Economic evaluations of AI-assisted technology in healthcare: how are we assessing the cost-effectiveness of these new therapies?

Yi Y[1], Subash V[1], Meiwald A[1], Hirst A[1], Heron L[1]
[1]Adelphi Values PROVE™, Adelphi Mill, Bollington, Cheshire, SK10 5JB, UK

ADELPHI VALUES | PROVE Expertise in Access and Value Evidence Outcomes

EE841

Introduction

> Artificial Intelligence (AI)-assisted technologies are being applied across a wide range of healthcare areas promising to enhance clinical outcomes, efficiency, and decision-making. As the deployment of these technologies accelerates, understanding their economic implications will become increasingly critical.
> Economic evaluations (EEs) of AI-assisted technologies in healthcare are growing in number, with a parallel increase in systematic literature reviews (SLRs) that synthesize and assess these evaluations. This aligns with recent developments, such as the introduction of the CHEERS-AI (Consolidated Health Economic Evaluation Reporting Standards for AI) guidelines and updates to the NICE (National Institute for Health and Care Excellence) evidence standards framework (ESF) for digital health technologies, which now include evidence standards specific to AI-based technologies.[1,2]
> The objective of this study was to understand the current value proposition of AI-interventions in healthcare, to identify gaps between existing evaluations and the requirements outlined by CHEERS-AI and NICE, and to highlight key considerations for future economic evaluations of emerging AI-driven technologies.

Methods

> An umbrella review, a comprehensive synthesis of evidence from multiple systematic reviews, was performed on SLRs of EEs of AI-assisted interventions.
> The ISPOR good practice for critical appraisal for SLRs with cost and costs-effectiveness outcomes (SR-CCEO) was followed.[3]
> Medline, EMBASE, Cochrane Central and NICE HTA databases were searched to identify SLRs of EEs for AI-assisted interventions.
> Publications were included based on predefined inclusion and exclusion criteria (Table 1).

Table 1. Inclusion and exclusion criteria

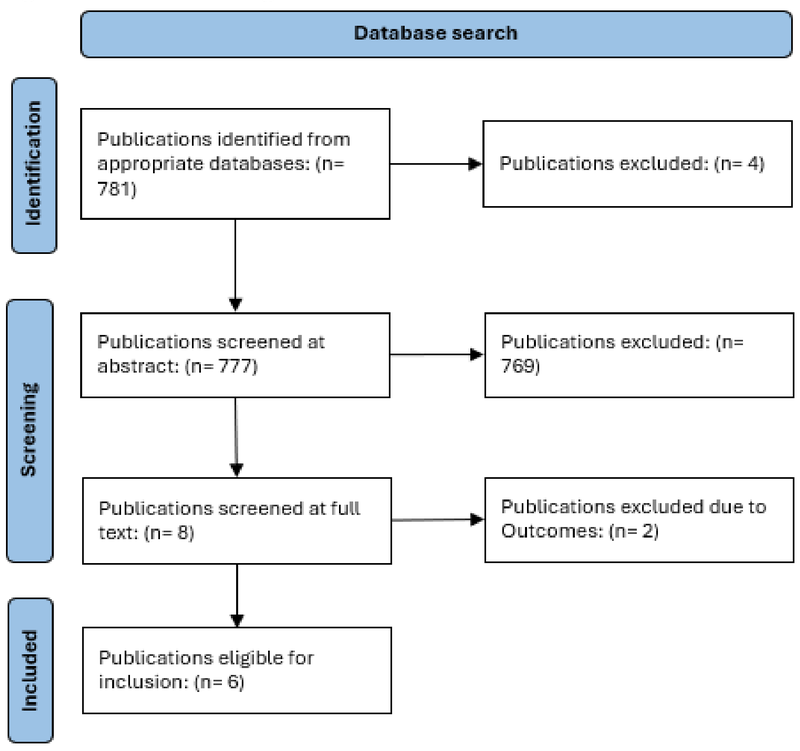
Types of criteria	Inclusion	Exclusion
Population	Actual or hypothetical patients who receive AI-assisted healthcare interventions	Patients receive healthcare without AI-technology involved
Intervention/ Comparator	AI-assisted healthcare interventions (for diagnosis, monitoring, surgery or other treatments) compared with standard/usual care without AI assistance	non-AI intervention; Healthcare intervention not compared to AI-assisted interventions
Outcome	Health economic evaluation methods or outcomes including cost effectiveness, cost utility, cost minimisation, cost benefit, cost consequences, costs, budget impact	Efficacy and/or safety only; No health economic evaluation outcomes
Study design	Systematic review of economic evaluations (model-based or trial-based studies, including cost effectiveness analysis (CEA), cost utility analysis (CUA), cost consequence analysis (CCA), cost minimisation analysis (CMA), budget impact analysis (BIA), health economic models)	Non-systematic reviews; Systematic reviews of clinical studies, abstract, posters, commentaries, letters, editorials and studies with no findings (e.g. protocol, methodology), animal studies, preprints
Study setting	No restrictions on country and healthcare settings	NA
Date of publication	2014 to 20 September 2024	Before 2014
Language of publication	English	Non-English

> Two independent reviewers conducted the screening, with any discrepancies resolved through discussion or by involving a third reviewer.
> Data extraction was carried out by one reviewer and verified by a second reviewer.
> Outcomes of interest included details of the interventions, disease area, EE methods, findings and quality assessment of the EEs included in the SLRs.
> The ISPOR CiCERO Checklist[3] was used to assess the quality of the included SLRs, and the EEs included in the SLRs.
> Due to heterogeneity in the SLRs and the original primary studies, narrative synthesis was performed.

Results

> 781 titles were identified, four removed for duplication, eight publications were screened at full text, with six SLRs being taken forward to data extraction (Figure 1).

Figure 1. PRISMA diagram



Results (continued)

> Six SLRs were included, all published after 2020. The key characteristics are summarised in Table 2.[4-9]

Table 2. Key characteristics of included studies

Search outcomes	Results (n=6)[4-9]
Databases searched	Embase (4), Cochrane Central (3), Web of Science (2), Scopus (2), Google search (2), HTA database (1)
Period searched	From inception (2), 5 years (3), unclear (1)
Language covered	English (6)
Funding source	Public organisation (3), No financial support (3); 1 not claimed
Conflict of interest	All 6 declared no conflict of interest
Disease area covered	General medicine/healthcare (3), Ophthalmology (3)
Type of AI-interventions	General medicines/healthcare focused search identified a breadth of AI-intervention including but not limited to pattern recognition, risk prediction models, and monitoring. Ophthalmology only identified EEs related to AI driven diagnosis and screening intervention.

A total 74 EEs were included in the six SLRs. After removing 21 duplications, 53 unique studies were identified, including 16 trial based and 37 model based EEs.

> Majority of the EEs were conducted in North America or Europe.
> Types of the EEs are shown in Figure 2. The category 'Other' (n=3, 6%) included profit analysis, regression analysis and financial performance analysis, which are not cost-effectiveness models.
> Perspectives taken by the EEs are shown in Figure 3. Decision trees (DT), Markov models or combinations were commonly used (Figure 4) with time horizons ranging from 28 days to lifetime.
> Key findings and limitations of the included SLRs are outlined in Table 3 and Figure 5.

Figure 2. Type of EEs (n=53)

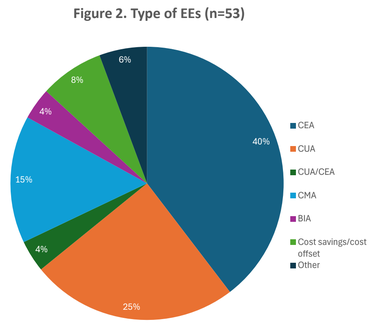


Figure 3. Perspectives used by the EEs (n=53)

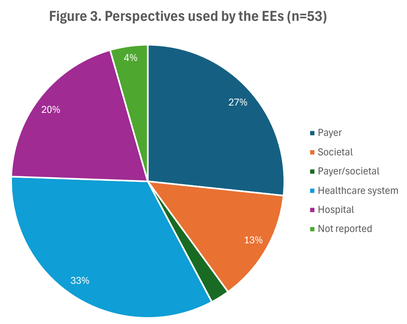


Figure 4. Types of models in CEA/CUA (n=32)

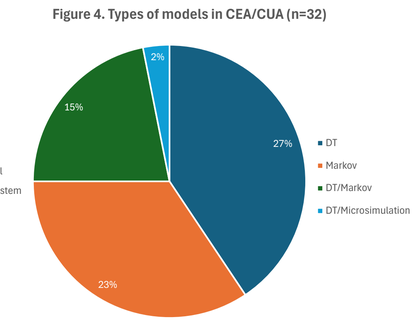
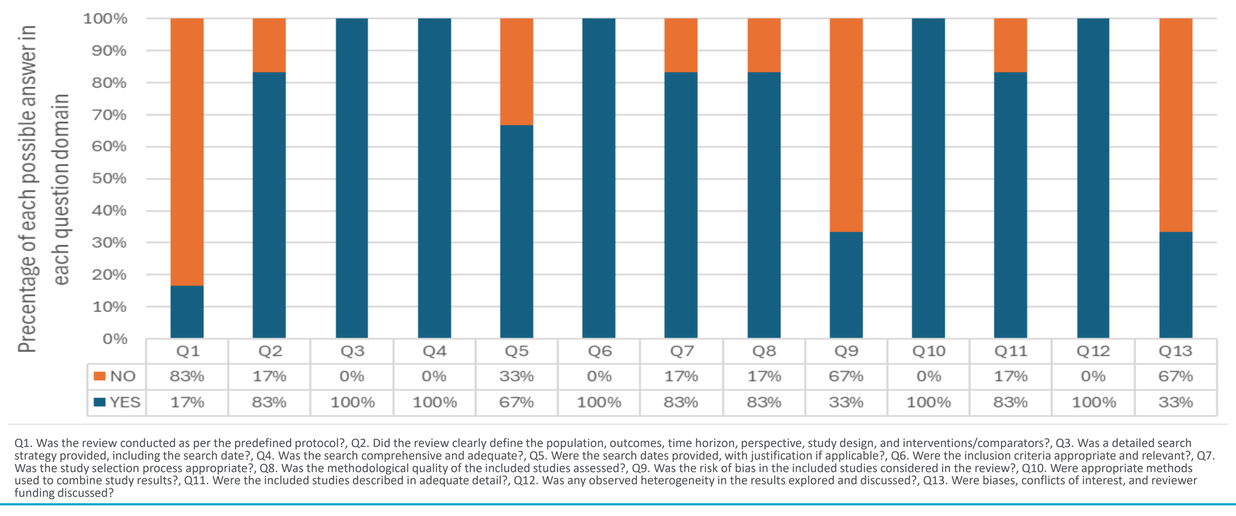


Table 3. Key findings and limitations of the included SLRs

Key findings for AI-interventions	Limitations of existing EEs
Economic Potential of AI for Cost Reduction and Efficiency in Screening: AI has substantial potential to lower costs and improve accessibility, particularly for screening in ophthalmology and diabetic retinopathy in low-resource settings. This is due to its ability to reduce the need for specialist involvement and enable early detection.[4,5,9] **Implementation Challenges:** Despite AI's promise, practical challenges exist, including regulatory barriers, high initial implementation costs, data privacy concerns, and the need for context-specific adaptations, particularly in LMICs.[4,5,8]	**Low Methodological Quality:** Many studies did not follow established best practices, making it challenging to compare results across different evaluations.[6,8] **Inadequate Real-World and Long-Term Data:** All the included EEs relied on short-term or hypothetical models, which might not reflect actual clinical and economic outcomes in practice and failed to accurately assess AI's impact over time.[5,7] **Lack of Comprehensive and Transparent Reporting:** The SLRs noted that EEs lack transparency and detail, with inadequate reporting on model structure, assumptions, and sensitivity analyses. This can reduce the studies' credibility and make it difficult for decision-makers to use the findings confidently.[4,6] **Gap Between AI Development and EE:** Rapid advancements in AI technology are not matched by the pace of EEs, which can hinder timely policy decisions. A faster, more agile approach to evaluation is recommended to keep up with AI's dynamic development.[7,8] The need for standardization in EEs of AI-interventions with consistent methods and reporting standards are highlighted in several SLRs to allow for more reliable comparisons across studies, ultimately supporting more informed decision-making.[6,7,8]

Figure 5. Application of the CiCERO checklist on the identified publications



Q1. Was the review conducted as per the predefined protocol?, Q2. Did the review clearly define the population, outcomes, time horizon, perspective, study design, and interventions/comparators?, Q3. Was a detailed search strategy provided, including the search date?, Q4. Was the search comprehensive and adequate?, Q5. Were the search dates provided, with justification if applicable?, Q6. Were the inclusion criteria appropriate and relevant?, Q7. Was the study selection process appropriate?, Q8. Was the methodological quality of the included studies assessed?, Q9. Was the risk of bias in the included studies considered in the review?, Q10. Were appropriate methods used to combine study results?, Q11. Were the included studies described in adequate detail?, Q12. Was any observed heterogeneity in the results explored and discussed?, Q13. Were biases, conflicts of interest, and reviewer funding discussed?

Conclusion

> As AI is increasingly used in healthcare, EEs of AI-assisted interventions and SLRs of these EEs are also increasing, but both are still limited in numbers, coverage, and quality.
> Although the existing SLRs are not without limitations, they recognize that AI has substantial potential to lower costs and improve accessibility, particularly for screening in ophthalmology and diabetic retinopathy in low-resource settings.
> Common gaps between existing EEs and CHEERS-AI reporting guideline[1] and NICE's ESF[2] included a lack of reporting transparency, inadequate real-world and long-term data, and adherence to standardized evaluation practices.
> CHEERS-AI reporting guideline[1] and NICE's ESF[2] both directly address weaknesses identified in current AI EEs, offering comprehensive, context-specific, and standardized guidelines that could substantially improve future evaluations' quality, transparency, and relevance in healthcare settings. They should be closely followed to improve the quality of the future EEs and ensure consistent and standardised methodologies, with transparent and reproducible reporting.

References:

1. Elvidge et al., Consolidated Health Economic Evaluation Reporting Standards for Interventions That Use Artificial Intelligence (CHEERS-AI), VALUE HEALTH. 2024; 27(9):1196–1205. **2**. NICE (2022) Evidence standards framework for digital health techonlogies, Corporate document, published 10 December 2018, last updated 9 August 2022. https://www.nice.org.uk/corporate/ecd7/resources/evidence-standards-framework-for-digital-health-technologies-pdf-1124017457605 **3**. Mandrik et al., Critical Appraisal of Systematic Reviews With Costs and Cost-Effectiveness Outcomes: An ISPOR Good Practices Task Force Report, VALUE HEALTH. 2021; 24(4):463–472 **4.** Pietris J, Lam A, Bacchi S, Gupta A.K, Kovoor J.G, Chan W.O (2022), Health Economics Implication of Artificial Intelligence Implementation for Ophthalmology in Australia: A Systematic Review, Asia Pac J Ophthalmol (Philia). 2022;11:554–562 **5**. Rizvi A, Rizvi F, Lalakia P, et al. (September 19, 2023) Is Artificial Intelligence the Cost-Saving Lens to Diabetic Retinopathy Screening in Low- and Middle-Income Countries?. Cureus 15(9): e45539. DOI 10.7759/cureus.45539 **6**. Vithlani J, Hawksworth C, Elvidge J, Ayiku L and Dawoud D (2023), Economic evaluations of artificial intelligence-based healthcare interventions: a systematic literature review of best practices in their conduct and reporting. Front. Pharmacol. 14:1220950. doi: 10.3389/fphar.2023.1220950 **7.** Voets M.M, Veltman J, Slump C, Siesling S, Koffijberg H (2022), Systematic Review of Health Economic Evaluations Focused on Artificial Intelligence in Healthcare: The Tortoise and the Cheetah, VALUE IN HEALTH. 2022; 25(3):340–349 **8.** Wolff J, Pauling J, Keck A, Baumbach J. The Economic Impact of Artificial Intelligence in Health Care: Systematic Review. J Med Internet Res. 2020 Feb 20;22(2):e16866. doi: 10.2196/16866. PMID: 32130134;PMCID: PMC7059082. **9.** Wu H, Jin K, Yip CC, Koh V, Ye J. A systematic review of economic evaluation of artificial intelligence-based screening for eye diseases: F rom possibility to reality. Surv Ophthalmol. 2024 Jul-Aug;69(4):499-507. doi: 10.1016/j.survophthal.2024.03.008. Epub 2024 Mar 15. PMID: 38492584.