

Review

Artificial Intelligence Platform Architecture for Hospital Systems: Systematic Review

Musitapa Maimaitiaili^{1,2*}, MBBS; Yiershatijiang Jiamaliding^{1,2*}, MBBS; Guangle Dai³, MBBS; Hui Xiao³, MSCS; Warisijiang Kuerbanjiang^{1,2}, MBBS; Yuexiong Yi¹, MD, PhD

¹Department of Gynecology, Zhongnan Hospital of Wuhan University, Wuhan, Hubei, China

²The Second Clinical Hospital, Wuhan University, Wuhan, China

³Information Center, Zhongnan Hospital of Wuhan University, Wuhan, China

*these authors contributed equally

Corresponding Author:

Yuexiong Yi, MD, PhD
Department of Gynecology
Zhongnan Hospital of Wuhan University
#169, Donghu Road, Wuchang District
Wuhan, Hubei 430071
China
Phone: 86 15671669885
Fax: 86 02767813142
Email: yiyuexiong@163.com

Abstract

Background: The construction of artificial intelligence (AI) platforms in hospitals is the backbone of the revolution in health care. While traditional hospital information systems have facilitated digitalization, they are still limited by data silos, fragmented workflows, and insufficient clinical intelligence that impede organizations from realizing the promise of data-led decision-making.

Objective: This study aimed to derive a hospital-specific 5-layer architecture (infrastructure, data, algorithm, application, and security and compliance) and to systematically review the evidence mapped onto the 5-layer framework to assess its applicability.

Methods: A systematic literature search was performed in Web of Science, Embase, PubMed, and Scopus from inception to May 2025. The review followed the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines. Studies were screened and selected for full-text review by two independent reviewers. We included peer-reviewed empirical studies describing hospital-based AI implementations across clinical domains. Reviews, commentaries, and purely technical bench studies without hospital context and non-English literature were excluded. Quality assessment of the identified papers was conducted using the Critical Appraisal Skills Programme tool. Using a 0 to 5 point ordinal maturity scale of 5 layers, we conducted a structured mapping with quantitative mapping, weighted co-occurrence analysis, weighted Jaccard similarity, and thematic synthesis with examples.

Results: In total, 29 studies met the eligibility criteria and included work specifically in emergency, radiology, routine imaging, chronic disease, and multihospital platform work, conducted in 11 countries. On average, the application (mean 3.17, SD 0.85) and data (mean 3.00, SD 0.76) layers demonstrated the highest maturity, followed by algorithm (mean 2.79, SD 0.77) and infrastructure (mean 2.79, SD 1.70). The security and compliance layer showed the lowest and most variable maturity (mean 1.69, SD 1.89). Weighted co-occurrence and Jaccard analyses revealed strong interconnections among data, algorithm, and application (Jaccard=0.80-0.89), forming a technical core, whereas security and compliance exhibited weak alignment (0.43-0.46).

Conclusions: Our review excluded non-English and gray literature, which may limit comprehensiveness. The ordinal maturity scoring may still simplify the contextual complexity of hospital AI implementations. Our synthesis validates a 5-layer hospital AI platform architecture, grounded in both theoretical frameworks and empirical evidence. The findings highlight that while clinical feasibility is achievable, sustainable hospital-wide AI requires stronger investment in infrastructure, data governance, and compliance.

Trial Registration: PROSPERO CRD420251133590; <https://www.crd.york.ac.uk/PROSPERO/view/CRD420251133590>

J Med Internet Res 2025;27:e79788; doi: [10.2196/79788](https://doi.org/10.2196/79788)

Keywords: artificial intelligence; hospital AI platform; 5-layer architecture; health care digitization; AI implementation framework

Introduction

Background

Artificial intelligence (AI) is often seen as a game changer for health care, yet many hospitals face challenges in scaling it beyond pilot projects due to limited resources and infrastructure. Rising chronic disease burden due to aging populations and increased demand for personalization has exacerbated these challenges [1]. Traditional hospital information systems (Hospital Information System [HIS], Laboratory Information Systems [LIS], and Picture Archiving and Communication System [PACS]) have become administrative tools without analytic or decision-support value [2]. Wide varieties of multimodal health data are now available along with advanced AI methods, such as image recognition, natural language processing, and predictive analytics, that are maturing rapidly [3]. However, most hospitals face difficulties in integrating these approaches within routine clinical workflows [4].

Research Gap

Many hospitals have not built a system-level roadmap for scaling up and governance structure to support the use of AI pilot projects [5]. Fragmented infrastructures worsen this challenge: siloed PACS and LIS [6] limited interoperability across electronic health record (EHR) vendors due to inconsistent Health Level 7 (HL7) Fast Healthcare Interoperability Resources (FHIR) adoption [7], and heterogeneous Internet of Things (IoT) data streams [8].

Reported implementations typically target isolated tasks, such as radiology computer-aided diagnosis [9] or triage algorithms. Focusing on isolated tasks does not facilitate enterprise-wide implementation of AI, as it fails to address the need for integrated and scalable solutions. In this context, hospital executives do not have sensible roadmaps to translate fragmented digital environments into integrated AI platforms that are based on evidence [10].

Existing Frameworks and Development of a Hospital-Specific Architecture

The hospital system-wide AI platform should be built over an existing framework and not a new one. To build a strong theory base, 4 representative frameworks were reviewed: (1) the Healthcare Information and Management Systems Society Digital Health Framework [11], (2) the World Health Organization (WHO) HIS framework [12], (3) sociotechnical theory [13], and (4) the National Institute of Standards and Technology (NIST) Big Data Reference Architecture (BDRA) [14]. Each offers distinct perspectives. The Healthcare Information and Management Systems Society emphasizes digital maturity and interoperability; the World Health Organization HIS highlights governance and service delivery; sociotechnical theory stresses the interaction of people, processes, and technologies; and BDRA provides a detailed blueprint for big data infrastructures. Comparison shows overlap in infrastructure and data integration, but divergences in analytics, governance, and algorithm lifecycle (Table 1).

Table 1. Original layer structures of 4 foundational frameworks.

Layer	HIMSS ^a digital health framework	WHO ^b HIS ^c framework	Sociotechnical theory	NIST ^d BDRA ^e
Layer 1	Infrastructure and interoperability	ICT ^f resources	Technical subsystem	Infrastructure and platform services
Layer 2	Data exchange and integration	Health data sources and data flow	Information subsystem	Data pipeline
Layer 3	Analytics and predictive insights	Health services delivery	Organizational processes and workflows	Analytics and big data applications
Layer 4	Person-enabled health	Policy, ethics, and regulation	People–organization interactions and accountability	Application services and outputs
Layer 5	Governance and workforce	— ^g	—	Governance, security, compliance services

^aHIMSS: Healthcare Information and Management Systems Society.

^bWHO: World Health Organization.

^cHIS: hospital information system.

^dNIST: National Institute of Standards and Technology.

^eBDRA: Big Data Reference Architecture.

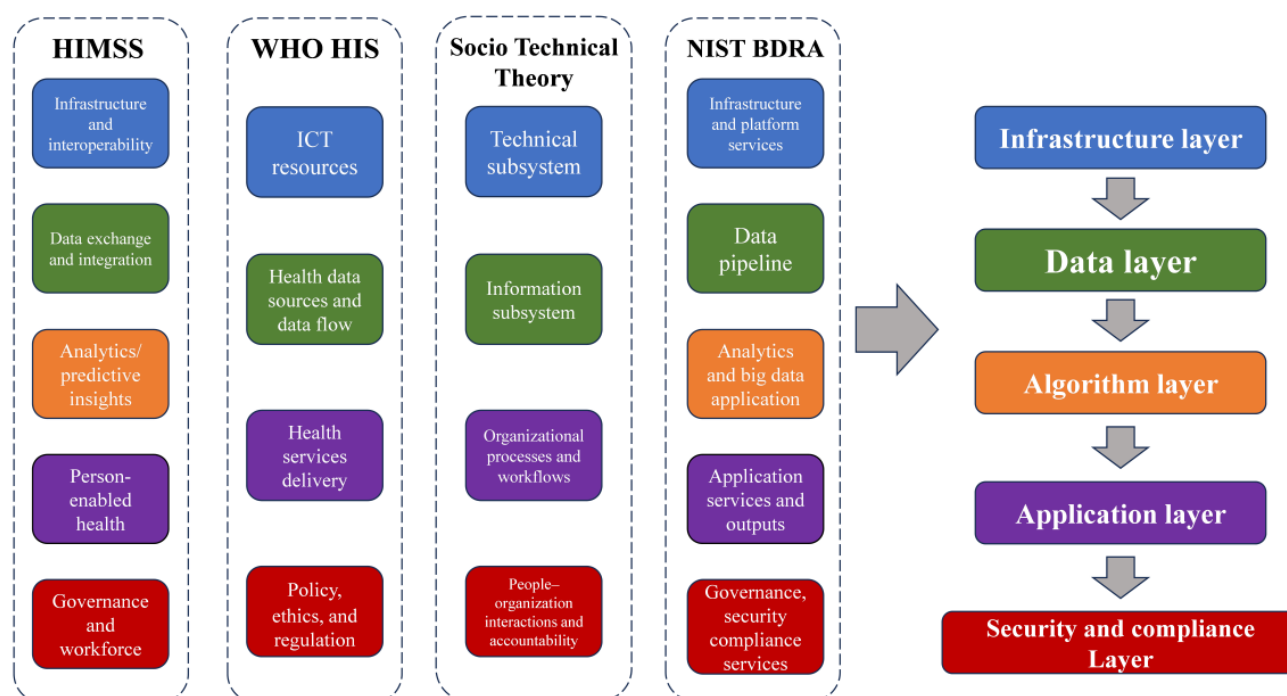
^fICT: information and communication technology.

^gNot available.

Despite important similarities, none of these frameworks fully address the unique requirements for hospital-wide AI implementation, particularly in terms of comprehensive integration and lifecycle management. To bridge this gap, we applied a merge-and-normalize approach by extracting core constructs, clustering overlapping elements, and adding components related to AI where the original was missing, especially on algorithm lifecycle management. This process produced 5 interoperable layers tailored to hospital AI contexts: infrastructure, data, algorithm, application, and security and compliance (Figure 1). Specifically, (1) the

infrastructure layer covers compute, storage, and network foundations; (2) the data layer ensures standards such as HL7 FHIR and Digital Imaging and Communications in Medicine, multimodal integration across EHR, LIS, and IoT, and quality management; (3) the algorithm layer introduces the AI lifecycle, including development, validation, monitoring, and operations; (4) the application layer emphasizes workflow integration, such as decision support, triage, and patient-facing tools; and (5) the security and compliance layer ensures privacy, accountability, and governance.

Figure 1. Derivation of the 5-layer architecture from 4 foundational frameworks. Constructs from the HIMSS digital health framework, the WHO HIS framework, sociotechnical theory, and the NIST BDRA were compared, merged, and normalized to generate a unified 5-layer model (infrastructure, data, algorithm, application, and security and compliance). Color coding highlights conceptually equivalent constructs across frameworks: blue represents infrastructure; green represents data; orange represents algorithm and analytics; purple represents applications; and red represents governance, security, and compliance. BDRA: Big Data Reference Architecture; HIMSS: Healthcare Information and Management Systems Society; HIS: Health Information System; NIST: National Institute of Standards and Technology; WHO: World Health Organization.



Objectives

On the basis of identified gaps and synthesized frameworks, this study intends to propose a 5-layer AI platform architecture specific to hospitals, to systematically review the empirical studies by mapping evidence onto the framework, and to synthesize theoretical and empirical insights to offer practical advice for hospital administrators, policymakers, and developers who want to build a scalable, secure, and clinically integrated AI ecosystem. Two research questions (RQs) guide this process: (1) RQ1: What empirical evidence supports each layer of the proposed 5-layer architecture? and (2) RQ2: How do the interrelationships among layers reveal strengths and gaps in hospital-level AI research?

Methods

Study Design

This systematic review followed the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) 2020 and PRISMA-P (Preferred Reporting Items for Systematic Review and Meta-Analysis Protocols) guidelines [15] and was prospectively registered in PROSPERO (CRD420251133590). A completed PRISMA checklist is provided in Checklist 1. The 5-layer architecture derived from existing frameworks was adopted as an a priori coding structure for evidence mapping for identifying studies on hospital AI implementation and aligning their findings with this architecture to link theoretical constructs with real-world practice.

Search Strategy

A comprehensive search was conducted in Web of Science, Embase, PubMed, and Scopus from their inception to May 23, 2025. To maximize sensitivity, the strategy combined free-text keywords with controlled vocabulary terms. Search terms included “Hospital Management,” “Healthcare Management,” “Hospital Operations,” “Healthcare Administration,” “Medical Administration,” “Hospital Information Systems,” “AI Deployment,” “AI Implementation,” “AI Integration,” “Artificial Intelligence,” “AI Applications,” “Large Language Models,” “Transformer Applications,” “Hospital Operations Optimization,” “Clinical Workflow Improvement,” “Resource Allocation,” and “Patient Flow Management.” The full search strings for all databases are detailed in [Multimedia Appendix 1](#).

In addition, reference lists of all included articles were manually screened to identify further eligible studies not captured by the electronic search.

Eligibility Criteria

Included studies were in a hospital setting and documented real-world AI use (eg, machine learning, deep learning, natural language processing, predictive analytics, and expert systems) that presented measurable outcomes, including enhanced efficiency, decreased costs/errors, optimized resource usage, improved patient experience, or accurate diagnosis/decision-making enhancement. Studies that did not use AI were excluded, along with reviews, commentaries, conference abstracts, non-English studies, and studies without an abstract. The summary of inclusion and exclusion criteria can be seen in [Textbox 1](#).

Textbox 1. Inclusion and exclusion criteria

Inclusion criteria	
<ul style="list-style-type: none">• Studies conducted in hospital environments (operation rooms, decision-making, emergency services, patient flow management, resource allocation)• Studies involving artificial intelligence technologies such as machine learning, deep learning, natural language processing, predictive analytics, or expert systems• Studies describing real-world, pilot, or simulation-based evaluations conducted in a hospital context using real or representative institutional data• Studies reporting outcomes related to efficiency improvement, cost reduction, error reduction, resource utilization, patient satisfaction, or decision-making accuracy• Applications targeting hospital operations, planning, scheduling, cost control, risk prediction, patient care, or workflow optimization	
Exclusion criteria	
<ul style="list-style-type: none">• Studies set entirely outside of hospitals• Studies that mention artificial intelligence (AI) only conceptually or theoretically, without implementation details• Studies focusing only on general information systems without AI integration• Studies without measurable or reported outcomes• Non-English study• Review or commentaries or meeting abstract or letters• No abstract exists	

Screening

Two reviewers independently screened titles and abstracts, removed duplicates, and assessed preliminary eligibility. The full-text evaluation of the relevant studies was conducted based on the inclusion and exclusion criteria. Reviewers were not blinded to study authorship or outcomes. Discrepancies were resolved by discussion and consensus. The reliability of screening was evaluated through the Cohen κ statistic, and the results were summarized in a 2×2 contingency table.

objectives, clinical applications, and limitations. Explicit coding definitions were used to enable mapping of each study to the 5-layer architecture (infrastructure, data, algorithm, application, and security and compliance). To promote transparency, the coding criteria for each layer are given in [Table 2](#). Each layer was scored on a 0 to 5 point ordinal maturity scale conceptually aligned with the Capability Maturity Model Integration framework [16] with level definitions and examples detailed in [Table 3](#). Two reviewers independently performed the scoring, and discrepancies were resolved by consensus.

Data Extraction and Synthesis

Data were extracted using a standardized form, including bibliographic details, study design, clinical domain,

Table 2. Coding definitions for the 5-layer architecture.

Layer	Coding definition
Infrastructure	Assign when the evidence describes compute, storage, network, cloud/edge topology, containerization/orchestration, or system-level integration

Layer	Coding definition
Data	capacity that hosts or executes AI ^a workloads (eg, data centers, GPU pools, hybrid cloud, and enterprise integration with HIS ^b /PACS ^c /LIS ^d). Assign when the evidence concerns data sources and flows, standards/interoperability (HL7 ^e FHIR ^f , DICOM ^g , and terminologies), identity/linkage (EMPI ^h), multimodal integration (EHR ⁱ , imaging, monitors/IoT ^j , notes, and omics), data quality, lineage/provenance, or deidentification that make data AI ready.
Algorithm	Assign when the evidence covers AI/ML ^k methods and lifecycle: model development/training, internal or external validation, performance metrics, monitoring/drift, retraining, explainability, or federated or edge learning, irrespective of where the model will later be used.
Application	Assign when the evidence demonstrates embedding AI into clinical or operational workflows, including CDS ^l , triage/priority, worklist optimization, patient-flow/bed management, scheduling, or patient-facing tools; focus is on use in work (UI ^m /UX ⁿ , pathway location, task changes).
Security and compliance	Assign when the evidence addresses privacy/security controls (access control, encryption, OAuth/HTTPS, audit logs, blockchain audit), consent and data use governance, regulatory/ethical compliance (HIPAA ^o , GDPR ^p , local policies), or model governance/accountability.

^aAI: artificial intelligence.

^bHIS: hospital information system.

^cPACS: picture archiving and communication system.

^dLIS: laboratory information systems.

^eHL7: Health Level 7.

^fFHIR: Fast Healthcare Interoperability Resources.

^gDICOM: Digital Imaging and Communications in Medicine.

^hEMPI: Enterprise Master Patient Index.

ⁱEHR: electronic health record.

^jIoT: Internet of Things.

^kML: machine learning.

^lCDS: clinical decision support.

^mUI: user interface.

ⁿUX: user experience.

^oHIPAA: Health Insurance Portability and Accountability Act.

^pGDPR: General Data Protection Regulation.

Table 3. The 0 to 5 point maturity scale and the Capability Maturity Model Integration (CMMI) framework.^a

Score	CMMI level	CMMI descriptor	Hospital AI ^b platform context
0	— ^c	—	No evidence of this layer in the study; the layer is not addressed or discussed.
1	Level 1: initial	Processes are ad hoc, chaotic, and unstructured. Success depends on individual effort.	Conceptual or pilot-level mention; ad hoc implementation without governance or integration.
2	Level 2: managed	Basic project management processes established to track cost, schedule, and functionality.	Layer partially implemented or managed within a limited scope (eg, single department).
3	Level 3: defined	Processes are documented, standardized, and integrated into organizational practice.	Layer implemented with defined workflows, policies, or institutional governance structures.
4	Level 4: quantitatively managed	Organization uses quantitative data to control and monitor processes.	Layer performance is monitored with metrics; multiple departments coordinate and share data.
5	Level 5: optimizing	Focus on continuous improvement and innovation based on quantitative feedback.	Fully institutionalized, hospital-wide, or cross-site implementation with feedback loops and continuous optimization.

^aThe 0-5 maturity scoring system was conceptually aligned with the CMMI framework (Software Engineering Institute, Carnegie Mellon University), with “0” added to capture the absence of evidence in individual studies.

^bAI: artificial intelligence.

^cNot applicable.

Quantitative synthesis summarized the average maturity scores and SDs for each layer across studies and by study design. Weighted co-occurrence matrices were computed to quantify the cumulative maturity shared between layers, whereas weighted Jaccard similarity indices measured the strength of cross-layer coupling. These metrics were visualized in heatmaps to illustrate maturity distribution and interlayer relationships. Qualitative synthesis identified thematic patterns, gaps, and illustrative cases to demonstrate interactions among layers in real-world contexts.

Quality Assessment

The Critical Appraisal Skills Programme (CASP) tool [17] was applied to assess the quality of included studies because it can be used across various study types. Using checklists for qualitative, quantitative, and mixed methods research, the Critical Appraisal Tools (CATS) tool is used more often than other important assessments, such as the Joanna Briggs Institute, Cochrane, and GRADE. All articles were evaluated for methodology, reliability, interpretation, and usability.

Statistical Analysis

Descriptive statistics were used to summarize the maturity distribution of the 5 layers, expressed as mean scores and SDs across all included studies. Studies were further stratified by methodological category to compare average maturity levels within and across study designs.

Cross-layer relationships were analyzed using weighted co-occurrence and weighted Jaccard similarity metrics derived from the 0 to 5 maturity scores.

The weighted co-occurrence is defined as follows:

$$W_{co}(A, B) = \sum_{i=1}^n \min(A_i, B_i)$$

where A_i and B_i represent the maturity scores (0-5) assigned to layers A and B in study i .

The weighted Jaccard similarity is defined as follows:

$$J_w(A, B) = \frac{\sum_{i=1}^n \min(A_i, B_i)}{\sum_{i=1}^n \max(A_i, B_i)}$$

where A_i and B_i represent the maturity scores (0-5) assigned to layers A and B in study i .

Interrater reliability during study selection was quantified using Cohen κ statistic, defined as follows:

$$\kappa = \frac{p_o - p_e}{1 - p_e}$$

where p_o is the observed agreement, and p_e is the expected agreement by chance.

Ethical Considerations

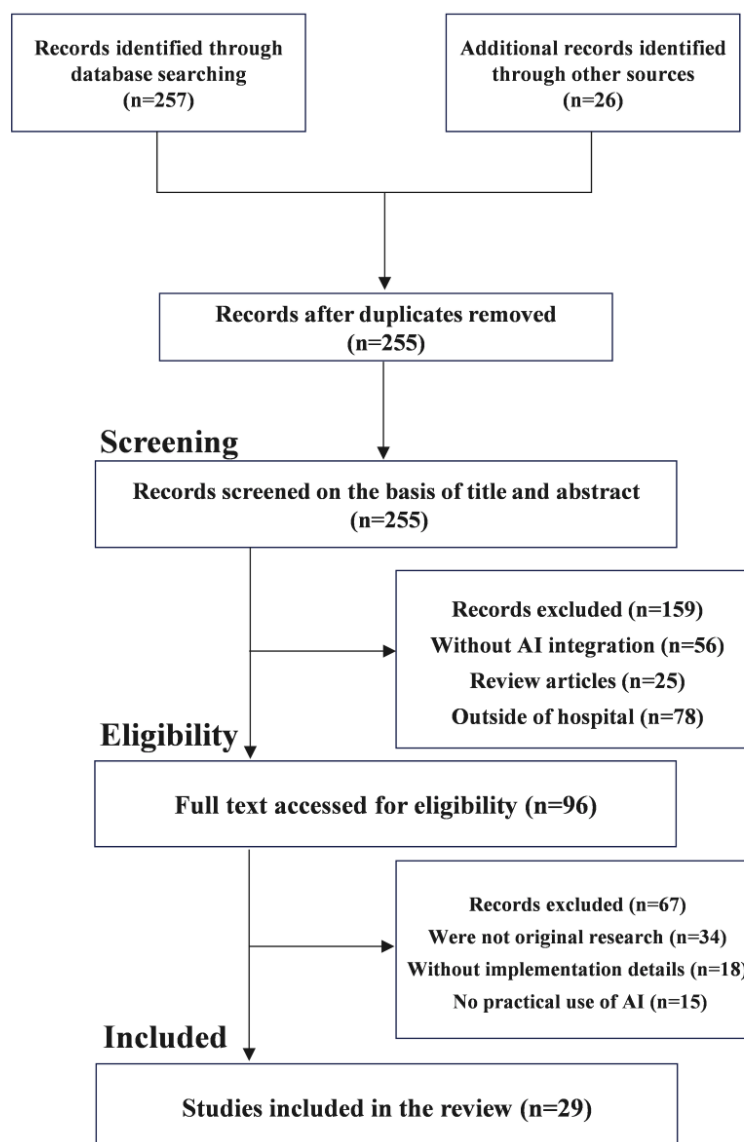
This systematic review does not involve human participants, identifiable patient data, or protected health information. All data analyzed in this review were obtained from publicly accessible publications. Therefore, an ethical review was not required under Zhongnan Hospital of Wuhan University's secondary research policies. The study complied with the Declaration of Helsinki and institutional guidelines for secondary data use.

Results

Study Selection

A total of 283 records were identified, of which 257 were from databases and 26 from other sources. After removing duplicates, 255 records were available to undergo title and abstract screening. At this stage, 159 records were excluded, comprising 56 non-AI cases, 25 review articles, and 78 cases not completed in the hospital. Of 96 full texts assessed for eligibility, 67 were ruled out for the following reasons: not original research ($n=34$), lacked implementation details ($n=18$), or practical AI use was not described ($n=15$). In the end, a total of 29 studies [18-46] were included in the review, representing 10.2% of the 283 records. Figure 2 illustrates the selection process. Screening was done by 2 reviewers independently, who demonstrated high interrater reliability (Cohen $\kappa=0.98$), suggesting almost perfect agreement. The 2x2 contingency table is shown in Multimedia Appendix 2.

Figure 2. The search strategy for study inclusion is based on PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines. AI: artificial intelligence.



Study Characteristics

The characteristics of the 29 included studies are summarized in [Tables 4](#) and [5](#). Clinical coverage was broad: nonspecific or cross-specialty applications were most frequent (10/29, 34.5%) [[18-21,25-27,30,40,45](#)], followed by radiology (6/29,

20.7%) [[22,24,29,43,44,46](#)], emergency medicine (5/29, 17.2%) [[28,31,33,37,39](#)], cardiology (2/29, 6.9%) [[34,41](#)], and gynecology (2/29, 6.9%) [[23,38](#)]. Surgery [[36](#)], chronic disease [[35](#)], nursing [[32](#)], and psychiatry [[42](#)] each contributed 1 study.

Table 4. Characteristics of included studies.

Author (year)	Country	Clinical domain	Study design	Objective	Clinical application	Limitations
Ahsen et al (2025) [29]	United States	Radiology	Economic evaluations	Provide guidance on integrating AI ^a into mammography workflows by balancing tasks between radiologists and algorithms	Breast cancer detection and risk assessment	Costs associated with algorithms, false assessments, and litigation expenses from false negatives
Boussen et al (2024) [33]	France	Intensive care medicine	Clinical prediction	To evaluate the performance of SAPS 2 PLUS model compared to the original SAPS 2	Predicting ICU ^b patient survival	Potential biases in datasets and limited generalizability due to single-center data

Author (year)	Country	Clinical domain	Study design	Objective	Clinical application	Limitations
				model by incorporating heart rate complexity and diastolic blood pressure measurements		
Alam et al (2025) [18]	United States	Nonspecific	Diagnostic test study	To assess the accuracy and reliability of ChatGPT 4.0 in interpreting 24 h ABPM ^c data in clinical settings	ChatGPT 4.0 for interpreting 24 h ABPM data	Limited research validating AI models against expert interpretations in real-world clinical scenarios
Areias et al (2024) [20]	United States	Nonspecific	Cohort study	To explore the impact of scaling care through AI on patient outcomes, engagement, satisfaction, and adverse events	AI tool integrated into the physical therapist clinical portal to streamline workflow and support decision-making	Limited research on the impact of AI scalability approaches in clinical outcomes for MSK ^d rehabilitation
Chen and Miao (2025) [25]	China	Nonspecific	Cross-sectional study	To evaluate the impact and effectiveness of DeepSeek, an AI-driven diagnostic tool, deployed across 90 tertiary hospitals in China	Improving diagnostic accuracy, enhancing clinical decision support, automating medical image analysis, streamlining workflow processes	High initial investment costs, requirement for robust data infrastructure, potential resistance from health care professionals, variability in model performance across different settings
Farghaly and Deshpande (2024) [44]	United States	Radiology	Diagnostic test study	Develop a novel classification model to distinguish COVID-19 from viral pneumonia using chest x-ray images	Automated classification of chest x-ray images into normal, COVID-19, and viral pneumonia categories to assist in early detection and diagnosis	Dataset bias, model generalizability, interpretability. The dataset used may not fully capture the diversity of real-world clinical settings. Imaging protocol variations could affect model performance
Fairbairn et al [34] (2025)	United Kingdom	Cardiovascular science	Cohort study	To evaluate the impact of a national AI technology program on cardiovascular outcomes and its broader effects on the health system	Predicting cardiovascular risk, optimizing treatment strategies, improving patient management and follow-up, enhancing clinical decision support systems	Data privacy concerns, initial implementation costs, variability in data quality across different regions, potential resistance from health care professionals
Jaganathan and Natesan (2025) [23]	India	Gynecology	Diagnostic study	To develop and evaluate an integrated system using blockchain technology and explainable AI for the detection and management of polycystic ovary syndrome	Early detection of PCOS ^e , personalized treatment recommendations, secure data sharing, enhanced patient privacy	Initial setup costs, need for robust data infrastructure, potential resistance from health care professionals, complexity in integrating blockchain with existing systems
Muntasir et al (2023) [40]	United States	Nonspecific	Qualitative study	To evaluate the impact of AI-assisted technologies on optimizing laboratory workflows in	Automating sample processing, optimizing test ordering and prioritization, predicting equipment	Automating sample processing, optimizing test ordering and prioritization, predicting equipment

Author (year)	Country	Clinical domain	Study design	Objective	Clinical application	Limitations
				hospitals and improving overall efficiency	maintenance needs, enhancing data management and reporting	maintenance needs, enhancing data management and reporting
Ju et al [32] (2025)	Korea	Nursing	Cohort study	To develop and evaluate a generative AI system that provides nursing diagnosis and documentation recommendations using virtual patient electrocardiogram data	Assisting in nursing diagnoses, automating documentation processes, improving clinical decision support, enhancing patient care quality	Initial setup costs, need for high-quality training data, potential resistance from nursing staff, variability in model performance across different settings
Klumpp et al (2021) [21]	Germany	Nonspecific	Qualitative study	To explore various application cases of AI in hospital health care settings and address the challenges faced during implementation in European hospitals	Predictive analytics for patient outcomes, clinical decision support systems, automated diagnostic tools, workflow optimization	Data privacy regulations (eg, GDPR ^f), initial investment costs, need for high-quality data, resistance from health care professionals, variability in model performance across different institutions
Le et al (2024) [39]	United States	Emergency medicine	Cohort study	To evaluate the impact of a ML ^h -enabled automated system for detecting LVO ^g on transfer times and patient outcomes in primary stroke centers	Early detection of LVO, optimizing patient transfer protocols, improving clinical decision-making, reducing time to treatment	Initial setup costs, need for high-quality training data, potential resistance from health care professionals, variability in model performance across different populations
Li et al (2024) [45]	United States	Nonspecific	Qualitative study	To develop and evaluate TrajVis, a visual clinical decision support system that translates AI trajectory models into actionable insights for health care	Predictive analytics for patient trajectories, personalized treatment recommendations, workflow optimization, enhancing communication between clinicians	Initial setup costs, need for high-quality training data, potential resistance from health care professionals, complexity in interpreting AI-generated insights
Lin et al (2025) [43]	United States	Radiology	Cohort study	To evaluate the effectiveness of risk-stratified screening schedules using AI models in optimizing daily mammography recalls and improving patient outcomes	Risk stratification for personalized screening, optimizing recall scheduling, reducing unnecessary follow-ups, enhancing early detection of breast cancer	Initial setup costs, need for high-quality training data, potential resistance from health care professionals, variability in model performance across different populations
Novak et al (2021) [26]	United States	Nonspecific	Qualitative study	To explore how design thinking methodologies can be applied to health informatics projects, using insights from Project Health Design as a case study	Enhancing patient-centered care, improving user experience, fostering innovation in health IT solutions, promoting interdisciplinary collaboration	Limited generalizability due to case-specific nature, potential resistance from traditional health care structures, need for ongoing stakeholder engagement, challenges in integrating with existing systems

Author (year)	Country	Clinical domain	Study design	Objective	Clinical application	Limitations
Nsubuga et al (2025) [31]	Uganda	Emergency medicine	Diagnostic test study	To evaluate the performance of ML models for trauma triage in low-resource settings and compare it with traditional triage methods	Automated trauma triage, predictive analytics for patient outcomes, optimizing resource allocation, improving clinical decision-making	Initial setup costs, need for high-quality training data, potential resistance from health care professionals, variability in model performance across different populations, challenges in low-resource settings
Pariso et al (2025) [19]	Italy	Nonspecific	Cross-sectional study	To evaluate the impact of integrating AI into energy management systems in Italian hospitals, focusing on efficiency improvements and cost savings	Energy consumption optimization, predictive maintenance, demand response, and reducing carbon footprint	Initial setup costs, need for high-quality data, potential resistance from facility managers, variability in model performance across different facilities, and integration with existing systems
Vignapiano et al (2025) [42]	Italy	Psychiatry	Cross-sectional study	To evaluate proximity-based solutions that integrate clinical and technological advances to optimize treatment for autism spectrum disorder	Personalized treatment plans, behavior monitoring, predictive analytics for symptom progression, and enhancing communication and social skills	Data privacy concerns, initial setup costs, need for high-quality training data, variability in model performance across different individuals, resistance from health care professionals and caregivers
Roppelt et al (2025) [30]	Germany	Nonspecific	Qualitative study	To explore the effective adoption of AI technologies in health care settings through multiple case studies, highlighting best practices and challenges	Diagnostic support, personalized medicine, patient monitoring, predictive analytics, and improving clinical workflows	Data privacy concerns, initial setup costs, need for high-quality training data, variability in model performance across different settings, resistance from health care professionals
Xie et al (2021) [35]	China	Chronic disease management	Qualitative study	To explore the integration of AI, blockchain, and wearable technology in managing chronic diseases, focusing on improving patient outcomes and optimizing health care delivery	Continuous monitoring, predictive analytics for disease progression, personalized treatment plans, secure data sharing, enhancing patient engagement	Data privacy concerns, initial setup costs, need for high-quality training data, variability in model performance across different populations, resistance from healthcare professionals
Yang et al (2022) [38]	United Kingdom	Gynecology	Cohort study	To develop and validate a ML-based risk stratification model for gestational diabetes management, aiming to improve patient outcomes through personalized risk assessment and intervention	Risk prediction, early detection, personalized treatment plans, improving clinical workflows, and enhancing patient engagement	Data privacy concerns, initial setup costs, need for high-quality training data, variability in model performance across different populations, resistance from health care professionals
Yoo et al (2022) [27]	Korea	Nonspecific	Clinical prediction	To develop and evaluate an interoperable and easily transferable	Diagnostic support, treatment planning, predictive analytics, improving clinical	Data privacy concerns, initial setup costs, need for high-quality training data,

Author (year)	Country	Clinical domain	Study design	Objective	Clinical application	Limitations
				clinical decision support system that can be effectively deployed across various health care settings, aiming to improve clinical workflows and patient outcomes	workflows, and enhancing patient safety	variability in model performance across different settings, resistance from health care professionals
Wshah et al (2022) [24]	United States	Radiology	Diagnostic test study	To develop and validate a ML model for classifying intravascular volume status using point-of-care ultrasound, aiming to improve the accuracy of clinical assessments in critical care settings	Intravascular volume classification, diagnostic support, improving clinical decision-making, and enhancing patient outcomes	Data privacy concerns, initial setup costs, need for high-quality training data, variability in model performance across different populations, resistance from health care professionals
Tam et al (2021) [22]	United Kingdom	Radiology	Diagnostic test study	To evaluate how AI can assist radiologists as the first reader of chest x-rays, improving accuracy and efficiency in lung cancer diagnosis by triaging HCT ^I cases before standard reporting	AI-based triage workflow; detection of lung nodules, masses, and hilar enlargement; reduction of missed cancers; standardization of radiologist performance; improved diagnostic consistency; enhanced performance on difficult or distracting findings	Increase in false positives, performance drop with distracting findings (eg, COPD ^J , pleural effusion), requires threshold tuning for HCT classification, algorithm not trained on local data
Seyam et al (2022) [46]	Switzerland	Radiology	Diagnostic test study	To evaluate the diagnostic performance and impact on clinical workflow of an AI-based tool for detecting ICH on emergent noncontrast head CT ^L scans	Detection of various types of ICH ^K (eg, intraparenchymal, subdural, subarachnoid, intraventricular), improved prioritization of critical findings, reduction in report turnaround times and ED ^M length of stay	Lower detection rates for specific subtypes of ICH (eg, subdural and acute subarachnoid hemorrhage), false-positive findings (eg, postoperative defects, artifacts, tumors), need for clear standard operating procedures to ensure optimal functioning in patient care workflows
Raven et al (2022) [28]	The Netherlands	Emergency medicine	Clinical prediction	To evaluate whether ML combined with clinical judgment outperforms clinical judgment alone in predicting in-hospital mortality in both older and younger patients suspected of infection presenting to the ED	Risk stratification of ED patients with suspected infections, rapid initiation of appropriate treatment and disposition based on risk prediction, enhanced decision-making support for clinicians by integrating ML models into clinical workflows	Potential bias in training datasets. Need for validation in diverse populations. Integration challenges within existing clinical workflows. Possible overfitting if not properly validated
Hond et al (2021) [37]	The Netherlands	Emergency medicine	Clinical prediction	To develop and validate a ML model for predicting hospital admission in	Prediction of hospital admission likelihood, optimization of patient triage and	Risk of overfitting the model to specific datasets, limited generalizability

Author (year)	Country	Clinical domain	Study design	Objective	Clinical application	Limitations
				ED patients, aiming to improve patient flow and resource allocation	resource management, enhanced decision-making support for clinicians by providing real-time predictive analytics	across different ED settings, potential bias due to missing data or unrepresentative sample, integration challenges with existing hospital information systems
Tuwatananurak et al (2019) [36]	United States	Surgery	Diagnostic test study	To evaluate whether ML models can improve the estimation of surgical case duration compared to traditional methods, aiming to optimize operating room scheduling and resource allocation	Accurate prediction of surgical case durations, optimization of operating room schedules, enhanced decision-making support for surgical planning and resource management, improved efficiency in hospital operations	Potential overfitting to specific datasets, limited generalizability across different surgical specialties or hospitals, Data quality issues such as missing or inaccurate data entries, integration challenges with existing hospital information systems
Bertsimas et al (2020) [41]	United States	Cardiology	Clinical prediction	To develop and validate a ML model to provide personalized treatment recommendations for patients with coronary artery disease, aiming to improve patient outcomes by optimizing treatment strategies	Personalized treatment recommendations based on individual patient characteristics, prediction of treatment effectiveness and adverse events, enhanced decision-making support for cardiologists, improved patient outcomes through optimized therapy selection	Potential overfitting to specific datasets, limited generalizability across different populations or health care systems, data quality issues such as missing or inaccurate data entries, ethical considerations regarding the use of ML in clinical decision-making, integration challenges with existing electronic health records systems

^aAI: artificial intelligence.

^bICU: intensive care unit.

^cABPM: ambulatory blood pressure monitoring.

^dMSK: musculoskeletal.

^ePCOS: polycystic ovary syndrome.

^fGDPR: General Data Protection Regulation.

^gLVO: large vessel occlusion.

^hML: machine learning.

ⁱHCT: high-confidence tumor.

^jCOPD: chronic obstructive pulmonary disease.

^kICH: intracranial hemorrhage.

^lCT: computed tomography.

^mED: emergency department.

Table 5. Distribution of included studies by clinical domain, study design, and country of origin.

Category	Studies, n (%)
Clinical domain	
Nonspecific applications	10 (34.5)
Radiology	6 (20.7)
Emergency medicine	5 (17.2)
Gynecology	2 (6.9)
Cardiology	2 (6.9)
Surgery	1 (3.4)

Category	Studies, n (%)
Psychiatry	1 (3.4)
Nursing	1 (3.4)
Chronic disease	1 (3.4)
Study design	
Diagnostic test	8 (27.6)
Qualitative	6 (20.7)
Cohort	6 (20.7)
Clinical prediction	5 (17.2)
Cross-sectional	3 (10.3)
Economic evaluation	1 (3.4)
Country of origin	
United States	12 (41.4)
United Kingdom	3 (10.3)
The Netherlands	2 (6.9)
Korea	2 (6.9)
Italy	2 (6.9)
Germany	2 (6.9)
China	2 (6.9)
Uganda	1 (3.4)
Switzerland	1 (3.4)
India	1 (3.4)
France	1 (3.4)

In terms of study design, diagnostic test studies (8/29, 27.6%) [18,22-24,31,36,44,46] were the most common. Other study designs included cohort studies (6/29, 20.7%) [20,32,34,38,39,43], qualitative studies (6/29, 20.7%) [21,26,30,35,40,45], clinical prediction studies (5/29, 17.2%) [27,28,33,37,41], cross-sectional studies (3/29, 10.3%) [19,25,42], and economic evaluation studies (1/29, 3.4%) [29].

Geographically, the majority of studies originated from high-income countries (Table 6). The United States accounted for nearly half (12/29, 41.3%) [18,20,24,26,29,36,39-41,43-45], reflecting strong emphasis on EHR-linked AI, radiology, cardiology, and workflow optimization. European contributions included the United Kingdom (3/29, 10.3%) [22,34,38], Germany (2/29, 6.9%) [21,30], Italy (2/29, 6.9%) [19,42], France (1/29, 3.4%) [33], the Netherlands (2/29, 6.9%)

[28,37], and Switzerland (1/29, 3.4%) [46], with common emphases on General Data Protection Regulation compliance, data sharing, and national cardiovascular initiatives. China contributed 2 studies (2/29, 6.9%) [25,35], reporting multicenter deployments (eg, DeepSeek across 90 tertiary hospitals) and AI integration with blockchain and wearable technologies. Korea contributed 2 studies (2/29, 6.9%) [27,32], focusing on nursing decision support and interoperable clinical decision support systems. Emerging economies were also represented: India (blockchain-enabled gynecology) [23] and Uganda (trauma triage in low-resource settings) [31]. Collectively, these geographic patterns demonstrate United States and European dominance but also highlight distinct implementation trajectories and challenges from Asia and other regions.

Table 6. National-level patterns of hospital AI^a Implementation.

Country/region	Studies, n	National/institutional strategy	Clinical focus	Reported benefits and barriers
China	2	Smart hospital initiatives, large-scale deployment	Multihospital imaging (DeepSeek); chronic disease management (AI+block-chain + wearables)	Improved decision support, high upfront costs, privacy/security concerns
United States	12	Data governance, interoperability, FDA ^b oversight	Radiology, cardiology, emergency stroke, laboratory optimization, surgery	Strong EHR ^c -linked AI, adoption fragmented, dataset bias, explainability needs
United Kingdom	3	National programs, ethics, and privacy focus	Cardiovascular AI program, radiology triage, cohort, and risk models	Privacy governance emphasized, limited scale, implementation costs

Country/region	Studies, n	National/institutional strategy	Clinical focus	Reported benefits and barriers
The Netherlands	2	Hospital innovation pilots, workflow optimization	Emergency department prediction models (mortality, admission risk)	Improved triage and flow, generalizability limited, integration challenges
Korea	2	Interoperable CDSS ^d , generative AI in nursing	Nursing documentation, cross-setting CDSS	Workflow support, staff resistance, training data requirements
Italy	2	Efficiency and specialty focused	Energy management and psychiatry (ASD ^e treatment)	Cost savings potential, privacy and acceptance challenges
Germany	2	Hospital AI adoption studies	Case-based AI adoption analysis	GDPR ^f compliance and high-quality data needs
France, Switzerland	2	Specialty-specific pilots	ICU ^g prediction (France); ICH ^h detection (Switzerland)	Single-center focus and generalizability limits
India, Uganda	2	Low-/middle-income strategies	PCOS ⁱ management with blockchain (India); trauma triage (Uganda)	Infrastructure limits and workforce adaptation

^aAI: artificial intelligence.

^bFDA: US Food and Drug Administration.

^cEHR: electronic health record.

^dCDSS: clinical decision support system.

^eASD: autism spectrum disorder.

^fGDPR: General Data Protection Regulation.

^gICU: intensive care unit.

^hICH: intracranial hemorrhage.

ⁱPCOS: polycystic ovary syndrome.

Quality Assessment

The quality of the 29 studies was evaluated using the CASP standard. The studies comprised 6 study types: economic evaluations (n=1) [29], clinical prediction studies (n=5) [27, 28,33,37,41], diagnostic test studies (n=8) [18,22-24,31,36, 44,46], cohort studies (n=6) [20,32,34,38,39,43], qualitative studies (n=6) [21,26,30,35,40,45], and cross-sectional studies (n=3) [19,25,42]. Three studies (3/29, 10.3%) [20, 23,34] scored below 40% of CASP items due to insufficient methodological descriptions and unclear recruitment or analysis procedures. Fifteen studies (15/29, 51.7%) [18, 21,22,24,25,27,28,33,37,41-46] met between 50% and 70% of the criteria, whereas 11 studies (11/29, 37.9%) [19,26,29-32,35,36,38-40] exceeded 80%. In diagnostic and prediction studies, common limitations included incomplete reporting of recruitment processes, lack of external validation, and absence of blinding. Despite these weaknesses, most studies demonstrated clear research aims and appropriate methodological choices. Detailed CASP scores for each study are presented in [Multimedia Appendix 3](#).

RQ1 Findings

Overall Maturity Across the 5 Layers

To validate the proposed 5-layer architecture, all 29 studies were systematically mapped to the framework based on

the maturity levels ([Table 7](#)). The application layer (mean 3.17, SD 0.85) and data layer (mean 3.00, SD 0.76) demonstrated the highest maturity, followed by the algorithm layer (mean 2.79, SD 0.77) and infrastructure (mean 2.79, SD 1.70) layers, the latter showing considerable variability across hospitals. Security and compliance layer (mean 1.69, SD 1.89) remained the least mature and most inconsistently addressed across studies. These findings suggest that research has higher maturity on data readiness, model development, and workflow integration, whereas infrastructure and governance considerations showed both lower maturity and greater variability, suggesting that technical capacity, institutional governance, and compliance mechanisms remain unevenly developed and inconsistently reported in primary studies. Detailed evidence for each mapping decision is provided in [Multimedia Appendix 4](#).

Table 7. Five-layer evidence matrix.

Study	Infrastructure	Data	Algorithm	Application	Security and compliance
Ahsen et al [29]	0	2	2	2	0
Boussen et al [33]	0	3	3	2	0

Study	Infrastructure	Data	Algorithm	Application	Security and compliance
Alam et al [18]	1	2	2	2	2
Areias et al [20]	4	3	3	4	0
Chen et al [25]	5	4	4	5	4
Farghaly and Deshpande [44]	2	2	3	2	0
Fairbairn et al [34]	5	4	4	4	4
Jaganathan and Natesan [23]	4	4	4	3	5
Muntasir et al [40]	4	4	3	4	3
Ju et al [32]	2	3	3	3	0
Klump et al [21]	1	2	2	2	3
Le et al [39]	4	3	3	4	0
Li et al [45]	4	3	3	4	0
Lin et al [43]	0	2	2	2	0
Novak et al [26]	0	2	0	3	2
Nsubuga et al [31]	0	2	3	2	0
Pariso et al [19]	4	3	3	4	0
Vignapiano et al [42]	3	3	2	3	0
Roppelt et al [30]	1	2	2	3	3
Xie et al [35]	4	4	3	3	5
Yang et al [38]	4	3	3	3	0
Yoo et al [27]	4	4	3	4	4
Wshah et al [24]	3	3	3	3	0
Tam et al [22]	4	3	3	4	4
Seyam et al [46]	4	3	3	4	4
Raven et al [28]	4	3	3	4	3
De Hond et al [37]	4	4	3	3	3
Tuwatananurak et al [36]	2	3	3	3	0
Bertsimas et al [41]	4	4	3	3	0
Mean (SD)	2.79 (1.70)	3.00 (0.76)	2.79 (0.77)	3.17 (0.85)	1.69 (1.89)

Evidence Stratified by Study Design

Given the methodological heterogeneity of the included studies, we conducted a stratified synthesis by study design (Table 8). The results of this analysis show that studies of clinical prediction, diagnostic test, and cohort studies achieved higher maturity in the data, algorithm, and application layers, particularly the clinical prediction studies, which showed the most consistent and advanced technical implementation. In contrast, qualitative and cross-sectional studies exhibited greater maturity variation, contributing more

substantially to the infrastructure and security and compliance layers. The single economic evaluation demonstrated moderate maturity, limited mainly to the technical layers. This stratified synthesis highlights how methodological design shapes the visibility of different layers, with quantitative evaluation studies emphasizing technical robustness and data integration, whereas qualitative designs better capture infrastructural and governance maturity essential for sustainable AI platform development.

Table 8. Five-layer mapping stratified by study design.^a

Design	Infrastructure, mean (SD)	Data, mean (SD)	Algorithm, mean (SD)	Application, mean (SD)	Security and compliance, mean (SD)
Clinical prediction	4.00 (0.00)	3.75 (0.50)	3.00 (0.00)	3.50 (0.58)	2.50 (1.73)
Cohort	3.17 (1.83)	3.00 (0.63)	3.00 (0.63)	3.33 (0.82)	0.67 (1.63)
Cross-sectional	4.00 (1.00)	3.33 (0.58)	3.00 (1.00)	4.00 (1.00)	1.33 (2.31)
Diagnostic test	2.50 (1.51)	2.75 (0.71)	3.00 (0.53)	2.88 (0.83)	1.88 (2.17)
Economic evaluations	0	2	2	2	0
Qualitative	2.00 (1.87)	2.80 (1.10)	2.00 (1.22)	3.00 (0.71)	3.20 (1.10)

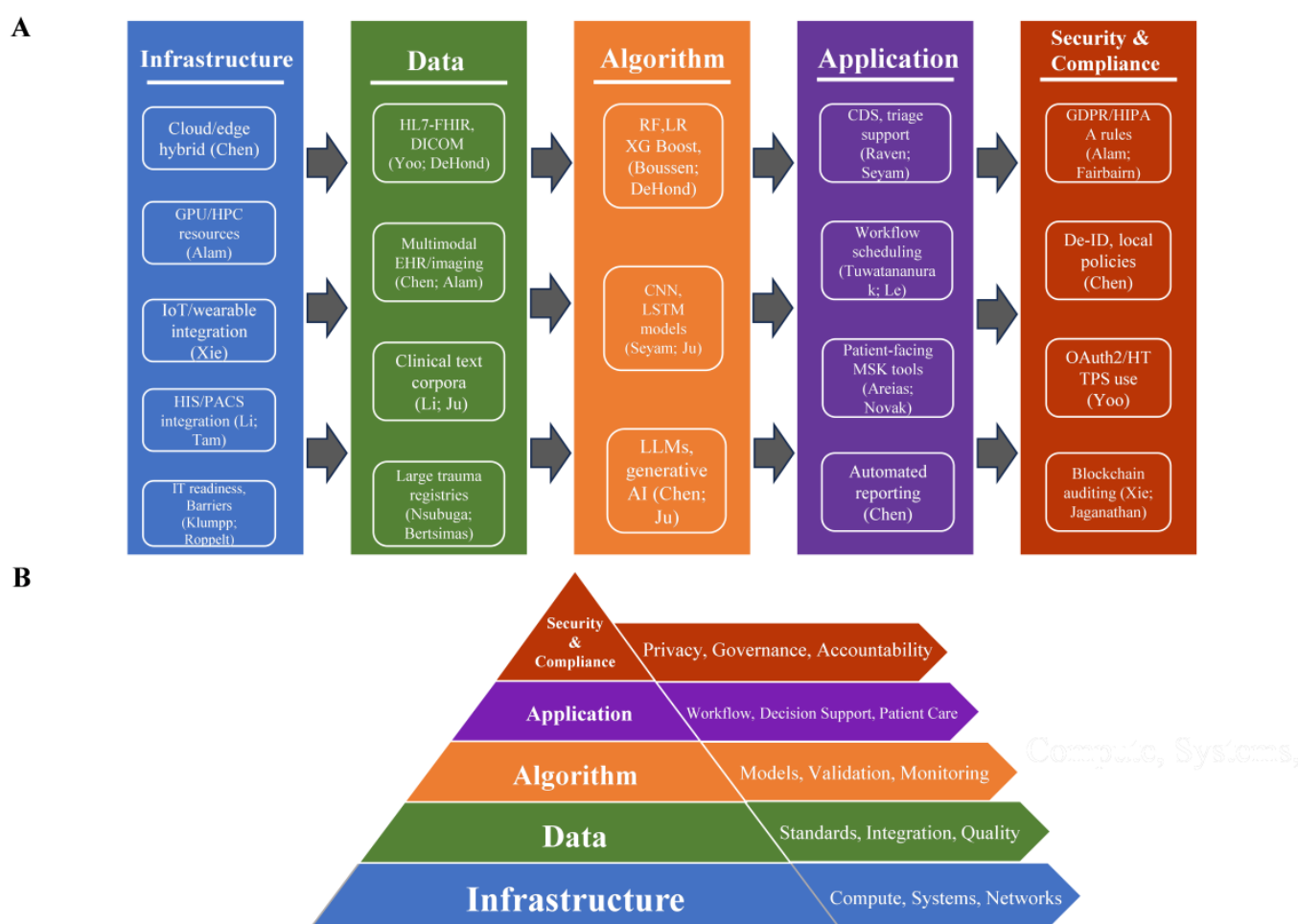
^aValues represent the mean (SD) of weighted maturity scores across studies of the same design type, calculated within each of the 5 layers of the proposed hospital artificial intelligence platform architecture: infrastructure, data, algorithm, application, and security and compliance. For the economic evaluations category, only 1 study was available; therefore, only the mean score is reported without SD.

Mapping the Evidence to the 5-Layer Framework

We mapped the identified evidence to the 5-layer framework: (1) infrastructure layer, (2) data layer, (3) algorithm layer, (4) application layer, and (5) security and compliance layer. To

achieve this, we reviewed the specific evidence in [Multimedia Appendix 4](#) and selected high-frequency examples to map into the architecture. [Figure 3A](#) shows the detailed evidence mapping, where each layer is populated by distinct categories of evidence, and [Figure 3B](#) presents the simplified conceptual pyramid as an overview.

Figure 3. Overview and evidence mapping of the hospital AI platform architecture. (A) Extracted study-level findings were synthesized and organized within the 5-layer AI platform architecture. Each box summarizes commonly reported evidence elements, with exemplar studies cited in parentheses. (B) Simplified 5-layer pyramid showing broad categories (infrastructure: compute, systems, networks; data: standards, integration, quality; algorithm: models, validation, monitoring; application: workflow, decision support, patient care; security and compliance: privacy, governance, accountability). AI: artificial intelligences; CDS: clinical decision support; CNN: convolutional neural network; De-ID: deidentification; DICOM: Digital Imaging and Communications in Medicine; EHR: electronic health record; EMPI: Enterprise Master Patient Index; FHIR: Fast Healthcare Interoperability Resources; GDPR: General Data Protection Regulation; HIPAA: Health Insurance Portability and Accountability Act; HIS: health information system; HL7: Health Level Seven; IoT: Internet of Things; LLM: large language model; LR: logistic regression; LSTM: long short-term memory; MSK: musculoskeletal; PACS: picture archiving and communication system; RF: random forest.



RQ2 Findings

While RQ1 examined the maturity of individual layers, RQ2 explored the interrelationships among layers in hospital AI systems. The weighted co-occurrence heatmap ([Figure 4](#)) clearly shows that data, algorithm, and application exhibited the strongest interconnections in the center. The weighted

Jaccard similarity heatmap ([Figure 5](#)) further confirmed this pattern, showing strong maturity overlap among the 3 layers (data-application=0.85, data-algorithm=0.89, and algorithm-application=0.80). This “core triad” indicates that data preparation, model development, and workflow integration are inseparable steps in most implementations.

Figure 4. Weighted co-occurrence heatmap across the 5-layer architecture.

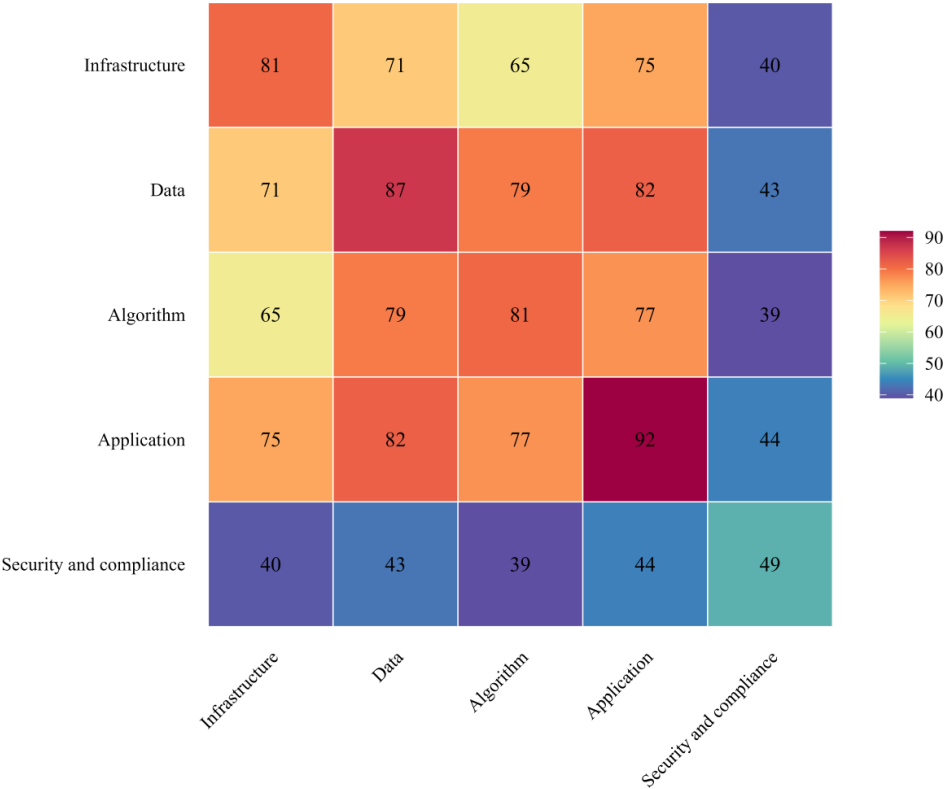
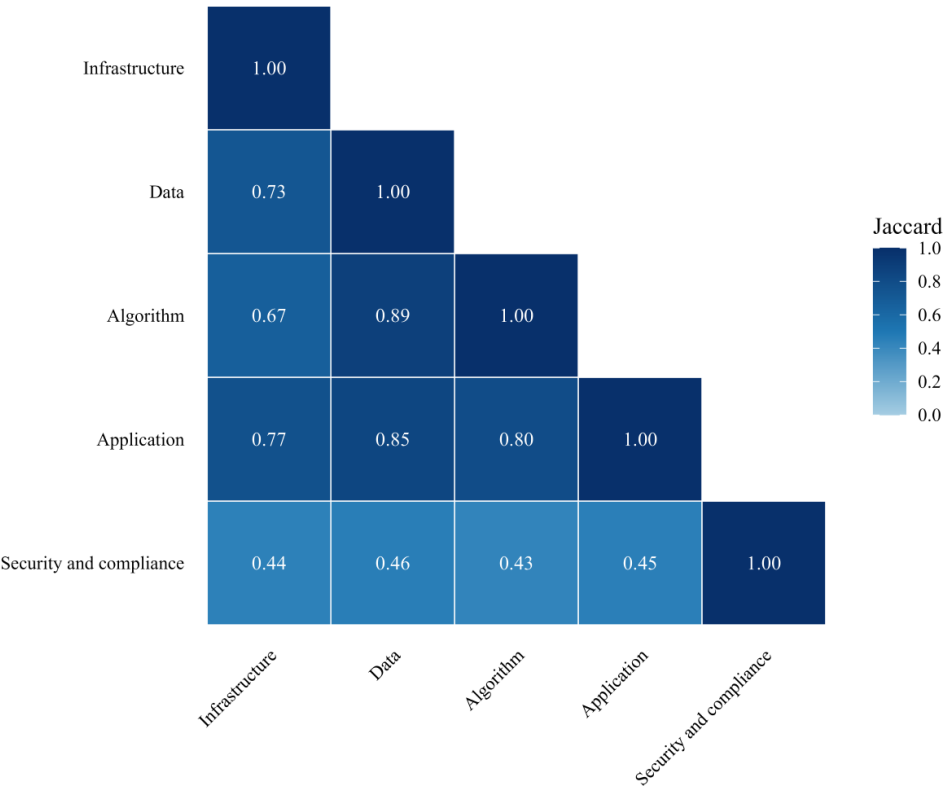


Figure 5. Weighted Jaccard similarity heatmap across the 5-layer architecture. Each cell represents the weighted Jaccard similarity between two layers across 29 included studies. Values were computed as the sum of the minimum maturity scores for each layer pair divided by the sum of their maximum scores. Higher values (darker shades) indicate stronger maturity alignment between layers. The data, algorithm, and application layers form a highly cohesive core (Jaccard=0.80-0.89), whereas security and compliance exhibit weaker associations (Jaccard=0.43-0.46), highlighting its peripheral integration in current hospital artificial intelligence implementations.



Cells show weighted co-occurrence scores (sum of per-study maturity shared by each pair of layers, 0-5 scale), so larger values indicate stronger cross-layer coupling. Diagonal cells give the cumulative maturity for each layer. The central

triad, data, algorithm, and application, shows the strongest coupling (eg, data-application=82; data-algorithm=79; and algorithm-application=77), whereas security and compliance is consistently weaker with other layers. Warmer colors denote higher weighted coupling.

In contrast, security and compliance appeared peripheral, with lower scores when coupled with the other layers (Jaccard=0.43-0.46). Infrastructure demonstrated moderate connectivity with application (Jaccard=0.77) and data (Jaccard=0.73), implying that infrastructural maturity often co-develops with technical capability but is not systematically aligned with governance or oversight mechanisms. These patterns point to the need for earlier integration of governance and compliance into platform design, ensuring they function as core rather than peripheral components.

Implementation Examples

To assess if the proposed 5-layer hospital AI platform (infrastructure, data, algorithm, application, and security and compliance) can work in practice, we narratively synthesize 4 fielded deployments from our 29 included studies, in settings with explicit clinical integration and measurable end points.

Example 1

An AI application for noncontrast head computed tomography (CT) [46] was deployed in an emergency radiology setting to flag multiple intracranial hemorrhage subtypes and reprioritize critical cases in the radiology worklist. The deployed system, using PACS infrastructure (layer 1: infrastructure), which focused on curated CT datasets, engaged difficult cases, such as postoperative changes and artifacts (layer 2: data). A multisubtype detection algorithm (layer 3: algorithm) applied with balanced sensitivity and false positives allowing for better triage. This led to automated elevation of priority levels and reordering of queues clinically, which reduced turnaround times for report generation and emergency department length of stay (layer 4 application). To reduce the chances of false alarms being raised through human verification and not to miss out on any subtype, standard operating procedures were framed at this layer (layer 5: security and compliance).

Example 2

An AI-generated first-reader system was integrated into current workflows of chest radiographs to flag high-confidence tumor cases for earlier review [44]. The deployment was able to link with the digital radiography acquisition, PACS, Radiology Information System, reporting system for batch inference, and rule-based triggers (layer 1: infrastructure). Through the use of large-scale chest X-ray datasets in Digital Imaging and Communications in Medicine format with structured reads to calibrate thresholds and monitor downstream confirmations (layer 2: data). The software gave triage scores for nodules, masses, and hilar enlargement. These scores formed the high-confidence queue (layer 3: algorithm). At the application level, this facilitated reprioritization and consistency-oriented quality management across radiologists (layer 4: application), whereas false-positive

audits and drift monitoring ensured compliance and performance stability over time (layer 5: security and compliance).

Example 3

A machine learning-enabled system at primary stroke centers aided in the early detection of large-vessel occlusion and accelerated transfer protocols [39]. Through CT and computed tomography angiography (CTA) acquisition, PACS, alerting on-call, and interhospital coordination platforms, the system was tightly coupled for smooth integration of infrastructure (layer 1: infrastructure). It combined imaging with time-stamped process data (arrival, transfer, reperfusion) to monitor pathway performance (layer 2: data). The algorithm automatically stratified large vessel occlusion cases, triggering alerts and activating predefined stroke pathways (layer 3: algorithm). At the application level, this enabled the rapid activation of “green channel,” transport prioritization, and standardization of decision points in acute stroke care (layer 4: application). As per the established transfer policies and accountability along the care chain, cross-site data sharing took place in a manner that governed the deployment according to the governance and compliance expectations (layer 5: security and compliance).

Example 4

A platform-scale AI deployment was rolled out across 90 tertiary hospitals to provide image analysis and clinical decision support at scale [25]. The infrastructure consisted of multisite compute and networking resources with containerized inference engines with full-stack EHR and PACS integration for seamless updates (layer 1: infrastructure). Data was curated across institutions, combining different types and standardization procedures as part of (layer 2: data) for use in other contexts. A portfolio of task-specific models was built as a platform asset with continuous iteration (layer 3: algorithm), providing triage, detection, recommendation, quality control, and more capabilities. The platform created common clinical entry points for decision support (layer 4: application) to streamline workflows across sites. Security and compliance were formalized through access control, audit trails, change management, and staff training, underscoring its readiness for large-scale operations (layer 5: security and compliance).

Across these examples, 3 cross-cutting themes emerged. First, workflow-native integration, such as worklist reprioritization, automated alerts, and transfer coordination, was the critical pathway for translating algorithmic outputs into measurable clinical benefits. Second, infrastructure and data robustness are critical to sustaining application-level improvements. Deployments without standard pipelines are fragile. Third, governance was put into action through standard operating procedures, threshold calibration, and audit mechanisms to control false positives, drift, and intersite variation. The 5-layer model is practically viable and also highlights where further investment is needed, including in infrastructure resilience, data governance, and clinician adoption.

Discussion

Principal Findings

A 5-layer hospital AI platform model (infrastructure, data, algorithm, application, and security and compliance) was developed by synthesizing 4 reference frameworks. From 283 records screened, 29 studies (29/283, 10.2%) were included with high interrater reliability ($\kappa=0.98$). Most studies were diagnostic test (8/29, 27.6%) or qualitative (6/29, 20.7%), and almost half were conducted in the United States (12/29, 41.4%). Quality varied: only 37.9% (11/29) of studies achieved more than 80% of CASP items, whereas 10.3% (3/29) scored below 40%.

Evidence mapping showed the application (mean 3.17, SD 0.85), data (mean 3.00, SD 0.76), and algorithm (mean 2.79, SD 0.77) layers have the highest and most balanced maturity, forming a tightly integrated core triad. Weighted co-occurrence analysis demonstrated the strongest interconnections among these 3 layers (data-application=82, data-algorithm=79, and algorithm-application=77), and weighted Jaccard similarity indices confirmed substantial maturity overlap (data-algorithm=0.89, data-application=0.85, and algorithm-application=0.80). The infrastructure layer (mean 2.79, SD 1.70) displayed moderate maturity but high variability. The security and compliance layer (mean 1.69, SD 1.89) remained the least mature and weakly connected to the others (Jaccard=0.43-0.46). Stratification by study design showed that technical studies (diagnostic, prediction, and cohort) achieved higher maturity within the technical core, whereas qualitative and cross-sectional studies more often addressed infrastructure and governance that remain underdeveloped in most technical evaluations.

Comparison With Prior Work

There was a maturity imbalance across the layers, which were dominated by technical layers. There were similar trends observed in earlier reviews, which showed that algorithm performance and workflow integration achieved higher maturity, whereas infrastructure, ethics, and compliance were less developed [47,48]. Several factors may explain this. The focus of studies and publication bias has favored predictive accuracy and model validation [47,49]. Most studies were led by clinical and technical teams. The governance and IT planning teams were featured less due to disciplinary divides [50]. Many studies described pilot projects, where compliance and infrastructure emerged later in scaling.

Domain-specific differences were noted. In emergency imaging [39,46], infrastructure and application layers were highly mature due to the urgency of real time, whereas compliance was limited to operating procedures. In routine imaging [22,44], data-algorithm dominated, with application relying on threshold tuning. In chronic disease management [38,41], data governance and compliance were critical for privacy-preserving integration, and application focused on longitudinal risk stratification. In deployments across multiple hospitals [25], infrastructure and data pipelines were key bottlenecks, and compliance was institutionalized through

audit trails and training. Negative outcomes were also reported: chest radiograph triage increased false positives in patients with comorbidities [22], ICH detection struggled with specific subtypes [46], and multihospital systems faced high costs and low adoption [25]. These findings confirm that success depends on context, workflow fit, and governance.

Implications for Practice and Policy

Administrators can use the 5-layer model for phased investments, particularly in the areas of infrastructure and compliance, to eradicate blockages at a later stage [51,52]. Policymakers can promote compliance-by-design standards, requiring privacy, accountability, and explainability from the outset [53,54]. Vendors can align products with hospital priorities by embedding interoperability and monitoring tools [52]. Funding organizations and health systems should support the underdeveloped layers, especially governance and infrastructure through training and cross-institutional collaboration [51].

Challenges and Limitations of AI Deployment in Hospitals

Technical and infrastructural barriers were widely reported. Data were often fragmented across EHR, LIS, and PACS, and many systems lacked application programming interface support or sufficient computing capacity [55,56]. These weaknesses limited scalability, whereas bandwidth bottlenecks reduced real-time performance and inadequate monitoring allowed model drift [22,25,46]. Solutions that have been proposed include HL7 FHIR-based data lakes [57], hybrid cloud-edge architectures [58], and continuous monitoring with automated retraining pipelines [59,60].

Organizational and adoption barriers also affected implementation. Clinician skepticism, workflow misalignment [22,46], and lack of structured feedback were common problems [30,35,61,62]. Better outcomes were described when interdisciplinary teams were formed, clinical champions were engaged early, and projects targeted low-risk and high-value use cases [63,64].

Regulatory and compliance gaps were identified as the weakest layer. Privacy safeguards, governance frameworks, and liability protocols were rarely embedded into early projects [21]. Compliance-by-design approaches have been recommended, supported by federated learning [65], explainability mechanisms, and blockchain-based audit trails [66], together with early engagement of regulators [67].

Economic and resource barriers were another critical concern. High upfront and maintenance costs, uncertain return on investment, and misalignment with hospital budget cycles were repeatedly described [25,68]. Strategies such as phased investments, AI-as-a-service models, and shared consortia were suggested to reduce costs and support sustainable deployment [69].

Limitations of This Review and Future Research

This review has several limitations. A limitation is that only peer-reviewed English-language publications were included, which may have excluded valuable implementation reports published in other languages or as gray literature. This restriction was applied to maintain consistency and reliability in the 0 to 5 ordinal maturity scoring, as translation variability could compromise coding accuracy and interrater agreement. In addition, most gray literature lacks formal peer review or standardized reporting, which may introduce methodological inconsistency and compromise the overall reliability of evidence synthesis. Therefore, it was intentionally excluded to preserve data quality and comparability. Another limitation is that studies were mapped to the 5-layer framework using an ordinal maturity scoring, which may not fully reflect the complexity of AI applications. A further limitation is that the framework was validated only through literature synthesis and has not been prospectively tested in real hospital environments, such as in resource-limited settings or smaller hospitals.

Future research should extend evidence retrieval to include non-English and gray literature through calibrated multilingual screening and curated institutional sources. Further refinement of maturity metrics and cross-layer evaluation methods is warranted to better represent the dynamic evolution of hospital AI systems. Furthermore, prospective multicenter studies are also required to validate the framework in practice and to test its scalability across diverse hospital settings.

Conclusions

The integration of hospital information systems with AI is essential for the digital transformation of health care. Evidence from 29 empirical studies was synthesized with established frameworks to validate a 5-layer architecture. This model provides both theoretical and practical guidance for platform-level AI in hospitals. The framework can be applied by researchers, practitioners, and policymakers to support the development of scalable, secure, and clinically integrated AI platforms.

Acknowledgments

The authors did not use any generative artificial intelligence tools for this research paper.

Funding

This work was supported by Teaching Research Project of Wuhan University Medical School, 2025 (grant 2025YB34) and Science and Technology Innovation Cultivation Funding of Zhongnan Hospital of Wuhan University (grant CXPY2022049).

Data Availability

All data generated or analyzed during this study are included in this published article and its multimedia appendices.

Authors' Contributions

Conceptualization: MM (lead), YY (equal)

Writing – original draft: MM (lead), YY (equal)

Formal analysis: YJ

Investigation: YJ

Data curation: HX

Methodology: HX (lead), YY (equal)

Visualization: GD (lead), YY (equal)

Validation: WK

Funding acquisition: YY

Supervision: YY

Writing – review & editing: YY (lead), MM (supporting), YJ (supporting), HX (supporting), GD (supporting), WK (supporting)

Conflicts of Interest

None declared.

Multimedia Appendix 1

Search strategy and strings for all databases.

[\[DOCX File \(Microsoft Word File\), 18 KB-Multimedia Appendix 1\]](#)

Multimedia Appendix 2

The 2x2 contingency table used for interrater reliability calculation.

[\[DOCX File \(Microsoft Word File\), 18 KB-Multimedia Appendix 2\]](#)

Multimedia Appendix 3

Detailed results of the Critical Appraisal Skills Programme quality assessment for included studies.

[\[DOCX File \(Microsoft Word File\), 54 KB-Multimedia Appendix 3\]](#)

Multimedia Appendix 4

Detailed evidence mapping of all included studies to the 5-layer framework.

[\[DOCX File \(Microsoft Word File\), 29 KB-Multimedia Appendix 4\]](#)

Checklist 1

PRISMA 2020 checklist.

[\[DOCX File \(Microsoft Word File\), 33 KB-Checklist 1\]](#)

References

1. Li D, Hu Y, Liu S, et al. A latent profile analysis of Chinese physicians' workload tethered to paperwork during outpatient encounters. *Front Public Health*. 2022;10. [doi: [10.3389/fpubh.2022.854772](https://doi.org/10.3389/fpubh.2022.854772)]
2. Abbasi R, Sadeqi Jabali M, Khajouei R, Tadayon H. Investigating the satisfaction level of physicians in regards to implementing medical Picture Archiving and Communication System (PACS). *BMC Med Inform Decis Mak*. Aug 5, 2020;20(1):180. [doi: [10.1186/s12911-020-01203-0](https://doi.org/10.1186/s12911-020-01203-0)] [Medline: [32758220](https://pubmed.ncbi.nlm.nih.gov/32758220/)]
3. Younis HA, Eisa TAE, Nasser M, et al. A systematic review and meta-analysis of artificial intelligence tools in medicine and healthcare: applications, considerations, limitations, motivation and challenges. *Diagnostics (Basel)*. Jan 4, 2024;14(1):109. [doi: [10.3390/diagnostics14010109](https://doi.org/10.3390/diagnostics14010109)] [Medline: [38201418](https://pubmed.ncbi.nlm.nih.gov/38201418/)]
4. Li YH, Li YL, Wei MY, Li GY. Innovation and challenges of artificial intelligence technology in personalized healthcare. *Sci Rep*. Aug 16, 2024;14(1):18994. [doi: [10.1038/s41598-024-70073-7](https://doi.org/10.1038/s41598-024-70073-7)]
5. Nair M, Svedberg P, Larsson I, Nygren JM. A comprehensive overview of barriers and strategies for AI implementation in healthcare: mixed-method design. *PLoS ONE*. 2024;19(8):e0305949. [doi: [10.1371/journal.pone.0305949](https://doi.org/10.1371/journal.pone.0305949)] [Medline: [39121051](https://pubmed.ncbi.nlm.nih.gov/39121051/)]
6. Alhajeri M, Shah SGS. Limitations in and solutions for improving the functionality of picture archiving and communication system: an exploratory study of PACS professionals' perspectives. *J Digit Imaging*. Feb 2019;32(1):54-67. [doi: [10.1007/s10278-018-0127-2](https://doi.org/10.1007/s10278-018-0127-2)] [Medline: [30225824](https://pubmed.ncbi.nlm.nih.gov/30225824/)]
7. Torab-Miandoab A, Samad-Soltani T, Jodati A, Rezaei-Hachesu P. Interoperability of heterogeneous health information systems: a systematic literature review. *BMC Med Inform Decis Mak*. Jan 24, 2023;23(1):18. [doi: [10.1186/s12911-023-02115-5](https://doi.org/10.1186/s12911-023-02115-5)] [Medline: [36694161](https://pubmed.ncbi.nlm.nih.gov/36694161/)]
8. Pournik O, Mukherjee T, Ghalichi L, Arvanitis TN. How interoperability challenges are addressed in healthcare IoT projects. *Stud Health Technol Inform*. Oct 20, 2023;309:121-125. [doi: [10.3233/SHTI230754](https://doi.org/10.3233/SHTI230754)] [Medline: [37869820](https://pubmed.ncbi.nlm.nih.gov/37869820/)]
9. Farič N, Hinder S, Williams R, et al. Early experiences of integrating an artificial intelligence-based diagnostic decision support system into radiology settings: a qualitative study. *J Am Med Inform Assoc*. Dec 22, 2023;31(1):24-34. [doi: [10.1093/jamia/ocad191](https://doi.org/10.1093/jamia/ocad191)] [Medline: [37748456](https://pubmed.ncbi.nlm.nih.gov/37748456/)]
10. Petersson L, Larsson I, Nygren JM, et al. Challenges to implementing artificial intelligence in healthcare: a qualitative interview study with healthcare leaders in Sweden. *BMC Health Serv Res*. Jul 1, 2022;22(1):850. [doi: [10.1186/s12913-022-08215-8](https://doi.org/10.1186/s12913-022-08215-8)] [Medline: [35778736](https://pubmed.ncbi.nlm.nih.gov/35778736/)]
11. Snowdon DA. Digital health: a framework for healthcare transformation. *Healthcare Information and Management Systems*. URL: <https://www.himss.org/sites/hde/files/media/file/2022/12/21/dhi-white-paper.pdf> [Accessed 2025-12-08]
12. Framework and standards for country health information systems, 2 ed. World Health Organization. Apr 24, 2023. URL: <https://www.who.int/publications/i/item/9789241595940> [Accessed 2025-12-08]
13. Sony M, Naik S. Industry 4.0 integration with socio-technical systems theory: a systematic review and proposed theoretical model. *Technol Soc*. May 2020;61:101248. [doi: [10.1016/j.techsoc.2020.101248](https://doi.org/10.1016/j.techsoc.2020.101248)]
14. Iglesias CA, Favenza A, Carrera Á. A big data reference architecture for emergency management. *Information*. Dec 2020;11(12):569. [doi: [10.3390/info11120569](https://doi.org/10.3390/info11120569)]
15. Page MJ, Moher D, Bossuyt PM, et al. PRISMA 2020 explanation and elaboration: updated guidance and exemplars for reporting systematic reviews. *BMJ*. Mar 29, 2021;372:n160. [doi: [10.1136/bmj.n160](https://doi.org/10.1136/bmj.n160)] [Medline: [33781993](https://pubmed.ncbi.nlm.nih.gov/33781993/)]
16. Gomes J, Romão M. Evaluating maturity models in healthcare information systems: a comprehensive review. *Healthcare (Basel)*. Jul 29, 2025;13(15):1847. [doi: [10.3390/healthcare13151847](https://doi.org/10.3390/healthcare13151847)] [Medline: [40805880](https://pubmed.ncbi.nlm.nih.gov/40805880/)]
17. Long HA, French DP, Brooks JM. Optimising the value of the Critical Appraisal Skills Programme (CASP) tool for quality appraisal in qualitative evidence synthesis. *Res Methods Med Health Sci*. Sep 2020;1(1):31-42. [doi: [10.1177/2632084320947559](https://doi.org/10.1177/2632084320947559)]
18. Alam SF, Thongprayoon C, Miao J, et al. Advancing personalized medicine in digital health: the role of artificial intelligence in enhancing clinical interpretation of 24-h ambulatory blood pressure monitoring. *Digit Health*. 2025;11:20552076251326014. [doi: [10.1177/20552076251326014](https://doi.org/10.1177/20552076251326014)] [Medline: [40093710](https://pubmed.ncbi.nlm.nih.gov/40093710/)]

19. Pariso P, Picariello M, Marino A. AI integration in energy management: enhancing efficiency in Italian hospitals. *Health Econ Rev*. May 19, 2025;15(1):40. [doi: [10.1186/s13561-025-00638-3](https://doi.org/10.1186/s13561-025-00638-3)]
20. Areias AC, Janela D, Moulder RG, et al. Applying AI to safely and effectively scale care to address chronic MSK conditions. *J Clin Med*. Jul 26, 2024;13(15):4366. [doi: [10.3390/jcm13154366](https://doi.org/10.3390/jcm13154366)] [Medline: [39124635](https://pubmed.ncbi.nlm.nih.gov/39124635/)]
21. Klumpp M, Hintze M, Immonen M, et al. Artificial intelligence for hospital health care: application cases and answers to challenges in European hospitals. *Healthcare (Basel)*. Jul 29, 2021;9(8):961. [doi: [10.3390/healthcare9080961](https://doi.org/10.3390/healthcare9080961)] [Medline: [34442098](https://pubmed.ncbi.nlm.nih.gov/34442098/)]
22. Tam M, Dyer T, Dissez G, et al. Augmenting lung cancer diagnosis on chest radiographs: positioning artificial intelligence to improve radiologist performance. *Clin Radiol*. Aug 2021;76(8):607-614. [doi: [10.1016/j.crad.2021.03.021](https://doi.org/10.1016/j.crad.2021.03.021)] [Medline: [33993997](https://pubmed.ncbi.nlm.nih.gov/33993997/)]
23. Jaganathan G, Natesan S. Blockchain and explainable-AI integrated system for polycystic ovary syndrome (PCOS) detection. *PeerJ Comput Sci*. 2025;11:e2702. [doi: [10.7717/peerj-cs.2702](https://doi.org/10.7717/peerj-cs.2702)] [Medline: [40062287](https://pubmed.ncbi.nlm.nih.gov/40062287/)]
24. Wshah S, Xu B, Steinharter J, Reilly C, Morrisette K. Classification of clinically relevant intravascular volume status using point of care ultrasound and machine learning. *J Med Imag*. 2022;9(5):054502. [doi: [10.1117/1.JMI.9.5.054502](https://doi.org/10.1117/1.JMI.9.5.054502)]
25. Chen J, Miao C. DeepSeek deployed in 90 Chinese tertiary hospitals: how artificial intelligence is transforming clinical practice. *J Med Syst*. Apr 24, 2025;49(1):53. [doi: [10.1007/s10916-025-02181-4](https://doi.org/10.1007/s10916-025-02181-4)] [Medline: [40272650](https://pubmed.ncbi.nlm.nih.gov/40272650/)]
26. Novak LL, Harris JW, Koonce TY, Johnson KB. Design thinking in applied informatics: what can we learn from Project HealthDesign? *J Am Med Inform Assoc*. Aug 13, 2021;28(9):1858-1865. [doi: [10.1093/jamia/ocab081](https://doi.org/10.1093/jamia/ocab081)] [Medline: [34142141](https://pubmed.ncbi.nlm.nih.gov/34142141/)]
27. Yoo J, Lee J, Min JY, et al. Development of an interoperable and easily transferable clinical decision support system deployment platform: system design and development study. *J Med Internet Res*. Jul 27, 2022;24(7):e37928. [doi: [10.2196/37928](https://doi.org/10.2196/37928)] [Medline: [35896020](https://pubmed.ncbi.nlm.nih.gov/35896020/)]
28. Raven W, de Hond A, Bouma LM, Mulder L, de Groot B. Does machine learning combined with clinical judgment outperform clinical judgment alone in predicting in-hospital mortality in old and young suspected infection emergency department patients? *Eur J Emerg Med*. Jun 1, 2023;30(3):205-206. [doi: [10.1097/MEJ.0000000000000996](https://doi.org/10.1097/MEJ.0000000000000996)] [Medline: [37103898](https://pubmed.ncbi.nlm.nih.gov/37103898/)]
29. Ahsen ME, Ayvaci MUS, Mookerjee R, Stolovitzky G. Economics of AI and human task sharing for decision making in screening mammography. *Nat Commun*. Mar 7, 2025;16(1):2289. [doi: [10.1038/s41467-025-57409-1](https://doi.org/10.1038/s41467-025-57409-1)] [Medline: [40055356](https://pubmed.ncbi.nlm.nih.gov/40055356/)]
30. Roppelt JS, Jenkins A, Kanbach DK, Kraus S, Jones P. Effective adoption of artificial intelligence in healthcare: a multiple case study. *J Decis Syst*. Jan 2, 2025;34(1):2458883. [doi: [10.1080/12460125.2025.2458883](https://doi.org/10.1080/12460125.2025.2458883)]
31. Nsubuga M, Kintu TM, Please H, Stewart K, Navarro SM. Enhancing trauma triage in low-resource settings using machine learning: a performance comparison with the Kampala Trauma Score. *BMC Emerg Med*. Jan 23, 2025;25(1):14. [doi: [10.1186/s12873-025-01175-2](https://doi.org/10.1186/s12873-025-01175-2)] [Medline: [39849342](https://pubmed.ncbi.nlm.nih.gov/39849342/)]
32. Ju H, Park M, Jeong H, et al. Generative AI-based nursing diagnosis and documentation recommendation using virtual patient electronic nursing record data. *Healthc Inform Res*. Apr 2025;31(2):156-165. [doi: [10.4258/hir.2025.31.2.156](https://doi.org/10.4258/hir.2025.31.2.156)] [Medline: [40384067](https://pubmed.ncbi.nlm.nih.gov/40384067/)]
33. Boussen S, Benard-Tertrais M, Ogéa M, et al. Heart rate complexity helps mortality prediction in the intensive care unit: a pilot study using artificial intelligence. *Comput Biol Med*. Feb 2024;169:107934. [doi: [10.1016/j.combiomed.2024.107934](https://doi.org/10.1016/j.combiomed.2024.107934)] [Medline: [38183707](https://pubmed.ncbi.nlm.nih.gov/38183707/)]
34. Fairbairn TA, Mullen L, Nicol E, et al. Implementation of a national AI technology program on cardiovascular outcomes and the health system. *Nat Med*. Jun 2025;31(6):1903-1910. [doi: [10.1038/s41591-025-03620-y](https://doi.org/10.1038/s41591-025-03620-y)] [Medline: [40186078](https://pubmed.ncbi.nlm.nih.gov/40186078/)]
35. Xie Y, Lu L, Gao F, et al. Integration of artificial intelligence, blockchain, and wearable technology for chronic disease management: a new paradigm in smart healthcare. *Curr Med Sci*. Dec 2021;41(6):1123-1133. [doi: [10.1007/s11596-021-2485-0](https://doi.org/10.1007/s11596-021-2485-0)] [Medline: [34950987](https://pubmed.ncbi.nlm.nih.gov/34950987/)]
36. Tuwatananurak JP, Zadeh S, Xu X, et al. Machine learning can improve estimation of surgical case duration: a pilot study. *J Med Syst*. Jan 17, 2019;43(3):44. [doi: [10.1007/s10916-019-1160-5](https://doi.org/10.1007/s10916-019-1160-5)] [Medline: [30656433](https://pubmed.ncbi.nlm.nih.gov/30656433/)]
37. De Hond A, Raven W, Schinkelshoek L, et al. Machine learning for developing a prediction model of hospital admission of emergency department patients: hype or hope? *Int J Med Inform*. Aug 2021;152:104496. [doi: [10.1016/j.ijmedinf.2021.104496](https://doi.org/10.1016/j.ijmedinf.2021.104496)] [Medline: [34020171](https://pubmed.ncbi.nlm.nih.gov/34020171/)]
38. Yang J, Clifton D, Hirst JE, et al. Machine learning-based risk stratification for gestational diabetes management. *Sensors (Basel)*. Jul 2022;22(13):4805. [doi: [10.3390/s22134805](https://doi.org/10.3390/s22134805)]
39. Le NM, Iyyangar AS, Kim Y, et al. Machine learning-enabled automated large vessel occlusion detection improves transfer times at primary stroke centers. *SVIN*. May 2024;4(3). [doi: [10.1161/SVIN.123.001119](https://doi.org/10.1161/SVIN.123.001119)]

40. Muntasir J, Rahman T, Rauf Z, Arif L. Optimizing laboratory workflow in hospitals using AI-assisted technologies. *J Bras Patol Med Lab*. 2023;24-31. [doi: [10.1900/JBPML.2023.59.005](https://doi.org/10.1900/JBPML.2023.59.005)]
41. Bertsimas D, Orfanoudaki A, Weiner RB. Personalized treatment for coronary artery disease patients: a machine learning approach. *Health Care Manag Sci*. Dec 2020;23(4):482-506. [doi: [10.1007/s10729-020-09522-4](https://doi.org/10.1007/s10729-020-09522-4)] [Medline: [33040231](https://pubmed.ncbi.nlm.nih.gov/33040231/)]
42. Vignapiano A, Monaco F, Landi S, et al. Proximity-based solutions for optimizing autism spectrum disorder treatment: integrating clinical and process data for personalized care. *Front Psychiatry*. 2024;15:1512818. [doi: [10.3389/fpsyt.2024.1512818](https://doi.org/10.3389/fpsyt.2024.1512818)] [Medline: [39911557](https://pubmed.ncbi.nlm.nih.gov/39911557/)]
43. Lin Y, Hoyt AC, Manuel VG, et al. Risk-stratified screening: a simulation study of scheduling templates on daily mammography recalls. *J Am Coll Radiol*. Mar 2025;22(3):297-306. [doi: [10.1016/j.jacr.2024.12.010](https://doi.org/10.1016/j.jacr.2024.12.010)] [Medline: [40044308](https://pubmed.ncbi.nlm.nih.gov/40044308/)]
44. Farghaly O, Deshpande P. Texture-based classification to overcome uncertainty between COVID-19 and viral pneumonia using machine learning and deep learning techniques. *Diagnostics (Basel)*. May 15, 2024;14(10):1017. [doi: [10.3390/diagnostics14101017](https://doi.org/10.3390/diagnostics14101017)] [Medline: [38786315](https://pubmed.ncbi.nlm.nih.gov/38786315/)]
45. Li Z, Liu X, Tang Z, et al. TrajVis: a visual clinical decision support system to translate artificial intelligence trajectory models in the precision management of chronic kidney disease. *J Am Med Inform Assoc*. Nov 1, 2024;31(11):2474-2485. [doi: [10.1093/jamia/ocae158](https://doi.org/10.1093/jamia/ocae158)] [Medline: [38916922](https://pubmed.ncbi.nlm.nih.gov/38916922/)]
46. Seyam M, Weikert T, Sauter A, Brehm A, Psychogios MN, Blackham KA. Utilization of artificial intelligence-based intracranial hemorrhage detection on emergent noncontrast CT images in clinical workflow. *Radiol Artif Intell*. Mar 2022;4(2):e210168. [doi: [10.1148/ryai.210168](https://doi.org/10.1148/ryai.210168)] [Medline: [35391777](https://pubmed.ncbi.nlm.nih.gov/35391777/)]
47. Oke F, Bolaji O, Umakor M. Building AI-ready infrastructure for U.S. healthcare: a product management perspective. *World J Adv Res Rev*. Aug 30, 2025;27(2):588-603. URL: <https://journalwjarr.com/ArchiveIssue-2025-Vol27-Issue2> [Accessed 2025-06-30] [doi: [10.30574/wjarr.2025.27.2.2892](https://doi.org/10.30574/wjarr.2025.27.2.2892)]
48. Batool A, Zowghi D, Bano M. AI governance: a systematic literature review. *arXiv*. Preprint posted online on Jul 25, 2023. [doi: [10.21203/rs.3.rs-4784792/v1](https://doi.org/10.21203/rs.3.rs-4784792/v1)]
49. Celi LA, Cellini J, Charpignon ML, et al. Sources of bias in artificial intelligence that perpetuate healthcare disparities—a global review. *PLOS Digit Health*. 2022;1(3):e0000022. [doi: [10.1371/journal.pdig.0000022](https://doi.org/10.1371/journal.pdig.0000022)] [Medline: [36812532](https://pubmed.ncbi.nlm.nih.gov/36812532/)]
50. Moldovan AMN, Vescan A, Grosan C. Healthcare bias in AI: a systematic literature review. Presented at: 20th International Conference on Evaluation of Novel Approaches to Software Engineering; Apr 4-6, 2025:835-842; Porto, Portugal. URL: <http://www.scitepress.org/DigitalLibrary/ProceedingLink.aspx?ID=1893> [Accessed 2025-12-03] [doi: [10.5220/0013480300003928](https://doi.org/10.5220/0013480300003928)]
51. Kim JY, Hasan A, Kueper J, et al. Establishing organizational AI governance in healthcare: a case study in Canada. *NPJ Digit Med*. Aug 15, 2025;8(1):522. [doi: [10.1038/s41746-025-01909-3](https://doi.org/10.1038/s41746-025-01909-3)] [Medline: [40817280](https://pubmed.ncbi.nlm.nih.gov/40817280/)]
52. Gebler R, Reinecke I, Sedlmayr M, Goldammer M. Enhancing clinical data infrastructure for AI research: comparative evaluation of data management architectures. *J Med Internet Res*. Aug 1, 2025;27:e74976. [doi: [10.2196/74976](https://doi.org/10.2196/74976)] [Medline: [40749197](https://pubmed.ncbi.nlm.nih.gov/40749197/)]
53. Amann J, Blasimme A, Vayena E, Frey D, Madai VI, Precise4Q consortium. Explainability for artificial intelligence in healthcare: a multidisciplinary perspective. *BMC Med Inform Decis Mak*. Nov 30, 2020;20(1):310. [doi: [10.1186/s12911-020-01332-6](https://doi.org/10.1186/s12911-020-01332-6)] [Medline: [33256715](https://pubmed.ncbi.nlm.nih.gov/33256715/)]
54. Nouis SC, Uren V, Jariwala S. Evaluating accountability, transparency, and bias in AI-assisted healthcare decision-making: a qualitative study of healthcare professionals' perspectives in the UK. *BMC Med Ethics*. Jul 8, 2025;26(1):89. [doi: [10.1186/s12910-025-01243-z](https://doi.org/10.1186/s12910-025-01243-z)] [Medline: [40629303](https://pubmed.ncbi.nlm.nih.gov/40629303/)]
55. Onyeabor US, Onwuasoigwe O, Okenwa WO, Schaaf T, Pinkwart N, Balzer F. Exploring user experiences of clinicians engaged with the digital healthcare interventions across the referral and university teaching hospitals in Nigeria: a qualitative study. *Front Digit Health*. 2025;7:1488880. [doi: [10.3389/fdgh.2025.1488880](https://doi.org/10.3389/fdgh.2025.1488880)] [Medline: [40510415](https://pubmed.ncbi.nlm.nih.gov/40510415/)]
56. Jo BW, Khan RMA, Lee YS. Hybrid blockchain and internet-of-things network for underground structure health monitoring. *Sensors (Basel)*. Dec 4, 2018;18(12):4268. [doi: [10.3390/s18124268](https://doi.org/10.3390/s18124268)] [Medline: [30518124](https://pubmed.ncbi.nlm.nih.gov/30518124/)]
57. Genereaux BW, Dennison DK, Ho K, et al. Background and application of the web standard for medical imaging. *J Digit Imaging*. Jun 2018;31(3):321-326. [doi: [10.1007/s10278-018-0073-z](https://doi.org/10.1007/s10278-018-0073-z)] [Medline: [29748852](https://pubmed.ncbi.nlm.nih.gov/29748852/)]
58. Islam U, Alatawi MN, Alqazzaz A, Alamro S, Shah B, Moreira F. A hybrid fog-edge computing architecture for real-time health monitoring in IoMT systems with optimized latency and threat resilience. *Sci Rep*. Jul 15, 2025;15(1):25655. [doi: [10.1038/s41598-025-09696-3](https://doi.org/10.1038/s41598-025-09696-3)] [Medline: [40665167](https://pubmed.ncbi.nlm.nih.gov/40665167/)]
59. Ayaz M, Pasha MF, Alzahrani MY, Budiarto R, Stiawan D. The Fast Health Interoperability Resources (FHIR) standard: systematic literature review of implementations, applications, challenges and opportunities. *JMIR Med Inform*. Jul 30, 2021;9(7):e21929. [doi: [10.2196/21929](https://doi.org/10.2196/21929)] [Medline: [34328424](https://pubmed.ncbi.nlm.nih.gov/34328424/)]
60. Wiggins WF, Magudia K, Schmidt TMS, et al. Imaging AI in practice: a demonstration of future workflow using integration standards. *Radiol Artif Intell*. Nov 2021;3(6):e210152. [doi: [10.1148/ryai.2021210152](https://doi.org/10.1148/ryai.2021210152)] [Medline: [34870224](https://pubmed.ncbi.nlm.nih.gov/34870224/)]

61. Ahmed MI, Spooner B, Isherwood J, Lane M, Orrock E, Dennison A. A systematic review of the barriers to the implementation of artificial intelligence in healthcare. *Cureus*. Oct 2023;15(10):e46454. [doi: [10.7759/cureus.46454](https://doi.org/10.7759/cureus.46454)] [Medline: [37927664](https://pubmed.ncbi.nlm.nih.gov/37927664/)]
62. Hassan M, Kushniruk A, Borycki E. Barriers to and facilitators of artificial intelligence adoption in health care: scoping review. *JMIR Hum Factors*. Aug 29, 2024;11(1):e48633. [doi: [10.2196/48633](https://doi.org/10.2196/48633)] [Medline: [39207831](https://pubmed.ncbi.nlm.nih.gov/39207831/)]
63. Yildirim N, Zlotnikov S, Sayar D, et al. Sketching AI concepts with capabilities and examples: AI innovation in the intensive care unit. Presented at: CHI '24; May 11-16, 2024; Honolulu, HI. URL: <https://dl.acm.org/doi/proceedings/10.1145/3613904> [Accessed 2024-05-11] [doi: [10.1145/3613904.3641896](https://doi.org/10.1145/3613904.3641896)]
64. Crain N, Qiu CY, Moy S, et al. Implementation science for the adductor canal block: a new and adaptable methodology process. *World J Orthop*. Nov 18, 2021;12(11):899-908. [doi: [10.5312/wjo.v12.i11.899](https://doi.org/10.5312/wjo.v12.i11.899)] [Medline: [34888150](https://pubmed.ncbi.nlm.nih.gov/34888150/)]
65. Dayan I, Roth HR, Zhong A, et al. Federated learning for predicting clinical outcomes in patients with COVID-19. *Nat Med*. Oct 2021;27(10):1735-1743. [doi: [10.1038/s41591-021-01506-3](https://doi.org/10.1038/s41591-021-01506-3)] [Medline: [34526699](https://pubmed.ncbi.nlm.nih.gov/34526699/)]
66. Fang HSA, Tan TH, Tan YFC, Tan CJM. Blockchain personal health records: systematic review. *J Med Internet Res*. Apr 13, 2021;23(4):e25094. [doi: [10.2196/25094](https://doi.org/10.2196/25094)] [Medline: [33847591](https://pubmed.ncbi.nlm.nih.gov/33847591/)]
67. Aboy M, Minssen T, Vayena E. Navigating the EU AI Act: implications for regulated digital medical products. *NPJ Digit Med*. Sep 6, 2024;7(1):237. [doi: [10.1038/s41746-024-01232-3](https://doi.org/10.1038/s41746-024-01232-3)] [Medline: [39242831](https://pubmed.ncbi.nlm.nih.gov/39242831/)]
68. El Arab RA, Al Moosa OA. Systematic review of cost effectiveness and budget impact of artificial intelligence in healthcare. *NPJ Digit Med*. Aug 26, 2025;8(1):548. [doi: [10.1038/s41746-025-01722-y](https://doi.org/10.1038/s41746-025-01722-y)] [Medline: [40858882](https://pubmed.ncbi.nlm.nih.gov/40858882/)]
69. Parikh RB, Helmchen LA. Paying for artificial intelligence in medicine. *NPJ Digit Med*. May 20, 2022;5(1):63. [doi: [10.1038/s41746-022-00609-6](https://doi.org/10.1038/s41746-022-00609-6)] [Medline: [35595986](https://pubmed.ncbi.nlm.nih.gov/35595986/)]

Abbreviations

AI: artificial intelligence

CASP: Critical Appraisal Skills Programme

CT: computed tomography

EHR: electronic health record

FHIR: Fast Healthcare Interoperability Resources

HIS: Hospital information system

ICH: intracranial hemorrhage

IoT: Internet of Things

LIS: laboratory information system

ML: machine learning

PACS: picture archiving and communication system

PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses

Edited by Javad Sarvestan; peer-reviewed by Ankit Gupta, Kola Adegoke; submitted 28.Jun.2025; final revised version received 25.Oct.2025; accepted 30.Oct.2025; published 17.Dec.2025

Please cite as:

Maimaitiaili M, Jiamaliding Y, Dai G, Xiao H, Kuerbanjiang W, Yi Y

Artificial Intelligence Platform Architecture for Hospital Systems: Systematic Review

J Med Internet Res 2025;27:e79788

URL: <https://www.jmir.org/2025/1/e79788>

doi: [10.2196/79788](https://doi.org/10.2196/79788)

© Musitapa Maimaitiaili, Yiershatijiang Jiamaliding, Guangle Dai, Hui Xiao, Warisijiang Kuerbanjiang, Yuexiong Yi. Originally published in the Journal of Medical Internet Research (<https://www.jmir.org>), 17.Dec.2025. 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, first published in the Journal of Medical Internet Research (ISSN 1438-8871), is properly cited. The complete bibliographic information, a link to the original publication on <https://www.jmir.org/>, as well as this copyright and license information must be included.