatives, ensuring that individuals with RDR are not missed and receive timely intervention to prevent vision loss. Similarly, specificity, which measures the test's ability to correctly identify individuals who do not have the condition, is also crucial. A highly specific test minimizes false positives, reducing unnecessary referrals and the associated burden on healthcare systems and patients. While TO demonstrated strong diagnostic performance, it is important to acknowledge that the pooled estimates for TFP showed marginally higher accuracy, with a sensitivity of 92% and a specificity of 93%. This subtle difference suggests that, based on the synthesized evidence, TFP may possess a slightly superior diagnostic capability in detecting RDR. However, it is crucial to interpret this finding cautiously, considering the substantial heterogeneity observed in the accuracy estimates for both TO and TFP across the included studies. The presence of significant heterogeneity implies that the performance of both screening modalities is not uniform across different contexts and study populations. Factors such as variations in study design, patient populations, imaging protocols, grader expertise, and reference standards can all contribute to the observed variability in accuracy. Therefore, while the pooled estimates provide a general comparison, they should not be taken as definitive indicators of the superiority of one method over the other in all circumstances. The observed heterogeneity highlights the importance of considering contextual factors when selecting the most appropriate DR screening modality. In settings where access to specialized ophthalmic services is readily available and resources are not constrained, TFP may be the preferred option due to its potentially slightly higher accuracy. However, in resource-limited settings, geographically remote areas, or situations where patient access is a significant barrier, TO may offer a more practical and effective solution, even if its accuracy is slightly lower. The decision should be guided by a careful assessment of the specific needs and constraints of the target population and the healthcare system. It is also important to note that the accuracy of TO can be influenced by several factors, including the type of imaging technology used, the training and expertise of the personnel acquiring and interpreting the images, and the use of automated image analysis systems, such as artificial intelligence (AI). Variations in these factors can contribute to the observed heterogeneity in TO performance. Further research is needed to optimize TO protocols and technologies to maximize its diagnostic accuracy and consistency across different settings.[11-15]

The narrative synthesis of economic evaluations consistently suggested that TO pathways tend to be cost-effective and, in some scenarios, even dominant compared to TFP pathways. This finding carries significant implications for healthcare resource allocation and policy decisions related to DR screening programs. Cost-effectiveness analysis is a crucial tool for evaluating the economic efficiency of healthcare interventions. It compares the costs of different interventions with their health outcomes, allowing decision-makers to identify strategies that provide the greatest health benefit for the resources invested. In the context of DR screening, cost-effectiveness analyses can help determine whether TO or TFP represents a more efficient use of healthcare resources. The economic advantage of TO, as indicated by the synthesized evidence, is often driven by several factors. One key driver is the potential for reduced patient travel costs. TO can facilitate screening in remote or underserved areas, eliminating the need for patients to travel long distances to access specialized

ophthalmic services. This not only reduces the direct costs incurred by patients but also minimizes indirect costs associated with time off work and transportation arrangements. In addition, TO can optimize the use of specialized personnel. By enabling remote image interpretation, TO can centralize the expertise of ophthalmologists or trained graders, reducing the need for these specialists to be physically present at every screening site. This can lead to significant cost savings, particularly in settings where specialist availability is limited. Furthermore, TO can lead to efficiency gains in the screening process. For example, TO can streamline the workflow by enabling faster image acquisition and transmission, reducing patient waiting times, and improving the overall throughput of screening programs. These efficiency gains can contribute to cost savings and improved resource utilization. The finding that TO can be dominant in some scenarios, meaning that it is both less costly and more effective than TFP, further strengthens the economic argument for its adoption. Dominance implies that TO not only saves money but also leads to better health outcomes, making it a highly attractive option for DR screening. However, it is crucial to acknowledge that the cost-effectiveness of TO is not uniform across all settings and is sensitive to several factors. Initial investment costs associated with TO infrastructure, such as the purchase of fundus cameras, image transmission equipment, and information technology systems, can be substantial. These upfront costs may be a barrier to adoption, particularly in resource-constrained settings. Local cost structures, including personnel costs, equipment maintenance costs, and the costs of managing referrals and follow-up care, can also influence the overall cost-effectiveness of TO. Therefore, economic evaluations need to be context-specific, carefully considering local circumstances and resource availability. The specific program model implemented also plays a significant role in determining the cost-effectiveness of TO. Factors such as the screening protocols used, the involvement of AI in image interpretation, and the integration of TO into existing healthcare systems can all affect the economic profile of TO. For example, TO programs that utilize AI for initial image screening may achieve greater efficiency and cost savings compared to those that rely solely on human graders. Similarly, integrating TO into existing primary care settings can leverage existing infrastructure and resources, potentially reducing implementation costs. The finding that cost-effectiveness improves with scale is particularly important for policy planning. This suggests that initial investments in TO infrastructure can yield substantial long-term returns in large, organized screening programs. As the number of individuals screened increases, the fixed costs associated with TO implementation are spread across a larger population, leading to lower per-patient costs. This highlights the potential for TO to achieve significant cost savings in large-scale national or regional DR screening programs.[16-20]

5. Conclusion

In conclusion, this systematic review and meta-analysis provides a comprehensive evaluation of the comparative effectiveness, cost-effectiveness, and policy implications of tele-ophthalmology (TO) and traditional fundus photography (TFP) for diabetic retinopathy (DR) screening. The synthesized evidence indicates that TO demonstrates high diagnostic accuracy for detecting referable DR, achieving pooled sensitivity and specificity comparable to TFP. While TFP may offer slightly superior diagnostic capability, the substantial heterogeneity observed across studies suggests that the choice between TO and TFP should be context-dependent, considering factors such as resource availability and patient access. Economic evaluations largely favor TO, highlighting its potential for efficient resource allocation in DR screening programs. The cost-effectiveness of TO is often driven by reduced patient travel costs, optimized use of specialized personnel, and efficiency gains in the screening process. However, the cost-effectiveness of TO is influenced by factors such as initial investment costs, local cost structures, and the specific program

model implemented. Successful translation of TO into widespread, effective public health policy requires addressing implementation barriers related to infrastructure, workforce training, reimbursement parity, and quality assurance. Facilitators for TO implementation include strong project champions, partnerships between tertiary centers and peripheral clinics, and the potential for integration with existing primary care infrastructure. Barriers include a lack of sustainable reimbursement models, provider skepticism, challenges in integrating TO platforms with electronic health records, and infrastructure limitations in some settings. Future policies should focus on creating supportive ecosystems for TO integration to maximize its potential in preventing diabetes-related vision loss. This includes the development of clear reimbursement policies, strategies for better electronic health record integration, tailored workforce training programs, and the establishment of national tele-ophthalmology standards and quality assurance guidelines.

6. References

1. Fernández-Gutiérrez DÁ, Núñez-Marrero J, Martínez-Alberto CE, Rodríguez-Álvaro M, Retisalud-Tenerife Primary Healthcare Research Group, García-Hernández AM, et al. Knowledge of diabetic retinopathy among primary care nurses performing fundus photography and agreement with ophthalmologists on screening. Nurs Rep. 2023; 13(3): 1064–76.
2. Castro C, Marques JH, Silva N, Abreu AC, Furtado MJ, Lume M. Comparison of color fundus photography and multicolor fundus imaging for detection of lesions in diabetic retinopathy and retinal vein occlusion. Clin Ophthalmol. 2023; 17: 2515–24.
3. Avanesova TA, Oganezova JG, Anisimova VV, Baeva AB, Miaev DK. Prevalence of diabetic retinopathy assessed using two-field mydriatic fundus photography. Vestn Oftalmol. 2024; 140(4): 60–7.
4. Kabanarou S, Bitzanakis N, Bontzos G, Smoustopoulos G, Sgourou C, Garnavou-Xirou C, et al. Color fundus photography versus fundoscopy in diabetic retinopathy patients: a pilot study. Acta Ophthalmol. 2023; 102(S279).
5. Hu C-L, Wang Y-C, Wu W-F, Xi Y. Evaluation of AI-enhanced non-mydriatic fundus photography for diabetic retinopathy screening. Photodiagnosis Photodyn Ther. 2023; 49(104331): 104331.
6. Zhou W, Yuan X-J, Li J, Wang W, Zhang H-Q, Hu Y-Y, et al. Application of non-mydriatic fundus photography-assisted telemedicine in diabetic retinopathy screening. World J Diabetes. 2023; 15(2): 251–9.
7. Sadek NA, Al-Dahan ZT, Rattan SA. Comprehensive survey of the state-of-the-art Deep Learning models for Diabetic Retinopathy detection and grading using retinal fundus photography. Al-Nahrain J Eng Sci. 2023; 27(2): 155–63.
8. Tan R, Teo KYC, Husain R, Tan NC, Lee QX, Hamzah H, et al. Evaluating the outcome of screening for glaucoma using colour fundus photography-based referral criteria in a teleophthalmology screening programme for diabetic retinopathy. Br J Ophthalmol. 2023; 108(7): 933–9.
9. Szulborski KJ, Ramsey DJ. Enhancing diabetic retinopathy screening: Non-mydriatic fundus photography-assisted telemedicine for improved clinical management. World J Diabetes. 2023; 15(8): 1820–3.
10. Liu Z, Han X, Gao L, Chen S, Huang W, Li P, et al. Cost-effectiveness of incorporating self-imaging optical coherence tomography into fundus photography-based diabetic retinopathy screening. NPJ Digit Med. 2023; 7(1):225.
11. He X, Deng X, Lin Z, Wen L, Zhou W, Xu X, et al. SCREENING AND MONITORING OF DIABETIC RETINOPATHY IN COMMUNITY

CARE: The effectiveness of single-field versus multifield fundus photography. Retina. 2023; 45(2): 318–24.

12. Wroblewski JJ, Sanchez-Buenfil E, Inciarte M, Berdia J, Blake L, Wroblewski S, et al. Diabetic retinopathy screening using smartphone-based fundus photography and deep-learning artificial intelligence in the Yucatan Peninsula: a field study. J Diabetes Sci Technol. 2023; 19(2): 370–6.
13. Pareja-Ríos A, Bonaque-González S, Serrano-García M, Cabrera-López F, Abreu-Reyes P, Marrero-Saavedra MD. Tele-ophthalmology for diabetic retinopathy screening: 8 years of experience. Arch Soc Esp Oftalmol (Engl Ed). 2017; 92(2): 63–70.
14. Nunes F, Madureira P, Rego S, Braga C, Moutinho R, Oliveira T, et al. A mobile Tele-ophthalmology system for planned and opportunistic screening of diabetic retinopathy in primary care. IEEE Access. 2021; 9: 83740–50.
15. Abreu-González R, Alonso-Plasencia M. Tele-ophthalmology for diabetic retinopathy screening: Value of continuing education. Arch Soc Esp Oftalmol (Engl Ed). 2017; 92(7): e37–8.
16. Pareja-Ríos A, Ceruso S, Romero-Aroca P, Bonaque-González S. A new deep learning algorithm with activation mapping for diabetic retinopathy: Backtesting after 10 years of Tele-ophthalmology. J Clin Med. 2022; 11(17): 4945.
17. Kuklinski EJ, Henry RK, Shah M, Zarbin MA, Szirth B, Bhagat N. Screening of diabetic retinopathy using artificial intelligence and Tele-ophthalmology. J Diabetes Sci Technol. 2023; 17(6): 1724–5.
18. Cao J, Felfeli T, Merritt R, Brent MH. Sociodemographics associated with risk of diabetic retinopathy detected by Tele-ophthalmology: 5-year results of the Toronto Tele-Retinal Screening Program. Can J Diabetes. 2022; 46(1): 26–31.
19. Rajalakshmi R, UmaSankari G, Prathiba V, Anjana RM, Unnikrishnan R, Venkatesan U, et al. Tele-ophthalmology versus face-to-face retinal consultation for assessment of diabetic retinopathy in diabetes care centers in India: a multicenter cross-sectional study. Diabetes Technol Ther. 2022; 24(8): 556–63.
20. Boucher MC, Qian J, Brent MH, Wong DT, Sheidow T, Duval R, et al. Evidence-based Canadian guidelines for tele-retina screening for diabetic retinopathy: recommendations from the Canadian Retina Research Network (CR2N) Tele-Retina Steering Committee. Can J Ophthalmol. 2020; 55(1 Suppl 1): 14–24.