iS
CHANNEL

Revisiting Health Information System Evaluation in the Age of Artificial Intelligence: A Critical Literature Review of Theoretical Perspectives and Evaluation Purposes

Caroline Anna Marie Ziegler[1]

MSc Management of Information Systems and Digital Innovation
Department of Management
London School of Economics and Political Science

KEYWORDS

HIS Evaluation
Technological Ontologies
AI in Healthcare
Evaluation Purposes
Sociotechnical Systems
Institutional Theory

ABSTRACT

Despite decades of research, the evaluation of health information systems continues to be characterised by unresolved disagreements regarding evidence, success and accountability. These tensions are intensifying as artificial intelligence is increasingly integrated into clinical systems. The critical literature review examines current research on health information system evaluation from information systems and medical informatics journals published between 2000 and 2025 to analyse how controversies are shaped by underlying assumptions about technology and the expected goals of evaluation. Using a matrix framework that relates technology concepts to evaluation purposes, the analysis identifies connections between instrumental, sociotechnical, and institutional perspectives and corresponding evaluation logics, showing that controversies reflect ontological and normative differences rather than methodological pluralism. For AI-enabled health information systems, the study identifies a growing risk that governance-driven legitimisation could stabilise systems through standards and compliance before a practice-based understanding has been developed. The study proposes a socially embedded evaluation approach for future research that emphasises how AI is reshaping agency, accountability and learning in clinical practice.

1. Introduction

The increasing legitimation and mobilisation of artificial intelligence (AI) mark a new technological era, with healthcare among the most affected sectors (Davenport & Kalakota, 2019; Topol, 2019). In the United States alone, widespread AI adoption is projected to generate net savings of approximately 200 to 350 billion US dollars within the next five years (Sahni et al., 2023). When embedded in health information systems (HIS), AI is expected to transform not only diagnostic and therapeutic practices (Serin et al., 2025), but also how hospitals communicate, coordinate work, and distribute authority (Cross et al., 2024; Olawade et al., 2025). Understanding these shifts requires attention to evaluation as a mechanism through which control, learning, and legitimation are produced (Avgar et al., 2018; Mohsin Khan et al., 2025; Rozenblit et al., 2025).

Evaluation of HIS has a long tradition in information systems (IS) and medical informatics (MI), generating an extensive body of research (Neame et al., 2020; Yusof et al.,

[1] Corresponding Author
Email Address: Caroline.ziegler@ziegler-mueller.de

2007). Nevertheless, there is still considerable debate about how HIS should be evaluated, what constitutes success, and whose perspectives should inform such assessments (Adjerid et al., 2018; Venkatesh et al., 2011). As AI becomes increasingly embedded in HIS, these debates gain renewed urgency, as existing evaluation approaches are confronted with AI systems whose adaptive and probabilistic operation complicates interpretability in clinical practice (Lebovitz et al., 2021). Based on these considerations, this literature review examines how established approaches to HIS evaluation are shaped by underlying assumptions about technology and the expected goals of evaluation, and how these assumptions are reinforced as AI becomes increasingly integrated into clinical systems. Subsequently, the following research question arises: "How do divergent technological ontologies and evaluative purposes underpin persistent controversies in HIS evaluation, particularly considering the increasing implementation of AI-enabled systems?"

The literature review addresses this question by critically analysing the latest HIS evaluation research contributions from the last 25 years. The review adopts a critical rather than a systematic approach and is informed by targeted research in IS and MI journals, including empirical studies, conceptual framework contributions, and systematic reviews. To ground the analysis theoretically, a matrix-based framework is employed integrating two complementary theoretical perspectives. The first examines technology through instrumental, sociotechnical, and institutional lenses (Goodhue & Thompson, 1995; Kling & Iacono, 1989; Orlikowski, 2000). The second draws on evaluation theory to distinguish between control-, learning-, and legitimation-oriented purposes of evaluation (Chelimsky, 1997; Preskill & Torres, 1999; Dahler-Larsen, 2012). Structurally, the underlying theoretical frameworks will be first introduced, followed by the analysis of HIS evaluation research through the matrix framework, before areas of convergence and controversy will be synthesised and implications for evaluating AI-enabled HIS will be discussed.

2. Theoretical Framework

The theoretical framework integrates two analytical lenses, capturing how conceptions of technology shape understandings of HIS at a meta-level while distinguishing core evaluative purposes at a functional level.

2.1 Instrumental, Sociotechnical and Institutional Views of Technology

Research on information systems and organisations can be distinguished by three dominant perspectives on technology that imply different understandings of HIS and their effects. Instrumental perspectives, rooted in early IS success and task–technology fit models, conceptualise technology as a bounded functional artefact whose value is assessed through measurable performance, efficiency, and usage outcomes (DeLone & McLean, 1992; Goodhue & Thompson, 1995). Human and organisational factors are treated primarily as variables influencing effectiveness, reinforcing managerial logics of optimisation and control (Kling, 1987; Latour, 1987). Sociotechnical perspectives challenge this view by rejecting assumptions of technological neutrality and instead conceptualising technology as enacted through ongoing interactions between social structures and material artefacts (Orlikowski, 1992; Orlikowski, 2000). Technological effects are understood as emergent and contextual, shaped through use and adaptation, shifting analytical attention from system performance to technology-in-use within organisations. Institutional perspectives share a similar set of assumptions on social embeddedness but extend the analytical focus to macro-level conditions under which technologies and their evaluation

acquire legitimacy. Technologies are embedded in political and regulatory frameworks that determine what is defined as necessary and acceptable (Avgerou, 2000; Kling & Iacono, 1989). Thus, evaluation is examined as an institutional process through which technological trajectories are authorised and organisational actions stabilised (Klecun & Cornford, 2005; Klecun et al., 2014).

2.2 Evaluation as Control, Learning and Legitimation

Across evaluation theory and organisational research, evaluation is predominantly associated with three analytically distinct but overlapping purposes: control, learning, and legitimation. Control-oriented traditions conceptualise evaluation as a mechanism of accountability and performance steering, closely linked to audit practices, comparability, and managerial oversight (Behn, 2003; Chelimsky, 1997; Power, 1997). As captured in March's (1991) exploitation–exploration distinction, these approaches risk prioritising optimisation over adaptability. Learning-oriented perspectives respond to this tension by framing evaluation as a developmental process through which actors reflect on experience, revise assumptions, and adapt practices over time (Preskill & Torres, 1999). Legitimacy-oriented perspectives emphasise the role of evaluation in creating credibility, acceptance and consistency with institutional expectations, thereby supporting organisational action in political, regulatory and professional environments (Dahler-Larsen, 2012; Meyer & Rowan, 1977).

3. Critical Review Methodology

The review adopts a critical rather than systematic approach, reflecting a deliberate epistemological choice rather than a methodological limitation. Critical literature reviews are oriented toward conceptual interrogation, examining how a field constructs its objects of inquiry, what ontological and normative assumptions underpin dominant approaches, and where structural tensions and explanatory gaps persist across the accumulated body of scholarship (Hart, 1998; Snyder, 2019; Torraco, 2005). Rather than treating the literature as a repository of findings to be understood, critical surveys engage with it as an intellectual field whose controversies are analytically significant in themselves and require explanation. The evaluation of HIS represents precisely such an area, since despite decades of empirical contributions, framework developments and systematic reviews, there remain profound differences of opinion about what constitutes success, what counts as valid evidence, and whose perspectives should be taken into account in the evaluation (Ammenwerth et al., 2003; Greenhalgh et al., 2017; Neame et al., 2020). Thus, a critical approach is analytically necessary to move beyond cataloguing methodological variation toward explaining why controversies persist and what assumptions sustain them.

The literature corpus was constructed through targeted searches across Google Scholar, PubMed, and the ACM Digital Library, using search terms combining "health information system evaluation", "HIS evaluation framework", "digital health evaluation theory", and "AI health information system evaluation". Searches were concentrated on peer-reviewed contributions from journals central to IS and MI scholarship. Additional contributions were identified through backward citation tracing from highly cited empirical evaluation, framework and review papers, ensuring foundational theoretical works were incorporated even where they predated the primary search window. The search covered publications from 2000 to 2025, capturing contemporary developments in HIS evaluation, including emerging challenges associated with AI (Ammenwerth et al., 2003; Cresswell et al.,

2024; Neame et al., 2020). The final corpus comprised 50 empirical studies, conceptual contributions, and systematic reviews.

Table 1: Classification of the Reviewed Literature by Perspective and Type of Contribution

	Empirical Evaluations	**Evaluation Framework**	**Literature Reviews****
Instrumental perspective*	Adjerid et al. (2018) Barber et al. (2006) Farzandipour et al. (2022) Menon et al. (2000) Tsiknakis & Kouroubali (2009)	Ancker et al. (2012) Dixon et al. (2010) Ji et al. (2021)	Ahmadian et al. (2015) Andargoli et al. (2017) Ammenwerth et al. (2003) Ammenwert et al. (2004) Nykänen & Kaipio (2016) Wronikowska et al. (2021)
Sociotechnical perspective*	Greenhalgh et al. (2013) Hendy et al. (2025) Ostervang et al. (2024) Westbrook et al. (2007) Yusof (2015)	Effken (2002) Greenhalgh et al. (2017) Jacob et al. (2025) Kaplan & Harris-Salamone (2009) Kaplan & Shaw (2004) Lagsten (2011) Reddy et al. (2021) Shaw (2002) Sockolow et al. (2012) Yusof et al. (2008)	Cresswell & Sheikh (2014) Holden & Karsh (2010) Lilford et al. (2009) Neame et al. (2020) Sligo et al. (2017) Yusof et al. (2007)
Institutional perspective*	Abdelgawad et al. (2024) Addo & Aqyepong (2025) AlHussainan & Alhuwail (2025) Ha et al.(2025) Lebovitz et al. (2021) Woods et al. (2024)	Klecun & Cornford (2005) Klecun et al. (2014) Takian et al. (2012) Vassey et al. (2022)	Cresswell et al. (2024) Magrabi et al. (2016) Noël et al. (2022) Sacramento et al. (2025)

* The perspectives and corresponding allocation of papers displayed represent the dominant conceptual framing of each study, while recognising that many contributions incorporate multiple, overlapping theoretical logics.

** The allocation reflects the prevailing analytical orientation of each literature review rather than a rigid theoretical assignment, acknowledging that review studies in particular frequently integrate multiple perspectives on technology and evaluation.

Studies were selected to achieve theoretical breadth across instrumental, sociotechnical, and institutional perspectives rather than exhaustive coverage, prioritising contributions that explicitly engage with questions of evaluation purpose, methodology, or theoretical framing. Each study was analysed to identify the dominant theoretical perspective underlying its primary conceptual framing of technology and its stated or implied evaluative orientation, drawing on established distinctions in IS and evaluation theory to guide the interpretive process (Hart, 1998; Klecun & Cornford, 2005; Orlikowski, 2000). Where contributions drew on multiple perspectives, classification reflected the overarching analytical logic of the study rather than a rigid categorical assignment, acknowledging that theoretical boundaries in the field are frequently porous and contested. The interpretive dimensions of this process are consistent with the epistemological commitments of critical review methodology, in which analytical judgement is not a source of weakness but a condition of theoretically informed engagement with the literature (Snyder, 2019; Torraco, 2005). Table 1 provides an overview of the reviewed literature classified by perspective and contribution type.

4. Critical Analysis of Theoretical Perspective

4.1 Instrumental HIS Evaluation: Technology as Tool

Instrumental HIS evaluation is dominated by a logic of control that frames success as the optimisation of predefined objectives like efficiency, cost reduction, and adoption (Ancker et al., 2012; Barber et al., 2006; Menon et al., 2000). Despite methodological diversity, the instrumental strand shares the assumption that HIS system value can be rendered visible through quantifiable outcomes and standardised indicators that enable causal attribution between implementation and effects (Dixon et al., 2010; Ji et al., 2021). Economic evaluations operationalise success through productivity and expenditure effects (Adjerid et al., 2018; Menon et al., 2000), while safety and usability studies rely on standardised performance metrics (Ancker et al., 2012; Barber et al., 2006; Farzandipour et al., 2022; Wronikowska et al., 2021). Framework and review contributions reinforce this orientation by formalising what is evaluated and how, thereby privileging rigour, validity, and comparability as core evaluative criteria (Ahmadian et al., 2015; Ammenwerth et al., 2004; Andargoli et al., 2017; Nykänen & Kaipio, 2016).

The instrumental evaluative logic is grounded in an ontology that treats HIS as bounded technical artefacts whose effects can be analytically separated from their social context of use (Goodhue & Thompson, 1995). HIS are framed as tools acting upon organisations, with impacts assumed to be identifiable through performance indicators, economic models, or usability metrics (Dixon et al., 2010; Farzandipour et al., 2022; Menon et al., 2000). Fit based approaches extend the analytical scope to human and organisational factors while retaining the assumption that misalignments can be resolved through improved design and measurement (Tsiknakis & Kouroubali, 2009). Consequently, instrumental evaluation prioritises functions such as performance monitoring, risk management, and accountability, aligning evaluation closely with managerial oversight and external reporting requirements (Adjerid et al., 2018; Barber et al., 2006; Menon et al., 2000). Learning remains largely confined to incremental system refinement rather than reflective engagement with changing practices, while legitimation operates implicitly by stabilising narratives of efficiency and safety that justify continued investment and policy support without becoming explicit objects of evaluation (Ammenwerth et

al., 2004; Ji et al., 2021; Menon et al., 2000). Although instrumental approaches enable standardisation and comparability, they obscure how meanings, responsibilities and practices evolve in use. An omission that has consequences for AI-enabled HIS, where systems achieve good indicator values but simultaneously reshape clinical reasoning (Lebovitz et al., 2021).

4.2 Sociotechnical HIS Evaluation: Technology as Practice

In contrast to instrumental approaches, sociotechnical HIS evaluation treats outcomes as emergent rather than predefined, directing analytical attention to use and practice (Orlikowski, 2000). Thus, evaluation focuses on implementation trajectories, appropriation, and the reconfiguration of work practices, shifting attention away from retrospective assessments of system effects (Greenhalgh et al., 2013; Westbrook et al., 2007). Success is understood through integration into clinical routines, professional judgement, and organisational coordination instead of performance indicators alone (Cresswell & Sheikh, 2014; Westbrook et al., 2007). This understanding supports the use of qualitative, mixed method, and longitudinal designs that foreground situated use, stakeholder interpretations, and temporal dynamics (Kaplan & Shaw, 2004; Lagsten, 2011; Ostervang et al., 2024; Sligo et al., 2017). Methodological pluralism is therefore treated as necessary for capturing the emergent character of HIS practices, including variation across organisational layers (Hendy et al., 2025; Lilford et al., 2009; Neame et al., 2020).

At a conceptual level, sociotechnical evaluation rests on a relational view of technology that rejects assumptions of neutrality and fixed effects. HIS are understood as socio material arrangements whose meanings and consequences arise through enactment in practice rather than residing in technical artefacts alone (Effken, 2002; Orlikowski, 2000). Outcomes emerge from recursive interactions among material features, organisational routines, professional norms, and local adaptations, rendering variation in usage analytically informative rather than symptomatic of failure (Greenhalgh et al., 2017). Sociotechnically informed fit and acceptance models, including the HOT fit framework, reflect this view by incorporating human, organisational, and contextual dimensions into evaluation (Holden & Karsh, 2010; Yusof et al., 2007; Yusof et al., 2008; Yusof, 2015). Hence, sociotechnical work foregrounds learning through reflection, adaptation, and sensemaking, while control-oriented mechanisms are frequently questioned for obscuring the complexity of clinical work (Kaplan & Harris-Salamone, 2009; Shaw, 2002; Sockolow et al., 2012). Legitimation tends to remain implicit, as evaluations may reinforce confidence and continued investment by demonstrating adaptive implementation without being articulated as an explicit evaluative aim (Jacob et al., 2025; Reddy et al., 2021). By foregrounding learning and situated use, sociotechnical studies provide critical insight into why technologies succeed or fail in real clinical settings (Cresswell & Sheikh, 2014; Westbrook et al., 2007). At the same time, sociotechnical research often resists stabilised success criteria, which can limit its capacity to inform large scale decision making, regulation, and resource allocation (Lilford et al., 2009; Neame et al., 2020). In AI-enabled HIS, this tension can become particularly salient, as richly contextualised insights must compete with demands for standardisation, comparability, and actionable evaluation outputs (Greenhalgh et al., 2013; Kaplan & Shaw, 2004). Without mechanisms linking sociotechnical learning to governance, learning risks remaining localised.

4.3 Institutional HIS Evaluation: Technology as Power, Legitimacy and Political Artefact

Institutional perspectives conceptualise HIS evaluation as a constitutive practice through which technologies become authorised and stabilised within healthcare systems (Avgerou, 2000; Kling & Iacono, 1989). Rather than treating stability as the result of demonstrable technical effectiveness, the literature emphasises how alignment with regulatory expectations, professional norms, and policy agendas enables technologies to endure and expand (Klecun & Cornford, 2005; Klecun et al., 2014; Takian et al., 2012). Evaluation is therefore not positioned as an external judgement applied after implementation, but as an institutional process that shapes technological trajectories by defining what counts as acceptable evidence, appropriate risk, and responsible innovation (Avgerou, 2000; Dahler-Larsen, 2012; Meyer & Rowan, 1977). From this viewpoint, evidence itself is understood as institutionally produced and interpreted, meaning that evaluation participates directly in constructing the criteria by which technologies are judged and legitimised.

As a result, HIS operate as governance artefacts by embedding regulatory rules and professional hierarchies directly into technical systems (Klecun & Cornford, 2005; Abdelgawad et al., 2024; Addo & Aqyepong, 2025). Through this, they shape the organisation's actions by determining which activities can be monitored, compared or sanctioned, and thus often guide behaviour without the need for direct action on the part of management or supervisory authorities (Avgerou, 2000; Woods et al., 2024). Subsequently, evaluation plays a central role in this process by establishing evidentiary thresholds, compliance practices, and accountability mechanisms that support system level oversight (AlHussainan & Alhuwail, 2025; Noël et al., 2022; Sacramento et al., 2025). Legitimation emerges as the dominant evaluative function, particularly in AI-enabled HIS, where standards, reporting requirements, and early clinical evaluations constitute the conditions under which algorithmic systems are recognised as credible clinical actors (Cresswell et al., 2024; Lebovitz et al., 2021; Vasey et al., 2022). Control is exercised mainly through governance mechanisms, with evaluation enabling oversight by defining standards, accreditation requirements, and compliance procedures rather than by directly enforcing performance targets (Noël et al., 2022; Sacramento et al., 2025). Learning remains present but is largely channelled through institutional mechanisms, as safety incidents or failures prompt regulatory recalibration while being constrained by concerns over public trust, investment protection, and reform continuity (Ha et al., 2025; Magrabi et al., 2016). Subsequently, a central risk of institutional evaluation is that legitimacy may be secured faster than critical learning can occur. In AI-enabled HIS, systems may therefore become stabilised through compliance and standards before their organisational and clinical consequences are fully understood, making later contestation or redesign difficult (Avgerou, 2000; Dahler-Larsen, 2012; Lebovitz et al., 2021). Addressing this risk ultimately requires evaluation approaches that move beyond compliance-oriented legitimation toward socially embedded practices capable of keeping pace with the evolving organisational and clinical consequences of AI integration.

5. Critical Synthesis

Across three decades of HIS evaluation research, a set of shared orientations and persistent points of contention can be identified. Despite substantial theoretical and methodological diversity, the literature broadly converges on the recognition that HIS interventions unfold in complex, context-dependent, and multi-actor environments, limiting the plausibility of purely technical or decontextualised evaluation approaches (Ammenwerth et al., 2003; Cresswell & Sheikh, 2014; Klecun et al., 2014). There is also growing agreement that single-

method and purely cross-sectional designs are insufficient to capture temporal dynamics, unintended consequences, and adaptive processes associated with digital health interventions (Lilford et al., 2009; Neame et al., 2020; Sligo et al., 2017). Moreover, scholars increasingly emphasise the inclusion of multiple stakeholder perspectives, including clinicians, managers, patients, and regulators, to avoid systematically partial representations of HIS effects (Cresswell et al., 2024; Magrabi et al., 2016; Yusof et al., 2007). These areas of convergence indicate a gradual shift away from narrow technical notions of evaluation toward more relational and process-sensitive understandings.

At the same time, the literature remains structured by enduring controversies that resist resolution. Central among these is disagreement over what constitutes "successful" HIS evaluation and who is entitled to define it. Instrumental approaches prioritise efficiency, safety, and cost containment, sociotechnical studies emphasise situated usefulness and practice change, and institutional analyses foreground legitimacy, policy alignment, and regulatory credibility, producing evaluative logics that remain only partially commensurable (Ammenwerth et al., 2003; Greenhalgh et al., 2017). Closely related tensions concern whether standardised evaluation frameworks can be meaningfully transferred across heterogeneous organisational, national, and technological contexts (Sligo et al., 2017; Yusof et al., 2007; Sacramento et al., 2025). Further controversy surrounds the status of objectivity, as sociotechnical and institutional research repeatedly demonstrates how evaluation outcomes are shaped by governance arrangements, professional interests, and epistemic conventions rather than neutral measurement alone (Lebovitz et al., 2021; Magrabi et al., 2016). An overview of these areas of consensus and controversy is provided in Figure 1.

Figure 1: Areas of Consensus and Controversy in HIS Evaluation Literature

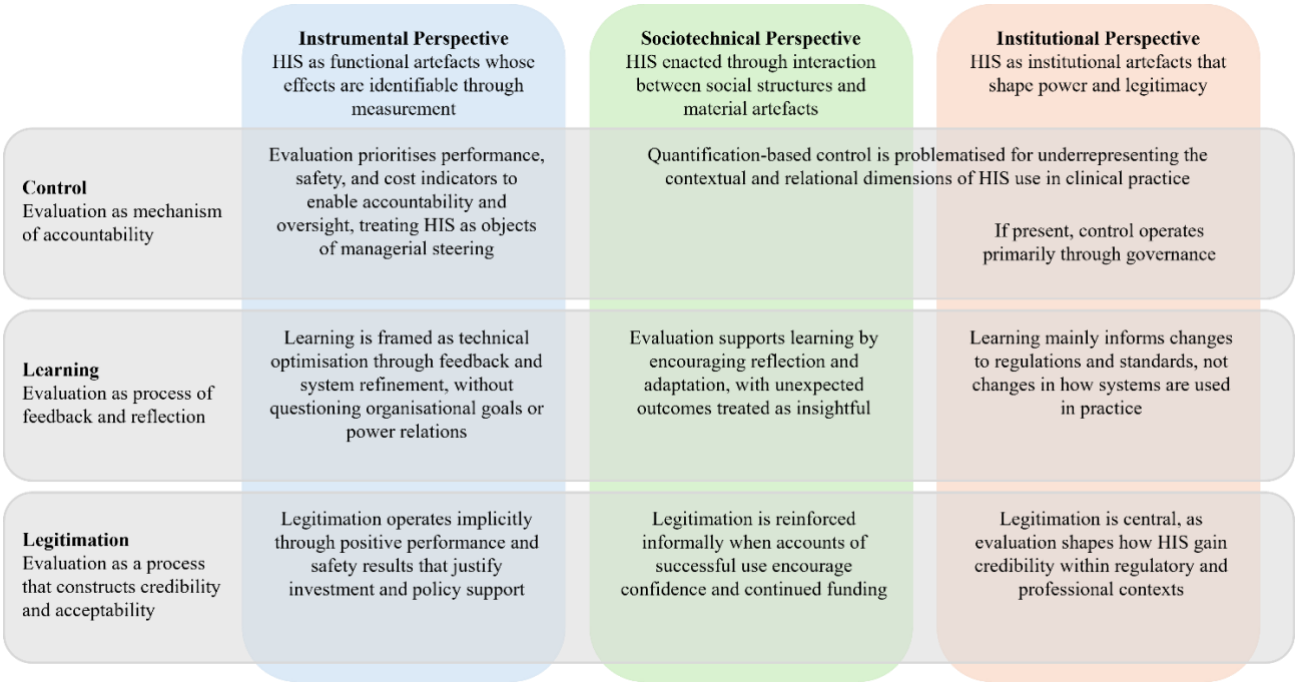
Theme	Consensus	Controversies
Complexity and Context-Dependence	HIS evaluation is widely understood as inherently complex and strongly shaped by organisational structures, professional routines, and institutional conditions (Kaplan & Shaw, 2004; Greenhalgh et al., 2017; Ammenwerth et al., 2003).	*Whether formalised evaluation approaches can adequately capture this complexity without reducing it to abstract variables remains disputed (Klecun & Cornford, 2005; Takian et al., 2012).*
Multi-Method and Multi-Stakeholder Evaluation	Meaningful HIS evaluation is broadly agreed to require multiple methods and the inclusion of diverse stakeholder perspectives, including clinicians, managers, and technical staff (Yusof et al., 2008; Yusof, 2015).	*The balance between methodological pluralism and demands for standardisation and comparability remains unresolved (Lilford et al., 2009; Neame et al., 2020).*
Longitudinal and Iterative Evaluation	Evaluation is widely conceptualised as a longitudinal and iterative process due to the evolving nature of system use and organisational adaptation (Venkatesh et al., 2011; Greenhalgh et al., 2017; Cresswell & Sheikh, 2014).	*Tension persists between the need for long-term learning and policy pressures for early, decisive evidence of success (Magrabi et al., 2016; Sligo et al., 2017).*
What Counts as "Success"	Success is generally acknowledged as multi-dimensional, encompassing technical performance, safety, usability, and organisational fit (DeLone & McLean, 1992; Yusof et al., 2008; Nykänen & Kaipio, 2016).	No agreement exists on which dimensions should dominate, with competing emphases on efficiency, practice change, and institutional legitimacy (Menon et al., 2000; Greenhalgh et al., 2017; Cresswell et al., 2024).
Role of Formal Evaluation	Evaluation is widely viewed as central for accountability, learning, and decision-making in HIS and AI implementation (Ammenwerth et al., 2004; Kaplan & Shaw, 2004; Cameron et al., 2011).	Critics argue that formal evaluation frameworks risk oversimplifying sociotechnical realities and marginalising informal, practice-based knowledge (Klecun & Cornford, 2005; Westbrook et al., 2007).
Generalisability of Frameworks	Evaluation frameworks are widely used to structure assessment practices across diverse settings (Yusof et al., 2007; Andargoli et al., 2017; Ahmadian et al., 2015).	The transferability of frameworks across heterogeneous organisational, institutional, and national contexts remains contested (Burton-Jones et al., 2020; Sligo et al., 2017).
Objectivity vs. Subjectivity	Quantitative indicators are widely valued for transparency, accountability, and comparability (Menon et al., 2000; Nykänen & Kaipio, 2016).	Sociotechnical and institutional perspectives challenge claims of objectivity by highlighting interpretation, power, and epistemic authority in evaluation (Klecun & Cornford, 2005; Lebovitz et al., 2021; Greenhalgh et al., 2017).

✗ Light grey shading indicates areas of relatively high conceptual convergence across perspectives, with differences primarily concerning emphasis rather than fundamental disagreement.

While Figure 1 summarises points of consensus and controversy, the matrix analysis offers a deeper explanation of why disagreements persist across the HIS evaluation literature. By systematically relating technological ontologies to evaluative purposes, the matrix, being illustratively summarised in Figure 2, reveals that debates are rooted in divergent assumptions about what technology is, how effects are produced, and what evaluation is meant to achieve (Greenhalgh et al., 2017; Klecun & Cornford, 2005). Disagreement therefore does not primarily arise from insufficient data or inadequate methods, but from structurally different configurations that shape what is rendered visible, relevant, and actionable through evaluation.

Instrumental configurations couple a technology-as-tool ontology with control-oriented evaluation, producing a logic centred on verification, optimisation, and risk management (Ammenwerth et al., 2003; Ahmadian et al., 2015). Within this configuration, success is framed in terms of efficiency, safety, and cost containment, while organisational dynamics and professional judgement are treated as secondary or confounding factors (Menon et al., 2000; Barber et al., 2006). The matrix clarifies that controversies surrounding narrow success definitions or claims of objectivity follow directly from these assumptions about causal attribution and evaluative authority (Nykänen & Kaipio, 2016).

Figure 2: HIS Evaluation Literature Matrix Analysis



Sociotechnical configurations align a technology-as-practice ontology with learning-oriented evaluation. Evaluation here is oriented toward understanding appropriation, adaptation, and sensemaking in use, rather than enforcing predefined performance targets (Westbrook et al., 2007; Cresswell & Sheikh, 2014; Greenhalgh et al., 2017). The matrix can help explain why sociotechnical studies keep questioning standardised indicators and cross-context comparability, since these criticisms originate from an ontology that sees results as emerging and context-dependent rather than stable effects (Lilford et al., 2009; Sligo et al., 2017). At the

same time, the analysis reveals internal tensions in which learning-oriented evaluation struggles to translate local findings into forms that are readable for policymakers and large-scale decision-making processes (Neame et al., 2020).

Institutional configurations link technology as an institutional artefact with legitimacy-oriented evaluation and redefine evaluation as a mechanism of governance rather than an instrument for performance assessment (Klecun & Cornford, 2005; Takian et al., 2012). Within this configuration, evaluation stabilises reform processes, approves investments and ensures the credibility of regulation, particularly in the case of large-scale digital programmes and AI-supported systems (Klecun et al., 2014; Cresswell et al., 2024; Lebovitz et al., 2021). The analysis illustrates that institutional control does not work through performance indicators, but rather through the establishment of standards, accreditations and evidence conventions, thereby shaping what is considered an acceptable risk, responsibility and knowledge (Magrabi et al., 2016; Noël et al., 2022).

Across all three configurations, the matrix analysis indicates that control, learning, and legitimation are rarely absent, but rather are hierarchically ordered and articulated differently depending on the prevailing technological ontology. These findings contribute to explaining why efforts to "improve" HIS assessment by adding methods or expanding indicator sets often fail to resolve the underlying controversies (Ammenwerth et al., 2003; Greenhalgh et al., 2017). Without considering the configuration of technological assumptions and evaluation purposes, interventions risk reproducing existing tensions rather than overcoming them. Therefore, the matrix offers a diagnostic perspective that goes beyond merely cataloguing approaches and explains the structural persistence of disagreements in HIS evaluation research.

6. Implications for AI in HIS

The analysis provides a basis for rethinking how AI-enabled HIS should be evaluated. AI exacerbates familiar evaluation problems by replacing stable system behaviour with adaptive and probabilistic processes that fundamentally complicate assessment in clinical practice (Cross et al., 2024; Lebovitz et al., 2021). Instrumental indicators such as accuracy and performance benchmarks remain necessary but are insufficient on their own, as relying on them exclusively risks reproducing the limitations of tool-oriented evaluation approaches that abstract technology from its conditions of use (Ji et al., 2021; Vasey et al., 2022). Quantitative metrics are therefore most analytically productive when embedded within sociotechnical evaluation designs that examine how algorithmic outputs are interpreted in clinical reasoning and how responsibility and trust are redistributed between clinicians and systems (Cresswell & Sheikh, 2014; Greenhalgh et al., 2017; Westbrook et al., 2007).

From an institutional perspective, AI evaluation increasingly functions as a site of legitimation, where standards, governance frameworks, and early clinical evaluation protocols shape the conditions under which AI becomes acceptable in healthcare (Cresswell et al., 2024; Olawade et al., 2025; Vasey et al., 2022). A key risk arising from the situation is that legitimisation creates a learning gap in which institutional legitimisation outpaces sociotechnical understanding. Compliance with standards, reporting requirements, and early trials may stabilise AI systems before their implications for clinical reasoning, professional authority, and work practices are adequately understood. Empirically, the gap manifests as rapid regulatory approval and organisational adoption alongside limited insight into how AI systems are enacted, adapted, or contested in everyday clinical practice (Magrabi et al., 2016; Mohsin Khan et al., 2025).

As identified in section 4.3, existing institutional frameworks are not structurally equipped to close this gap, precisely because their evaluative logic privileges legitimation over learning.

AI-enabled HIS therefore require a shift from universal evaluation models toward approaches that are explicitly sensitive to the organisational and social contexts in which algorithmic systems are embedded (Cross et al., 2024). For narrowly scoped applications with relatively stable workflows and limited discretion, traditional evaluation designs may remain sufficient. As AI systems increasingly reconfigure clinical decision-making, professional authority, and accountability structures, however, evaluation must move beyond outcome validation to examine how agency and responsibility are redistributed between human and algorithmic actors in practice (Anthony et al., 2023; Lebovitz et al., 2021). This demands an evaluative orientation that treats learning not as a terminal assessment of system performance but as an ongoing, reflexive process through which clinicians, organisations, and governance bodies collectively develop understanding of AI in use.

These considerations together point toward a socially embedded evaluation approach as a necessary corrective to the limitations identified across all three configurations. Such an approach would be distinguished by four analytical principles. First, evaluation is treated as constitutively embedded in social and organisational practice rather than applied to it from the outside (Klecun & Cornford, 2005; Orlikowski, 2000). Second, agency and accountability are analysed as distributed across human and algorithmic actors rather than assigned to systems or individuals in isolation (Anthony et al., 2023; Lebovitz et al., 2021). Third, learning is foregrounded as a continuous and reflexive process rather than a terminal judgement of system effectiveness (Preskill & Torres, 1999; Westbrook et al., 2007). Fourth, institutional legitimation mechanisms are subjected to critical scrutiny rather than adopted as evaluative endpoints, ensuring that governance processes remain responsive to emerging sociotechnical understanding (Dahler-Larsen, 2012; Meyer & Rowan, 1977). Operationalising these principles requires evaluation designs that integrate longitudinal, qualitative, and participatory methods with governance-level analysis, thus moving beyond the partial commensurability of existing configurations toward a more integrated and reflexive evaluative practice (Greenhalgh et al., 2017; Kaplan & Shaw, 2004; Neame et al., 2020).

7. Limitations

The critical literature review is subject to several limitations that must be considered. First, the analysis does not follow a completely systematic protocol for literature review. Although the selected literature is broad and theoretically sound, it inevitably remains selective and cannot claim to be exhaustive. The implication is that relevant contributions may not have been included. Second, the review deliberately focuses on literature from the last 25 years to reflect current developments in the field of HIS and digital transformation. Earlier fundamental work is therefore only considered selectively, which may limit the historical depth of the analysis. Third, the overview takes a conceptual perspective at the meta-level and does not provide a detailed examination of specific technologies, implementations, or individual evaluation frameworks. As a result, important domain-specific dynamics may be underrepresented. Fourth, the classification of studies according to instrumental, sociotechnical, and institutional perspectives, as well as the distinction between control-, learning-, and legitimation-oriented evaluation purposes, is interpretive in nature. Although grounded in theory, the categorisation inevitably simplifies more nuanced positions within the literature.

Fifth, the analysis relies exclusively on published research findings and is therefore subject to publication bias, including an overrepresentation of large-scale programs and positive results. Finally, the discussion of AI-related evaluations is limited by the nascent nature of this field of research and the currently limited empirical evidence base.

8. Conclusion

By critically examining the literature, the review demonstrates that evaluation practices in HIS are shaped not only by methodological choices but by underlying assumptions about technology and the purposes evaluation is expected to serve. By systematically linking instrumental, sociotechnical, and institutional concepts of technology with control-, learning-, and legitimacy-oriented evaluation, the analysis reveals that persistent disagreements over success, evidence, and accountability reflect structurally different couplings of technological ontology and evaluative purpose rather than methodological pluralism. This shifts attention from improving evaluation techniques to interrogating how evaluation itself produces authority, learning, and legitimacy. From this perspective, a central tension for AI-supported HIS becomes apparent as institutional legitimation may advance faster than the organisational learning emphasised by sociotechnical approaches, stabilising technologies through standards and compliance before their organisational and clinical consequences are fully understood.

The synthesis further suggests that these tensions become particularly consequential in the context of artificial intelligence, where established notions of evidence, agency, and accountability are increasingly destabilised. Under these conditions, evaluation cannot be reduced to technical validation alone but acquires heightened significance as a mechanism through which organisational learning, professional authority, and institutional trust are negotiated. Against this backdrop, the socially embedded evaluation approach proposed in this critical literature review offers a path forward that integrates sociotechnical learning with institutional analysis, treats agency and accountability as distributed across human and algorithmic actors, and positions evaluation as a continuous and reflexive practice rather than a terminal governance mechanism. Ultimately, evaluation itself emerges not only as a methodological activity but as a central arena in which the consequences of digital transformation in healthcare are critically shaped.

References

Abdelgawad, Y. H., Abd El Razik, M. S., Saleh, D. A., Abuelela, M. H., & Salem, M. R. (2024). Promoting health information system in guiding decisions for improving performance: An intervention study at the Research Institute of Ophthalmology, Giza, Egypt. ***Frontiers in Digital Health***, 6.

Addo, K., & Agyepong, P. K. (2025). ***Evaluating the Health Information system implementation and utilization in healthcare delivery.***

Adjerid, I., Adler-Milstein, J., & Angst, C. (2018). Reducing Medicare Spending Through Electronic Health Information Exchange: The Role of Incentives and Exchange Maturity. ***Information Systems Research***, 29(2), 341–361.

Ahmadian, L., Salehi Nejad, S., & Khajouei, R. (2015). Evaluation methods used on health information systems (HISs) in Iran and the effects of HISs on Iranian healthcare: A systematic review. ***International Journal of Medical Informatics***, 84(6), 444–453.

AlHussainan, A., & Alhuwail, D. (2025). Bridging Global Frameworks and Local

Practice: Quantitative Evaluation of Electronic Health Record Safety in Kuwait's Public Hospitals. ***JMIR Medical Informatics***, 13(1), e70782.

Ammenwerth, E., Brender, J., Nykänen, P., Prokosch, H.-U., Rigby, M., & Talmon, J. (2004). Visions and strategies to improve evaluation of health information systems: Reflections and lessons based on the HIS-EVAL workshop in Innsbruck. ***International Journal of Medical Informatics***, 73(6), 479–491.

Ammenwerth, E., Gräber, S., Herrmann, G., Bürkle, T., & König, J. (2003). Evaluation of health information systems—Problems and challenges. ***International Journal of Medical Informatics***, 71(2), 125–135.

Ancker, J. S., Kern, L. M., Abramson, E., & Kaushal, R. (2012). The Triangle Model for evaluating the effect of health information technology on healthcare quality and safety. ***Journal of the American Medical Informatics Association***, 19(1), 61–65.

Andargoli, A., Scheepers, H., Rajendran, D., & Sohal, A. (2017). Health information systems evaluation frameworks: A systematic review. ***International Journal of Medical Informatics***, 97, 195–209.

Avgar, A., Tambe, P., & Hitt, L. M. (2018). Built to Learn: How Work Practices Affect Employee Learning During Healthcare Information Technology Implementation. ***MIS Quarterly***, 42(2), 645–660.

Avgerou, C. (2000). IT and organizational change: An institutionalist perspective. ***Information Technology and People***, 13(4), 234–262.

Barber, N., Franklin, B. D., Cornford, T., Klecun, E., & Savage, I. (2006). ***Safer, faster, better? Evaluating electronic prescribing***.

Barrett, A. K. (2021). Healthcare workers' communicative constitution of health information technology (HIT) resilience. ***Information Technology & People***, 35(2), 781–801.

Behn, R. D. (2003). Why Measure Performance? Different Purposes Require Different Measures. ***Public Administration Review***, 63(5), 586–606.

Chelimsky, E., R.Shadish, W., & Chelimsky, E. (1997). The Coming Transformations in Evaluation. In Evaluation for the 21st Century: A Handbook (pp. 1–26). ***SAGE Publications, Inc***.

Cresswell, K., Keizer, N. de, Magrabi, F., Williams, R., Rigby, M., Prgomet, M., Kukhareva, P., Wong, Z. S.-Y., Scott, P., Craven, C. K., Georgiou, A., Medlock, S., McNair, J. B., & Ammenwerth, E. (2024). Evaluating Artificial Intelligence in Clinical Settings—Let Us Not Reinvent the Wheel. ***Journal of Medical Internet Research***, 26(1),

Cresswell, K. M., & Sheikh, A. (2014). Undertaking sociotechnical evaluations of health information technologies. ***Informatics in Primary Care***, 21(2), 78–83.

Cross, J. L., Choma, M. A., & Onofrey, J. A. (2024). Bias in medical AI: Implications for clinical decision-making. ***PLOS Digital Health***, 3(11), e0000651.

Dahler-Larsen, P. (2012). Evaluation as a situational or a universal good? Why evaluability assessment for evaluation systems is a good idea, what it might look like in practice, and why it is not fashionable. ***Scandinavian Journal of Public Administration***, 16(3), 29–46.

Davenport, T. H., & Kalakota, R. (2019). The potential for artificial intelligence in healthcare. ***Future Healthcare Journal***, 6(2),

94–98.

DeLone, W. H., & McLean, E. R. (1992). Information Systems Success: The Quest for the Dependent Variable. ***Information Systems Research***, 3(1), 60–95.

Dixon, B. E., Zafar, A., & Overhage, J. M. (2010). A Framework for evaluating the costs, effort, and value of nationwide health information exchange. ***Journal of the American Medical Informatics Association***, 17(3), 295–301.

Effken, J. A. (2002). Different lenses, improved outcomes: A new approach to the analysis and design of healthcare information systems. ***International Journal of Medical Informatics,*** 65(1), 59–74.

Farzandipour, M., Nabovati, E., & Sadeqi Jabali, M. (2022). Comparison of usability evaluation methods for a health information system: Heuristic evaluation versus cognitive walkthrough method. ***BMC Medical Informatics and Decision Making***, 22(1), 157.

Goodhue, D. L., & Thompson, R. L. (1995). Task-Technology Fit and Individual Performance. ***MIS Quarterly***, 19(2), 213–236.

Greenhalgh, T., Morris, L., Wyatt, J. C., Thomas, G., & Gunning, K. (2013). Introducing a nationally shared electronic patient record: Case study comparison of Scotland, England, Wales and Northern Ireland. ***International Journal of Medical Informatics***, 82(5), e125–e138.

Greenhalgh, T., Wherton, J., Papoutsi, C., Lynch, J., Hughes, G., A'Court, C., Hinder, S., Fahy, N., Procter, R., & Shaw, S. (2017). Beyond Adoption: A New Framework for Theorizing and Evaluating Nonadoption, Abandonment, and Challenges to the Scale-Up, Spread, and Sustainability of Health and Care Technologies. ***Journal of Medical Internet Research***, 19(11), e8775.

Ha, E., Choon-Kon-Yune, I., Murray, L., Luan, S., Montague, E., Bhattacharyya, O., & Agarwal, P. (2025). Evaluating the Usability, Technical Performance, and Accuracy of Artificial Intelligence Scribes for Primary Care: Competitive Analysis. ***JMIR Human Factors***, 12(1), e71434.

Hart, C. (1998). Doing a literature review: Releasing the social science research imagination. ***SAGE Publications***.

Hendy, J., Fulop, N., Reeves, B. C., Hutchings, A., & Collin, S. (2007). Implementing the NHS information technology programme: Qualitative study of progress in acute trusts. ***British Medical Journal***, 334(7608), 1360.

Jacob, C., Brasier, N., Laurenzi, E., Heuss, S., Mougiakakou, S.-G., Cöltekin, A., & Peter, M. K. (2025). AI for IMPACTS Framework for Evaluating the Long-Term Real-World Impacts of AI-Powered Clinician Tools: Systematic Review and Narrative Synthesis. ***Journal of Medical Internet Research***, 27, e67485.

Ji, M., Genchev, G. Z., Huang, H., Xu, T., Lu, H., & Yu, G. (2021). Evaluation Framework for Successful Artificial Intelligence–Enabled Clinical Decision Support Systems: Mixed Methods Study. ***Journal of Medical Internet Research***, 23(6), e25929.

Kaplan, B., & Harris-Salamone, K. D. (2009). Health IT Success and Failure: Recommendations from Literature and an AMIA Workshop. ***Journal of the American Medical Informatics Association***, 16(3), 291–299.

Kaplan, B., & Shaw, N. T. (2004). Future directions in evaluation research: People, organizational, and social issues. ***Methods of Information in Medicine***, 43(3), 215–231.

Klecun, E., & Cornford, T. (2005). A Critical Approach to Evaluation. ***EJIS***, 14, 229–243.

Klecun, E., Lichtner, V., Cornford, T., & Petrakaki, D. (2014). Evaluation as a multi-ontological endeavour: A case from the English National Programme for IT in health care. ***Journal of the Association of Information Systems***, 15.

Kling, R. (1987). Defining the boundaries of computing across complex organizations. In R. J. Boland Jr. & R. Hirschheim (Eds.), ***Critical issues in information systems research*** (pp. 307–362). John Wiley & Sons.

Kling, R., & Iacono, S. (1989). The Institutional Character of Computerized Information Systems. ***Office Technology and People***, 5(1), 7–28.

Lagsten, J. (2011). Evaluating Information Systems according to Stakeholders: A pragmatic perspective and method. ***Electronic Journal of Information Systems Evaluation***, 14(1), pp73-88-pp73-88.

Latour, B. (1987). Science in action: How to follow scientists and engineers through society. ***Harvard University Press.***

Lebovitz, S., Levina, N., & Lifshitz-Assaf, H. (2021). Is AI Ground Truth Really 'True'? The Dangers of Training and Evaluating AI Tools Based on Experts' Know-What (SSRN Scholarly Paper No. 3839601). ***Social Science Research Network***.

Lilford, R. J., Foster, J., & Pringle, M. (2009). ***Evaluating eHealth: How to Make Evaluation More Methodologically Robust.*** PLoS Med.

Magrabi, F., Ammenwerth, E., Hyppönen, H., de Keizer, N., Nykänen, P., Rigby, M., Scott, P., Talmon, J., & Georgiou, A. (2016). Improving Evaluation to Address the Unintended Consequences of Health Information Technology: A Position Paper from the Working Group on Technology Assessment & Quality Development. ***Yearbook of Medical Informatics***, 1, 61–69.

March, J. G. (1991). Exploration and Exploitation in Organizational Learning. ***Organization Science***, 2(1), 71–87.

Menon, N. M., Lee, B., & Eldenburg, L. (2000). Productivity of Information Systems in the Healthcare Industry. ***Information Systems Research***, 11(1), 83–92.

Meyer, J. W., & Rowan, B. (1977). Institutionalized Organizations: Formal Structure as Myth and Ceremony. ***American Journal of Sociology***, 83(2), 340–363.

Mohsin Khan, M., Shah, N., Shaikh, N., Thabet, A., alrabayah, T., & Belkhair, S. (2025). Towards secure and trusted AI in healthcare: A systematic review of emerging innovations and ethical challenges. ***International Journal of Medical Informatics***, 195, 105780.

Neame, M. T., Sefton, G., Roberts, M., Harkness, D., Sinha, I. P., & Hawcutt, D. B. (2020). Evaluating health information technologies: A systematic review of framework recommendations. ***International Journal of Medical Informatics***, 142, 104247.

Noël, R., Taramasco, C., & Márquez, G. (2022). Standards, Processes, and Tools Used to Evaluate the Quality of Health Information Systems: Systematic Literature Review. ***Journal of Medical Internet Research***, 24(3), e26577.

Nykänen, P., & Kaipio, J. (2016). Quality of Health IT Evaluations. ***Studies in Health Technology and Informatics***, 222, 291–303.

Olawade, D. B., Weerasinghe, K., Teke, J., Msiska, M., Boussios, S., & Hatzidimitriadou, E. (2025). Evaluating AI

adoption in healthcare: Insights from the information governance professionals in the United Kingdom. ***International Journal of Medical Informatics***, 199, 105909.

Orlikowski, W. J. (1992). The Duality of Technology: Rethinking the Concept of Technology in Organizations. ***Organization Science***, 3(3), 398–427.

Orlikowski, W. J. (2000). Using Technology and Constituting Structures: A Practice Lens for Studying Technology in Organizations. ***Organization Science***, 11(4), 404–428.

Power, M. (1997). ***The Audit Society: Rituals of Verification.*** OUP Oxford.

Preskill, H., & Torres, R. T. (1999). The Role of Evaluative Enquiry in Creating Learning Orgnizations. In ***Organizational Learning and the Learning Organization: Developments in Theory and Practice*** (pp. 92–114). SAGE.

Reddy, S., Rogers, W., Makinen, V.-P., Coiera, E., Brown, P., Wenzel, M., Weicken, E., Ansari, S., Mathur, P., Casey, A., & Kelly, B. (2021). Evaluation framework to guide implementation of AI systems into healthcare settings. ***BMJ Health & Care Informatics***, 28(1).

Rozenblit, L., Price, A., Solomonides, A., Joseph, A. L., Koski, E., Srivastava, G., Labkoff, S., Bray, D., Lopez-Gonzalez, M., Singh, R., deBronkart, D., Barr, P. J., Szolovits, P., Dattani, K., Jaffe, C., Fridsma, D., Baris, R., Leftwich, R., Stolper, R., … Quintana, Y. (2025). Toward responsible AI governance: Balancing multi-stakeholder perspectives on AI in healthcare. ***International Journal of Medical Informatics***, 203, 106015.

Sacramento, S. P. R., Ongkeko, A., Ejera, A. V., Mata, E., Romero, C., Catameo, A., & Mende, M. A. (2025). A systematic review of tools cited in health information systems evaluations of developing and low to middle income countries. ***Discover Public Health***, 22(1), 528.

Sahni, N., Stein, G., Zemmel, R., & Cutler, D. M. (2023). The Potential Impact of Artificial Intelligence on Healthcare Spending (SSRN Scholarly Paper No. 4334926). ***Social Science Research Network.***

Serin, O., Yıldırım, B. F., Duman, F., Ercorumlu, K., Yavas, R., Tasar, M. A., & Celik, M. (2025). Physician perspectives on Artificial Intelligence in healthcare: A cross-sectional study of child-focused care in a Turkish tertiary hospital. ***International Journal of Medical Informatics***, 203, 106003.

Snyder, H. (2019). Literature review as a research methodology: An overview and guidelines. ***Journal of Business Research***, 104, 333–339.

Sligo, J., Gauld, R., Roberts, V., & Villa, L. (2017). A literature review for large-scale health information system project planning, implementation and evaluation. ***International Journal of Medical Informatics***, 97, 86–97.

Sockolow, P. S., Crawford, P. R., & Lehmann, H. P. (2012). Health Services Research Evaluation Principles. ***Methods of Information in Medicine,*** 51(02), 122–130.

Takian, A., Petrakaki, D., Cornford, T., Sheikh, A., Barber, N., & Evaluat, N. (2012). Building a house on shifting sand: Methodological considerations when evaluating the implementation and adoption of national electronic health record systems. ***BMC Health Services Research***, 12.

Topol, E. (2019). ***Deep medicine: How artificial intelligence can make healthcare human again.*** Basic Books.

Torraco, R. J. (2005). Writing integrative

literature reviews: Guidelines and examples. ***Human Resource Development Review***, 4(3), 356–367.

Tsiknakis, M., & Kouroubali, A. (2009). Organizational factors affecting successful adoption of innovative eHealth services: A case study employing the FITT framework. ***International Journal of Medical Informatics***, 78(1), 39–52.

Vasey, B., Nagendran, M., Campbell, B., Clifton, D. A., Collins, G. S., Denaxas, S., Denniston, A. K., Faes, L., Geerts, B., Ibrahim, M., Liu, X., Mateen, B. A., Mathur, P., McCradden, M. D., Morgan, L., Ordish, J., Rogers, C., Saria, S., Ting, D. S. W., … McCulloch, P. (2022). ***Reporting guideline for the early stage clinical evaluation of decision support systems driven by artificial intelligence***: DECIDE-AI. BMJ, 377, e070904.

Venkatesh, V., Zhang, X., & Sykes, T. A. (2011). 'Doctors Do Too Little Technology': A Longitudinal Field Study of an Electronic Healthcare System Implementation. ***Information Systems Research***, 22(3), 523–546.

Westbrook, J. I., Braithwaite, J., Georgiou, A., Ampt, A., Creswick, N., Coiera, E., & Iedema, R. (2007). Multimethod evaluation of information and communication technologies in health in the context of wicked problems and sociotechnical theory. ***Journal of the American Medical Informatics Association***, 14(6), 746–755.

Woods, L., Eden, R., Green, D., Pearce, A., Donovan, R., McNeil, K., & Sullivan, C. (2024). Impact of digital health on the quadruple aims of healthcare: A correlational and longitudinal study (Digimat Study). ***International Journal of Medical Informatics***, 189, 105528.

Wronikowska, M. W., Malycha, J., Morgan, L. J., Westgate, V., Petrinic, T., Young, J. D., & Watkinson, P. J. (2021). Systematic review of applied usability metrics within usability evaluation methods for hospital electronic healthcare record systems. ***Journal of Evaluation in Clinical Practice***, 27(6), 1403–1416.

Yusof, M. Mohd. (2015). A case study evaluation of a Critical Care Information System adoption using the socio-technical and fit approach. ***International Journal of Medical Informatics***, 84(7), 486–499.

Yusof, M. Mohd., Kuljis, J., Papazafeiropoulou, A., & Stergioulas, L. K. (2008). An evaluation framework for Health Information Systems: Human, organization and technology-fit factors (HOT-fit). ***International Journal of Medical Informatics***, 77(6), 386–398.

Yusof, M. Mohd., Papazafeiropoulou, A., Paul, R. J., & Stergioulas, L. K. (2007). Investigating evaluation frameworks for health information systems. ***International Journal of Medical Informatics***, 77(6), 377–385.