LINK MEDICAL JOURNAL
of Health and Community Research | LMJHCR
Volume 4, Issue 1 | June 2026 | ISSN 3007-3448

Indexed in DOAJ | CC Bq4jc7c36Y 4.0
DOI: 10.61919/mx93my14

Original Article

Patient Experiences of AI-Assisted Diabetic Retinopathy Screening: Trust, Consent and Understanding of Results

Zhang Shuixin[1], Ali Napiah Nasution[1], Clarisa Lister[1]

[1] Universitas Prima Indonesia, Indonesia

***Corresponding author: Ali Napiah Nasution, alinapiahnasution@unprimdn.ac.id**

"Cite this Article" | **Received**: 11 March 2026; **Accepted**: 23 June 2026; **Published**: 30 June 2026.

ABSTRACT

Background: Artificial intelligence (AI)-assisted diabetic retinopathy screening can increase the speed and accessibility of retinal image interpretation, particularly in primary-care and underserved settings. However, diagnostic performance alone does not establish whether patients understand AI involvement, provide meaningful consent, trust automated decisions, or successfully navigate care after screening. Patient-centred evidence remains fragmented across studies of acceptance, satisfaction, trust, consent, data governance, communication, and referral. Objective: This critical integrative review synthesized patient-focused evidence on AI-assisted diabetic retinopathy screening to examine how patients develop or withhold trust in AI-supported screening, whether disclosure and consent processes support meaningful understanding and choice, and how patients comprehend results and navigate subsequent clinical care. Methods: Peer-reviewed literature published between January 2016 and July 2026 was identified using structured combinations of terms related to diabetic retinopathy, retinal screening, artificial intelligence, patient experience, trust, consent, privacy, understanding, communication, and referral, supplemented by citation chaining. Twenty publications informed the conceptual and theoretical context, while a separate corpus of 15 primary studies published between 2021 and 2026 informed the findings. Heterogeneous evidence from surveys, implementation studies, experiments, and patient-experience assessments was synthesized thematically, with design-sensitive critical appraisal. Results: Three interrelated themes emerged. Trust was conditional on visible human oversight and accountability, with patients generally preferring AI as an adjunct rather than an autonomous substitute for clinicians. Evidence concerning consent was substantially less developed than evidence concerning acceptance; disclosure did not necessarily establish comprehension, voluntariness, or permission for secondary data use. Speed, local access, and immediate results were valued, but favourable screening experiences did not consistently translate into referral completion or understanding of subsequent actions. Conclusion: Patient-centred AI-assisted diabetic retinopathy screening requires more than technical accuracy or high satisfaction. Sustainable implementation depends on identifiable human accountability, intelligible consent, practical explanation of results, transparent data governance, meaningful choice, and a navigable pathway from screening to follow-up care. Keywords: artificial intelligence; diabetic retinopathy; screening; patient experience; trust; informed consent; result communication

EDITORIAL INFORMATION

Author Contributions: Concept: ZS; Design: ZS, ANN; Data Collection: ZS, ANN, CL; Analysis: ZS, ANN; Drafting: ZS; Review and Editing: ANN, CL.
Ethical Approval: Universitas Prima Indonesia
Informed Consent: Written informed consent was obtained from all participants
Conflict of Interest: The authors declare no conflict of interest; **Funding**: No external funding; **Data Availability**: Available from the corresponding author on reasonable request; **Acknowledgments**: N/A.

INTRODUCTION

Artificial intelligence (AI) is increasingly being incorporated into diabetic retinopathy screening to support automated interpretation of retinal images, extend screening beyond specialist ophthalmology settings, and reduce dependence on limited human grading capacity. Deep-learning systems have demonstrated clinically relevant performance for detecting referable diabetic retinopathy, including autonomous use in primary-care environments and validation across geographically diverse populations (1-4). These developments have shifted the implementation question beyond whether an algorithm can classify retinal images accurately toward whether AI-assisted screening can be integrated into clinical pathways in ways that are acceptable, understandable, equitable, and accountable to patients. Real-world evaluations further indicate that clinical performance depends not only on algorithmic accuracy but also on image

quality, workflow integration, staff communication, referral processes, and the interaction between automated outputs and healthcare professionals (5-7).

This distinction is particularly important in diabetic retinopathy screening because the patient experience extends beyond image acquisition and automated classification. Individuals undergoing screening may need to understand why retinal images are being collected, whether an AI system is involved in interpretation, whether a clinician will review the result, how their data may be stored or reused, what an uncertain or ungradable result means, and who remains responsible if an automated output is incorrect. AI can redistribute interpretive and decision-making functions among primary-care teams, retinal graders, ophthalmologists, healthcare institutions, technology vendors, and automated systems, yet these changes may remain largely invisible to patients. Consequently, willingness to undergo AI-assisted screening or satisfaction with a rapid and convenient service cannot automatically be interpreted as informed trust in the technology or as evidence of meaningful consent.

Trust is especially consequential when patients cannot independently evaluate an algorithm's technical performance and therefore depend on clinicians and healthcare institutions to select, oversee, interpret, and correct AI-supported systems. Existing research across medical AI indicates that acceptance is often conditional: patients may value speed, accessibility, and standardized assessment while simultaneously preferring professional oversight and resisting the transfer of final clinical authority to an autonomous system (8-10). Ethical analyses likewise emphasize that transparency, privacy, accountability, and meaningful choice become increasingly important as algorithmic decision-making is incorporated into routine care (11-14). In this context, disclosure that AI is being used may be insufficient if patients do not understand its role, limitations, relationship to human review, or implications for their personal health information.

Result communication presents a related but distinct challenge. Diabetic retinopathy screening generally produces an action-oriented classification rather than a definitive clinical diagnosis: patients may need to distinguish between absence of referable disease and absence of any retinopathy, understand why some images are ungradable, recognize uncertainty in an automated assessment, and know what action is required after a positive screen. Immediate results can improve convenience, but rapid classification has limited patient-centred value if the meaning of the result is unclear or if referral pathways are inaccessible. Evidence from medical AI further suggests that patients commonly value professional involvement, accessible explanation, and identifiable responsibility even when they are receptive to automated technologies (15-18). Patient-centred implementation therefore requires attention to practical explainability—the ability to communicate what was assessed, what the result means, what uncertainty remains, and what should happen next—rather than expecting patients to understand the internal computational structure of an AI model.

Despite substantial literature on the diagnostic performance and feasibility of AI-based diabetic retinopathy screening, the patient-facing evidence remains fragmented. Technical validation studies primarily address sensitivity, specificity, workflow, or operational feasibility, whereas studies of patient attitudes frequently focus on willingness, satisfaction, or general acceptability. These outcomes do not necessarily establish whether consent was informed and voluntary, whether patients understood the role and limitations of AI, whether responsibility for decisions was clear, or whether a rapid screening result led to successful navigation of subsequent care. Direct evidence concerning consent quality, sustained comprehension, meaningful alternatives, responses to ungradable or uncertain results, and accountability after error remains comparatively limited. Differences in health and digital literacy, cultural context, prior institutional trust, and referral capacity may further influence how patients experience the same technology.

A focused synthesis is therefore needed to distinguish technical acceptance from ethically and clinically meaningful participation in AI-assisted diabetic retinopathy screening. Rather than treating trust, consent, and understanding as interchangeable manifestations of satisfaction, this review considers them as related but analytically distinct dimensions of patient-centred implementation. Accordingly, this critical

integrative review synthesizes recent patient-focused evidence to examine three questions: how patients develop or withhold trust in AI-assisted diabetic retinopathy screening; whether consent and disclosure processes support meaningful understanding and voluntary participation; and how patients comprehend screening results and navigate subsequent clinical action. The review further evaluates how human oversight, accountability, workflow, data governance, health literacy, and referral pathways shape these experiences, with the aim of identifying practical requirements for patient-centred implementation of AI-assisted retinal screening.

MATERIALS AND METHODS

This study was conducted as a critical integrative review with thematic synthesis. An integrative approach was selected because the relevant patient-experience literature comprised methodologically heterogeneous evidence, including cross-sectional surveys, prospective implementation studies, questionnaire-based feasibility studies, web-based experiments, and mixed-method evaluations. Such heterogeneity was not considered appropriate for statistical pooling because the studies differed in populations, clinical settings, degree of actual exposure to AI, measurement approaches, and outcomes of interest. Integrative review methodology permits evidence from heterogeneous empirical designs to be examined comparatively while retaining attention to methodological and contextual differences (19). The synthesis was interpretative and thematic rather than meta-analytic, with theme development informed by established principles of thematic analysis (20).

A structured literature search was undertaken for peer-reviewed publications issued between January 2016 and July 2026. Search concepts combined terms relating to the clinical setting and technology—*diabetic retinopathy*, *retinal screening*, *artificial intelligence*, and *autonomous AI*—with patient-centred concepts including *patient experience*, *acceptance*, *trust*, *consent*, *privacy*, *understanding*, *communication*, and *referral*. Citation chaining from relevant publications was additionally used to identify implementation, patient-experience, governance, and ethics literature not captured through the initial concept combinations. Only English-language peer-reviewed literature relevant to healthcare AI and at least one of the predefined patient-centred domains was considered. Because the review was integrative rather than a formal systematic review, the search was designed to identify conceptually relevant evidence for critical synthesis rather than to support statistical estimation of an exhaustive pooled effect.

To maintain a clear distinction between conceptual background and evidence used to generate the principal findings, two literature functions were defined. Twenty publications were used to establish the technical, theoretical, ethical, and implementation context presented in the Introduction and conceptual interpretation. A separate findings corpus comprised 15 primary empirical studies published between 2021 and 2026. Direct studies of patient experiences with AI-assisted diabetic retinopathy or retinal screening were prioritized. Evidence from neighbouring diagnostic contexts was retained only when it addressed patient-centred issues that were insufficiently examined in the retinal-screening literature, particularly informed consent, data sharing, physician oversight, accountability, or explanation of AI-supported results. This separation was intended to prevent theoretical and commentary sources from being treated as equivalent to direct patient evidence in the thematic findings.

Studies were eligible for the findings synthesis when they reported original patient-derived data relevant to at least one of the core domains of trust, consent, understanding, communication, acceptance, human oversight, data governance, or referral experience in AI-supported healthcare. Greater interpretive weight was placed on studies involving actual exposure to AI-assisted screening or clinical workflows than on studies measuring hypothetical willingness or attitudes toward imagined AI use. Review articles and commentaries were excluded from the primary findings corpus, as were technical algorithm-development or validation studies without patient-reported data and studies reporting only clinician perspectives. Studies from diagnostic specialties outside diabetic retinopathy were included selectively and only where they addressed a patient-experience construct for which direct retinal-screening evidence was limited.

For each study included in the primary findings corpus, information was extracted on the clinical setting, participant characteristics, sample, role of AI within the clinical pathway, whether participants had actual

or hypothetical exposure to the technology, patient-reported outcomes, communication processes, human oversight, consent or data-governance issues, referral experiences, and study limitations. Extraction was structured around the review questions while retaining methodological features necessary to distinguish direct experience from hypothetical attitudes and immediate satisfaction from evidence of comprehension or subsequent behaviour.

The synthesis combined deductive and inductive interpretation. Trust, consent, and understanding were used as sensitising concepts because they represented the prespecified domains of the review. Within these domains, recurrent concepts were identified inductively, including expectations of human oversight, allocation of responsibility, convenience and immediacy of results, AI familiarity, data-sharing preferences, result explanation, and navigation of referral pathways. Codes were compared across studies, clinical contexts, and types of AI exposure. Particular attention was given to differences between studies examining hypothetical preferences and those evaluating patients who had undergone AI-assisted screening. Candidate themes were retained when they were supported across multiple studies and could accommodate both convergent and contradictory findings rather than representing isolated positive or negative observations. Theme development and refinement followed the conceptual principles of thematic analysis described by Braun and Clarke (20).

Methodological appraisal was incorporated into interpretation rather than converted into a single numerical quality score. Studies were examined in relation to relevance to the review question, study design, representativeness of the recruited population, validity and limitations of patient-experience measures, degree of participant exposure to the AI system, and the extent to which reported outcomes represented attitudes, immediate experience, comprehension, or subsequent behaviour. Hypothetical willingness surveys were therefore interpreted as measures of anticipated attitudes rather than direct evidence of behaviour, whereas high post-screening satisfaction was not treated as proof of informed consent, durable trust, or understanding. Evidence based on actual screening experience and downstream behaviour was given greater interpretive weight when addressing implementation questions. This design-sensitive appraisal was used to prevent methodologically different forms of evidence from being treated as equivalent.

The review did not undertake quantitative meta-analysis because substantial heterogeneity in design, exposure, measures, and patient outcomes precluded a clinically meaningful pooled estimate. Formal statistical heterogeneity testing and publication-bias analyses were consequently not applicable to the thematic synthesis. As an integrative review of previously published literature, the study involved no recruitment of human participants, collection of identifiable patient information, or primary intervention and therefore did not require institutional research ethics approval.

RESULTS

The primary findings corpus comprised 15 empirical studies published between 2021 and 2026 that examined patient attitudes, experiences, preferences, or behaviours relevant to artificial intelligence in diabetic retinopathy screening or closely related diagnostic settings. Direct evidence from diabetic retinopathy and retinal-screening contexts was prioritized, while studies from adjacent diagnostic settings were retained when they addressed patient-centred domains for which retinal-specific evidence remained limited, particularly informed consent, data sharing, physician oversight, and explanation. The evidence base was methodologically heterogeneous and included patient surveys, questionnaire-based feasibility evaluations, implementation studies, web-based experimental research, and patient-experience assessments. Accordingly, findings were synthesized thematically rather than quantitatively. Three interrelated themes emerged: conditional trust and the importance of visible human accountability; procedural consent and uneven understanding of AI; and the value of speed and accessibility only when screening results were connected to an understandable pathway to care.

Theme 1: Trust Was Conditional on Visible Human Accountability

Across the evidence base, patients generally did not reject AI-assisted screening simply because an algorithm participated in image interpretation. Instead, acceptance depended substantially on the role assigned to AI within the clinical pathway. Yap et al. reported favourable patient perceptions of AI in diabetic eye screening, but the pattern of responses suggested that general willingness to use an AI-enabled service could coexist with more qualified trust in the technology itself (21). Fritsch et al. similarly identified broadly positive attitudes toward healthcare AI despite limited patient knowledge and substantial expectations that physicians would remain involved in oversight and decision-making (30). Lennartz et al. extended this distinction across the clinical workflow, demonstrating greater patient comfort when AI performed an assistive function than when it occupied the position of final clinical decision-maker (31). Collectively, these studies indicated that acceptance was role-dependent rather than equivalent to unrestricted confidence in autonomous AI.

Table 1. Characteristics and Contribution of Studies Included in the Primary Findings Synthesis

Study	Year	Clinical/Evidence Context	Principal Patient Centred Domain	Contribution to Synthesis
Yap et al. (21)	2022	Diabetic eye screening	Perceptions, willingness, trust	Demonstrated generally favourable perceptions of AI-assisted screening while indicating that willingness and trust were not equivalent constructs
Jin et al. (22)	2025	AI-assisted diabetic retinopathy screening	Acceptance, usefulness, ease of use, trust	Examined determinants of acceptance; usefulness and ease contributed to acceptance, whereas trust did not function as a significant moderator
Rustam et al. (23)	2026	AI-based diabetic retinopathy screening, urban US setting	Satisfaction, clinician involvement, autonomy	Showed high comfort and satisfaction alongside reluctance to replace physician involvement with AI
Aggarwal et al. (24)	2021	Healthcare AI and health-data use	Data sharing, privacy, secondary use	Demonstrated that willingness to share health data varied according to recipient, purpose, and perceived protections
Park (25)	2024	Medical AI, web-based experiment	Disclosure and informed consent	Showed that the framing of AI disclosure and consent could influence patient preferences
Malerbi et al. (26)	2024	Real-world deep-learning-assisted diabetic retinopathy screening	AI familiarity, attitudes, human collaboration	Identified limited AI familiarity despite favourable attitudes and preference for AI-human collaboration
Nolan et al. (27)	2023	Rural primary-care diabetic retinopathy screening	Experience, accessibility, repeat-use preference	Demonstrated favourable experience with locally delivered AI screening and willingness to use the service again
Whitestone et al. (28)	2024	AI-based diabetic retinopathy screening in Rwanda	Feasibility, satisfaction, referral	Demonstrated high screening acceptance while revealing a discrepancy between positive screening experience and subsequent referral attendance
Bhambhwani et al. (29)	2025	AI-based diabetic retinopathy screening in Northern Ontario	Satisfaction, screening completion, referral	Reported favourable patient experience and comparatively strong ophthalmology attendance among referred participants
Fritsch et al. (30)	2022	Healthcare AI	Knowledge, attitudes, physician control	Identified positive attitudes toward AI alongside limited knowledge and a strong expectation of physician oversight
Lennartz et al. (31)	2021	AI across the medical workflow	Control, decision authority, workflow role	Indicated greater patient comfort with AI in assistive roles than with AI assuming final decision-making authority
Wahlich et al. (32)	2025	English NHS diabetic eye screening programme	Acceptability, safety, equity, professional oversight	Demonstrated conditional support for AI when professional safeguards, safety, and equity were maintained
Krogh et al. (33)	2025	AI-assisted diabetic retinopathy screening in primary care	Acceptance and patient-related determinants	Demonstrated overall acceptance of AI-assisted screening while identifying variation in factors contributing to that acceptance
Baghdadi et al. (34)	2024	AI as a diagnostic tool in radiology	Knowledge, attitudes, professional communication	Showed interest in AI together with continued preference for communication with healthcare professionals
Shah et al. (35)	2022	AI-based retinal screening	Acceptability, satisfaction, perceived understanding	Reported high willingness and satisfaction and perceived improvement in knowledge, while direct testing of comprehension remained limited

Note: The table summarizes each study's contribution to the present thematic synthesis rather than constituting a formal quantitative comparison. Studies outside diabetic retinopathy screening were used only for patient-centred domains for which direct retinal evidence was comparatively limited.

Retinal-screening studies reinforced this pattern. Wahlich et al., examining perceptions within the English NHS diabetic eye screening programme, found support for AI use in a context where safety, equity, and professional controls remained central to implementation (32). Krogh et al. likewise found acceptance of AI-assisted diabetic retinopathy screening in primary care, although acceptance reflected multiple patient-related considerations rather than undifferentiated trust in automation (33). Jin et al. found that perceived usefulness and ease of use contributed to acceptance of AI for diabetic retinopathy screening, whereas trust did not significantly moderate the relationship examined in their model (22). These results further distinguish practical acceptance of an accessible screening technology from complete confidence in its independent clinical authority.

The distinction became particularly apparent in Rustam et al., where patients reported high comfort and satisfaction with AI-based diabetic retinopathy screening but remained reluctant to regard AI as a substitute for physician involvement (23). Patients continued to value medical supervision and identifiable professional responsibility even when their immediate experience of automated screening was positive. Across these studies, willingness, satisfaction, trust, and delegation of decision-making authority were related but non-equivalent outcomes.

The synthesis consequently identified visible human accountability as the principal condition connecting favourable attitudes toward AI with sustained trust. Patients appeared more receptive when AI functioned within an identifiable clinical structure in which a healthcare professional remained available to interpret results, answer questions, review uncertain findings, and retain responsibility for subsequent clinical decisions. Conversely, evidence of convenience or satisfaction alone did not establish that patients were willing to transfer final authority to an autonomous system.

Theme 2: Consent Was Often Procedural, While Understanding of AI Remained Uneven

Evidence addressing consent and comprehension was less developed than evidence addressing general acceptance. Park's experimental study demonstrated that patient preferences could vary according to how AI involvement and consent were framed, indicating that disclosure is not a neutral informational event and that the clinical role assigned to AI can influence responses to consent (25). This finding suggested that merely identifying the presence of AI does not necessarily establish meaningful understanding of how the technology affects the screening or diagnostic process.

Data governance represented a related dimension of patient choice. Aggarwal et al. demonstrated substantial variation in willingness to share health data for AI-related purposes depending on who would receive the data, why the data were being used, and what protections were perceived to be in place (24). The findings therefore did not support treating participation in clinical imaging as equivalent to unrestricted authorization for secondary use of health information in algorithm development or other downstream applications. Direct-care participation and willingness to permit additional data uses emerged as conceptually separate decisions.

Understanding of AI was also uneven in studies conducted during actual or anticipated clinical use. Malerbi et al. found limited familiarity and self-reported knowledge of AI among participants attending a real-world deep-learning-assisted diabetic retinopathy screening event, despite generally positive attitudes and preference for collaboration between AI and healthcare professionals (26). In a neighbouring diagnostic context, Baghdadi et al. similarly found that patients were interested in AI while continuing to value communication with healthcare professionals (34). Together, these findings indicated that favourable attitudes can develop even when technical or procedural understanding remains incomplete.

Shah et al. reported high willingness and satisfaction with AI-based retinal screening, and participants commonly perceived that screening improved their knowledge of their eye condition (35). However, perceived improvement in knowledge was not equivalent to demonstrated comprehension of the limitations of screening. The available evidence did not establish whether patients could reliably distinguish screening from definitive diagnosis, understand false-positive or false-negative possibilities, interpret an ungradable image, or describe the degree and nature of human review. Similarly, evidence that patients possessed a practically accessible alternative to AI-assisted screening, or could decline its use without disadvantage, was limited.

Taken together, the studies showed a substantial asymmetry between evidence for acceptability and evidence for informed understanding. Acceptance and immediate satisfaction were commonly measured, whereas retained comprehension, voluntariness, understanding of data use, recognition of uncertainty, and awareness of alternatives were examined much less consistently. The synthesis therefore identified consent as the least empirically developed of the three patient-centred domains. Disclosure of AI involvement was necessary but could not, by itself, demonstrate informed and voluntary participation.

Theme 3: Speed and Access Created Value Only When Results Led to Care

A consistent advantage attributed to AI-assisted diabetic retinopathy screening was its capacity to provide retinal assessment within more accessible settings and, in some programmes, to shorten the interval between image acquisition and receipt of a screening result. Nolan et al. reported favourable experiences among patients receiving AI screening in a rural primary-care setting, with many participants expressing a preference to undergo future screening within the same practice (27). The attractiveness of the model therefore appeared to arise not solely from the algorithm but from the combination of local access, familiar clinical surroundings, and simplified delivery of retinal screening.

Bhambhwani et al. similarly reported favourable patient experience with a pilot AI-based diabetic retinopathy screening programme in Northern Ontario (29). Most participants successfully completed screening, and ophthalmology attendance among referred participants was comparatively strong. This evidence suggested that AI-assisted screening could be perceived positively when rapid assessment was linked to a functioning downstream pathway rather than operating as an isolated technological encounter.

However, favourable screening experience did not consistently predict completion of subsequent care. Whitestone et al. reported very high satisfaction and substantial acceptance of AI-based diabetic retinopathy screening in Rwanda while referral adherence remained considerably lower than the level of satisfaction with the screening encounter itself (28). This divergence was one of the most important findings of the synthesis because it demonstrated that patient satisfaction at the point of screening could coexist with failure to complete clinically necessary follow-up.

The comparison across implementation settings indicated that the value of immediate classification was contingent on what occurred after the result was produced. Screening outcomes could generate a need for further ophthalmic evaluation, but patients might still encounter transportation difficulties, cost, uncertainty about the urgency of referral, specialist availability, or insufficient explanation of the next step. Accordingly, speed of result generation and ease of initial access could remove one barrier while leaving subsequent barriers unresolved.

Understanding the result itself was part of this pathway. A patient-centred result required more than a positive or negative automated classification; its clinical meaning depended on whether patients understood what had been assessed, whether the image was gradable, whether additional evaluation was required, and how that evaluation could be obtained. Across the evidence base, the strongest implementation value therefore appeared to arise when AI shortened and simplified the whole screening-to-care pathway, rather than merely accelerating image classification.

Table 2. Integrated Thematic Synthesis of Patient Experiences With AI-Assisted Diabetic Retinopathy Screening

Theme	Principal Evidence Pattern	Evidence Supporting the Theme	Contrasting or Qualifying Evidence	Critical Synthesis
Conditional trust and visible human accountability	Patients generally accepted AI more readily as an adjunct than as an autonomous replacement for clinicians	Favourable perceptions, usefulness, convenience, and satisfaction were reported across retinal and broader healthcare AI studies (21-23,30-33)	Positive attitudes frequently coexisted with limited AI knowledge, preference for physician control, and reluctance to transfer final decision authority to AI (23,30-32)	Acceptance was role-specific. Willingness to use AI did not establish unconditional trust in autonomous decision-making
Procedural consent and uneven AI understanding	Disclosure, familiarity, and willingness to share data varied considerably; direct evidence of comprehension and voluntariness was sparse	Consent framing influenced preferences; data-sharing choices varied by purpose and protection; real-world screening studies identified favourable attitudes (24-26)	Low AI familiarity persisted, patients continued to value professional communication, and studies rarely demonstrated retained understanding of uncertainty, alternatives, or secondary data use (26,34,35)	Disclosure of AI was not equivalent to informed understanding. Evidence supporting acceptance was stronger than evidence supporting meaningful consent
Speed and access created value only when results led to care	Local screening, convenience, and rapid feedback were generally valued	Rural primary-care and remote screening programmes reported favourable experiences and willingness to use local services (27-29)	High satisfaction could coexist with substantially poorer referral completion, demonstrating a gap between screening acceptance and downstream action (28)	Patient experience extended beyond image capture and classification; clinical value depended on understandable results and a navigable referral pathway

Note: Themes were generated from the separate 15-study primary findings corpus. The table distinguishes the dominant evidence pattern from contradictory or qualifying findings to avoid interpreting uniformly positive satisfaction measures as equivalent to trust, informed consent, comprehension, or successful completion of care.

Cross-Theme Synthesis

The three themes were closely related but represented distinct dimensions of patient experience. Trust concerned who or what patients were prepared to rely on; consent concerned whether patients understood AI involvement and retained meaningful choice; and result understanding concerned whether an automated output could be translated into appropriate subsequent action. Across all three domains, the evidence favoured a relational rather than technology-centred interpretation of acceptability. Patients commonly valued convenience, speed, and expanded access, but these advantages were repeatedly embedded within expectations of clinician involvement, understandable communication, and a functioning healthcare pathway.

The evidence also revealed a recurring difference between proximal experience measures and more demanding patient-centred outcomes. Willingness, satisfaction, and favourable attitudes were comparatively common outcomes in the included studies, whereas retained comprehension, voluntariness, understanding of data reuse, error response, and completed referral were examined less consistently. Consequently, a positive screening encounter could not be assumed to demonstrate informed consent, durable trust, or successful clinical follow-through.

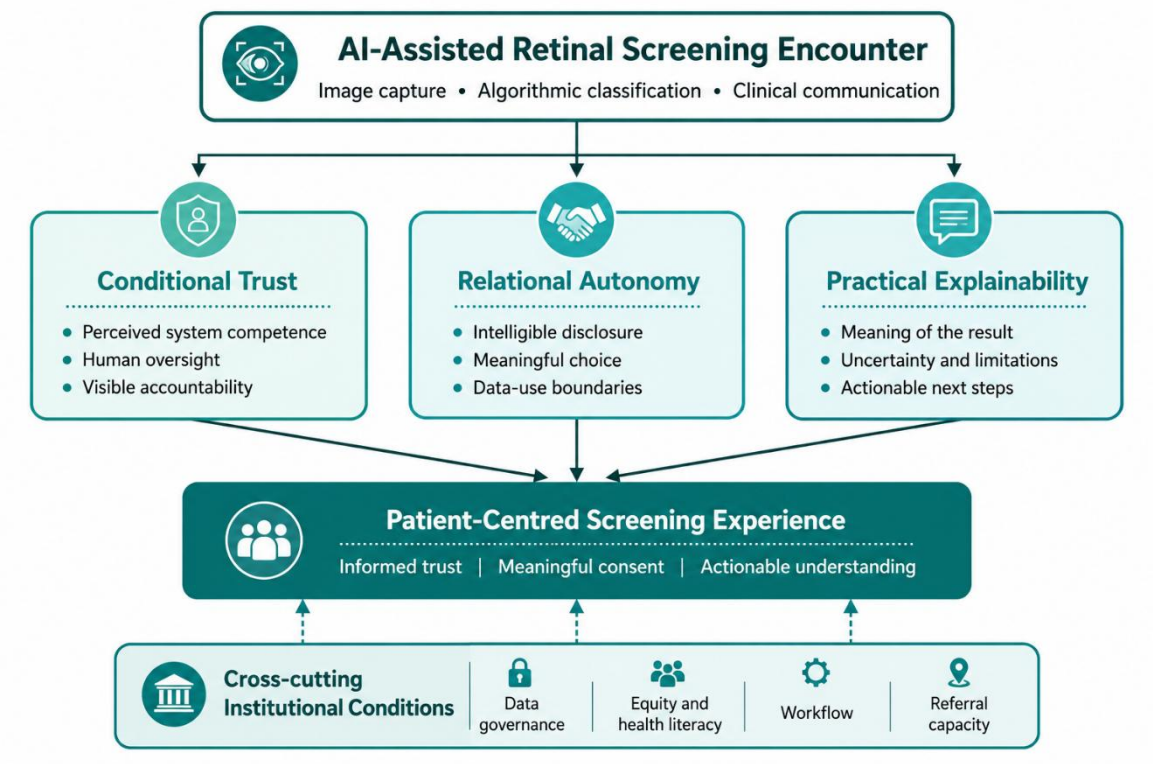

Figure 1 Integrated conceptual framework for patient experiences of AI-assisted diabetic retinopathy screening.

Direct retinal-screening studies provided the strongest evidence concerning accessibility, satisfaction, and real-world workflow, whereas neighbouring diagnostic studies contributed principally to domains in which retinal-specific research remained sparse, particularly consent framing, data-sharing preferences, physician control, and expectations for explanation (24,25,30,31,34). This distinction was retained during synthesis so that evidence derived from general medical AI was not interpreted as though it arose directly from diabetic retinopathy screening.

Overall, the findings supported a pathway-oriented interpretation of patient experience. AI-assisted screening was most positively experienced when the technology operated as one component of an

understandable and professionally supported clinical process. The evidence was substantially weaker for models in which favourable satisfaction was interpreted as sufficient evidence of informed trust, meaningful consent, or successful care completion. The synthesis therefore identified human accountability, intelligible disclosure, understandable result communication, and continuity from screening to referral as the central patient-facing conditions emerging from the available evidence.

DISCUSSION

This critical integrative review indicates that positive patient responses to AI-assisted diabetic retinopathy screening should not be interpreted as evidence of unconditional trust, fully informed consent, or adequate understanding of screening results. Across the 15-study findings corpus, three closely connected patterns were evident. First, patients generally appeared more comfortable with AI when it functioned as an adjunct to clinicians rather than as an autonomous replacement for professional judgment. Second, evidence concerning willingness and satisfaction was considerably more developed than evidence examining comprehension, voluntariness, or understanding of secondary data use. Third, rapid and locally accessible screening was consistently valued, but its patient-centred benefit depended on whether results were understandable and connected to a feasible referral pathway. Together, these findings suggest that the acceptability of AI-assisted screening is determined less by the presence of automation itself than by the clinical, relational, and organizational arrangements surrounding its use.

The finding that trust was conditional and role-specific is consistent with broader evidence on medical AI. Glikson and Woolley distinguished confidence in AI performance from the social and relational cues through which trust is established, emphasizing that individuals often rely on surrounding institutions and professionals when they cannot independently judge algorithmic competence (8). Longoni et al. similarly showed that resistance to medical AI may arise when patients perceive algorithmic systems as insufficiently responsive to individual circumstances (9). Hatherley further argued that AI lacks the reciprocal moral agency associated with interpersonal trust, reinforcing the importance of human responsibility in clinical deployment (10). The present synthesis extends these arguments into diabetic retinopathy screening: patients could value automated efficiency while simultaneously expecting clinicians to remain available, retain oversight, and assume responsibility for consequential decisions. Positive perceptions of AI therefore reflected responsible delegation more consistently than replacement of professional authority.

This distinction is particularly important because screening systems are frequently evaluated using measures such as willingness, satisfaction, convenience, or intention to reuse the service. Early retinal-screening studies demonstrated the feasibility and acceptability of AI-supported models, while implementation research emphasized the operational advantages of bringing automated interpretation closer to the point of care (5,6). However, favourable encounter-level experience does not necessarily establish that patients understand how AI contributed to the result or how responsibility is distributed between the technology and healthcare professionals. The findings of Yap et al., Rustam et al., Fritsch et al., Lennartz et al., Wahlich et al., and Krogh et al. collectively support this distinction: patients often expressed favourable views while continuing to prefer physician involvement or professional safeguards (21,23,30-33). Consequently, implementation evaluations should avoid treating acceptance as a single construct and should distinguish willingness to use a service from trust in the algorithm, confidence in the institution, and willingness to delegate final clinical authority.

Consent emerged as the least empirically developed component of patient experience. Ethical frameworks for healthcare AI have repeatedly emphasized transparency, privacy, accountability, governance, and meaningful patient choice (11-14), yet the empirical literature identified in this review more often measured attitudes toward AI than the quality of consent itself. Park demonstrated that consent and disclosure framing can influence patient preferences, indicating that the manner in which AI involvement is communicated is ethically consequential rather than merely administrative (25). Aggarwal et al. further showed that willingness to share health data varies according to recipient, purpose, and safeguards, underscoring the distinction between agreeing to clinical imaging and permitting secondary use of

information for algorithm development or other purposes (24). These findings suggest that AI consent should not be reduced to notification that an algorithm is involved.

The review also identified an important gap between disclosure and comprehension. Malerbi et al. reported favourable attitudes toward AI-human collaboration despite limited familiarity with AI, while Baghdadi et al. found continued interest in professional communication among patients considering AI-supported diagnosis (26,34). Shah et al. reported high satisfaction and perceived improvement in knowledge following AI-based retinal screening, but immediate self-reported understanding cannot establish whether patients retained information about screening limitations, false classifications, ungradable images, human review, or subsequent management (35). This distinction is important because informed consent requires more than awareness of a technological label. Patients need sufficient practical information to understand what the system is doing, the degree of clinician involvement, the limits of the screening result, relevant data uses, and whether meaningful alternatives are available.

Practical explainability may therefore be more relevant to patient-centred screening than technical explainability alone. London described the tension between accuracy and explainability in black-box medical decision systems, while Ploug et al. demonstrated that individuals may make trade-offs between performance and explainability depending on context (17,18). Patients do not need to reconstruct a neural network's internal computation to participate meaningfully in screening. They do, however, need to understand what was assessed, whether the image was interpretable, what the reported category means, whether the result was reviewed by a healthcare professional, and what action is required. Explainability at the patient level is consequently better conceptualized as actionable clinical intelligibility than as detailed technical transparency.

The third theme demonstrates why patient experience must be evaluated across the entire screening-to-care pathway. Nolan et al. reported favourable experiences with AI screening delivered within rural primary care, while Bhambhwani et al. described positive patient experience and comparatively strong ophthalmology attendance among referred patients in Northern Ontario (27,29). In contrast, Whitestone et al. observed very high satisfaction with AI-based screening in Rwanda despite substantially lower referral adherence (28). This divergence is clinically important. A screening programme may be technically accurate, rapid, and well received while still failing to deliver its intended benefit if patients cannot complete subsequent evaluation or treatment.

AI may therefore remove one bottleneck while exposing another. Technical studies have demonstrated that automated retinal interpretation can expand screening capacity and reduce dependence on specialist graders (1-4), but identification of referable disease increases the importance of downstream ophthalmic capacity, transport, affordability, communication, and navigation. Grauslund highlighted the evolving role of AI within diabetic retinopathy screening systems, while Bellemo et al. demonstrated the potential value of automated screening where specialist capacity is limited (3,4). The present synthesis suggests that those advantages should be assessed in conjunction with the ability of healthcare systems to convert screening classifications into completed clinical action. Detection without an accessible referral pathway may improve throughput without proportionately improving patient outcomes.

These findings also have implications for equity. AI-assisted screening is frequently promoted as a means of extending services into underserved settings, but the benefits of decentralization may not be evenly distributed. Patients with lower health or digital literacy may experience greater difficulty understanding automated results, while those facing transportation costs, language barriers, limited specialist availability, or distrust of healthcare institutions may remain disadvantaged after the screening encounter. Beede et al. demonstrated that image quality, clinic workflow, staff behaviour, and local workarounds influence real-world AI deployment (5), and Kelly et al. similarly emphasized the importance of generalisability and human factors in translating AI performance into clinical impact (7). Patient-centred implementation must therefore address not only algorithmic performance across demographic groups but also whether communication, referral, and follow-up processes are accessible to the populations for whom decentralized screening is intended.

From a service-design perspective, the findings support a layered approach to patient communication. Initial disclosure should explain in accessible language that AI is being used, whether the system is assistive or autonomous, whether a clinician reviews the result, and who remains responsible for subsequent care. More detailed written or visual information may then address image storage, secondary data use, uncertainty, and options for questions or refusal. Direct clinical consent should be distinguished from permission for optional secondary data use. Result communication should explicitly differentiate screening from diagnosis, explain ungradable images and uncertainty, identify the next required action, and provide a named clinical contact or referral mechanism. These recommendations arise from the recurrent gaps identified across the synthesis rather than from an assumption that a single disclosure model will be appropriate in every healthcare context.

Governance arrangements should similarly preserve visible lines of responsibility. The findings do not support a model in which accountability becomes diffuse because an automated system generated the classification. Healthcare organizations introducing AI-assisted retinal screening should be able to identify who is responsible for system selection, performance monitoring, interpretation of unexpected results, handling of errors, communication with patients, and ensuring referral completion. Vendor responsibility, clinical responsibility, and institutional governance may differ, but those boundaries should not become invisible to patients. Procurement and monitoring should therefore extend beyond diagnostic performance to include subgroup performance, error review, comprehension, referral completion, and patient access to human review.

Several limitations of the evidence base should temper interpretation. First, the included studies were methodologically heterogeneous and examined different forms of AI exposure, ranging from hypothetical attitudes and questionnaire-based assessments to actual participation in AI-assisted screening. These study types cannot be assumed to measure the same construct. Hypothetical willingness may not predict behaviour, while immediate post-screening satisfaction may reflect convenience, politeness, relief, or confidence in the healthcare setting rather than informed trust in AI. Second, direct evidence on consent quality, retained comprehension, voluntariness, error experiences, and long-term responses to screening remained sparse compared with evidence on acceptance and satisfaction. Third, some patient-centred domains required evidence from diagnostic settings outside diabetic retinopathy, which improves conceptual breadth but limits condition-specific inference. Fourth, convenience sampling and the underrepresentation of people who declined screening, received ungradable results, failed to attend referral, or disengaged from care may have biased the available evidence toward more favourable patient experiences.

The review itself also has methodological limitations. As a critical integrative review, it was designed to synthesize heterogeneous patient-centred evidence rather than to provide the exhaustive study identification, dual independent screening, standardized risk-of-bias assessment, and statistical certainty estimation expected of a formal systematic review. The inclusion of English-language peer-reviewed literature may have excluded relevant evidence from other languages or unpublished implementation programmes. Selective inclusion of neighbouring diagnostic studies was conceptually justified where retinal-specific evidence was limited, but inevitably introduced clinical heterogeneity. The thematic framework also used trust, consent, and understanding as sensitising concepts, which strengthened analytical coherence but may have foregrounded these dimensions over other patient-experience constructs. These limitations mean that the themes should be interpreted as a critical synthesis of the available evidence rather than as prevalence estimates or exhaustive measures of patient opinion.

Future research should therefore move beyond general acceptability surveys. Prospective implementation studies should test whether patients can accurately explain the role of AI, distinguish screening from diagnosis, understand ungradable or uncertain results, identify the responsible clinician, and describe what happens to their data. Studies should assess voluntariness and document whether patients can realistically decline AI-assisted screening or optional secondary data use without losing access to care. Longitudinal follow-up should examine referral attendance, treatment initiation, responses to false or discordant results, and whether trust changes after adverse or uncertain experiences. Research should

intentionally include individuals with limited health literacy, those who decline screening, patients with ungradable images, and those who fail to complete referral, as these populations may reveal implementation problems that conventional satisfaction surveys systematically overlook.

Overall, the evidence indicates that patient-centred AI-assisted diabetic retinopathy screening is fundamentally a sociotechnical care pathway rather than an isolated diagnostic technology. Algorithmic performance is necessary, but its practical legitimacy depends on visible professional accountability, meaningful disclosure and choice, understandable result communication, and a referral pathway capable of converting detection into care. Evaluation frameworks should therefore assess comprehension, voluntariness, responsibility, and downstream action alongside traditional measures of accuracy, throughput, and satisfaction.

CONCLUSION

Patients generally appear receptive to AI-assisted diabetic retinopathy screening when it improves accessibility, reduces delay, and remains embedded within a trusted clinical pathway. However, the available evidence shows that willingness and satisfaction should not be treated as proxies for informed trust, meaningful consent, or comprehension. Trust was strongest when human oversight and responsibility remained visible; consent evidence was comparatively underdeveloped and often limited to disclosure or stated preferences; and the value of rapid screening depended on whether results were understandable and linked to achievable referral and treatment. Patient-centred implementation should therefore combine technical performance with clear accountability, proportionate and intelligible consent, practical result explanation, protection of choice and data governance, and active support from screening through follow-up care. Future evaluations should prioritize comprehension, voluntariness, equity, error experience, and referral completion rather than relying predominantly on immediate acceptance or satisfaction.

REFERENCES

1. Abràmoff MD, Lavin PT, Birch M, Shah N, Folk JC. Pivotal trial of an autonomous AI-based diagnostic system for detection of diabetic retinopathy in primary care offices. NPJ Digit Med. 2018;1:39. doi:10.1038/s41746-018-0040-6.

2. Gulshan V, Peng L, Coram M, Stumpe MC, Wu D, Narayanaswamy A, et al. Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. JAMA. 2016;316(22):2402-2410. doi:10.1001/jama.2016.17216.

3. Bellemo V, Lim ZW, Lim G, Nguyen QD, Xie Y, Yip MYT, et al. Artificial intelligence using deep learning to screen for referable and vision-threatening diabetic retinopathy in Africa: a clinical validation study. Lancet Digit Health. 2019;1(1). doi:10.1016/S2589-7500(19)30004-4.

4. Grauslund J. Diabetic retinopathy screening in the emerging era of artificial intelligence. Diabetologia. 2022;65:1415-1423. doi:10.1007/s00125-022-05727-0.

5. Beede E, Baylor E, Hersch F, Iurchenko A, Wilcox L, Ruamviboonsuk P, et al. A human-centered evaluation of a deep learning system deployed in clinics for the detection of diabetic retinopathy. Proc 2020 CHI Conf Hum Factors Comput Syst. 2020:1-12. doi:10.1145/3313831.3376718.

6. Keel S, Lee PY, Scheetz J, Li Z, Kotowicz MA, MacIsaac RJ, et al. Feasibility and patient acceptability of a novel artificial intelligence-based screening model for diabetic retinopathy at endocrinology outpatient services: a pilot study. Sci Rep. 2018;8:4330. doi:10.1038/s41598-018-22612-2.

7. Kelly CJ, Karthikesalingam A, Suleyman M, Corrado G, King D. Key challenges for delivering clinical impact with artificial intelligence. BMC Med. 2019;17:195. doi:10.1186/s12916-019-1426-2.

8. Glikson E, Woolley AW. Human trust in artificial intelligence: review of empirical research. Acad Manag Ann. 2020;14(2):627-660. doi:10.5465/annals.2018.0057.

9. Longoni C, Bonezzi A, Morewedge CK. Resistance to medical artificial intelligence. J Consum Res. 2019;46(4):629-650. doi:10.1093/jcr/ucz013.

10. Hatherley JJ. Limits of trust in medical AI. J Med Ethics. 2020;46(7):478-481. doi:10.1136/medethics-2019-105935.

11. Vayena E, Blasimme A, Cohen IG. Machine learning in medicine: addressing ethical challenges. PLoS Med. 2018;15(11). doi:10.1371/journal.pmed.1002689.

12. Grote T, Berens P. On the ethics of algorithmic decision-making in healthcare. J Med Ethics. 2020;46(3):205-211. doi:10.1136/medethics-2019-105586.

13. Morley J, Floridi L, Kinsey L, Elhalal A. From what to how: an initial review of publicly available AI ethics tools, methods and research to translate principles into practices. Sci Eng Ethics. 2020;26:2141-2168. doi:10.1007/s11948-019-00165-5.

14. Reddy S, Allan S, Coghlan S, Cooper P. A governance model for the application of AI in health care. J Am Med Inform Assoc. 2020;27(3):491-497. doi:10.1093/jamia/ocz192.

15. Ongena YP, Haan M, Yakar D, Kwee TC. Patients' views on the implementation of artificial intelligence in radiology: development and validation of a standardized questionnaire. Eur Radiol. 2020;30:1033-1040. doi:10.1007/s00330-019-06486-0.

16. Esmaeilzadeh P. Use of AI-based tools for healthcare purposes: a survey study from consumers' perspectives. BMC Med Inform Decis Mak. 2020;20:170. doi:10.1186/s12911-020-01191-1.

17. Ploug T, Sundby A, Moeslund TB, Holm S. Population preferences for performance and explainability of artificial intelligence in health care: choice-based conjoint survey. J Med Internet Res. 2021;23(12). doi:10.2196/26611.

18. London AJ. Artificial intelligence and black-box medical decisions: accuracy versus explainability. Hastings Cent Rep. 2019;49(1):15-21. doi:10.1002/hast.973.

19. Snyder H. Literature review as a research methodology: an overview and guidelines. J Bus Res. 2019;104:333-339. doi:10.1016/j.jbusres.2019.07.039.

20. Braun V, Clarke V. Conceptual and design thinking for thematic analysis. Qual Psychol. 2022;9(1):3-26. doi:10.1037/qup0000196.

21. Yap A, Wilkinson B, Chen E, Han L, Vaghefi E, Galloway C, et al. Patients' perceptions of artificial intelligence in diabetic eye screening. Asia Pac J Ophthalmol (Phila). 2022;11(3):287-293. doi:10.1097/APO.0000000000000525.

22. Jin L, Tao Y, Liu Y, Liu G, Lin L, Chen Z, et al. SEM model analysis of diabetic patients' acceptance of artificial intelligence for diabetic retinopathy. BMC Med Inform Decis Mak. 2025;25:175. doi:10.1186/s12911-025-03008-5.

23. Rustam Z, Xie Y, Amezcua Moreno J, Tran D, Zhu G, Yu SE, et al. Patient perspectives on artificial intelligence-based diabetic retinopathy screening at an urban US medical center. Clin Ophthalmol. 2026;20:581564. doi:10.2147/OPTH.S581564.

24. Aggarwal R, Farag S, Martin G, Ashrafian H, Darzi A. Patient perceptions on data sharing and applying artificial intelligence to health care data: cross-sectional survey. J Med Internet Res. 2021;23(8). doi:10.2196/26162.

25. Park HJ. Patient perspectives on informed consent for medical AI: a web-based experiment. Digit Health. 2024;10:20552076241247938. doi:10.1177/20552076241247938.

26. Malerbi FK, Ventura BM, Fischer M, Penha FM. Patients' perceptions of artificial intelligence in a deep learning-assisted diabetic retinopathy screening event: a real-world assessment. J Diabetes Sci Technol. 2024;18(3):750-751. doi:10.1177/19322968241234378.

27. Nolan BM, DayBranch ER, Barton K, Korsen N. Patient and provider experience with artificial intelligence screening technology for diabetic retinopathy in a rural primary care setting. J Maine Med Cent. 2023;5(2):2. doi:10.46804/2641-2225.1144.

28. Whitestone N, Nkurikiye J, Patnaik JL, Jaccard N, Lanouette G, Cherwek DH, et al. Feasibility and acceptance of artificial intelligence-based diabetic retinopathy screening in Rwanda. Br J Ophthalmol. 2024;108(6):840-845. doi:10.1136/bjo-2022-322683.

29. Bhambhwani V, Whitestone N, Patnaik JL, Ojeda A, Scali J, Cherwek DH. Feasibility and patient experience of a pilot artificial intelligence-based diabetic retinopathy screening program in Northern Ontario. Ophthalmic Epidemiol. 2025;32(5):518-524. doi:10.1080/09286586.2024.2434738.

30. Fritsch SJ, Blankenheim A, Wahl A, Hetfeld P, Maassen O, Deffge S, et al. Attitudes and perception of artificial intelligence in healthcare: a cross-sectional survey among patients. Digit Health. 2022;8:20552076221116772. doi:10.1177/20552076221116772.

31. Lennartz S, Dratsch T, Zopfs D, Persigehl T, Maintz D, Große Hokamp N, et al. Use and control of artificial intelligence in patients across the medical workflow: single-center questionnaire study of patient perspectives. J Med Internet Res. 2021;23(2). doi:10.2196/24221.

32. Wahlich C, Chandrasekaran L, Chaudhry UAR, Willis K, Chambers R, Bolter L, et al. Patient and practitioner perceptions around use of artificial intelligence within the English NHS diabetic eye screening programme. Diabetes Res Clin Pract. 2025;219:111964. doi:10.1016/j.diabres.2024.111964.

33. Krogh M, Germund Nielsen M, Byskov Petersen G, Jensen MSA, Jensen MB, Vorum H, et al. Patient acceptance of AI-assisted diabetic retinopathy screening in primary care: findings from a questionnaire-based feasibility study. Front Med (Lausanne). 2025;12:1610114. doi:10.3389/fmed.2025.1610114.

34. Baghdadi LR, Mobeirek AA, Alhudaithi DR, Albenmousa FA, Alhadlaq LS, Alaql MS, et al. Patients' attitudes toward the use of artificial intelligence as a diagnostic tool in radiology in Saudi Arabia: cross-sectional study. JMIR Hum Factors. 2024;11. doi:10.2196/53108.

35. Shah P, Mishra D, Shanmugam M, Vighnesh MJ, Jayaraj H. Acceptability of artificial intelligence-based retina screening in the general population. Indian J Ophthalmol. 2022;70(4):1140-1144. doi:10.4103/ijo.IJO_1840_21.