Website: http: jebhpme.ssu.ac.ir
EISSN: 2538-4716
Evid Based Health Policy Manag Econ 2026; 10: 42-68
doi: 10.18502/jebhpme.v10i6.23195

SYSTEMATIC REVIEW

Cost-Effectiveness of AI-Assisted Diagnostic Decision Support: Systematic Review

Salman Abbasian-Naghneh

Department of Finance, Na.C., Islamic Azad University, Najafabad, Iran

Corresponding Author: Salman Abbasian-Naghneh, Email: Salman_abasian@iau.ir, Tel: +98 913 0779427

Received: 25 June 2026 **Revised:** 2 September 2026 **Accepted:** 18 September 2026

A B S T R A C T

Introduction: Artificial Intelligence (AI)-assisted Clinical Decision Support Systems (CDSS) are increasingly integrated into diagnostic care, but their economic value remains uncertain. This systematic review aimed to evaluate the cost-effectiveness of AI-assisted CDSS in diagnostic settings and identify factors influencing their economic outcomes.

Materials and Methods: Primary economic evaluations of AI-assisted CDSS used for diagnosis, screening, or clinical triage were identified through searches of PubMed/MEDLINE, Embase, Scopus, Web of Science, the Cochrane Library, the NHS Economic Evaluation Database archive, and Google Scholar up to 15 June 2026. Reporting was guided by PRISMA 2020. Methodological appraisal used the Philips checklist for model-based evaluations and the Consensus Health Economic Criteria for empirical evaluations, while reporting quality was assessed using CHEERS 2022. A separate formal risk-of-bias assessment was not conducted. Study selection, data extraction, and appraisal were performed by one reviewer, with subsequent verification. Because of substantial clinical and methodological heterogeneity, findings were synthesized narratively. The review protocol was not prospectively registered.

Results: Nine primary economic evaluations published between 2021 and 2026 were included, covering cancer screening, ophthalmological screening, multispecialty diagnostic support, and emergency triage. Five studies identified at least one economically dominant AI-assisted strategy. Favorable economic outcomes were most frequently associated with reduced diagnostic workload, replacement of existing workflow components, and prevention of downstream healthcare costs. Other interventions increased expenditure without sufficient clinical benefit. Economic outcomes were sensitive to AI acquisition and maintenance costs, diagnostic performance, disease prevalence, implementation strategy, and healthcare perspective.

Conclusion: AI-assisted CDSS may be cost-effective in selected diagnostic settings, but economic value depends on clinical context and implementation conditions. The evidence remains limited by heterogeneous study designs, predominantly model-based evaluations, restricted real-world validation, and single-reviewer assessment. The absence of prospective protocol registration and complete original screening records further limits methodological transparency.

Keywords: Artificial Intelligence; Biomedical; Clinical; Cost-Benefit Analysis; Decision Support Systems; Technology Assessment

Citation: Abbasian-Naghneh S. **Cost-Effectiveness of AI-Assisted Diagnostic Decision Support: Systematic Review.** Evidence Based Health Policy, Management & Economics 2026; 10: 42-68.

[Downloaded from jebhpme.ssu.ac.ir on 2026-11-12]

[DOI: 10.18502/jebhpme.v10i6.23195]

Copyright: ©2026 The Author(s); Published by Shahid Sadoughi University of Medical Sciences. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/ licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Introduction

Diagnostic errors remain an important source of preventable patient harm and inefficient healthcare expenditure. They may lead to delayed or inappropriate treatment, unnecessary diagnostic procedures, avoidable hospital admissions, disease progression, and increased mortality. In addition to their clinical consequences, diagnostic errors impose a substantial economic burden through increased healthcare resource utilization, avoidable expenditures, and malpractice-related costs (1–3). Improving diagnostic accuracy is therefore relevant not only to patient safety but also to the long-term sustainability of healthcare systems.

Artificial intelligence-assisted clinical decision support systems (AI-assisted CDSS) have emerged as potentially valuable tools for improving diagnostic processes. These systems apply machine learning, deep learning, computer vision, natural language processing, and hybrid computational methods to assist clinicians with disease detection, diagnostic classification, medical-image interpretation, risk stratification, population screening, and clinical triage. AI-based diagnostic technologies have been developed across multiple clinical specialties, including radiology, pathology, cardiology, ophthalmology, dermatology, oncology, gastroenterology, and emergency medicine (4-6).

For this review, AI-assisted clinical decision support was operationally defined to include AI-based technologies that directly inform or perform diagnostic tasks within clinical care pathways. This definition encompassed clinician-assistive systems, including computer-aided detection and diagnostic recommendations, as well as autonomous AI systems used to replace components of established diagnostic workflows. These configurations were distinguished in the interpretation of economic findings according to clinical oversight, workflow integration, and resource use. However, this broad operational definition should not be interpreted as implying that all autonomous image-analysis technologies meet regulatory definitions of clinical decision support software.

The clinical value of an AI-assisted diagnostic system cannot be determined solely by its technical accuracy. A system with high sensitivity or specificity may provide limited practical value if it creates excessive false-positive findings, increases clinician workload, requires substantial infrastructure investment, or fails to improve patient-relevant outcomes. Conversely, AI may provide meaningful clinical and economic benefits when it reduces unnecessary diagnostic procedures, improves patient throughput, supports earlier disease detection, decreases specialist workload, or replaces part of an existing diagnostic pathway.

The implementation of AI-assisted CDSS may require expenditure on software acquisition or licensing, computing infrastructure, integration with electronic health records, local validation, cybersecurity, maintenance, staff training, regulatory compliance, and continuing performance monitoring. Economic evaluation is therefore necessary to determine whether the expected clinical benefits justify the additional implementation and operating costs. Moreover, cost-effectiveness analysis, cost-utility analysis, cost-benefit analysis, cost-consequence analysis, cost-offset analysis, and budget-impact analysis provide different but complementary forms of evidence for healthcare policymakers, payers, hospital administrators, and technology-procurement committees.

Primary economic evaluations have examined AI-assisted diagnostic technologies in melanoma detection, dental-caries assessment, diabetic-retinopathy screening, colorectal-cancer screening, hepatocellular-carcinoma detection, breast-cancer screening, population-based glaucoma screening, and emergency-department triage (7-15). Several studies have reported cost savings or favorable incremental cost-effectiveness ratios, whereas others have shown that economic value depends strongly on the price of the AI system, disease prevalence, diagnostic specificity, clinician-adoption rates, downstream treatment costs, and the extent to which AI substitutes for rather than supplements existing clinical resources.

Previous systematic reviews have examined the

[Downloaded from jebhpme.ssu.ac.ir on 2026-11-12]

[DOI: 10.18502/jebhpme.v10i6.23195]

economic implications of artificial intelligence across different healthcare applications, highlighting potentially favorable outcomes alongside substantial methodological and contextual heterogeneity (4–6). However, these broader findings do not fully clarify how the economic value of AI-assisted diagnostic technologies varies according to their clinical role, implementation strategy, and integration into existing diagnostic workflows. The present review addresses this issue by focusing on primary economic evaluations of diagnostic applications, distinguishing cost-saving, economically dominant, cost-effective, and unfavorable strategies, and examining the influence of workflow substitution, healthcare perspective, and implementation costs on economic outcomes. Such heterogeneity limits direct comparison between studies and generally prevents meaningful statistical pooling of incremental cost-effectiveness ratios, costs, or quality-adjusted life years. In particular, an ICER calculated for one healthcare jurisdiction cannot be interpreted independently of its local prices, clinical pathways, reimbursement arrangements, and decision threshold.

Previous systematic reviews have examined the economic implications of AI in healthcare, although their scope and evidence base differ. El Arab and Al Moosa (4) reviewed 19 studies across healthcare applications, Wu et al. (5) included 52 economic evaluations across healthcare interventions, and Brin and Tau (6) focused on 10 studies in radiology. In contrast, the present review specifically examines primary economic evaluations of AI-based diagnostic and screening technologies, incorporating evidence published through June 2026. Its contribution lies in distinguishing economically dominant, cost-saving, and cost-effective strategies and identifying how clinical workflow integration, implementation costs, and healthcare perspectives influence economic outcomes.

An additional limitation of the existing literature is that many evaluations are based on decision-analytic models rather than prospective real-world implementation studies. Model-based studies can estimate long-term costs and health outcomes, but their conclusions may be sensitive to assumptions about AI performance, implementation costs, disease progression, clinician behavior, and the durability of diagnostic improvements. Evidence from low- and middle-income countries also remains limited, reducing the generalizability of findings across healthcare systems.

Accordingly, the present review focused exclusively on primary economic evaluations of AI-assisted CDSS used in diagnostic care. Secondary reviews, narrative articles, and other non-original publications were used only for background information and citation tracking. Because the eligible studies evaluated different diseases, interventions, comparators, economic outcomes, and healthcare systems, their findings were synthesized using a structured narrative approach rather than a statistically pooled estimate.

The objective of this systematic review was to identify and critically summarize primary economic evaluations of AI-assisted CDSS in diagnostic care. The review examined the clinical settings, patient populations, AI interventions, comparators, economic perspectives, analytical approaches, costs, health outcomes, ICERs, cost savings, and budgetary consequences reported in the included studies. It also assessed reporting quality and methodological limitations, classified the economic findings using explicit criteria, and identified the principal factors influencing the economic value of AI-assisted diagnostic technologies.

Materials and Methods

Review Design and Protocol

This review was reported with reference to PRISMA 2020 (16). Eligibility criteria, information sources, search concepts, data items, appraisal methods, and the narrative synthesis approach are specified below. Available study-selection information is reported separately from database-level search counts, which could not be independently verified.

The review was not prospectively registered, and

[Downloaded from jebhpme.ssu.ac.ir on 2026-11-12]

[DOI: 10.18502/jebhpme.v10i6.23195]

CC BY 4.0

no dated preregistered protocol was available. Accordingly, the review cannot be described as prospectively protocol-driven. The eligibility framework applied to the reported evidence synthesis is described below, and the date and sequence of any earlier eligibility refinements cannot be independently established from the available documentation. Such uncertainty may introduce selective-inclusion bias and is acknowledged as a limitation.

Eligibility Criteria

Eligibility was determined using the population, intervention, comparator, outcomes, and study design framework. Eligible populations included patients undergoing diagnostic investigation, disease detection, risk stratification, clinical triage, or population-based screening in primary care, specialist clinics, hospitals, emergency departments, or community screening programs. Simulated patient cohorts used in decision-analytic models were also eligible when the model population and clinical pathway were adequately described.

No restriction on publication year was applied to the eligibility criteria. Studies were considered based on their relevance to AI-assisted diagnostic and screening pathways, the availability of original economic evaluations, and their compliance with the specified population, intervention, comparator, outcome, and study-design criteria. Therefore, the publication period of the included studies should not be interpreted as a predefined temporal restriction.

Eligible interventions were AI-assisted CDSS designed to support diagnostic care. These included machine-learning, deep-learning, computer-vision, natural-language-processing, and hybrid AI systems used for disease detection, diagnostic classification, medical-image interpretation, screening, risk prediction, or triage. The AI system could function as an independent reader, an additional reader, a replacement for part of an existing workflow, or a companion tool used alongside clinicians. Eligible comparators included standard clinical practice without AI assistance, clinician-only interpretation, conventional diagnostic or screening pathways, no-screening strategies, alternative diagnostic technologies, and non-AI decision-support approaches.

Both clinician-assistive and autonomous AI diagnostic systems were eligible when their outputs were incorporated into a defined clinical diagnostic or screening pathway, and an original economic evaluation was available. Moreover, AI systems used exclusively for technical performance benchmarking, without an evaluated clinical pathway or economic outcome, were excluded. Eligibility of autonomous image-reading systems was determined by their evaluated role in the clinical pathway rather than by a presumption that they met a narrower regulatory definition of clinical decision support software.

Studies were required to report at least one original economic outcome. Eligible outcomes included total or incremental costs, quality-adjusted life years, life-years gained, incremental cost-effectiveness ratios, cost per clinical event prevented, cost savings, cost offsets, net monetary benefit, budget impact, healthcare-resource use, implementation costs, revenue, operating expenditure, or operating margin. Clinical outcomes were extracted when they formed part of the economic evaluation, including diagnostic accuracy, false-positive and false-negative findings, disease cases detected or prevented, mortality, disease progression, clinician workload, interpretation time, and patient throughput.

Primary full or partial economic evaluations were eligible, including cost-effectiveness analyses, cost-utility analyses, cost-benefit analyses, cost-consequence analyses, cost-offset analyses, budget-impact analyses, and early health technology assessments. Both model-based and empirical evaluations using prospective, retrospective, trial-based, or routinely collected clinical data were also considered.

Systematic reviews, scoping reviews, narrative reviews, general literature reviews, book chapters, editorials, commentaries, letters, study protocols, and conference abstracts without sufficient economic data were excluded. Furthermore, purely technical AI-performance studies without an

[Downloaded from jebhpme.ssu.ac.ir on 2026-11-12]

[DOI: 10.18502/jebhpme.v10i6.23195]

CC BY 4.0

original economic component were excluded. Secondary reviews were used only to identify potentially relevant primary studies and to provide contextual background; they were not counted as included economic evaluations.

Information Sources and Search Strategy

The literature search covered PubMed/MEDLINE, Embase, Scopus, Web of Science, the Cochrane Library, Google Scholar, and the NHS Economic Evaluation Database (NHS EED) archive, which was consulted only as a supplementary historical source because it had not been updated since March 2015. Reference lists of eligible studies and relevant systematic reviews were also examined to identify additional primary evaluations. Relevant health technology assessment reports and grey-literature sources were considered when they contained sufficient methodological and numerical information for appraisal.

The search strategy combined controlled vocabulary and free-text terms relating to artificial intelligence, clinical decision support, diagnostic care, and economic evaluation. The principal concepts included "artificial intelligence," "machine learning," "deep learning," "computer vision," "clinical decision support system," "cost-effectiveness," "economic evaluation," "cost utility," "budget impact," "incremental cost-effectiveness ratio," "diagnosis," "screening," "detection," and "triage." Medical Subject Headings were used in PubMed/MEDLINE, and Emtree terms were used in Embase. The search strategy was adapted to the syntax and indexing system of each database. The main database searches were completed on 15 June 2026. The available database-specific search strategies are presented in Supplementary Appendix 2. Limitations in the documentation and sensitivity of the original searches are acknowledged.

No language filter was applied during the electronic searches. However, because reliable translation resources were unavailable, full-text eligibility assessment was restricted to reports available in English. Complete database-specific search strategies, including field tags, controlled vocabulary, Boolean operators, search dates, and restrictions, are provided in Supplementary Appendix 2.

Only studies with full-text reports available in English were eligible for inclusion because reliable translation resources were unavailable. No language restrictions were applied during database searching; however, the English-language requirement was applied during full-text eligibility assessment.

Study Selection and Data Extraction

All stages of study selection were conducted by the sole author. Retrieved records were managed using Rayyan, and the available records were assessed against the eligibility criteria described above. Titles and abstracts were screened, followed by full-text assessment of potentially relevant publications. To reduce the possibility of selection errors, inclusion decisions and available exclusion information were re-examined during a separate verification pass conducted by the same author. This verification procedure was intended as an additional consistency check and was not considered equivalent to independent duplicate screening. The absence of a second independent reviewer is acknowledged as a methodological limitation. Because the complete original screening records were not retained, Figure 1 presents the documented candidate-publication assessment rather than a complete PRISMA 2020 flow diagram.

Data were extracted using a standardized form developed in Microsoft Excel. Extracted information included publication year, country, clinical setting, target population, sample size or modeled cohort size, AI technology, diagnostic application, comparator, type of economic evaluation, economic perspective, analytical model, time horizon, discount rate, costing year, currency, willingness-to-pay threshold, implementation and operating costs, clinical outcomes, incremental costs, incremental effectiveness, QALYs, life-years, ICERs, cost savings, budgetary outcomes, and sensitivity-analysis results. The data-extraction framework is presented in Appendix 3

[Downloaded from jebhpme.ssu.ac.ir on 2026-11-12]

[DOI: 10.18502/jebhpme.v10i6.23195]

CC BY 4.0

(Supplementary Table S3), while study-specific economic parameters, including costing years, currencies, discount rates, and willingness-to-pay thresholds, are summarized in Appendix 4 (Supplementary Table S4). Each extracted entry was checked against the corresponding full-text article, tables, figures, and supplementary materials during a separate verification stage. When multiple publications described the same dataset or economic model, they were linked and treated as a single evaluation to prevent double counting. No unreported sample sizes, clinical events, costs, QALYs, or ICERs were imputed.

Reporting Quality and Methodological Appraisal

Reporting quality and methodological quality were assessed separately. The Consolidated Health Economic Evaluation Reporting Standards (CHEERS) 2022 checklist was used to evaluate the completeness and transparency of economic reporting (17). CHEERS was treated as a reporting standard and was not interpreted as a direct risk-of-bias instrument. Model-based economic evaluations were appraised using the Philips checklist (18). The assessment considered the appropriateness of the model structure, justification of assumptions, selection of input data, internal consistency, external validation, treatment of uncertainty, and transparency of model implementation.

Trial-based, observational, and other empirical economic evaluations were assessed using the Consensus Health Economic Criteria list (19). The appraisal considered the clarity of the research question, appropriateness of the comparator, identification and measurement of relevant costs and outcomes, time horizon, discounting, incremental analysis, uncertainty analysis, and generalizability. The Cochrane Risk of Bias 2 tool was not used because it was developed primarily for the assessment of randomized clinical-trial outcomes rather than for evaluating the methodological quality of health economic studies. All appraisal decisions were completed by the author and checked during a separate verification pass. Findings were also reported at the study level, with emphasis on specific strengths and limitations rather than unsupported overall labels such as low, moderate, or high risk. The study-level appraisal findings were summarized qualitatively in Table 2, with emphasis on the principal methodological strengths and limitations rather than numerical quality scores.

Data Synthesis and Economic Classification

Study characteristics and outcomes were organized in structured evidence tables and synthesized narratively. Studies were compared according to clinical application, AI function, comparator, economic design, perspective, time horizon, outcome measure, and principal economic conclusion.

A quantitative meta-analysis was not performed because the included studies differed substantially in patient populations, clinical conditions, AI technologies, comparators, currencies, costing years, perspectives, model structures, time horizons, and measures of effectiveness. Clinical outcomes also varied widely and included QALYs, life-years, disease cases prevented, diagnostic accuracy, blindness, clinician workload, interpretation time, and patient throughput. These outcomes could not be converted into a common effect measure without unsupported assumptions. In addition, economic outcomes were not statistically pooled because ICERs are context-dependent ratios determined by local costs, healthcare pathways, effectiveness measures, and willingness-to-pay thresholds. Calculating a pooled mean ICER or pooled percentage cost reduction would therefore have produced a potentially misleading estimate.

Incremental cost was defined as:

$$\Delta C = C_{\mathrm{AI}} - C_{\mathrm{comparator}}$$

Incremental effectiveness was defined as:

$$\Delta E = E_{\mathrm{AI}} - E_{\mathrm{comparator}}$$

The incremental cost-effectiveness ratio was interpreted as:

$$ICER = \frac{\Delta C}{\Delta E} = \frac{C_{\mathrm{AI}} - C_{\mathrm{comparator}}}{E_{\mathrm{AI}} - E_{\mathrm{comparator}}}$$

[Downloaded from jebhpme.ssu.ac.ir on 2026-11-12]

[DOI: 10.18502/jebhpme.v10i6.23195]

An AI-assisted intervention was classified as dominant when it was both less costly and more effective than the comparator ($\Delta C < 0$ and $\Delta E > 0$), and as dominated when it was more costly and less effective ($\Delta C > 0$ and $\Delta E < 0$). An intervention was classified as cost-saving when it reduced costs but greater clinical effectiveness was absent, unclear, or not measured. Interventions that increased both costs and effectiveness were considered cost-effective when their ICERs were below the relevant jurisdiction-specific willingness-to-pay threshold and not cost-effective when they exceeded that threshold. Interventions that reduced both costs and effectiveness (southwest quadrant) were evaluated by considering whether the cost savings justified the loss of effectiveness at the applicable willingness-to-pay threshold. Negative ICERs were not interpreted independently of the direction of incremental costs and effects. Findings were classified as uncertain when conclusions depended heavily on inadequately reported data or unstable assumptions.

Costs were reported in their original currencies and price years. Cross-study pooling was not undertaken because currency conversion alone could not account for differences in healthcare-system structure, reimbursement, clinical pathways, economic perspective, and analytical assumptions. Because no pooled statistical outcome was produced, subgroup meta-analysis, meta-regression, funnel-plot analysis, and Egger's regression test were not performed. Potential reporting bias and methodological heterogeneity were considered qualitatively when interpreting the findings.

Results

Study Selection and Available Selection Records

A documented set of 22 candidate publications was assessed for eligibility as primary economic evaluations of AI-based diagnostic interventions. Eight publications met the eligibility requirements; 13 were secondary or other non-primary reports, and one lacked sufficiently verifiable original economic results. One additional eligible primary evaluation was subsequently identified through supplementary literature assessment, yielding nine included evaluations. Secondary reviews were used for contextual interpretation and reference checking but were not counted as primary studies. These figures describe the available candidate-publication assessment, not the complete database-to-inclusion flow. Source-level record exports, duplicate-removal counts, and the original title-and-abstract screening log were unavailable for verification; therefore, a complete PRISMA 2020 flow diagram could not be reliably reconstructed. Figure 1 depicts only the documented candidate-publication assessment.

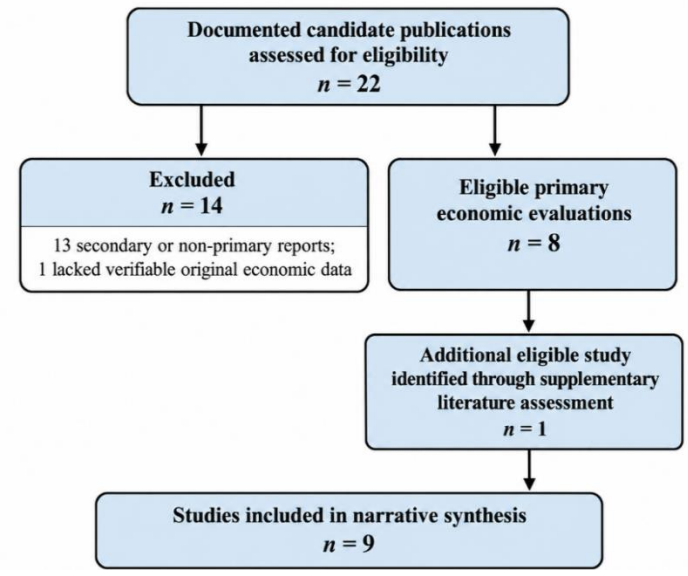


Figure 1. Documented Candidate-Publication Assessment and Eligibility Decisions.

[Downloaded from jebhpme.ssu.ac.ir on 2026-11-12]

[DOI: 10.18502/jebhpme.v10i6.23195]

 CC BY 4.0

Characteristics of the Included Studies

The 9 included studies were published between 2021 and 2026. Seven studies used decision-analytic economic models, one used a cost-consequence analysis based on several clinical pilot applications, and one combined real-world preimplementation and postimplementation data with economic modeling. Three studies evaluated AI in breast cancer screening. Two studies examined multiple diagnostic applications, and the remaining studies evaluated AI-assisted colorectal cancer screening, hepatocellular carcinoma detection, glaucoma screening, and emergency-department triage.

The studies represented healthcare settings in the United States, Germany, Brazil, China, Hong Kong (China), Italy, Sweden, Singapore, the United Kingdom, and several European clinical pilot sites. The economic evaluation by Wong et al. (10) modeled a hypothetical Asian screening population using epidemiological and screening data primarily from Hong Kong. Economic perspectives included healthcare-system, payer, hospital-management, public-policy, and national health-service perspectives. Time horizons ranged from two 180-day observation periods to lifetime modeling. The complete characteristics of the included studies are presented in Table 1.

Reporting Quality and Methodological Appraisal

The included studies generally reported their objectives, target populations, comparators, economic perspectives, principal assumptions, and base-case findings. Reporting completeness varied across the included studies, particularly regarding model assumptions, validation, uncertainty analysis, and economic reporting. The principal methodological limitations were dependence on modeled rather than directly observed long-term outcomes, uncertainty regarding AI licensing and implementation costs, limited external validation, and assumptions that improvements in diagnostic accuracy would translate directly into downstream clinical benefits.

The cost-consequence evaluation reported heterogeneous outcomes separately and did not convert all clinical consequences into QALYs or a common ICER. The emergency-triage evaluation used a large real-world dataset but employed a non-randomized preimplementation–postimplementation design and measured financial and operational outcomes rather than health utility. No overall low-, moderate-, or high-risk labels were assigned because such categories were not predefined by the appraisal instruments. CHEERS was interpreted as a reporting guideline and not as a direct measure of risk of bias. The main strengths and limitations of each study are summarized in Table 2.

Economic Outcomes

The economic outcomes were synthesized narratively because the studies differed substantially in clinical setting, intervention, comparator, economic perspective, currency, costing year, time horizon, and outcome measure. Outcomes included QALYs, life-years saved, tooth-retention years, disease cases prevented, cost offsets, operating margins, and cost consequences. These outcomes could not be meaningfully combined into a single pooled estimate. Table 3 presents the base-case clinical and economic findings of all included studies. Additional information on the costing years, currencies, discount rates, and willingness-to-pay thresholds of the included evaluations is provided in Appendix 4 (Supplementary Table S4).

[Downloaded from jebhpme.ssu.ac.ir on 2026-11-12]

[DOI: 10.18502/jebhpme.v10i6.23195]

CC BY 4.0

Table 1. Characteristics of the Included Primary Economic Evaluations

Study	Country or setting	Population	AI-assisted intervention	Comparator	Economic design and perspective	Time horizon
Gomez Rossi et al., 2022 (7)	United States, Germany, and Brazil	Three simulated cohorts: adults undergoing melanoma assessment, children undergoing dental-caries assessment, and adults with diabetes undergoing retinopathy screening	AI-assisted detection of melanoma, dental caries, or diabetic retinopathy	Standard specialist diagnosis without AI support	Three Markov-model economic evaluations; payer perspectives	Lifetime
Xiao et al., 2021 (8)	Changjiang County, China	19,395 residents aged 65 years or older	AI-assisted population-based glaucoma screening	No organized screening	Markov cost-offset analysis; healthcare-system perspective	15 years
Chatzikou et al., 2025 (9)	European Hosmart AI pilot settings	Patients undergoing assessment in obstetrics, coronary CT angiography, echocardiography, or capsule endoscopy	Interpretable AI-assisted diagnostic systems	Existing diagnostic workflows	Cost-consequence analysis; perspective and horizon varied by clinical application	Application-specific
Wong et al., 2025 (10)	Hong Kong, China (hypothetical Asian screening population)	Simulated cohort of 100,000 average-risk adults aged 50–75 years	FIT followed by AI-assisted colonoscopy or direct AI-assisted colonoscopy	No screening, FIT followed by conventional colonoscopy, or direct conventional colonoscopy	Markov cost-effectiveness analysis; population-screening perspective	Screening between 50 and 75 years
Maas et al., 2025 (11)	Italy	Simulated cohort of 1,000 patients with cirrhosis at risk of hepatocellular carcinoma	AI-enhanced MRI detection of liver lesions	Usual diagnostic care	Decision tree followed by a Markov cost-utility model; Italian healthcare perspective	Lifetime
Lyth et al., 2026 (12)	Sweden	Simulated cohort of 1,000 women aged 40–74 years	AI-assisted digital mammography with risk-dependent human reading	Conventional double human reading	State-transition cost-utility model; healthcare-system perspective	Lifetime
Goh et al., 2026 (13)	Singapore	Simulated cohort of 10,000 women aged 50–69 years	One radiologist plus AI or AI-only mammography interpretation	Double reading by two radiologists	Early health technology assessment using a Markov model; healthcare-system perspective	50 years
Levin et al., 2026 (14)	United States; three emergency departments	170,723 emergency-department visits across preimplementation and postimplementation periods	AI-assisted triage clinical decision support	Practice before implementation	Multisite economic evaluation; hospital-management and public-policy perspectives	Two 180-day periods
Hill and Roadevin, 2026 (15)	United Kingdom	Simulated cohort of 100,000 women invited to breast screening at age 50 years	AI-only reading, one human reader plus AI, or two human readers plus AI	Standard double human reading	Discrete-event cost-utility model; NHS and Personal Social Services perspective	Lifetime

Table note: CT, computed tomography; FIT, fecal immunochemical test; MRI, magnetic resonance imaging; NHS, National Health Service. Additional study-specific economic parameters, including costing years, currencies, discount rates, and willingness-to-pay thresholds, are provided in Appendix 4 (Supplementary Table S4).

[Downloaded from jebhpme.ssu.ac.ir on 2026-11-12]

[DOI: 10.18502/jebhpme.v10i6.23195]

CC BY 4.0

Table 2. Methodological Strengths and Limitations of the Included Economic Evaluations

Study	Principal methodological strengths	Principal methodological limitations
Gomez Rossi et al., 2022 (7)	Evaluated three distinct diagnostic applications; used lifetime Markov models and probabilistic and deterministic sensitivity analyses	Adapted previously developed models; health gains were very small; findings were sensitive to AI fees and assumptions about treatment after diagnosis
Xiao et al., 2021 (8)	Used local screening and treatment costs; reported long-term budgetary consequences; evaluated parameter uncertainty	Restricted to primary angle-closure glaucoma; relied partly on externally derived progression data; did not estimate QALYs
Chatzikou et al., 2025 (9)	Evaluated AI in four clinical pilot applications; reported clinical and economic consequences separately	Economic perspective and time horizon were not uniform; no common health-utility outcome; long-term consequences were incompletely modeled
Wong et al., 2025 (10)	Evaluated several competing screening strategies in a large cohort; reported extensive sensitivity analyses	Used a hypothetical cohort; did not estimate QALYs; conclusions depended on Asian screening assumptions and adherence rates
Maas et al., 2025 (11)	Used a lifetime cost-utility model and incremental analysis; conducted one-way and two-way sensitivity analyses	Incremental QALY gain was small; results were sensitive to AI cost, sensitivity, and specificity
Lyth et al., 2026 (12)	Used a lifetime horizon, probabilistic sensitivity analysis, and several alternative implementation scenarios	Findings were specific to the Swedish screening pathway and depended on modeled long-term cancer outcomes and AI implementation costs
Goh et al., 2026 (13)	Used a long time horizon, deterministic and probabilistic sensitivity analyses, and alternative human–AI configurations	Early health technology assessment based on modeled implementation; specificity and false-positive results substantially influenced findings
Levin et al., 2026 (14)	Included a large multisite real-world dataset; compared two relevant economic perspectives; reported break-even prices	Non-randomized before-and-after design; no QALY outcome; temporal and institutional factors may have influenced throughput and revenue
Hill and Roadevin, 2026 (15)	Used prospective diagnostic-performance evidence, external model validation, probabilistic analysis, and detailed NHS costing	UK-specific model; incremental QALY gains were small; AI licensing and implementation costs involved uncertainty

[Downloaded from jebhpme.ssu.ac.ir on 2026-11-12]

[DOI: 10.18502/jebhpme.v10i6.23195]

Table 3. Principal Economic Outcomes of the Included Primary Evaluations

Study and application	Clinical outcome	Cost outcome	ICER or principal economic finding	Economic interpretation
Gomez Rossi et al. (7): melanoma detection	Approximately 86.6 QALYs with both strategies	AI: US$750.35; standard diagnosis: US$759.03	Lower costs with a minimal reported difference in effectiveness; the negative ICER was not interpreted independently.	Cost-saving with essentially equivalent health outcomes; highly sensitive to AI price
Gomez Rossi et al. (7): dental-caries detection	62.4 versus 60.9 tooth-retention years	AI: €320.40; standard diagnosis: €342.24	Lower costs and greater tooth retention compared with standard diagnosis (dominant).	Dominant in the base case
Gomez Rossi et al. (7): diabetic-retinopathy screening	Approximately 8.42 QALYs with both strategies	AI: R$1,321; standard care: R$1,260.28	AI increased costs by approximately R$61 without a meaningful QALY gain	Standard care was economically preferred
Xiao et al. (8): glaucoma screening	Reduced PACG, primary angle closure, and blindness; 297 disease cases of any stage were prevented over 15 years	Fifteen-year incremental cost: US$434,903.20	US$1,464.30 per disease case prevented	Clinically beneficial but not cost-offsetting
Chatzikou et al. (9): obstetric diagnosis	Earlier detection and reduced severity of preterm birth	Estimated saving of €99,840 among 195 preterm births	No QALY-based ICER reported	Cost-saving
Chatzikou et al. (9): other diagnostic applications	CCTA was reported as cost-saving; echocardiography and capsule endoscopy improved accuracy, interpretation time, or productivity	Economic effects differed by application	No common ICER was reported	Mixed economic findings
Wong et al. (10): FIT followed by AI-assisted colonoscopy	132 CRC cases prevented and 308 life-years saved, compared with 120 cases and 280 life-years for FIT followed by conventional colonoscopy	Approximately US$263 million versus US$264 million	ICER versus no screening: US$122,539 per life-year saved; incremental ICER versus conventional FIT–colonoscopy: −US$36,462	Dominant relative to the corresponding conventional strategy
Wong et al. (10): direct AI-assisted colonoscopy	1,499 CRC cases prevented and 2,639 life-years saved, compared with 1,354 cases and 2,384 life-years for conventional colonoscopy	Approximately US$702 million versus US$712 million	Dominant relative to conventional colonoscopy (lower costs and greater effectiveness).	Dominant relative to direct conventional colonoscopy

[Downloaded from jebhpme.ssu.ac.ir on 2026-11-12]

[DOI: 10.18502/jebhpme.v10i6.23195]

CC BY 4.0

Study and application	Clinical outcome	Cost outcome	ICER or principal economic finding	Economic interpretation
Maas et al. (11): AI-enhanced HCC detection	Incremental gain of 0.55 QALY per 1,000 patients	Incremental lifetime cost of approximately €5,450 per 1,000 patients	€9,888 per QALY gained; threshold €33,000 per QALY	Cost-effective
Lyth et al. (12): AI-assisted mammography	Incremental gain of 10.8 QALYs per 1,000 women	Saving of €59,320; total costs €3,468,967 versus €3,528,288	Lower costs and greater effectiveness	Dominant
Goh et al. (13): one radiologist plus AI	Incremental gain of 15.9 QALYs per 10,000 women	Saving of US$316,090	Lower costs and greater effectiveness	Dominant
Goh et al. (13): AI-only screening	Incremental gain of 72.0 QALYs per 10,000 women	Additional cost of approximately US$1.35 million	US$18,743 per QALY gained	Cost-effective at a US$50,000-per-QALY threshold
Levin et al. (14): hospital-management perspective	Emergency-department visit volume increased by 9.6%	Revenue increased by US$15.4 million; costs increased by US$2.9 million; operating margin increased by US$12.6 million	Break-even AI price: US$66.02 per visit at a 5% increase in throughput	Financially favorable from the hospital-management perspective
Levin et al. (14): public-policy perspective	Same observed throughput increase	Costs increased by US$14.5 million; operating margin increased by US$0.9 million	Break-even AI price: US$4.69 per visit	Modestly favorable and highly perspective dependent
Hill and Roadevin (15): AI-only reading	Incremental gain of 0.00009 QALY per invited woman	Saving of £159.55 per woman	Highest net monetary benefit; 100% probability of being most cost-effective at £20,000 per QALY	Dominant
Hill and Roadevin (15): one human reader plus AI	Incremental gain of 0.00019 QALY per invited woman	Saving of £31.07 per woman	Greater effectiveness and lower cost than double human reading	Dominant
Hill and Roadevin (15): two human readers plus AI	Incremental gain of 0.00023 QALY per invited woman	Additional cost of £72.79 per woman	£317,423 per QALY gained	Not cost-effective at a £20,000-per-QALY threshold

Table note: CCTA, coronary computed tomography angiography; CRC, colorectal cancer; FIT, fecal immunochemical test; HCC, hepatocellular carcinoma; ICER, incremental cost-effectiveness ratio; PACG, primary angle-closure glaucoma; QALY, quality-adjusted life-year. Costing years, currencies, discount rates, and willingness-to-pay thresholds are summarized in Appendix 4 (Supplementary Table S4). Economic outcomes are presented in their original study-specific units and should not be directly compared across different healthcare settings, time horizons, or costing bases. Finally, negative ICERs were interpreted according to the direction of incremental costs and effectiveness rather than their numerical sign alone.

[Downloaded from jebhpme.ssu.ac.ir on 2026-11-12]

[DOI: 10.18502/jebhpme.v10i6.23195]

CC BY 4.0

Synthesis of the Economic Findings

Five of the nine studies identified at least one AI-assisted strategy that was economically dominant. Dominant strategies were reported for dental-caries detection, AI-assisted colonoscopy, AI-assisted mammography in Sweden, the AI-companion breast-screening strategy in Singapore, and the AI-only or one-human-plus-AI screening strategies in the United Kingdom. The economic findings were nevertheless not uniformly favorable. Within the Gomez Rossi et al. study, AI-assisted diabetic-retinopathy screening increased costs without producing a meaningful QALY gain. Similarly, AI-assisted glaucoma screening reduced disease progression and blindness but did not offset the additional screening expenditure during the 15-year horizon. Triple reading with two human readers plus AI produced the largest health gain in the UK breast-screening model but was not cost-effective because its ICER substantially exceeded the applicable threshold.

The evaluation by Maas et al. showed that AI-assisted detection of liver lesions represented a cost-effective strategy; however, the intervention did not reduce overall healthcare costs. Similarly, the AI-only breast-screening strategy evaluated in Singapore achieved favorable cost-effectiveness because it generated additional health benefits, although these benefits were accompanied by higher total expenditure. Cost-consequence analyses also demonstrated that the economic impact of AI differed substantially across clinical applications, depending on the type of clinical benefit achieved, resource utilization, and implementation requirements. AI-assisted obstetric diagnosis and coronary CT angiography were reported as cost-saving, whereas echocardiography and capsule endoscopy involved additional expenditure in exchange for improvements in diagnostic accuracy, interpretation time, productivity, or clinician experience. The emergency-triage evaluation demonstrated the importance of economic perspective. The same observed improvement in patient throughput produced a US$12.6 million increase in operating margin under the hospital-management framework but only a US$0.9 million increase under the public-policy framework.

Overall, favorable findings were most common when AI replaced part of an existing diagnostic workflow, reduced human reading requirements, increased patient throughput, or prevented downstream treatment costs through earlier disease detection. Findings were less favorable when AI was added to existing care without reducing other resource use, when AI specificity generated additional false-positive investigations, or when capital and implementation costs were high.

The direction of the economic findings across the included studies is summarized descriptively in Figure 2. Because several studies assessed multiple AI configurations, clinical applications, or economic perspectives, an individual study could contribute to more than one economic category. The figure is intended to demonstrate the distribution and variability of study-level conclusions and does not represent a quantitative pooling of costs, effectiveness outcomes, or incremental cost-effectiveness ratios.

[Downloaded from jebhpme.ssu.ac.ir on 2026-11-12]

[DOI: 10.18502/jebhpme.v10i6.23195]

CC BY 4.0

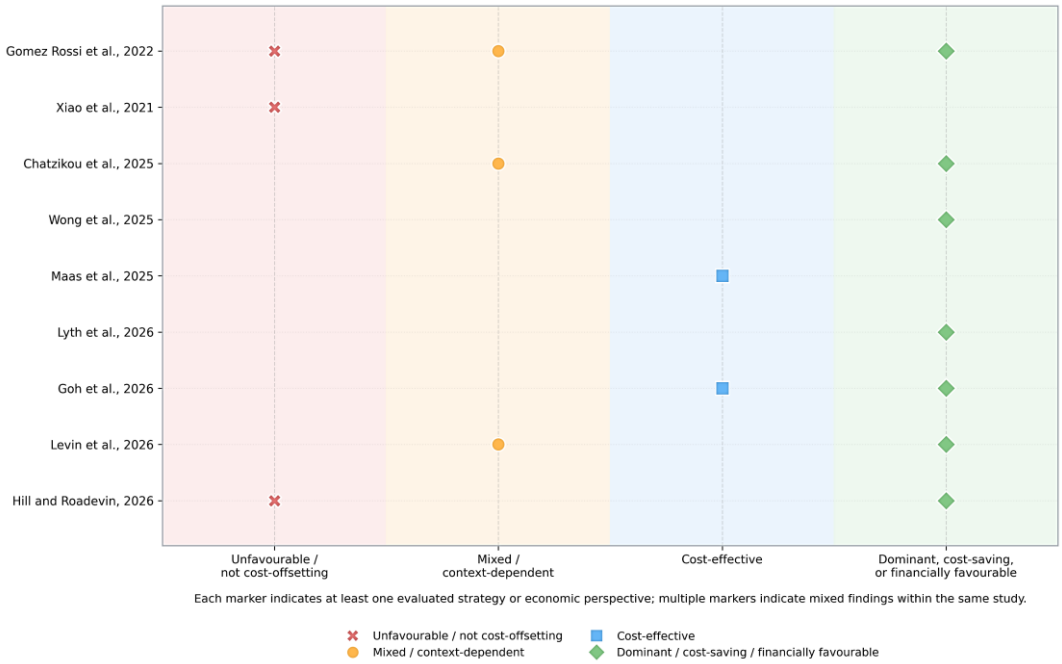


Figure 2. Descriptive evidence map of study-level economic findings.

As shown in Figure 2, favorable economic findings were common but not universal. Several evaluations identified dominant, cost-saving, or financially favorable AI-assisted strategies, particularly when AI replaced part of an existing workflow or reduced human-reading requirements. However, some studies reported mixed results across clinical applications or perspectives, and selected strategies were not cost-effective or failed to offset their additional costs. These findings reinforce the conclusion that the economic value of AI-assisted diagnostic systems is dependent on the intervention design, implementation pathway, comparator, and healthcare-system context.

Sources of Economic Uncertainty

Across the included studies, the most influential determinants of economic value were AI acquisition or licensing cost, implementation and maintenance expenditure, diagnostic sensitivity and specificity, disease prevalence, clinician-adoption rates, patient adherence, screening volume, and the extent to which AI replaced rather than supplemented existing clinical resources.

The Gomez Rossi et al. models were sensitive to relatively small changes in the fee charged for AI and to assumptions about treatment after diagnosis. In the liver-lesion model, the cost and diagnostic performance of the AI system were the principal drivers of the ICER. In the Singapore breast-screening model, specificity strongly influenced false-positive investigations and total costs. In the glaucoma model, equipment investment was a major determinant of the cumulative budget impact. The emergency-triage evaluation demonstrated that the estimated economic benefit of AI can vary considerably depending on the perspective adopted in the analysis. An intervention that appears financially attractive from the viewpoint of an individual healthcare provider may generate a different conclusion when evaluated from a broader healthcare-system perspective. These findings suggest that the economic value of AI-assisted CDSS depends not only on diagnostic performance but also on implementation context, resource allocation, and downstream effects on healthcare delivery.

Feasibility of Quantitative Synthesis

A quantitative meta-analysis was not performed. The included studies did not report a common economic or clinical outcome suitable for statistical pooling. Effectiveness was expressed as

[Downloaded from jebhpme.ssu.ac.ir on 2026-11-12]

[DOI: 10.18502/jebhpme.v10i6.23195]

QALYs, life-years saved, tooth-retention years, disease cases prevented, reductions in blindness, diagnostic and workflow consequences, or emergency-department throughput. Several outcomes were generated by decision-analytic models rather than observed patient-level comparative data. The studies also differed substantially in disease area, population, comparator, jurisdiction, economic perspective, currency, costing year, and time horizon. Converting these findings into a common effect measure would therefore have required unsupported assumptions.

Reporting Bias

Formal tests for publication bias were not undertaken because no outcome was statistically pooled. Funnel plots and Egger's regression are intended for meta-analyses containing sufficiently comparable independent effect estimates and were not appropriate for the present structured narrative synthesis. The possibility of reporting bias cannot be excluded. Studies with favorable economic findings may be more likely to be published, and several included evaluations were early or model-based assessments in which assumptions about AI price and diagnostic performance materially influenced the conclusions. The predominance of favorable findings should therefore be interpreted cautiously.

Discussion

Principal Findings

This systematic review included 9 primary economic evaluations assessing AI-assisted CDSS in diagnostic care. The included studies covered a wide range of clinical applications, including cancer screening, dermatological and dental diagnosis, ophthalmological screening, liver lesion detection, multispecialty diagnostic support, and emergency-department triage. This diversity reflects the broad clinical scope in which AI-assisted diagnostic technologies are currently being evaluated. Most studies used decision-analytic models, while only a small number incorporated real-world clinical or operational data.

The available evidence suggests that AI-assisted diagnostic technologies may offer economic benefits in selected clinical settings; however, these findings should be interpreted cautiously given the small number of included studies, substantial methodological heterogeneity, and predominance of model-based economic evaluations. Five of the nine included studies reported at least one AI-assisted strategy that was economically dominant under their respective base-case assumptions, indicating lower costs and improved effectiveness compared with the relevant comparator (7, 10, 12, 13, 15). These favorable findings were primarily observed in scenarios involving the replacement of existing diagnostic workflow components, reduced human-reading requirements, or improvements in screening efficiency. Nevertheless, the economic advantages were dependent on study-specific assumptions, implementation costs, clinical pathways, and healthcare-system perspectives. Furthermore, the absence of independent duplicate review and complete original screening records limits confidence in the comprehensiveness and reproducibility of the evidence synthesis. Therefore, the findings should not be interpreted as establishing the general cost-effectiveness of AI-assisted diagnostic technologies across healthcare settings.

Other AI-assisted strategies improved outcomes but increased costs. AI-enhanced detection of liver lesions and AI-only breast cancer screening in Singapore were considered cost-effective because their incremental cost-effectiveness ratios remained below the willingness-to-pay thresholds applied in the respective studies. In contrast, AI-assisted diabetic-retinopathy screening increased costs without producing a meaningful improvement in QALYs, and population-based glaucoma screening reduced disease progression and blindness but did not offset its additional costs over the evaluated time horizon. Triple reading with two human readers plus AI also produced additional health benefits in the UK breast-screening model but was not considered cost-effective because its ICER exceeded the relevant threshold.

These findings demonstrate that AI-assisted CDSS

[Downloaded from jebhpme.ssu.ac.ir on 2026-11-12]

[DOI: 10.18502/jebhpme.v10i6.23195]

 CC BY 4.0

cannot be considered economically valuable merely because they improve diagnostic performance. Their economic value depends on the relationship between clinical benefit, implementation cost, workflow design, downstream resource use, and the extent to which the AI system replaces or supplements existing clinical activities.

Interpretation of the Economic Evidence

The most consistent favorable findings were observed in high-volume screening and diagnostic pathways. In such settings, even relatively small improvements in efficiency may be applied across large populations and may therefore generate substantial cumulative savings. Breast cancer screening provided the clearest example. Moreover, the Swedish, Singaporean, and UK evaluations found that selected human–AI or AI-only strategies could reduce reading workload, lower costs, or produce favorable ICERs. However, the preferred configuration differed across healthcare systems, illustrating that results cannot be transferred directly from one jurisdiction to another.

The colorectal cancer screening model similarly suggested that AI-assisted colonoscopy could improve the prevention of colorectal cancer and increase life-years while reducing costs relative to the corresponding conventional colonoscopy strategies. The economic advantage was partly attributable to improved lesion detection and the resulting prevention of future cancer treatment costs. Nevertheless, these results were generated within a modeling framework and depended on assumptions concerning screening adherence, adenoma detection, disease progression, and local treatment expenditure.

The liver-lesion evaluation produced a favorable ICER but only a small incremental QALY gain. In such circumstances, apparently favorable cost-effectiveness results should be interpreted carefully because modest changes in AI price, diagnostic sensitivity, specificity, or disease prevalence may materially alter the conclusion. A low ICER does not necessarily indicate a large clinical benefit; it may instead reflect a relatively small additional cost divided by a very small improvement in effectiveness.

The emergency-triage evaluation illustrated the importance of analytical perspective. The AI-assisted system appeared strongly favorable from the hospital-management perspective because increased patient throughput generated additional revenue and operating margin. The benefit was substantially smaller from the public-policy perspective in which a broader range of costs was considered. Therefore, an intervention may be financially attractive to an individual hospital without necessarily generating equivalent savings for the wider healthcare system.

The cost-consequence study also demonstrated that AI may produce outcomes that are difficult to express through a single ICER. Reductions in interpretation time, improvements in clinician productivity, enhanced diagnostic confidence, and greater user satisfaction may be relevant to implementation decisions even when QALYs are unavailable. However, cost-consequence findings should be interpreted carefully because they present multiple clinical and operational outcomes rather than a single summary measure. Consequently, healthcare decision-makers must determine whether the observed improvements in diagnostic performance, workflow efficiency, or service delivery justify the additional resources required for implementation.

Comparison with Previous Literature

The findings of this review are broadly consistent with previous systematic reviews reporting that AI may provide economic value in selected healthcare applications while emphasizing the limited maturity and substantial heterogeneity of the evidence base (4-6). Previous reviews have identified potentially favorable findings in radiology, ophthalmology, cancer screening, and clinical workflow optimization, but have also noted that most available evaluations are model based and highly dependent on local assumptions.

Unlike the broader healthcare reviews by El Arab

[Downloaded from jebhpme.ssu.ac.ir on 2026-11-12]

[DOI: 10.18502/jebhpme.v10i6.23195]

CC BY 4.0

and Al Moosa and Wu et al. (4, 5), the present review focuses specifically on primary economic evaluations of AI-assisted diagnostic and screening technologies, distinguishing clinician-assisted from autonomous AI configurations. It additionally considers non-radiological diagnostic applications and emergency triage, complementing the radiology-focused review by Brin and Tau (6). The principal contribution of this review is its structured interpretation of economic findings across nine primary evaluations, particularly the distinction between cost savings, economic dominance, and cost-effectiveness based on jurisdiction-specific willingness-to-pay thresholds. Furthermore, the review examines how workflow substitution, implementation costs, and healthcare-system perspectives may influence the economic value of AI-assisted diagnostic technologies.

The present review extends this literature in several ways. First, only primary economic evaluations were included in the evidence set. Systematic reviews, narrative reviews, scoping reviews, and book chapters were used for contextual purposes but were not treated as primary studies. This approach reduced the risk of double counting evidence from the same underlying evaluations.

Second, economic conclusions were classified according to explicit incremental criteria. Interventions were described as dominant only when they were both less costly and more effective than their comparators. Cost-saving interventions without evidence of superior effectiveness were distinguished from dominant strategies, while interventions producing additional benefits at additional cost were evaluated against the relevant willingness-to-pay threshold.

Third, the review did not calculate a pooled mean ICER, pooled percentage cost reduction, or pooled diagnostic-error estimate. The included studies differed too substantially in their clinical populations, AI applications, currencies, economic perspectives, outcome measures, and time horizons for such pooling to be meaningful. Avoiding an artificial summary estimate reduced the risk of presenting a statistically precise but clinically misleading conclusion.

Implications for Policy and Clinical Implementation

Based on the findings of this review and the authors' interpretation of their potential policy implications, a cautious and context-specific approach to implementing AI-assisted CDSS may be appropriate rather than a universal adoption strategy. Although the included studies did not directly evaluate specific implementation or procurement policies, their findings highlight the importance of diagnostic performance, implementation costs, clinical workflow, and healthcare-system context in determining economic value. Accordingly, the author suggests that healthcare organizations consider local validation, workflow assessment, and comprehensive cost evaluation, including software acquisition, integration, cybersecurity, maintenance, staff training, performance monitoring, and regulatory compliance, before adopting AI-assisted diagnostic technologies.

The degree of workflow substitution is likely to be a major determinant of value. AI may generate savings when it safely replaces one human reader, reduces specialist interpretation time, prioritizes high-risk cases, or prevents unnecessary investigations. In contrast, adding AI to an unchanged clinical pathway may increase expenditure without reducing staff requirements or other resource use. Economic evaluations should therefore state clearly whether AI replaces, supplements, or reorganizes existing care.

Local disease prevalence and diagnostic specificity also require careful consideration. In population screening, a small reduction in specificity may produce large numbers of additional false-positive findings, resulting in further imaging, biopsies, consultations, anxiety, and treatment-related expenditure. The economic consequences of false-positive and false-negative results should be incorporated into both modeling and implementation decisions.

A gradual implementation approach may be more

[Downloaded from jebhpme.ssu.ac.ir on 2026-11-12]

[DOI: 10.18502/jebhpme.v10i6.23195]

CC BY 4.0

appropriate than immediate large-scale deployment. Initial pilot projects can provide valuable information regarding local diagnostic performance, clinician acceptance, workflow integration, equity implications, and actual patterns of resource utilization. Establishing predefined clinical and economic criteria before implementation would allow organizations to evaluate whether the expected benefits are achieved and whether continued investment remains justified as clinical conditions and technology characteristics evolve.

Procurement strategies should also consider the uncertainty associated with emerging AI technologies. Flexible contractual approaches, such as performance-based agreements, limited-term licensing, and requirements for local validation, may help reduce financial risks when long-term effectiveness and economic value are uncertain. In addition, decision-makers should account for future costs associated with algorithm updating, retraining, validation, and replacement as patient populations, clinical practices, imaging technologies, and data environments change.

Strengths and Limitations

A principal strength of this review was the restriction of the evidence set to primary economic evaluations, which reduced the risk of double counting evidence from secondary reviews. Reporting quality was distinguished from methodological appraisal, and design-appropriate tools were used for model-based and empirical economic evaluations.

Furthermore, economic classifications were based on incremental costs and effects rather than solely on the conclusions stated by the original investigators. The review also preserved the original perspectives, currencies, time horizons, and willingness-to-pay thresholds, thereby avoiding inappropriate cross-jurisdictional pooling. The decision not to perform a meta-analysis was methodologically appropriate because the clinical and economic outcomes were not sufficiently comparable.

Several limitations must nevertheless be acknowledged. First, study selection, data extraction, and methodological appraisal were conducted by a single reviewer. Although all decisions and extracted data were checked during a separate verification pass, the absence of independent duplicate assessment increases the possibility of selection, extraction, and interpretation errors.

Second, the review was not prospectively registered, and a dated protocol specifying the original eligibility criteria and any amendments were unavailable. The timing and reasons for possible eligibility refinements therefore cannot be independently established, creating a risk of post hoc selection decisions.

Third, source-level database exports and the complete screening log were unavailable. The counts shown in Figure 1 relate only to the documented candidate publications, not the full search process, and must not be interpreted as a complete PRISMA 2020 flow. The inability to verify retrieval, deduplication, and screening counts reduces reproducibility and means that complete identification of all eligible studies cannot be assured.

Fourth, eligibility assessment was restricted to English-language full-text reports because reliable translation resources were unavailable. Relevant studies published in other languages may therefore have been missed.

Fifth, most included evaluations were based on decision-analytic models. Model-based studies are valuable for estimating long-term outcomes, but their conclusions depend on assumptions about clinical effectiveness, disease progression, AI cost, implementation, and downstream care. Only limited evidence was derived from routine real-world implementation.

Sixth, the included studies represented a small number of healthcare systems, predominantly high-income settings. Economic results from Sweden, Singapore, the United Kingdom, Italy, or the United States may not be directly applicable

[Downloaded from jebhpme.ssu.ac.ir on 2026-11-12]

[DOI: 10.18502/jebhpme.v10i6.23195]

to countries with different workforce costs, diagnostic pathways, reimbursement systems, digital infrastructure, and willingness-to-pay thresholds.

Finally, formal publication-bias assessment was not possible because no outcome was statistically pooled. Studies reporting favorable economic results may be more likely to be published, and industry involvement or optimistic assumptions in early health technology assessments may influence reported conclusions.

Recommendations for Future Research

Future research should prioritize prospective, multicenter, real-world economic evaluations conducted alongside implementation studies. Such studies should compare AI-assisted pathways with clearly defined current practice and should measure both patient-relevant health outcomes and actual changes in resource use.

Economic evaluations should include the full lifecycle cost of AI, including acquisition, integration, validation, training, maintenance, cybersecurity, monitoring, updates, and eventual replacement. Studies should clearly state whether the AI system substitutes for or supplements existing staff and should measure changes in clinician workload rather than assuming that improved algorithmic efficiency automatically produces labor savings.

Standardized reporting would also improve comparability. Future studies should report the costing year, currency, economic perspective, time horizon, discount rate, incremental costs, incremental effects, ICERs, willingness-to-pay threshold, and deterministic and probabilistic sensitivity analyses. In addition, model-based evaluations should provide sufficient information to permit external validation and independent replication.

Future research should also evaluate the equity implications of AI implementation. AI-assisted systems may improve access to specialist-level diagnostic support in regions with limited healthcare resources; however, unequal access to digital infrastructure or reduced algorithm performance in under-represented populations could potentially increase existing disparities. Distributional cost-effectiveness approaches may provide additional insight into how the costs and benefits of AI technologies are distributed among different socioeconomic, geographic, demographic, and clinical groups.

Additional evidence from low- and middle-income countries is particularly important. Although limited specialist availability in these settings may increase the potential value of AI-assisted diagnostic systems, infrastructure limitations, affordability constraints, workforce capacity, and implementation challenges may influence their feasibility and economic value. Future evaluations should therefore incorporate local healthcare priorities, budget constraints, and implementation conditions when assessing AI technologies in these environments. Evaluations in these settings should consider affordability and budget impact in addition to conventional cost-effectiveness thresholds.

Finally, economic value should be reassessed throughout the lifecycle of an AI system. Algorithm performance, clinical workflows, licensing prices, disease prevalence, and available comparators may change over time. Continuous postimplementation monitoring and periodic economic reassessment are therefore necessary to determine whether initially favorable value is maintained in routine clinical practice.

Conclusion

This systematic review indicates that AI-assisted clinical decision support systems may provide economic value in selected diagnostic settings, particularly when they reduce specialist workload, replace part of an existing clinical workflow, increase patient throughput, or support earlier disease detection. Across the 9 primary economic evaluations included in the present synthesis, several AI-assisted strategies were reported to be dominant or cost-effective. The most favorable

[Downloaded from jebhpme.ssu.ac.ir on 2026-11-12]

[DOI: 10.18502/jebhpme.v10i6.23195]

 CC BY 4.0

findings were observed in high-volume applications such as breast cancer screening, colorectal cancer screening, and selected diagnostic pathways in which AI reduced human-reading requirements or prevented downstream healthcare expenditure.

However, the evidence does not support the conclusion that all AI-assisted diagnostic systems are inherently cost-effective. Economic outcomes varied according to the clinical application, AI price, diagnostic sensitivity and specificity, disease prevalence, implementation scale, clinician-adoption rate, healthcare perspective, and the extent to which AI replaced rather than supplemented existing resources. Some interventions improved clinical outcomes but increased costs, while others generated additional expenditure without producing a meaningful improvement in QALYs. Economic findings were also strongly dependent on jurisdiction-specific clinical pathways, prices, and willingness-to-pay thresholds.

The available evidence should therefore be interpreted cautiously. Most included studies were model-based evaluations, and relatively little evidence was derived from prospective real-world implementation. The small evidence base, concentration of studies in high-income healthcare systems, absence of independent duplicate review, lack of prospective protocol registration, and inability to reconstruct the original database screening records further limit confidence in the overall findings.

Before widespread adoption, AI-assisted CDSS should undergo local clinical validation, workflow assessment, budget-impact analysis, and prospective economic evaluation. Future studies should include the full lifecycle costs of AI, measure actual changes in healthcare-resource use, assess patient-relevant outcomes, and examine equity and generalizability across different populations and healthcare systems.

In conclusion, AI-assisted CDSS may represent economically valuable diagnostic technologies when implemented in carefully selected clinical pathways and supported by appropriate workflow redesign. Their value is context dependent and should be demonstrated through transparent, locally relevant, and methodologically rigorous economic evaluations rather than assumed from diagnostic performance alone.

Authors' Contributions

S.A designed the research, developed the methodology (conceptualization and methodology), conducted the literature search, selected relevant studies, and extracted and verified the data (investigation and data curation). S.A also performed methodological appraisal, analyzed and synthesized the evidence, interpreted the findings (formal analysis), prepared the original manuscript, and critically revised its content (writing – original draft and writing – review and editing). S.A was responsible for the accuracy and integrity of the work, read and approved the final manuscript, and assumed full responsibility for its content.

Acknowledgements

The author received no external assistance requiring acknowledgement.

Competing Interests

The author declared no competing interests.

Data Availability

The study-level characteristics, extracted economic outcomes, and methodological strengths and limitations of the nine included evaluations are summarized in Tables 1–3. The original database search exports and complete screening logs were not retained and are therefore unavailable for public deposition. This limitation is explicitly acknowledged in the Strengths and Limitations section.

Ethical Considerations

Ethical approval and informed consent were not required because this review used published aggregate data and involved no recruitment of participants or collection of identifiable data.

Funding

The authors have no support or funding to report.

[Downloaded from jebhpme.ssu.ac.ir on 2026-11-12]

[DOI: 10.18502/jebhpme.v10i6.23195]

Artificial Intelligence Use Disclosure Statement

During the preparation of this manuscript, the author used ChatGPT (GPT-5, OpenAI; September 2026) for language-related assistance, including improving grammar, readability, sentence structure, and clarity of expression. All AI-assisted outputs were critically reviewed and verified by the author. The author retains full responsibility for the accuracy, originality, scientific integrity, and final content of the manuscript.

References

1. Newman-Toker DE, Nassery N, Schaffer AC, Yu-Moe CW, Clemens GD, Wang Z, et al. Burden of serious harms from diagnostic error in the USA. BMJ Quality & Safety. 2024;33(2):109-20. doi: 10.1136/bmjqs-2021-014130
2. Saber Tehrani AS, Lee H, Mathews SC, Shore A, Makary MA, Pronovost PJ, et al. 25-Year summary of US malpractice claims for diagnostic errors 1986-2010: an analysis from the National Practitioner Data Bank. BMJ Quality & Safety. 2013 Aug;22(8):672-80. doi: 10.1136/bmjqs-2012-001550
3. Slawomirski L, Kelly D, de Bienassis K, Kallas KA, Klazinga N. The economics of diagnostic safety. OECD Health Working Papers. No. 176. Paris: OECD Publishing; 2025. doi:10.1787/fc61057a-en
4. El Arab RA, Al Moosa OA. Systematic review of cost effectiveness and budget impact of artificial intelligence in healthcare. NPJ Digit Med. 2025;8(1):548. doi: 10.1038/s41746-025-01722-y
5. Wu WT, Chao YW, Lin TK, Huang CK, Hsieh PH. Economic evaluation of AI-assisted technologies in healthcare: a systematic review. J Food Drug Anal. 2025;33(4):487-500. doi: 10.38212/2224-6614.3570
6. Brin D, Tau N. Cost-effectiveness of artificial intelligence tools in radiology: a systematic review. Eur Radiol. 2026;36(5):3755-65. doi: 10.1007/s00330-025-12242-4
7. Gomez Rossi J, Rojas-Perilla N, Krois J, Schwendicke F. Cost-effectiveness of artificial intelligence as a decision-support system applied to the detection and grading of melanoma, dental caries, and diabetic retinopathy. JAMA Netw Open. 2022;5(3):e220269. doi: 10.1001/jamanetworkopen.2022.0269
8. Xiao X, Xue L, Ye L, Li H, He Y. Health care cost and benefits of artificial intelligence-assisted population-based glaucoma screening for the elderly in remote areas of China: a cost-offset analysis. BMC Public Health. 2021;21(1):1065. doi: 10.1186/s12889-021-11097-w
9. Chatzikou M, Latsou D, Apostolidis G, Billis A, Charisis V, Rigas ES, et al. Economic evaluation of artificially intelligent (AI) diagnostic systems: cost consequence analysis of clinician-friendly interpretable computer-aided diagnosis (ICADX) tested in cardiology, obstetrics, and gastroenterology, from the HosmartAI Horizon 2020 project. Healthcare (Basel). 2025;13(14):1661. doi: 10.3390/healthcare13141661
10. Wong MC, Huang J, Lam TY, Lau LH, Chiu PW. The cost-effectiveness of AI-assisted colonoscopy as a primary or secondary screening test in a population-based colorectal cancer screening program: Markov modeling–based cost effectiveness analysis. J Med Internet Res. 2025;27:e67762. doi: 10.2196/67762
11. Maas L, Contreras-Meca C, Ghezzo S, Belmans F, Corsi A, Cant J, et al. Cost-effectiveness analysis of artificial intelligence (AI) in earlier detection of liver lesions in cirrhotic patients at risk of hepatocellular carcinoma in Italy. J Med Econ. 2025;28(1):1023-36. doi: 10.1080/13696998.2025.2525006
12. Lyth J, Gialias P, Husberg M, Bernfort L, Bjerner T, Wiberg MK, et al. Results from a Swedish model-based analysis of the cost-effectiveness of AI-assisted digital mammography. Eur Radiol. 2026;36(1):754-64. doi: 10.1007/s00330-025-11821-9
13. Goh SS, Lim YZ, Ong C, Hartman M, Wang Y. Cost effectiveness analysis of an AI-assisted breast cancer screening programme in Singapore: an early health technology assessment. Cancers (Basel). 2026;18(5):836. doi: 10.3390/cancers18050836
14. Levin S, Steinhart B, Sangal R, Meng L, Frick KD, McGinn C, et al. AI-based triage decision support: multisite economic evaluation in the United States. J Med Internet Res. 2026;28:e95213. doi: 10.2196/95213
15. Hill H, Roadevin C. Economic evaluation of artificial intelligence for cancer detection in the UK breast screening programme. Br J Cancer. 2026;135(3):453-60. PMID: 42069897; PMCID: PMC13373195. doi: 10.1038/s41416-026-03465-3
16. Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ. 2021;372:n71. doi: 10.1136/bmj.n71
17. Husereau D, Drummond M, Augustovski F, de Bekker-Grob E, Briggs AH, Carswell C, et al.

[Downloaded from jebhpme.ssu.ac.ir on 2026-11-12]

[DOI: 10.18502/jebhpme.v10i6.23195]

 CC BY 4.0

Consolidated health economic evaluation reporting standards 2022 (CHEERS 2022) statement: updated reporting guidance for health economic evaluations. MDM Policy Pract. 2022;7(1):23814683211061097. doi: 10.1177/23814683211061097

18. Philips Z, Bojke L, Sculpher M, Claxton K, Golder S. Good practice guidelines for decision-analytic modelling in health technology assessment: a review and consolidation of quality assessment. Pharmacoeconomics. 2006;24(4):355-71. doi: 10.2165/00019053-200624040-00006
19. Evers S, Goossens M, De Vet H, Van Tulder M, Ament A. Criteria list for assessment of methodological quality of economic evaluations: consensus on health economic criteria. Int J Technol Assess Health Care. 2005;21(2):240-45. doi: 10.1017/S0266462305050324

[Downloaded from jebhpme.ssu.ac.ir on 2026-11-12]

[DOI: 10.18502/jebhpme.v10i6.23195]

Appendix 1. PRISMA 2020 Checklist

Supplementary Table S1. PRISMA 2020 Checklist and Location of Reported Information

Section/topic	Item	PRISMA requirement	Location in manuscript
Title	1	Identify the report as a systematic review	Title page
Abstract	2	Provide a structured summary of the review	Abstract
Rationale	3	Describe the rationale in the context of existing knowledge	Introduction
Objectives	4	Provide an explicit statement of the review objectives	Introduction – Final paragraph
Eligibility criteria	5	Specify inclusion and exclusion criteria	Materials and Methods – Eligibility Criteria
Information sources	6	Specify all databases and other sources searched	Materials and Methods – Information Sources and Search Strategy
Search strategy	7	Present the complete search strategies for all databases	Appendix 2 – Complete Database Search Strategies (Supplementary Table S2)
Selection process	8	Describe how studies were screened and selected	Materials and Methods – Study Selection and Data Extraction
Data-collection process	9	Describe how data were extracted and verified	Materials and Methods – Study Selection and Data Extraction
Data items	10a–b	Define the outcomes and other variables extracted	Materials and Methods – Eligibility Criteria; Study Selection and Data Extraction
Methodological appraisal	11	Describe the methods used to appraise included studies	Materials and Methods – Reporting Quality and Methodological Appraisal
Effect measures	12	Specify the measures used to interpret economic results	Materials and Methods – Data Synthesis and Economic Classification
Synthesis methods	13a–f	Describe study grouping, synthesis, heterogeneity, and sensitivity methods	Materials and Methods – Data Synthesis and Economic Classification
Reporting-bias assessment	14	Describe the assessment of possible reporting bias	Materials and Methods – Data Synthesis and Economic Classification; Results – Reporting Bias
Certainty assessment	15	Describe any assessment of confidence in the evidence	Materials and Methods – Data Synthesis and Economic Classification; Discussion – Strengths and Limitations (no formal GRADE assessment)
Study selection	16a	Report the results of the selection process	Results – Study Selection and Available Selection Records
Excluded reports	16b	Cite reports that appeared eligible but were excluded and explain why	Results – Study Selection and Available Selection Records
Study characteristics	17	Present characteristics of every included study	Results – Characteristics of the Included Studies; Table 1
Appraisal results	18	Present methodological-appraisal findings for each study	Results – Reporting Quality and Methodological Appraisal; Table 2
Individual-study results	19	Present results for every included study	Results – Economic Outcomes; Table 3
Results of synthesis	20a–d	Present the findings of all syntheses	Results – Economic Outcomes; Synthesis of the Economic Findings; Sources of Economic Uncertainty; Feasibility of

[Downloaded from jebhpme.ssu.ac.ir on 2026-11-12]

[DOI: 10.18502/jebhpme.v10i6.23195]

 CC BY 4.0

			Quantitative Synthesis
Reporting biases	21	Present the assessment of reporting bias	Results – Reporting Bias
Certainty of evidence	22	Present confidence in the overall evidence	Discussion – Strengths and Limitations; Conclusion
Discussion	23a–d	Interpret findings, limitations, implications, and future research	Discussion – Principal Findings; Interpretation of the Economic Evidence; Comparison with Previous Literature; Implications for Policy and Clinical Implementation; Strengths and Limitations; Recommendations for Future Research
Registration and protocol	24a–c	Report registration and protocol information	Materials and Methods – Review Design and Protocol
Support	25	Describe financial and non-financial support	Funding
Competing interests	26	Declare competing interests	Conflict of Interest
Availability of materials	27	State where data and review materials are available	Data Availability

[Downloaded from jebhpme.ssu.ac.ir on 2026-11-12]

[DOI: 10.18502/jebhpme.v10i6.23195]

Appendix 2. Complete Database Search Strategies

Supplementary Table S2. Database-Specific Search Strategies

Database	Complete search strategy
PubMed/MEDLINE	("Artificial Intelligence"[Mesh] OR "Machine Learning"[Mesh] OR "Decision Support Systems, Clinical"[Mesh] OR "artificial intelligence"[tiab] OR "machine learning"[tiab] OR "deep learning"[tiab] OR "neural network"[tiab] OR "computer vision"[tiab] OR "natural language processing"[tiab] OR "clinical decision support system"[tiab] OR "AI-assisted"[tiab]) AND ("Costs and Cost Analysis"[Mesh] OR "Cost-Benefit Analysis"[Mesh] OR "Economics, Medical"[Mesh] OR "cost-effectiveness"[tiab] OR "economic evaluation"[tiab] OR "cost utility"[tiab] OR "cost benefit"[tiab] OR "cost consequence"[tiab] OR "cost offset"[tiab] OR "budget impact"[tiab] OR ICER[tiab] OR QALY[tiab] OR "quality-adjusted life year"[tiab] OR "cost saving"[tiab]) AND ("Diagnosis"[Mesh] OR "Mass Screening"[Mesh] OR diagnos[tiab] OR screening[tiab] OR detection[tiab] OR triage[tiab])
Embase	('artificial intelligence'/exp OR 'machine learning'/exp OR 'clinical decision support system'/exp OR 'artificial intelligence':ti,ab,kw OR 'machine learning':ti,ab,kw OR 'deep learning':ti,ab,kw OR 'neural network':ti,ab,kw OR 'computer vision':ti,ab,kw OR 'natural language processing':ti,ab,kw OR 'clinical decision support system':ti,ab,kw OR 'AI-assisted':ti,ab,kw) AND ('economic evaluation'/exp OR 'cost effectiveness analysis'/exp OR 'cost utility analysis'/exp OR 'budget impact analysis'/exp OR 'cost-effectiveness':ti,ab,kw OR 'economic evaluation':ti,ab,kw OR 'cost utility':ti,ab,kw OR 'cost benefit':ti,ab,kw OR 'cost consequence':ti,ab,kw OR 'cost offset':ti,ab,kw OR 'budget impact':ti,ab,kw OR ICER:ti,ab,kw OR QALY:ti,ab,kw OR 'cost saving':ti,ab,kw) AND ('diagnosis'/exp OR 'screening'/exp OR diagnos:ti,ab,kw OR screening:ti,ab,kw OR detection:ti,ab,kw OR triage:ti,ab,kw)
Scopus	TITLE-ABS-KEY(("artificial intelligence" OR "machine learning" OR "deep learning" OR "neural network" OR "computer vision" OR "natural language processing" OR "clinical decision support system" OR "AI-assisted") AND ("cost-effectiveness" OR "economic evaluation" OR "cost utility" OR "cost benefit" OR "cost consequence" OR "cost offset" OR "budget impact" OR ICER OR QALY OR "quality-adjusted life year" OR "cost saving") AND (diagnos OR screening OR detection OR triage))
Web of Science Core Collection	TS=(("artificial intelligence" OR "machine learning" OR "deep learning" OR "neural network" OR "computer vision" OR "natural language processing" OR "clinical decision support system" OR "AI-assisted") AND ("cost-effectiveness" OR "economic evaluation" OR "cost utility" OR "cost benefit" OR "cost consequence" OR "cost offset" OR "budget impact" OR ICER OR QALY OR "quality-adjusted life year" OR "cost saving") AND (diagnos OR screening OR detection OR triage))
Cochrane Library	([mh "Artificial Intelligence"] OR [mh "Decision Support Systems, Clinical"] OR "artificial intelligence":ti,ab,kw OR "machine learning":ti,ab,kw OR "deep learning":ti,ab,kw OR "clinical decision support system":ti,ab,kw) AND ([mh "Costs and Cost Analysis"] OR [mh "Cost-Benefit Analysis"] OR "cost-effectiveness":ti,ab,kw OR "economic evaluation":ti,ab,kw OR "cost utility":ti,ab,kw OR "budget impact":ti,ab,kw OR ICER:ti,ab,kw OR QALY:ti,ab,kw) AND (diagnos:ti,ab,kw OR screening:ti,ab,kw OR detection:ti,ab,kw OR triage:ti,ab,kw)
Google Scholar	"artificial intelligence" OR "machine learning" "clinical decision support" diagnosis screening "cost effectiveness" OR "economic evaluation" OR ICER OR QALY
Additional searching	Reference lists of all eligible primary studies and relevant systematic reviews were examined. Forward citation searching was conducted for the included primary economic evaluations.

[Downloaded from jebhpme.ssu.ac.ir on 2026-11-12]

[DOI: 10.18502/jebhpme.v10i6.23195]

 CC BY 4.0

Appendix 3. Data Extraction Form and Piloting Documentation

Supplementary Table S3. Standardized Data Extraction Fields for Primary Economic Evaluations

Domain	Data extraction items	Recording instructions
Study identification	First author; publication year; country	Record bibliographic details from the original publication
Population and setting	Clinical setting; target population; sample size or modeled cohort size	Specify whether the study population was observed or simulated
AI intervention	AI technology; diagnostic application; role in the clinical pathway	Identify whether AI assisted, supplemented, or replaced existing diagnostic activities
Comparator	Standard care; alternative diagnostic strategy; no screening	Record the comparator used for each economic comparison
Economic evaluation	Evaluation type; perspective; analytical model; time horizon	Record the design and analytic assumptions reported by the authors
Economic inputs	Costing year; currency; discount rate	Record values separately and distinguish input-price year from currency-conversion year
Decision criteria	Willingness-to-pay threshold	Record the numerical threshold, currency, and outcome unit where applicable
Cost outcomes	Total costs; incremental costs; AI implementation costs; operating costs; cost savings	Preserve original currencies and population denominators
Clinical outcomes	Diagnostic performance; QALYs; life-years; clinical events; workflow outcomes	Specify outcome units, time horizons, and population denominators
Incremental economic findings	Incremental effectiveness; ICER; economic classification	Interpret dominance using the direction of incremental costs and effects
Uncertainty analysis	Deterministic sensitivity analysis; probabilistic sensitivity analysis; scenario analysis	Record principal assumptions and factors influencing economic conclusions
Reporting completeness	Missing or unclear information; supporting source location	Identify unavailable data without imputing numerical values
Verification	Source article; relevant table, figure, or supplementary material	Check extracted findings against the original publication

[Downloaded from jebhpme.ssu.ac.ir on 2026-11-12]

[DOI: 10.18502/jebhpme.v10i6.23195]

Appendix 4. Study-Specific Economic Parameters of the Included Evaluations

Supplementary Table S4. Costing Years, Currencies, Discount Rates, and Willingness-to-Pay Thresholds of the Included Economic Evaluations

Study	Costing year / price basis	Currency	Discount rate	WTP threshold
Gomez Rossi et al. (2022) (7)	Original price years not fully verified; 2020 PPP conversion reported	USD (melanoma); EUR (dental caries); BRL (diabetic retinopathy)	3% reported; application-specific verification required	Not verified
Xiao et al. (2021) (8)	Not verified	USD-equivalent outcomes reported in review; source currency basis requires checking	5% for projected healthcare costs	Not established for the cost-offset analysis
Chatzikou et al. (2025) (9)	Not verified; application-specific estimates	EUR	3% reported; application-specific verification required	No common QALY-based threshold established
Wong et al. (2025) (10)	Not verified	USD as reported in source results	3% annually	No QALY-based threshold established for reported life-year outcomes
Maas et al. (2025) (11)	2023	EUR	Base-case rate not yet confirmed; 0% and 5% assessed in sensitivity analyses	€33,000/QALY
Lyth et al. (2026) (12)	2023 price basis reported in previous extraction; final source check required	EUR as reported in review	3% reported; final source check required	Not verified
Goh et al. (2026) (13)	2023	SGD for cost inputs; USD also used in published economic results	3% annually for costs and outcomes	Approximately SGD 50,000/QALY stated in Methods; USD 50,000/QALY also appears in the published results
Levin et al. (2026) (14)	Model benchmarks inflation-adjusted to 2025	USD	Not reported for the observed 180-day comparisons	Not applicable; financial break-even analysis
Hill and Roadevin (2026) (15)	2023	GBP	3.5%	£20,000/QALY

[Downloaded from jebhpme.ssu.ac.ir on 2026-11-12]

[DOI: 10.18502/jebhpme.v10i6.23195]

 CC BY 4.0