Journal of Information Systems Engineering and Management
2025, 10(62s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

Trustworthy AI Governance and Validation for Healthcare and Pharmaceutical Systems

Suresh Babu Narra
Solutions Architect – AI, Machine Learning & Generative AI
Cincinnati, Ohio, USA
narra.sureshbabu@gmail.com
https://orcid.org/0009-0002-5429-0202

ARTICLE INFO

Received: 03 Oct 2025
Revised: 18 Nov 2025
Accepted: 28 Nov 2025

ABSTRACT

Artificial intelligence (AI) can be integrated into healthcare and pharmaceutical systems to transform clinical decision-making and drug development, as well as patient care. Nevertheless, the use of AI in such stakes areas requires a sound system of governance and validation in order to maintain safety, reliability and adherence to regulatory measures. This essay offers a comprehensive analysis of the most prominent principles and issues related to reliable AI in healthcare and pharmaceuticals, with a specific focus on the governance model, regulatory adherence, data integrity and models validation procedures. The article discusses the ethical and social consequences of the use of AI such as equity, openness, and the necessity to have human supervision. It further talks about the changing regulatory environment, the need to have constant monitoring, and the multidisciplinary approach in making sure that the AI systems are dependable and safe throughout their life cycle. Good Machine Learning Practices (GMLP), explainable artificial intelligence (XAI) and sound validation strategies are suggested as key ingredients of a credible AI framework. Lastly, the paper has proposed the directions of research in future and best practices to reinforce AI accountability and validation measures in these areas.

Keywords: Trustworthy AI, Healthcare AI, Pharmaceutical AI, AI Governance, AI Validation, Data Integrity, Regulatory Compliance, Explainable AI, Model Performance Monitoring, Fairness and Bias Detection, Good Machine Learning Practices (GMLP).

Introduction

The introduction of artificial intelligence into the clinical and pharmaceutical pipeline is rapid, which means that there must be a solid framework to incorporate high-speed algorithmic innovation with accepted GxP requirements (Agafonov et al., 2024; Nair, 2025). More importantly, the data-driven and adaptive nature of these systems inherently threatens classical validation models that are targeted at the deterministic code, which should be replaced by the risk-based lifecycle management (Sonawane & Baviskar, 2025). With such a paradigm shift, total product lifecycle approaches must be deployed, involving continuous monitoring of models and performance assessment of these models to maintain the level of compliance (Huysentruyt et al., 2021). Moreover, this practice implies the employment of the model risk tiering and the regular re-evaluations of the issues in order to address the problem of the performance drift and the changing manufacturing environments (Suksaeree, 2025).

To overcome these systemic complexities, the stakeholders should fill the knowledge gap between AI engineering practices and the accepted regulatory standards (Higgins and Johner, 2022). It is necessary to create a unified regulatory environment to reduce the risk of a fragmented policy environment and make sure that validation strategies do not compromise data integrity (Niazi, 2025; Somara et al., 2025). To do that, multidimensional aspects of trustworthiness, such as safety, robustness, and explainability, need to be operationalized at institutions across the full lifecycle (Scaramuzza et al., 2025). This quality governance strategy conforms to the ICH Q8 11 model that applies design spaces and quality risk management to guarantee system reliability and manufacturability (Sonawane & Baviskar, 2025). Through combining real-time monitoring of performance, which is included in the current Quality Management Systems, organizations can establish two-way communication between safety surveillance activities and clinical outcomes (Overgaard et al., 2023).

1.Integration of Artificial Intelligence in Healthcare Systems

The following section will discuss how AI technologies can revolutionize the healthcare sector, including their use in patient care and healthcare administration, including predictive analytics, medical image analysis, and personalized treatment recommendations (Atılgan, 2023; Fahim et al., 2025). AI-based tools are able to make diagnosis accuracy better, simplify the process and make treatment plans more appropriate and precise when using large datasets of different modalities (Abatal and Korchi, 2023; Varnosfaderani and Forouzanfar, 2024). As an example, AI algorithms used in the field of medical imaging, such as the X-rays, MRIs, and CT scans, enable the earlier detection of diseases and increase the efficiency of the diagnostic process (Udegbe et al., 2024). These are powerful deep learning-based models that are highly accurate at detecting anomalies and can even perform better than human clinicians in speed and accuracy at performing some diagnostic tasks (Chaparala et al., 2025; Faiyazuddin et al., 2025).

2.History and Importance of AI in Healthcare

Implementation of machine learning in the two industries is transformative in relation to pharmacovigilance and drug development as evidenced by application of natural language processing to improve safety signals detection (Jiao et al., 2025). Nevertheless, the development is limited by the lack of uniformity in regulations, which cannot harmonize the assessment of non-deterministic models in different international jurisdictions (Mirakhori and Niazi, 2025). As a result, the stakeholders need to balance this volatility nature with the inflexible demands of the established GxP frameworks by implementing ready-made change control strategies that enable controlled model evolution (Niazi, 2023, 2025).

3.Artificial Intelligence-based Patient Risk Stratification

AI systems use a wide range of data, such as electronic health records, genetic data, and real-time physiological data, to create advanced predictive systems to identify patient risks and optimize outcomes (Syed, 2025). The former models combine past data with patient demographics and clinical data to forecast their future health risks and outcomes, allowing the healthcare provider to prioritize patients and provide tailored care (Charan et al., 2023). As an example, AI can be used to identify the risk of a patient developing chronic diseases like diabetes or cardiovascular diseases, thereby preventing them in time (Zeb et al., 2024).

4.Medical Imaging Analysis Artificial Intelligence

One of the most promising fields in AI revolution of healthcare is medical imaging analysis. AI algorithms, in particular, deep learning models, analyze the X-rays, MRIs, CT scans, and other medical images with high precision (Li et al., 2024). Such systems have the ability to identify anomalies like

Journal of Information Systems Engineering and Management
2025, 10(62s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/
Research Article

tumors, fractures, and organ anomalies among others, which in most cases have high diagnostic accuracy than human radiologists and save on diagnostic time (Chatterjee et al., 2024). This is made possible by their ability to analyze images continuously without exhaustion, variability of interpretations, or disregard of the subtle indicators of the disease, which improves the consistency and accuracy of diagnoses (Zeb et al., 2024). The skillfulness of these AI-based diagnostic tools is additionally improved because it can calculate the uncertainty levels of their interpretation offering clinicians with the most important information on how reliable the algorithmic output can be used to make more sound decisions (Abatal and Korchi, 2023).

5.Challenges and Risks of Untrustworthy AI

There is no shared standard on the performance of these models, which has severe weaknesses, especially when it comes to spreading false/irrelevant content in bedrock language models (Saenz et al., 2024). Moreover, the lack of strong validation strategies contributes to a drop in the performance of models in conditions of clinical environments or working processes that do not match their initial training data (Hellmeier et al., 2024). The scarcity of open-source code and ineffective alignment with given principles of medical imaging combined with this environmental inflexibility contributes to the diminishing reproducibility and regulatory consistency of these systems (Yousef et al., 2025). Furthermore, the lack of transparency in black box algorithms has been identified as a significant impediment to clinical interpretability that may cover algorithmic biases that will promote the detection of adverse events differently among different groups of patients (Ahire et al., 2024; Shwetali, 2025). Also, the current certification processes have been relying on retrospective studies, which does not give prospective performance stability, which disguises the possibility of gradual decline (González et al., 2024).

6.objective of the Paper

This paper identifies a full system governance of the medical equipment validation and maintenance of AI-driven systems and the importance of ensuring that these systems are resistant to adversarial vulnerabilities (data poisoning and model evasion) (Saboury et al., 2022). Moreover, the offered model suggests that algorithmic activities and domain logic are separated, and it is the medical societies or other independent institutions that control clinical performance metrics (Wellnhofer, 2022). The stakeholders can prevent the so-called responsibility vacuum that is commonly produced when the algorithm malfunctions in post-deployment monitoring by institutionalizing the systematic repair procedures instead of viewing them as an anomaly (Owens et al., 2025). The solution to these emergent issues is the introduction of live monitoring structures in organizations to identify the change in data distribution in real-time and anomalies in operations. These specifications need to be backed with a strict adversarial stress testing to reveal latent brittle natures and general competency deficiencies that often escape traditional validation (Gu et al., 2025).

7. AI Integration in Healthcare Applications

This segment will consider the application of AI in the main healthcare service applications that include patient risk stratification, medical image analysis, clinical notes summary, virtual patient assistants, and claims review. It is aimed to learn how AI enhances accuracy, speed, and outcomes in those spheres through processing large volumes of health data to increase diagnostic accuracy and allow personalized treatment plans (Chinta et al., 2024). AI analytical functions allow detecting faint trends and deviations in complex patient data, which remain unnoticed by human practitioners in many cases and enhances diagnostic accuracy and proactive treatment (Mennella et al., 2024). In medical imaging and diagnostics, AI algorithms are used to process complex visual information based on X-rays, MRIs, and

Journal of Information Systems Engineering and Management
2025, 10(62s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

CT scans to identify abnormalities and help to detect the disease in its early stages, such as diabetic retinopathy and some forms of cancer (Loku and Malsia, 2024; Ram, 2024).

8.Impact on Healthcare Efficiency and Patient Care

Artificial intelligence is also bringing in a revolutionary period in the field of healthcare and it is having a significant impact on the diagnostics field, the paradigm of personalized treatment, and the work efficiency (Faiyazuddin et al., 2025). In particular, AI supports the decision-making of clinicians as they offer powerful diagnostic, disease anticipation, and treatment customization tools (Kshetri et al., 2023). Deep learning and machine learning solutions enable AI to read complex patient data in detail, detect the slightest trends, and predict disease progression, which may promote proactive and tailored treatment of patients (Faiyazuddin et al., 2025; Kshetri et al., 2023). Such an ability is the opportunity to identify diseases in time and choose more specific treatment options, resulting in better patient outcomes (Hudda et al., 2024).

9.Structure of the Paper

The subsequent parts outline a methodology foundation of AI assurance which begins with the development of an overall assessment metrics and the incorporation of algorithmic auditing to determine limitations in the model (Byfield et al., 2023). Next, the paper touches on the incorporation of individual benchmarking and clinical audits in order to make sure that diagnostic applications have high-grade reliability not only after approval by the applicable regulatory bodies but also after the baseline (Ross et al., 2024). Also, the analysis examines the implementation of continuous monitoring systems aimed at detecting model drifts and other operational irregularities to maintain the health of the system in the long term (Ahmed, 2025; Guan et al., 2025). Last but not least, the document makes suggestions regarding the formation of specific quality control divisions in medical organizations to manage these ongoing quality control efforts and provide the required updates in the algorithms (Ueda et al., 2023). Such efforts ought to be further supported by the use of local validation measures, as well as the fine-tuning of models to consider local variations in clinical information (Hellmeier et al., 2024). Additionally, these systems should have a strict adherence to the set data privacy laws so that the information about patients could be safe throughout the iterative learning process (Jha et al., 2023).

10.Trustworthy AI Conceptual Framework

This design revolves around the paradigm of the continuous feedback loop involving real-time clinical evidence and expert consensus that ensures that it is aligned with the changing medical standards. Such an adaptive strategy requires the introduction of external and internal audits to conduct a systematic assessment of performance so that the stakeholders can coordinate model adjustments with the centralized quality assurance procedures (Bragazzi and Garbarino, 2024; Gonzalo et al., 2021). Organizations can create verifiable standards of accuracy, sensitivity, and specificity by integrating regular audits and reviews of diagnostic outcomes by other professionals, which reflect the set clinical quality benchmarks (Li et al., 2025; Mahmood et al., 2024). These frameworks should also be able to apply post-market surveillance systems to respond to performance changes in clinical practices and clinical situations to ensure that diagnostic tools continue to be safe and equitable as populations of patients and sources of data change (Abdelwanis et al., 2025; Lundström and Lindvall, 2022). Moreover, the integration of future monitoring initiatives, including monitoring metrics including accuracy, calibration, and fairness will allow clinical teams to identify the performance shortfalls prior to them affecting patient safety (Li et al., 2025).

10.1 .Defining Trustworthy AI

In this regard, trustful AI is characterized by a multi-dimensional combination of technical robustness, ethical correspondence, and transparency in a way that systems behave effectively in the context of

Journal of Information Systems Engineering and Management
2025, 10(62s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/
Research Article

sophisticated healthcare systems (Isaacks and Borkowski, 2024). It involves a holistic model where the structures of governance should be multidisciplinary, flexible to new technologies and changing societal expectations, and where the systems used must be safe, just, and transparent at all stages of their lives (Baradwaj et al., 2024). This definition will require models to be flexible to new clinical practices and data distributions, so that they can continuously improve their performance without forgetting the underlying knowledge (Gonzalez-Gonzalo et al., 2021). Moreover, the realization of this trust demands a cybernetic process of model improvement that brings practical clinical experience and new medical discoveries on board to keep the system both pertinent and precise to changing patient demographics (Nasarian et al., 2023). At its core, this requirement includes the introduction of open-source knowledge bases and public auditability to avoid silent failures, which occurred in the case of collaborative systems such as The Blue-Scrubs where it is possible to continuously verify and reintegrate validated knowledge with expert help (Fernández et al., 2025).

Table 1: Trustworthy AI Frameworks in Healthcare

AI Framework Component	Description	Examples/Implementation
Technical Robustness	Ensures consistent performance of AI models in diverse clinical settings.	AI models for diagnostic tools that adapt to various conditions.
Ethical Correspondence	Ethical AI practices that promote fairness, transparency, and patient safety.	Incorporating bias detection algorithms in clinical AI.
Explainability	Methods ensuring AI decisions are interpretable and understandable by healthcare providers.	Use of SHAP and LIME methods in model interpretation.
Systematic Accountability	Ensures human oversight, making clinicians accountable for final decisions.	Implementation of "human-in-the-loop" in diagnostic AI systems.
Continuous Stewardship	Active monitoring and adjustment of models throughout their life cycle to ensure long-term reliability.	Regular audits and updates to AI systems post-deployment.

The table above represents the most important pillars of a Trustworthy AI Framework in healthcare. These pillars will make AI Systems reliable, ethical, transparent, and constantly updated in accordance with clinical requirements and regulatory standards.

Journal of Information Systems Engineering and Management
2025, 10(62s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

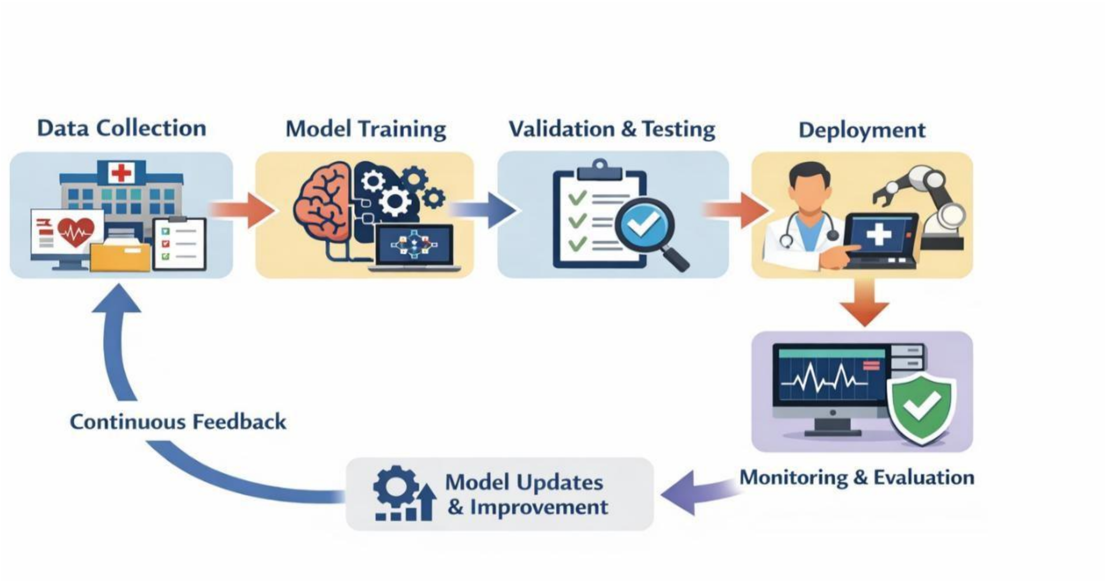

Figure 1: Trustworthy AI in Healthcare Workflow

As shown in Figure 1, a trustworthy AI workflow in healthcare is organized as a continuous cycle of data collection, model training, validation, deployment, monitoring, and feedback driven improvement.

10.2 Major Principles of Reliable AI

These principles can be operationalized on four pillars, namely technical robustness, explainability of its algorithms, systematic accountability, and continuous steward-ship (Baradwaj et al., 2024; Olumuyiwa et al., 2024). The technical robustness guarantees the consistency of model performance in different clinical scenarios, whereas explainability enables the interpretation of the complex diagnostic results human practitioners (Göktaş and Grzybowski, 2025; Labkoff et al., 2024). Systematic responsibility creates set standards regarding human control to make sure that clinicians are the ultimate decision makers regarding critical decision making (Adibi et al., 2025). Lastly, active governance models are essential in continuous stewardship to put automated outputs in line with the overall ethical principles to reduce risks and make them reliable in the long term (Solanki et al., 2022). All these pillars help to create a culture of reasonable dependency, when the systems work safely only under the conditions of the internal uncertainty and vulnerability that can be supported by the evidence of rigorous clinical validation and the transparent governance mechanisms (Lopes et al., 2025; Nasarian et al., 2024).

10.3 AI Healthcare Ethical Reflections

The incorporation of the ethical imperatives implies the need to follow the fundamental principles such as respect of human autonomy, avoidance of harm, fairness, and transparency. These codes of ethics require the AI systems to serve as a helping guide that upholds the privacy and autonomy of patients and reduces the possibility of bias in the diagnostic information (Alelyani, 2025; Moreno-Sánchez et al., 2025). Additionally, such frameworks should be characterized by the absence of the black-box problem and the focus on the transparency of algorithms and the responsibility of the latter, to make sure that clinicians are able to explain the automated insights (Weiner et al., 2024). To solve these dilemmas, an effective ethical system should be created, which balances clinical beneficence and non-maleficence, as well as makes patient care fair and free of systematic algorithmic discrimination (Pham, 2025). The inclusion of the paradigm of the clinician-in-the-loop is a necessary protective measure in this case because it helps to empower doctors to challenge, interpret, and monitor the outputs of the automated

Journal of Information Systems Engineering and Management
2025, 10(62s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/
Research Article

systems and make sure that they do not diverge from the objective clinical knowledge (Almotairi et al., 2022).

11. Artificial Intelligence Governance in Medical and Pharmaceutical Services

To institute these frameworks, the establishment should be in line with international regulatory standards like the risk-based classification of the EU AI Act, and the WHO standards, to guarantee that safety and compliance would be systematic across health jurisdictions globally (Bouderhem, 2024). These regulatory frameworks are further supported by the recent policy of the U.S. Food and Drug Administration that states the need to implement life-cycle management and clinical validation studies to guarantee the reliability of diagnoses (Oyeniyi, 2024; Pakhale, 2025). Outside these regulatory requirements, organizations need to adopt the so-called Regulatory Genome an agile regulatory tool that reflects the alignment between the local compliance requirements and the dynamic technological changes to fill significant gaps in the existing governance (Göktaş and Grzybowski, 2025). Such institutional governance approaches should be associated with the creation of standardized ethical auditing guidelines that go beyond compliance with regulations to tackle the socio-technical complexities of clinical settings (Mökander & Floridi, 2022).

12. Survey of the Current AI Governance Models

New models also tend to be organized around six fundamental ethical concepts such as safeguarding human autonomy and advancing inclusivity, which can be used to ensure that technical implementation is based on expanded public interest imperatives (Kim et al., 2024). These frameworks also include "Healthcare AI Datasheets" in order to record demographic information and the possible sampling bias, it creates a clear record that may help to reduce the historical disparity (Chen et al., 2025). Moreover, these documentation activities should be complemented with institutional guidelines regarding the regular bias correction, according to which periodic assessments of the model performance against the stratified subsets should be conducted and the promotion of equitable clinical results (Akter et al., 2023). It is necessary to document AI interventions; in addition to it, formal oversight committees must be established that are actively involved in the monitoring, evaluation, and the refinement of AI interventions via the continuous feedback mechanism (Carini and Seyhan, 2024).

Table 2: Governance Models for Healthcare AI

Governance Model	Key Features	Examples/Applications
Centralized Oversight	A unified committee overseeing AI lifecycle management, from development to deployment.	Multidisciplinary team including clinicians, ethicists, and data scientists.
Decentralized Oversight	Distributed responsibility among various stakeholders for continuous monitoring and feedback.	Localized committees in hospitals ensuring AI alignment with clinical standards.
Regulatory Compliance	Adherence to regulatory standards, including real-time performance evaluation and transparency.	Implementation of GDPR, HIPAA in healthcare AI applications.

Governance Model	Key Features	Examples/Applications
Ethical Audits	Regular ethical evaluations to ensure fairness, privacy, and equity in AI healthcare applications.	Bias audits for AI algorithms used in healthcare diagnostics.
External Data Safety Monitoring	Involvement of third-party organizations to review AI performance in clinical settings.	Use of independent review boards for AI medical device validation.

In table 2, the various AI Governance Models that can be applied to healthcare organizations are demarcated. The models provide ethical conduct, regulation adherence, and strong oversight in the lifecycle of the AI system.

a. Regulatory Environment and Regulatory Compliance

Trying to navigate this landscape, companies have to combine the classical data governance features, including data lineage and quality assurance, with the unique AI-specific control mechanisms (Tsai, 2025). Such governance approaches must support distributed oversight models to effectively oversee various applications in addition to dealing with institutional issues, such as data privacy and fairness in algorithms (Hoffman-Peterson, 2025). In order to make these measures work, organizations should not only go beyond the autonomy of programs development but also use centralized, standardized governance mechanisms to assess all incoming AI products with the help of the unified, multidisciplinary committee (Kim et al., 2025).

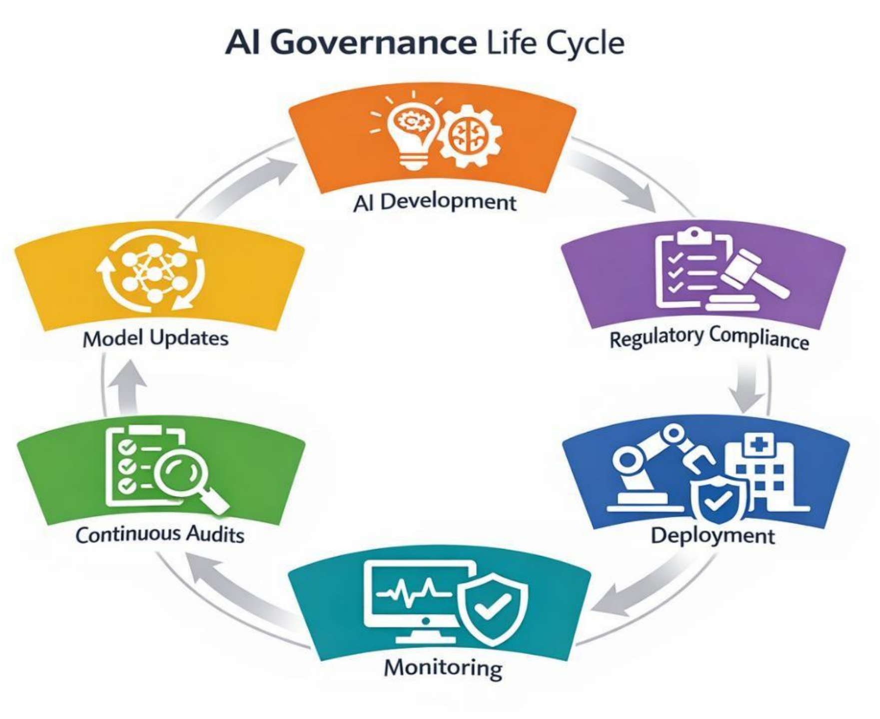

Figure 2: AI Governance Life Cycle

Figure 2 presents the AI Governance Life Cycle that illustrates the continuous AI system management process within the healthcare environment. It focuses on such critical steps like AI development, regulatory compliance, deployment, monitoring, audits and updating of models. The figure shows that

Journal of Information Systems Engineering and Management
2025, 10(62s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/
Research Article

governance is a way to make sure that AI models continuously undergo monitoring, verification, and modification to the standards of safety, ethics, and regulations.

b. AI Governance Organizational Structures

The latter are generally required to have a multidisciplinary governance committee including clinicians, ethicists data scientists and community advocates to achieve value-sensitive design and reduce systemic blind spots (Bodnari & Travis, 2025). The involvement of these stakeholders at the full algorithm life cycle will be essential in ensuring fairness and cultivating a sense of trust among the patient groups (Chin et al., 2023). Moreover, companies must also focus on continuous learning by such professionals to develop the profound awareness of ethical aspects, privacy concerns, and the need to provide human-centric care (Anyanwu et al., 2024). To develop this capacity, the equity literacy and standardized best practices should be institutionalized to make sure that automated tools actively reverse instead of reproduce structural inequities in the healthcare ecosystem (Nong, 2024).

c. Stakeholder Engagement in AI Governance

The success of stakeholder engagement is the development of official communication and collaboration between leaders of the health system and those who perform operations to ensure that the implementation of AI is aligned with the clinical goals of the organization and the goals of strategic management (Byberg & Crimi, 2025). The use of this inclusive method should also include patient advocates and community representatives in order to make sure that the deployment strategies should incorporate the lived experiences of the people who are directly concerned with the automated care interventions (Chin et al., 2023). In addition, the effective implementation of these technologies also depends on the creation of collaborative, patient-centered control frameworks, including special consumer consortia, capable of offering the necessary balances to the effectiveness of algorithms and their ethical integrity (Rozenblit et al., 2025). The said organizational changes require shifting away from the classic siloed departmental operations into the matrix forms, which help to create the specialized and cross-functional teams that will be able to navigate the multifaceted interplay of technical, clinical, and ethical supervision (Tiwari, 2025). These cross-functional teams help close the gap between technical and practical clinical use by forming a single overall oversight structure by bringing together knowledge experts, policy-implementing decision-makers, and end-users (Dixit et al., 2021).

13. Trustworthy AI Systems validation Methodologies

The strong validation may entail the replacement of the fixed, pre-deployment performance measures with the dynamic and continuous monitoring systems that can estimate the stability of algorithms in the real-life clinical setting (Peluso, 2025). This requires implementing clinician-informed assessment guidelines and standardized reporting resources - e.g., METRICS checklist - to guarantee that model performance is stable and clinically applicable to different environments with changing environments (Khalifa & Hussein, 2025). To ensure that these kinds of validation processes eliminate the proliferation of vendor-created tools that might not have an excellent performance in different groups of patients, equity audits and population-representative testing must be introduced specifically (Nong et al., 2024; Rashid, 2024). Moreover, the adoption of ecological validity testing is a way of ensuring that the algorithms are rigorously tested in an uncontrolled clinical setting, so that technical design is, again, reconciled with the actual, multifaceted reality of bedside decision-making (Choudhury, 2022).

a. Data Quality and Integrity Assessment

The assessment of data quality and integrity should be provided (including the assessment of the data model, see Table 1 below). Good data quality management implies the creation of strict data provenance management standards, makes the training datasets representative, high-fidelity, and free of any

Journal of Information Systems Engineering and Management
2025, 10(62s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/ **Research Article**

systematic errors that may threaten the reliability of the models (Tilala et al., 2024). In addition to these prerequisites, institutions are required to introduce computational verification procedures that involve scrutinizing the outcomes of algorithmic generalizability and accuracy by means of real-world clinical performance trials (Quazi et al., 2024). Such validation strategies must also consist of the so-called shadow mode when new tools are used simultaneously with the current clinical processes to ensure operational safety and build trust in the users before large-scale implementation (Tiwari, 2025).

b. This is the explainability and interpretability of models, also known as model XAI

The primary focus on transparency makes sure that healthcare professionals will be able to understand the logic behind the automated recommendations, without which it will be impossible to make informed clinical decisions (Alelyani, 2024). In order to do that, developers need to incorporate explainable AI structures that offer practical information about the underlying characteristics that underlie certain clinical predictions (Arafat et al., 2025). Moreover, as clinicians and various stakeholders are asked about their feedback, it is possible to incorporate human knowledge into the learning process and this further fills the gap between the output of the algorithm and clinical intuition (Ramírez. et al., 2024). Moreover, they must focus on extensive data protection through the use of federated learning and pseudonymization to ensure patient privacy but at the same time preserve the accessibility needed to train the model (Khalifa and Hussein, 2025).

c. Testing of Robustness and Security

This stage requires the establishment of strict encryption processes, access restrictions, and constant monitoring of any possible cyber threats to ensure that sensitive patient records are not exposed to the unauthorized parties (Guerra, 2024). Additionally, the system should be able to test its adversarial robustness against data poisoning and input manipulation to ensure that the integrity of clinical decision support is not compromised by the malicious input (Mallardi et al., 2025). Besides these security measures, the institutions ought to embrace standard auditing processes in order to ensure that models follow the set legal and regulatory frameworks that set out the patient safety (Nasarian et al., 2023).

d. Fairness and Bias Detection

Systematic inequalities cannot be resolved without incorporating standard bias identification devices that will detect and reduce demographic differences in training data (Prabhu, 2025). The adaptive weighting schemes used in these modules must guarantee that the prediction of algorithms are just in the face of various patient groups, i.e., the historical underrepresentation of biomedical knowledge bases should be fixed (Zaki et al., 2025). Such audits should be supplemented with the implementation of privacy-related approaches such as the use of differential privacy, which allows refining equitable models without violating the privacy of individual patients (Shafik et al., 2024).

e. Performance Evaluation and monitoring Performance

Such a step requires a gradual validation process starting with synthetic data validation, which is followed by retrospective studies and then proceeded to prospective clinical trials to determine long-term model effectiveness (Gambetti et al., 2025). Simultaneously with these experiments, organizations need to develop operational feedback mechanisms, which absorb the end-user experience to inform the refinement of their models and correct possible problems, including automation bias or alert fatigue (Abbas et al., 2025; Artsi et al., 2025).

14. The Strategies of Trustworthy AI Implementation in Healthcare

To implement it successfully, it is necessary to transition to centralized institutional structures which make human-AI interaction, reporting, and continuous validation in all clinical departments

Journal of Information Systems Engineering and Management
2025, 10(62s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/
Research Article

standardized (Hasanzadeh et al., 2025). These frames must focus on the interoperability with the current electronic health records, so that AI solutions can be incorporated into intricate clinical infrastructures (Quazi et al., 2024). Moreover, the operationalization of such systems must follow the international consensus guidelines which provide risk-proportional assessment in the full lifecycle, including the design stage and post-market monitoring (Lekadir et al., 2025). Besides these technical infrastructures, hospitals may also create special quality control departments that will check the development of performance drift and provide clinicians with continuous education on the inherent limitations and biases of deployed models (Ueda et al., 2023).

a. Ai Systems Life Cycle Management

This phase entails a detailed requirements table which puts together technical technique and dependencies in the development, implementation, and administration phases to sustain operational coherence (de-Manuel-Vicente et al., 2024). These frameworks enable the implementation of thirty various best practices that encompass the socioethical, legal, and clinical aspects of artificial intelligence and provide healthcare institutions with the opportunity to manage risk pragmatically over the period of functional duration of the tool (Lekadir et al., 2025). Additionally, a systematic method of identifying performance degradation can be applied in the form of structured protocols of monitoring and surveillance, which will guarantee the compliance of the state of regulatory procedures and clinical safety via the iterative audit reporting (Adibi et al., 2025).

b. Auditing and Accountability Mechanisms

To develop these mechanisms, it is necessary to institutionalize the cross-functional oversight committees that would conduct periodic evaluations of algorithmic performance by criteria of efficacy and safety. The committees must also include the external data safety monitoring boards as evaluators of the instances of treatment failure to encourage the culture of transparency and methodological improvement (Nair et al., 2024). Moreover, systematic approach to procurement will make sure that the technologies provided by the vendors will go through a stringent functionality test prior to being incorporated into clinical practices (Papadopoulou et al., 2025).

c. Artificial Intelligence Adoption Training and Education

The institutional programs should create an extensive set of competencies, which would cover the basic AI literacy of all employees as well as advanced training modules of specialists (Bignami et al., 2025). Those should promote interdisciplinary cooperation, which implies building communities of practice in which healthcare professionals of all tiers will be able to exchange knowledge and insights about the particular role-related AI applications (Byberg and Crimi, 2025).

d. Best Practices and Case Studies

The most important aspect of companies should be the documentation of so-called living frameworks, i.e., FAIR-AI, which require annual reviews to ensure that they align to the dynamically changing regulatory landscape and security threats (Wells et al., 2025). Such paradigms must be realized in terms of multidisciplinary monitoring where the institutions apply version-pinning and codified enablers to align the technical development with the dynamic regulatory needs (Gallifant et al., 2025).

15. Challenges and Future directions

Regardless of these developments, the shift toward scalable AI governance is still not facilitated due to the overwhelming interoperability shortages and the absence of universal performance assessment metrics (Saenz et al., 2024). In addition, all these obstacles require an innovation infrastructure that promotes a two-way channel of communication so that real-time safety monitoring and outcome data is closely linked with risk management activities (Overgaard et al., 2023). In addition, a system of

Journal of Information Systems Engineering and Management
2025, 10(62s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/
Research Article

licensed AI auditors and specialized positions, including a Chief AI Ethics Officer, can offer the required accountability to overcome such complicated sociotechnical issues (Hadley, 2022).

a. Technical Issues of AI Verification

The heterogeneity of clinical contexts and the effects of the variability in data quality, local processes, and disease incidence make the validation of models across heterogeneous clinical environments complicated and may require intensive re-testing at all stages of system scale-up (Tjondronegoro et al., 2022). To overcome this, companies need to use specialized validation tools to reduce the distance between theoretical risk management and operational practice (Chakraborty and Karhade, 2024). The initiatives of this nature are supported by the integration of sound ethical frameworks (with the focus on the autonomy of patients and their fairness) (As'ad and Faran, 2025), as well as tailored training programs that prepare various stakeholders, including purchase experts and clinicians, with the skills needed to overcome emerging algorithmic risks (Guerra, 2024; Xia et al., 2023).

b. Ethical and Societal Implications

In addition to the acceptable use of patient data, algorithmic equity and possible effects on healthcare equity, these considerations should be made without the risk of contributing to the unconscious challenge of systemic biases (Božić, 2023). In order to address such risks, organizations should institutionalize the algorithmic impact assessment that measures inequitable outcomes among patient demographics so that health equity could take the status of a background element of AI implementation (Du, 2025).

c. Policy and Regulatory Development

FDA and other global health organizations are moving towards adaptive models to govern the dynamic nature of machine learning systems (Fahim et al., 2025; Mullankandy, 2024). One of such efforts is the investigation of dynamic certification models that can be reconsidered and apply to medical devices that have undergone significant changes in their diagnostic protocols or intended usage (Farah et al., 2024). This development acknowledges that financial and administrative costs of re-validation tend to deter required changes to the systems, which may constrain the clinical utility of adaptive models (Raz et al., 2025).

Future Trends of Reliable AI

Explainable AI is also getting more prioritization in the field to boost the level of clinician confidence in unsupervised systems to overcome the existing lack of transparency that prevents widespread clinical adoption (Rahamtalla et al., 2025). Moreover, the application of human-in-the-loop control and repeated local validation procedures will make sure that models are resistant to data changes and to a clinically relevant degree of changes in healthcare settings (Behar et al., 2023). Practitioners should also support transparency to promote this efficacy through the delivery of datasheets that specify the nature of training data, which will prevent the existence of latent biases to enhance generalizability between different clinical settings (Dehghani et al., 2024).

Conclusion

The future sustainability of healthcare AI will be based on filling the gap between technical innovation and strong ethical regulation, where continued cooperation between clinicians, developers, and policymakers is necessary to provide equal patient outcomes (Weiner et al., 2025). In addition, privacy-saving solutions like federated learning and differential privacy must be embraced to ensure that patients retain the trust in the study as well as using sensitive information to constantly enhance the model.

Finally, to maintain such systems, there is a need to have common accountability systems that harmonize organizational interests and ethical motivation to reduce bias and protect the wellbeing of patients during the lifecycle of the implemented models (Gill et al., 2023; Sagona et al., 2025). Active decommissioning plans should therefore be made institutional to retire ineffective or outdated tools to avoid the accrual of non-sustainable technical debt which will undermine the performance of clinical processes and patient safety (Alsentzer et al., 2025). This paradigm shift requires the stakeholders to stop seeing the implementation of diagnostic and prognostic AI as a single- time validation but as a process of constant monitoring to control the unavoidable degradation of performance (Guan et al., 2025; Hellmeier et al., 2024).

Summary of Key Findings

Clinical-longitudinal validation of the system and rehearsal system monitoring contribute to the reliability of AI in dynamic and real-life clinical settings (Hellmeier et al., 2024)(Abbas et al., 2025). Clinicians must be able to interpret and validate model predictions with the use of explainable AI methods such as SHAP and LIME and thus reduce the gap between black-box predictions and useful medical information (Ahadian et al., 2025; Tuấn, 2024). Moreover, an institutionally based monitoring team to support systematic retraining guidelines is important to ensure that models keep up with the changing patient populations and treatment norms without jeopardizing the safety of clinical care (Agius et al., 2025; Balagopalan et al., 2024). Lastly, standardized datasheet documentation practices should be required so that they could ethically reuse data and have a clear audit trail of model provenance (Siddik & Pandit, 2025). Finally, the establishment of multidisciplinary governance entities that can be modified to match societal standards is crucial to the fact that AI implementations are oriented towards global legal and ethical requirements (Baradwaj et al., 2024). The institutionalization of these feedback loops will enable the healthcare systems to shift the focus on the reactive troubleshooting model to the proactive maintenance model, where the AI performance decline is viewed as a foreseeable, manageable variable, as opposed to an extraordinary failure (Owens et al., 2025).

Research and Practice Implications

The evolution of formalized accountability frameworks should be made a core priority of future studies in order to eliminate the current lack of accountability that has resulted in a lack of priority in maintaining the deployed AI tools by healthcare institutions (Owens et al., 2025). There is a need to set up special internal teams related to quality control of algorithms to track the continuity of performance and reduce the potential threat of dataset change (Owens et al., 2025; Ueda et al., 2023). Moreover, the main focus of current research should be put on the creation of the methodology that can differentiate the instances of the performance deterioration due to the dataset drift and due to the unintended effects of the interventions suggested by clinical models (Davis et al., 2022). In addition, the model must also focus on the creation of uniform performance monitoring standards and maintenance policies to make sure that the model ownership and license limit do not hinder the prompt resolution of the clinical accuracy deficit (Brady et al., 2024).

Future Research Avenues

The focus of the investigations must be placed on the standardization of model maintenance policies and the creation of systematic models of AI-enabled tool management at the level of the health system

(Davis et al., 2022). These frameworks should consider the logistical complexity of model drift detection, which implies the use of infrastructure that allows differentiating between expected changes in the environment and performance reduction that needs to be proactively retrained (Davis et al., 2022; Mashar et al., 2023). It is a viable way to overcome the shortcomings of the more traditional, fixed, and rigid regulatory frameworks, which tend to miss the effects of performance deterioration due to changing patient case profiles and clinical practice trends (Feng et al., 2022). Further, longitudinal studies that track the trends of failure and institutional reaction should be prioritized to the creation of collaborative, community practices of maintenance under the absence of central regulatory direction (Owens et al., 2025). Also, in further studies, temporal degradation testing must be incorporated to assess the stability of the model and the patterns of error distributions as the main signs of longevity of the system (Jarquin et al., 2022).

References:

[1] Abbas, Q., Jeong, W.-M., & Lee, S. W. (2025). Explainable AI in Clinical Decision Support Systems: A Meta-Analysis of Methods, Applications, and Usability Challenges [Review of *Explainable AI in Clinical Decision Support Systems: A Meta-Analysis of Methods, Applications, and Usability Challenges*]. *Healthcare*, *13*(17), 2154. Multidisciplinary Digital Publishing Institute. https://doi.org/10.3390/healthcare13172154

[2] Abdelwanis, M., Simsekler, M. C. E., Gabor, A. F., Sleptchenko, A., & Omar, M. (2025). Artificial intelligence adoption challenges from healthcare providers' perspectives: A comprehensive review and future directions. *Safety Science*, *193*, 107028. https://doi.org/10.1016/j.ssci.2025.107028

[3] Adibi, A., Xu, C., Ji, Z., Kaur, J., Chen, W., Healey, E., Nuwagira, B., Ye, W., Woollard, G., Xu, M., Cui, H. T., Xi, J., Chang, T., Bikia, V., Zhang, N., Noori, A., Yuan, X., Hossain, Md. B., Frank, H. A., … Jabbour, S. (2025). Recent Advances, Applications and Open Challenges in Machine Learning for Health: Reflections from Research Roundtables at ML4H 2024 Symposium. *arXiv (Cornell University)*. https://doi.org/10.48550/arxiv.2502.06693

[4] Agafonov, O., Babić, A., Sousa, S., & Alagaratnam, S. (2024). Editorial: Trustworthy AI for healthcare. *Frontiers in Digital Health*, *6*. https://doi.org/10.3389/fdgth.2024.1427233

[5] Agius, S., Cassar, V., Magri, C., Khan, W., & Topham, L. (2025). Position paper: advocating for a structured methodology in developing data-driven predictive models for healthcare – evidence from a large-scale national study. *Health and Technology*. https://doi.org/10.1007/s12553-025-01010-5

[6] Ahadian, P., Xu, W., Liu, D., & Guan, Q. (2025). Ethics of trustworthy AI in healthcare: Challenges, principles, and practical pathways. *Neurocomputing*, *661*, 131942. https://doi.org/10.1016/j.neucom.2025.131942

[7] Ahire, Y., Patil, J. H., Chordiya, H. N., Deore, R. A., & Bairagi, V. (2024). Advanced Applications of Artificial Intelligence in Pharmacovigilance: Current Trends and Future Perspectives. *Journal of Pharmaceutical Research*, *23*(1), 23. https://doi.org/10.18579/jopcr/v23.1.24

[8] Ahmed, I. (2025). NAVIGATING ETHICS AND RISK IN ARTIFICIAL INTELLIGENCE APPLICATIONS WITHIN INFORMATION TECHNOLOGY: A SYSTEMATIC REVIEW [Review of *NAVIGATING ETHICS AND RISK IN ARTIFICIAL INTELLIGENCE APPLICATIONS WITHIN INFORMATION TECHNOLOGY: A SYSTEMATIC REVIEW*]. *American Journal of Advanced Technology and Engineering Solutions*, *1*(1), 579. https://doi.org/10.63125/590d7098

[9] Akter, M. S., Sultana, N., Khan, M. A., & Mohiuddin, M. M. (2023). BUSINESS INTELLIGENCE-DRIVEN HEALTHCARE: INTEGRATING BIG DATA AND MACHINE LEARNING FOR STRATEGIC

COST REDUCTION AND QUALITY CARE DELIVERY. *American Journal of Interdisciplinary Studies*, *4*(2), 1. https://doi.org/10.63125/crv1xp27

[10] Alelyani, T. (2024). Trustworthiness of Artificial Intelligence – Powered Autonomous Healthcare Systems: An Assessment Framework Informed byExperts Insights. *Research Square (Research Square)*. https://doi.org/10.21203/rs.3.rs-3741717/v1

[11] Alelyani, T. (2025). A validated framework for responsible AI in healthcare autonomous systems. *Scientific Reports*, *15*(1), 44432. https://doi.org/10.1038/s41598-025-25266-z

[12] Almotairi, A. H., Alswegi, F. H., Almutairi, M., Almotiry, F. S. F., Rakhimi, M. N. H. A., Aljamily, A. H. M., Aldosari, N. M. M., & Alaoufi, R. Y. (2022). A REVIEW IN THE APPLICATION OF ARTIFICIAL INTELLIGENCE IN MANY MEDICAL SECTORS [Review of *A REVIEW IN THE APPLICATION OF ARTIFICIAL INTELLIGENCE IN MANY MEDICAL SECTORS*]. *Journal of Population Therapeutics and Clinical Pharmacology*. Codon Publications. https://doi.org/10.53555//jptcp.v29i04.5772

[13] Alsentzer, E., Charpignon, M., Chen, B. B., Dsouza, N., Fries, J., Jiang, Y., Kashyap, A., Kim, C., Lee, S., Mandyam, A., Mbilinyi, A. C., Mehandru, N., Nagesh, N., Nuwagira, B., Pierson, E., Pillai, A. S., Sano, A., Syeda-Mahmood, T., Yadav, S., … Thakoor, K. A. (2025). Reflections from Research Roundtables at the Conference on Health, Inference, and Learning (CHIL) 2025. *arXiv (Cornell University)*. https://doi.org/10.48550/arxiv.2510.15217

[14] Anyanwu, E. C., Okongwu, C. C., Olorunsogo, T. O., Ayo-Farai, O., Osasona, F., & Daraojimba, O. D. (2024). ARTIFICIAL INTELLIGENCE IN HEALTHCARE: A REVIEW OF ETHICAL DILEMMAS AND PRACTICAL APPLICATIONS [Review of *ARTIFICIAL INTELLIGENCE IN HEALTHCARE: A REVIEW OF ETHICAL DILEMMAS AND PRACTICAL APPLICATIONS*]. *International Medical Science Research Journal*, *4*(2), 126. Fair East Publishers. https://doi.org/10.51594/imsrj.v4i2.755

[15] Arafat, J., Tasmin, F., Poudel, S., Tareq, A. H., & Haider, I. (2025). Beyond Static Knowledge Messengers: Towards Adaptive, Fair, and Scalable Federated Learning for Medical AI. *arXiv (Cornell University)*. https://doi.org/10.48550/arxiv.2510.06259

[16] Abatal, A., & Korchi, A. (2023). Transforming Healthcare Systems with Artificial Intelligence: Revolutionizing Efficiency, Quality, and Patient Care. *Research Square (Research Square)*. https://doi.org/10.21203/rs.3.rs-3175341/v1

[17] Artsi, Y., Sorin, V., Glicksberg, B. S., Korfiatis, P., Nadkarni, G. N., & Klang, E. (2025). Large language models in real-world clinical workflows: a systematic review of applications and implementation [Review of *Large language models in real-world clinical workflows: a systematic review of applications and implementation*]. *Frontiers in Digital Health*, *7*. Frontiers Media. https://doi.org/10.3389/fdgth.2025.1659134

[18] As'ad, M., & Faran, N. (2025). Digital maturity scores as gatekeepers for health AI: useful proxy or false comfort? *Baylor University Medical Center Proceedings*, *38*(5), 779. https://doi.org/10.1080/08998280.2025.2524301

[19] Balagopalan, A., Baldini, I., Celi, L. A., Gichoya, J. W., McCoy, L. G., Naumann, T., Shalit, U., Schaar, M. van der, & Wagstaff, K. L. (2024). Machine learning for healthcare that matters: Reorienting from technical novelty to equitable impact. *PLOS Digital Health*, *3*(4). https://doi.org/10.1371/journal.pdig.0000474

[20] Baradwaj, S. S., Gilliland, D. G., Rincon, J., Hermjakob, H., Yan, Y., Adam, I., Lemaster, G., Wang, D., Watson, K. E., Bui, A., Wang, W., & Ping, P. (2024). Building an Ethical and Trustworthy Biomedical AI Ecosystem for the Translational and Clinical Integration of Foundational Models. *arXiv (Cornell University)*. https://doi.org/10.48550/arxiv.2408.01431

Journal of Information Systems Engineering and Management
2025, 10(62s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/
Research Article

[21] Behar, J. A., Levy, J., & Celi, L. A. (2023). Generalization in medical AI: a perspective on developing scalable models. *arXiv (Cornell University)*. https://doi.org/10.48550/arxiv.2311.05418

[22] Bignami, E., Darhour, L. J., Franco, G., Guarnieri, M., & Bellini, V. (2025). AI policy in healthcare: a checklist-based methodology for structured implementation. *Journal of Anesthesia Analgesia and Critical Care*, *5*(1). https://doi.org/10.1186/s44158-025-00278-3

[23] Bodnari, A., & Travis, J. (2025). Scaling enterprise AI in healthcare: the role of governance in risk mitigation frameworks [Review of *Scaling enterprise AI in healthcare: the role of governance in risk mitigation frameworks*]. *Npj Digital Medicine*, *8*(1), 272. Nature Portfolio. https://doi.org/10.1038/s41746-025-01700-4

[24] Bouderhem, R. (2024). Shaping the future of AI in healthcare through ethics and governance. *Humanities and Social Sciences Communications*, *11*(1). https://doi.org/10.1057/s41599- 024-02894-w

[25] Božić, V. (2023). Integrated Risk Management and Artificial Intelligence in Hospital. *Journal of AI*, *7*(1), 63. https://doi.org/10.61969/jai.1329224

[26] Brady, A. P., Allen, B., Chong, J., Kotter, E., Kottler, N., Mongan, J., Oakden-Rayner, L., Santos, D. P. dos, Tang, A., Wald, C., & Slavotinek, J. (2024). Developing, purchasing, implementing and monitoring AI tools in radiology: practical considerations. A multi-society statement from the ACR, CAR, ESR, RANZCR & RSNA. *Insights into Imaging*, *15*(1). https://doi.org/10.1186/s13244-023-01541-3

[27] Bragazzi, N. L., & Garbarino, S. (2024). Toward Clinical Generative AI: Conceptual Framework. *JMIR AI*, *3*. https://doi.org/10.2196/55957

[28] Byberg, E., & Crimi, M. (2025). Preparing hospitals and health organizations for AI: practical guidelines for the required infrastructure. *Frontiers in Digital Health*, *7*. https://doi.org/10.3389/fdgth.2025.1605006

[29] Byfield, A., Poulett, W., Wallace, B., Jose, A., Tyagi, S., Shembekar, S., Qayyum, A., Qadir, J., & Bilal, M. (2023). MedISure: Towards Assuring Machine Learning-based Medical Image Classifiers using Mixup Boundary Analysis. *arXiv (Cornell University)*. https://doi.org/10.48550/arxiv.2311.13978

[30] Carini, C., & Seyhan, A. A. (2024). Tribulations and future opportunities for artificial intelligence in precision medicine [Review of *Tribulations and future opportunities for artificial intelligence in precision medicine*]. *Journal of Translational Medicine*, *22*(1). BioMed Central. https://doi.org/10.1186/s12967-024-05067-0

[31] Carvalho, E., Mascarenhas, M., Pinheiro, F., Correia, R., Balseiro, S., Barbosa, G. P., Guerra, A., Oliveira, D., Moura, R., Santos, A., & Ramião, N. (2025). Predetermined Change Control Plans: Guiding Principles for Advancing Safe, Effective, and High-Quality AI-ML Technologies. *JMIR AI*, *4*. https://doi.org/10.2196/76854

[32] Chakraborty, A., & Karhade, M. (2024). Global AI Governance in Healthcare: A Cross-Jurisdictional Regulatory Analysis. *arXiv (Cornell University)*. https://doi.org/10.48550/arxiv.2406.08695

[33] Chen, Y.-J., Albarqawi, A., & Chen, C. (2025). Enhancing Clinical Decision-Making: Integrating Multi-Agent Systems with Ethical AI Governance. *arXiv (Cornell University)*. https://doi.org/10.48550/arxiv.2504.03699

[34] Chin, M. H., Afsarmanesh, N., Bierman, A. S., Chang, C., Colón-Rodríguez, C. J., Dullabh, P., Duran, D. G., Fair, M., Hernandez-Boussard, T., Hightower, M., Jain, A., Jordan, W. B., Konya, S., Moore, R. H., Moore, T. T., Rodriguez, R., Shaheen, G., Snyder, L. P., Srinivasan, M., ... Ohno-Machado, L. (2023). Guiding Principles to Address the Impact of Algorithm Bias on Racial and Ethnic Disparities in Health and Health Care. *JAMA Network Open*, *6*(12). https://doi.org/10.1001/jamanetworkopen.2023.45050

[35] Choudhury, A. (2022). Toward an Ecologically Valid Conceptual Framework for the Use of Artificial Intelligence in Clinical Settings: Need for Systems Thinking, Accountability, Decision-making, Trust,

and Patient Safety Considerations in Safeguarding the Technology and Clinicians. *JMIR Human Factors*, *9*(2). https://doi.org/10.2196/35421

[36] Chatterjee, I., -, R. G., -, S. S., Das, K. C., & -, M. K. (2024). Revolutionizing Innovations and Impact of Artificial Intelligence in Healthcare. *International Journal For Multidisciplinary Research*, *6*(3). https://doi.org/10.36948/ijfmr.2024.v06i03.19333

[37] Davis, S. E., Walsh, C. G., & Matheny, M. E. (2022). Open questions and research gaps for monitoring and updating AI-enabled tools in clinical settings. *Frontiers in Digital Health*, *4*. https://doi.org/10.3389/fdgth.2022.958284

[38] Dehghani, F., Dibaji, M., Anzum, F., Dey, L., Basdemir, A., Bayat, S., Boucher, J.-C., Drew, S., Eaton, S. E., Frayne, R., Ginde, G., Harris, A. E., Ioannou, Y., Lebel, C., Lysack, J., Arzuaga, L. S., Stanley, E. A. M., Souza, R., Santos, R. de S., … Bento, M. (2024). Trustworthy and Responsible AI for Human-Centric Autonomous Decision-Making Systems. *arXiv (Cornell University)*. https://doi.org/10.48550/arxiv.2408.15550

[39] de-Manuel-Vicente, C., Fernández-Narro, D., Blanes-Selva, V., García-Gómez, J. M., & Sáez, C. (2024). A Development Framework for Trustworthy Artificial Intelligence in Health with Example Code Pipelines. *medRxiv (Cold Spring Harbor Laboratory)*. https://doi.org/10.1101/2024.07.17.24310418

[40] Dixit, A., Quaglietta, J., & Gaulton, C. (2021). Preparing for the future: How organizations can prepare boards, leaders, and risk managers for artificial intelligence. *Healthcare Management Forum*, *34*(6), 346. https://doi.org/10.1177/08404704211037995

[41] Du, J. (2025). Toward Responsible and Beneficial AI: Comparing Regulatory and Guidance-Based Approaches -A Comprehensive Comparative Analysis of Artificial Intelligence Governance Frameworks across the European Union, United States, China, and IEEE. *arXiv (Cornell University)*. https://doi.org/10.48550/arxiv.2508.00868

[42] Edpuganti, S., Shamim, A., Gangolli, V. H., Weerasekara, R. A. D. K. N., & Yellamilli, A. (2025). Artificial Intelligence in Cardiovascular Imaging: Current Landscape, Clinical Impact, and Future Directions [Review of *Artificial Intelligence in Cardiovascular Imaging: Current Landscape, Clinical Impact, and Future Directions*]. *Discoveries*, *13*(2). https://doi.org/10.15190/d.2025.10

[43] Fahim, Y. A., Hasani, I. W., Kabba, S., & Ragab, W. M. (2025). Artificial intelligence in healthcare and medicine: clinical applications, therapeutic advances, and future perspectives [Review of *Artificial intelligence in healthcare and medicine: clinical applications, therapeutic advances, and future perspectives*]. *European Journal of Medical Research*, *30*(1). BioMed Central. https://doi.org/10.1186/s40001-025-03196-w

[44] Farah, L., Borget, I., Martelli, N., & Vallée, A. (2024). Suitability of the Current Health Technology Assessment of Innovative Artificial Intelligence-Based Medical Devices: Scoping Literature Review. *Journal of Medical Internet Research*, *26*. https://doi.org/10.2196/51514

[45] Feng, J., Phillips, R. V., Malenica, I., Bishara, A., Hubbard, A., Celi, L. A., & Pirracchio, R. (2022). Clinical artificial intelligence quality improvement: towards continual monitoring and updating of AI algorithms in healthcare [Review of *Clinical artificial intelligence quality improvement: towards continual monitoring and updating of AI algorithms in healthcare*]. *Npj Digital Medicine*, *5*(1), 66. Nature Portfolio. https://doi.org/10.1038/s41746-022-00611-y

[46] Fernández, C. I., Felipe, L., Shotande, M., Zitu, M., Tripathi, A., Rasool, G., Naqa, I. E., Rudrapatna, V. A., & Valdés, G. (2025). Trustworthy AI for Medicine: Continuous Hallucination Detection and Elimination with CHECK. *arXiv (Cornell University)*. https://doi.org/10.48550/arxiv.2506.11129

[47] Gallifant, J., Kellogg, K. C., Butler, M., Centi, A., Chen, S., Doyle, P. J., Dutta, S., Guo, J., Hadfield, M. J., Kim, E., Kozono, D., Aerts, H. J. W. L., Landman, A., Mak, R. H., Mishuris, R. G., Nelson, T. L., Savova, G., Sharon, E., Silverman, B. C., … Bitterman, D. S. (2025). Beyond the Algorithm: A Field Guide to Deploying AI Agents in Clinical Practice. *arXiv (Cornell University)*. https://doi.org/10.48550/arxiv.2509.26153

[48] Gambetti, A., Han, Q., Shen, H., & Soares, C. (2025). A Survey on Human-Centered Evaluation of Explainable AI Methods in Clinical Decision Support Systems. *arXiv (Cornell University)*. https://doi.org/10.48550/arxiv.2502.09849

[49] Gill, A. Y., Saeed, A., Rasool, S., Husnain, A., & Hussain, H. K. (2023). Revolutionizing Healthcare: How Machine Learning is Transforming Patient Diagnoses - a Comprehensive Review of AI's Impact on Medical Diagnosis [Review of *Revolutionizing Healthcare: How Machine Learning is Transforming Patient Diagnoses - a Comprehensive Review of AI's Impact on Medical Diagnosis*]. *Journal Of World Science*, *2*(10), 1638. https://doi.org/10.58344/jws.v2i10.449

[50] Göktaş, P., & Grzybowski, A. (2025). Shaping the Future of Healthcare: Ethical Clinical Challenges and Pathways to Trustworthy AI. *Journal of Clinical Medicine*, *14*(5), 1605. https://doi.org/10.3390/jcm14051605

[51] González, C., Fuchs, M., Santos, D. P. dos, Matthies, P., Trenz, M., Grüning, M., Chaudhari, A., Larson, D. B., Othman, A. E., Kim, M. I., Nensa, F., & Mukhopadhyay, A. (2024). Regulating radiology AI medical devices that evolve in their lifecycle. *arXiv (Cornell University)*. https://doi.org/10.48550/arxiv.2412.20498

[52] González-Gonzalo, C., Thee, E. F., Klaver, C. C. W., Lee, A., Schlingemann, R. O., Tufail, A., Verbraak, F. D., & Sánchez, C. I. (2021). Trustworthy AI: Closing the gap between development and integration of AI systems in ophthalmic practice [Review of *Trustworthy AI: Closing the gap between development and integration of AI systems in ophthalmic practice*]. *Progress in Retinal and Eye Research*, *90*, 101034. Elsevier BV. https://doi.org/10.1016/j.preteyeres.2021.101034

[53] Gu, Y., Fu, J., Liu, X., Valanarasu, J. M. J., Codella, N. C., Tan, R., Liu, Q., Jin, Y., Zhang, S., Wang, J., Wang, R., Song, L., Qin, G., Usuyama, N., Wong, C., Cheng, H., Lee, H., Sanapathi, P., Hilado, S., … Vozila, P. (2025). *The Illusion of Readiness in Health AI*. https://doi.org/10.48550/ARXIV.2509.18234

[54] Guan, H., Bates, D. W., & Zhou, L. (2025). Keeping Medical AI Healthy and Trustworthy: A Review of Detection and Correction Methods for System Degradation. *arXiv (Cornell University)*. https://doi.org/10.48550/arxiv.2506.17442

[55] Guerra, R. (2024). Enhancing risk management in hospitals: leveraging artificial intelligence for improved outcomes. *Italian Journal of Medicine*, *18*(2). https://doi.org/10.4081/itjm.2024.1721

[56] Hadley, E. (2022). Prioritizing Policies for Furthering Responsible Artificial Intelligence in the United States. *arXiv (Cornell University)*. https://doi.org/10.48550/arxiv.2212.00740

[57] Hasanzadeh, F., Josephson, C. B., Waters, G., Adedinsewo, D., Azizi, Z., & White, J. A. (2025). Bias recognition and mitigation strategies in artificial intelligence healthcare applications. *Npj Digital Medicine*, *8*(1), 154. https://doi.org/10.1038/s41746-025-01503-7

[58] Hellmeier, F., Brosien, K., Eickhoff, C., & Meyer, A. (2024). Beyond One-Time Validation: A Framework for Adaptive Validation of Prognostic and Diagnostic AI-based Medical Devices. *arXiv (Cornell University)*. https://doi.org/10.48550/arxiv.2409.04794

[59] Higgins, D., & Johner, C. (2022). Validation of artificial intelligence containing products across the regulated healthcare industries. *Research Square (Research Square)*. https://doi.org/10.21203/rs.3.rs-2153749/v1

[60] Hoffman-Peterson, A. (2025). Innovating with Trust: Regulation and Organizational Field Maturity in the Era of Emerging Health Technologies. *Deep Blue (University of Michigan)*. https://doi.org/10.7302/27116

[61] Huysentruyt, K., Kjoersvik, O., Dobracki, P., Savage, E., Mishalov, E., Cherry, M. J., Leonard, E. B., Taylor, R. C., Patel, B., & Abatemarco, D. (2021). Validating Intelligent Automation Systems in Pharmacovigilance: Insights from Good Manufacturing Practices. *Drug Safety*, *44*(3), 261. https://doi.org/10.1007/s40264-020-01030-2

[62] Isaacks, D. B., & Borkowski, A. (2024). Implementing Trustworthy AI in VA High Reliability Health Care Organizations. *Federal Practitioner*, *41*(2). https://doi.org/10.12788/fp.0454

[63] Jarquin, D. V., Sharp, A. J., Zhang, R., Nguyen, T., Hoang, A. T., & Pianykh, O. S. (2022). Temporal quality degradation in AI models. *Scientific Reports*, *12*(1). https://doi.org/10.1038/s41598-022-15245-z

[64] Jha, D., Rauniyar, A., Srivastava, A., Hagos, D. H., Tomar, N. K., Sharma, V., Keleş, E., Zhang, Z., Demir, U., Topcu, A. E., Yazidi, A., Håakegård, J. E., & Bağcı, U. (2023). A Conceptual Algorithm for Applying Ethical Principles of AI to Medical Practice. *arXiv (Cornell University)*. https://doi.org/10.48550/arxiv.2304.11530

[65] Jiao, J., Afroogh, S., Chen, K. J., Atkinson, D., & Dhurandhar, A. (2025). Generative AI and LLMs in Industry: A text-mining Analysis and Critical Evaluation of Guidelines and Policy Statements Across Fourteen Industrial Sectors. *arXiv (Cornell University)*. https://doi.org/10.48550/arxiv.2501.00957

[66] Khalifa, A., & Hussein, R. (2025). Artificial Intelligence For 6P Medicine: Consolidating AI Needs of Predictive, Preventive, Personalized, Participatory, Precision, and Public Health Trajectories. *Research Square (Research Square)*. https://doi.org/10.21203/rs.3.rs-7402571/v1

[67] Kim, J., Kim, S. Y., Kim, E.-A., Sim, J.-A., Lee, Y., & Kim, H. (2024). Developing a Framework for Self-regulatory Governance in Healthcare AI Research: Insights from South Korea. *Asian Bioethics Review*, *16*(3), 391. https://doi.org/10.1007/s41649-024-00281-w

[68] Kim, J. Y., Hasan, A., Kueper, J. K., Tang, T., Hayes, C., Fine, B., Balu, S., & Sendak, M. (2025). Establishing organizational AI governance in healthcare: a case study in Canada. *Npj Digital Medicine*, *8*(1). https://doi.org/10.1038/s41746-025-01909-3

[69] Labkoff, S. E., Oladimeji, B., Kannry, J., Solomonides, A., Leftwich, R., Koski, E., Joseph, A. L., López-González, M., Fleisher, L. A., Nolen, K., Dutta, S., Levy, D. R., Price, A., Barr, P., Hron, J. D., Lin, B., Srivastava, G., Pastor, N., Luque, U. S., … Quintana, Y. (2024). Toward a responsible future: recommendations for AI-enabled clinical decision support. *Journal of the American Medical Informatics Association*, *31*(11), 2730. https://doi.org/10.1093/jamia/ocae209

[70] Lekadir, K., Frangi, A. F., Porras, A. R., Glocker, B., Cintas, C., Langlotz, C. P., Weicken, E., Asselbergs, F. W., Prior, F., Collins, G. S., Kaissis, G., Tsakou, G., Buvat, I., Kalpathy-Cramer, J., Mongan, J., Schnabel, J. A., Kushibar, K., Riklund, K., Marias, K., … Starmans, M. P. A. (2025). FUTURE-AI: international consensus guideline for trustworthy and deployable artificial intelligence in healthcare. *BMJ*, *388*. https://doi.org/10.1136/bmj-2024-081554

[71] Li, Y., Yi, X., Fu, J., Yang, Y., Duan, C., & Wang, J. (2025). Reducing misdiagnosis in AI-driven medical diagnostics: a multidimensional framework for technical, ethical, and policy solutions. *Frontiers in Medicine*, *12*. https://doi.org/10.3389/fmed.2025.1594450

[72] Lopes, S., Mascarenhas, M., Fonseca, J., Fernandes, G., & Leite-Moreira, A. (2025). Unveiling the Algorithm: The Role of Explainable Artificial Intelligence in Modern Surgery. *Healthcare*, *13*(24), 3208. https://doi.org/10.3390/healthcare13243208

[73] Lundström, C., & Lindvall, M. (2022). Mapping the Landscape of Care Providers' Quality Assurance Approaches for AI in Diagnostic Imaging [Review of *Mapping the Landscape of Care Providers' Quality Assurance Approaches for AI in Diagnostic Imaging*]. *Journal of Digital Imaging*, *36*(2), 379. Springer Science+Business Media. https://doi.org/10.1007/s10278-022-00731-7

[74] Li, X., Zhang, L., Yang, J., & Teng, F. (2024). Role of Artificial Intelligence in Medical Image Analysis: A Review of Current Trends and Future Directions. *Journal of Medical and Biological Engineering*, *44*(2), 231. https://doi.org/10.1007/s40846-024-00863-x

[75] Mahmood, U., Shukla-Dave, A., Chan, H., Drukker, K., Samala, R. K., Chen, Q., Vergara, D., Greenspan, H., Petrick, N., Sahiner, B., Huo, Z., Summers, R. M., Kenny, H., Tourassi, G. D., Deserno, T. M., Grizzard, K., Näppi, J. J., Yoshida, H., Regge, D., … Hadjiiski, L. M. (2024). Artificial intelligence in medicine: mitigating risks and maximizing benefits via quality assurance, quality control, and acceptance testing. *BJR|Artificial Intelligence*, *1*(1). https://doi.org/10.1093/bjrai/ubae003

[76] Mallardi, G., Calefato, F., Quaranta, L., & Lanubile, F. (2025). Integrating AI into healthcare systems: A multivocal literature review. *Smart Health*, *39*, 100631. https://doi.org/10.1016/j.smhl.2025.100631

[77] Mashar, M., Chawla, S., Chen, F., Lubwama, B., Patel, K., Kelshiker, M. A., Bächtiger, P., & Peters, N. S. (2023). Artificial Intelligence Algorithms in Health Care: Is the Current Food and Drug Administration Regulation Sufficient? *JMIR AI*, *2*. https://doi.org/10.2196/42940

[78] Mirakhori, F., & Niazi, S. K. (2025). Harnessing the AI/ML in Drug and Biological Products Discovery and Development: The Regulatory Perspective. *Pharmaceuticals*, *18*(1), 47. https://doi.org/10.3390/ph18010047

[79] Mökander, J., & Floridi, L. (2022). Operationalising AI governance through ethics-based auditing: an industry case study. *AI and Ethics*, *3*(2), 451. https://doi.org/10.1007/s43681-022-00171-7

[80] Moreno-Sánchez, P. A., Ser, J. D., Gils, M. van, & Hernesniemi, J. (2025). A design framework for operationalizing trustworthy artificial intelligence in healthcare: Requirements, tradeoffs and challenges for its clinical adoption. *Information Fusion*, *127*, 103812. https://doi.org/10.1016/j.inffus.2025.103812

[81] Mullankandy, S. (2024). Navigating the Complexity of Regulations: Harnessing AI/ML for Precise Reporting. *Deleted Journal*, *3*(1), 49. https://doi.org/10.60087/jaigs.v3i1.65

[82] Nair, M., Svedberg, P., Larsson, I., & Nygren, J. (2024). A comprehensive overview of barriers and strategies for AI implementation in healthcare: Mixed-method design [Review of *A comprehensive overview of barriers and strategies for AI implementation in healthcare: Mixed-method design*]. *PLoS ONE*, *19*(8). Public Library of Science. https://doi.org/10.1371/journal.pone.0305949

[83] Nair, S. (2025). Explainable AI in Gxp Validation: Balancing Automation, Traceability, And Regulatory Trust in The Pharmaceutical Industry. *Clinical Medicine And Health Research Journal*, *5*(5), 1430. https://doi.org/10.18535/cmhrj.v5i05.509

[84] Nasarian, E., Alizadehsani, R., Acharya, U. R., & Tsui, K. (2023). Robust Interpretability Approach to Enhance Trustworthy AI in Healthcare: A Systematic Review of the Last Decade Toward Proposed Framework [Review of *Robust Interpretability Approach to Enhance Trustworthy AI in Healthcare: A Systematic Review of the Last Decade Toward Proposed Framework*]. *Research Square (Research Square)*. Research Square (United States). https://doi.org/10.21203/rs.3.rs-3626164/v1

[85] Nasarian, E., Alizadehsani, R., Acharya, U. R., & Tsui, K. (2024). Designing interpretable ML system to enhance trust in healthcare: A systematic review to proposed responsible clinician-AI-collaboration framework [Review of *Designing interpretable ML system to enhance trust in healthcare: A systematic review to proposed responsible clinician-AI-collaboration framework*]. *Information Fusion*, *108*, 102412. Elsevier BV. https://doi.org/10.1016/j.inffus.2024.102412

[86] Nasarian, E., Alizadehsani, R., Acharyac, U. R., & Tsui, D. (2023). Designing Interpretable ML System to Enhance Trust in Healthcare: A Systematic Review to Proposed Responsible Clinician-AI-Collaboration Framework [Review of *Designing Interpretable ML System to Enhance Trust in Healthcare: A Systematic Review to Proposed Responsible Clinician-AI-Collaboration Framework*]. *arXiv (Cornell University)*. Cornell University. https://doi.org/10.48550/arxiv.2311.11055

[87] Niazi, S. K. (2023). The Coming of Age of AI/ML in Drug Discovery, Development, Clinical Testing, and Manufacturing: The FDA Perspectives. *Drug Design Development and Therapy*, 2691. https://doi.org/10.2147/dddt.s424991

[88] Niazi, S. K. (2025). Regulatory Perspectives for AI/ML Implementation in Pharmaceutical GMP Environments [Review of *Regulatory Perspectives for AI/ML Implementation in Pharmaceutical GMP Environments*]. *Pharmaceuticals*, *18*(6), 901. Multidisciplinary Digital Publishing Institute. https://doi.org/10.3390/ph18060901

[89] Nong, P. (2024). Equity and AI Governance at Academic Medical Centers. *The American Journal of Managed Care*, *30*. https://doi.org/10.37765/ajmc.2024.89555

[90] Olumuyiwa, B. I., Han, T. A., & Shamszaman, Z. U. (2024). Enhancing Cancer Diagnosis with Explainable & Trustworthy Deep Learning Models. *arXiv (Cornell University)*. https://doi.org/10.48550/arxiv.2412.17527

[91] Overgaard, S. M., Graham, M., Brereton, T. A., Pencina, M., Halamka, J., Vidal, D., & Economou-Zavlanos, N. (2023). Implementing quality management systems to close the AI translation gap and facilitate safe, ethical, and effective health AI solutions. *Npj Digital Medicine*, *6*(1). https://doi.org/10.1038/s41746-023-00968-8
[92] Owens, K., Griffen, Z., & Damaraju, L. (2025). Managing a "responsibility vacuum" in AI monitoring and governance in healthcare: a qualitative study. *BMC Health Services Research*, *25*(1). https://doi.org/10.1186/s12913-025-13388-z
[93] Oyeniyi, J. (2024). The role of AI and mobile apps in patient-centric healthcare delivery. *World Journal of Advanced Research and Reviews*, *22*(1), 1897. https://doi.org/10.30574/wjarr.2024.22.1.1331
[94] Pakhale, Ms. D. D. (2025). Unbreakable Codes: Exploring Quantum Computing Applications in Cybersecurity. *International Journal for Research in Applied Science and Engineering Technology*, *13*(6), 900. https://doi.org/10.22214/ijraset.2025.72266
[95] Papadopoulou, E., Exarchos, T. P., Jenko, S., Krepelkova, K., & Namorado, J. (2025). Artificial intelligence (AI) in health systems: introducing a 'Do Good' approach. *AI and Ethics*. https://doi.org/10.1007/s43681-025-00757-x
[96] Peluso, A. (2025). Recent Advances, Applications and Open Challenges in Machine Learning for Health: Reflections from Research Roundtables at ML4H 2024 Symposium. *OSTI OAI (U.S. Department of Energy Office of Scientific and Technical Information)*. https://www.osti.gov/biblio/2538034
[97] Pham, T. D. (2025). Ethical and legal considerations in healthcare AI: innovation and policy for safe and fair use. *Royal Society Open Science*, *12*(5), 241873. https://doi.org/10.1098/rsos.241873
[98] Prabhu, N. (2025). Exploring the Effects of Artificial Intelligence on Healthcare: Emphasizing Patient Safety, Data Security and Fair Access. *International Journal for Research in Applied Science and Engineering Technology*, *13*(10), 1777. https://doi.org/10.22214/ijraset.2025.74683
[99] Quazi, F., Mohammed, A. S., & Gorrepati, N. (2024). *Transforming Treatment and Diagnosis in Healthcare through AI*. https://doi.org/10.21428/e90189c8.072ffbe8
[100] Rahamtalla, B., Medani, I. E., Abdelhag, M. E., Eltigani, S., Rajan, S. K., Falgy, E., Hassan, N. M., Fadailu, M. E., Khudhayr, H. A., & Abdalla, A. (2025). The AI-Powered Healthcare Ecosystem: Bridging the Chasm Between Technical Validation and Systemic Integration—A Systematic Review. *Future Internet*, *17*(12), 550. https://doi.org/10.3390/fi17120550
[101] Ramírez., J. G. C., Islam, Md. M., & Even, A. I. H. (2024). Machine Learning Applications in Healthcare: Current Trends and Future Prospects. *Deleted Journal*, *1*(1). https://doi.org/10.60087/jaigs.v1i1.33
[102] Rashid, A. (2024). *Untitled*. https://doi.org/10.55277/researchhub.vq5dnd6h
[103] Raz, A. E., Inbar, Y., Avnoon, N., Lifshitz-Milwidsky, L. B., Hofman, B., Honig, A. M., & Paz, Z. (2025). Who moved my scan? Early adopter experiences with pre- and post-market healthcare AI regulation challenges. *Frontiers in Medicine*, *12*. https://doi.org/10.3389/fmed.2025.1651934
[104] Ross, J., Hammouche, S., Chen, J. C. Y., Rockall, A., Alabed, S., Chen, M., Dwivedi, K., Fascia, D., Greenhalgh, R., Hall, M., Halliday, K., Harden, S., Ramsden, W., & Shelmerdine, S. C. (2024). Beyond regulatory compliance: evaluating radiology artificial intelligence applications in deployment [Review of *Beyond regulatory compliance: evaluating radiology artificial intelligence applications in deployment*]. *Clinical Radiology*, *79*(5), 338. Elsevier BV. https://doi.org/10.1016/j.crad.2024.01.026
[105] Rozenblit, L., Price, A., Solomonides, A., Joseph, A. L., Koski, E., Srivastava, G., Labkoff, S. E., Bray, D. A., López-González, M., Singh, R., deBronkart, D., Barr, P., Szolovits, P., Dattani, K., Jaffe, C. L., Fridsma, D. B., Baris, R., Leftwich, R., Stolper, R., … Quintana, Y. (2025). Toward responsible AI governance: Balancing multi-stakeholder perspectives on AI in healthcare. *International Journal of Medical Informatics*, *203*, 106015. https://doi.org/10.1016/j.ijmedinf.2025.106015
[106] Sabet, C., Tamirisa, K., Bitterman, D. S., & Gallifant, J. (2025). Regulating medical AI before midnight strikes: Addressing bias, data fidelity, and implementation challenges. *PLOS Digital Health*, *4*(8). https://doi.org/10.1371/journal.pdig.0000986
[107] Saboury, B., Bradshaw, T., Boellaard, R., Buvat, I., Dutta, J., Hatt, M., Jha, A. K., Li, Q., Liu, C., McMeekin, H., Morris, M., Scott, P. J. H., Siegel, E. L., Sunderland, J. J., Pandit-Taskar, N., Wahl, R.

L., Zuehlsdorff, S., & Rahmim, A. (2022). Artificial Intelligence in Nuclear Medicine: Opportunities, Challenges, and Responsibilities Toward a Trustworthy Ecosystem. *Journal of Nuclear Medicine*, *64*(2), 188. https://doi.org/10.2967/jnumed.121.263703

[108] Saenz, A., McCoy, T. P., Mantha, A. B., Martin, R. F., Rizzoni, D., Adair, D., Heaney, D., Sisodia, R., Park, L., Forsberg, R., Tuffy, G., Murphy, S. N., Dreyer, K. J., Jones, M. W., Cosier, H. J., Logan, M., Bundela, Y., Centi, A., Ting, D. S. -K., … Mishuris, R. G. (2024). Establishing responsible use of AI guidelines: a comprehensive case study for healthcare institutions. *Npj Digital Medicine*, *7*(1), 348. https://doi.org/10.1038/s41746-024-01300-8

[109] Sagona, M., Dai, T., Macis, M., & Darden, M. (2025). Trust in AI-assisted health systems and AI's trust in humans. *Npj Health Systems*, *2*(1). https://doi.org/10.1038/s44401-025-00016-5

[110] Scaramuzza, F., Quattrocchi, G., & Tamburri, D. A. (2025). Engineering Trustworthy Machine-Learning Operations with Zero-Knowledge Proofs. *arXiv (Cornell University)*. https://doi.org/10.48550/arxiv.2505.20136

[111] Shafik, W., Hidayatullah, A. F., Kalinaki, K., Gul, H., Zakari, R. Y., & Tufail, A. (2024). A Systematic Literature Review on Transparencyand Interpretability of AI models in Healthcare:Taxonomies, Tools, Techniques, Datasets, OpenResearch Challenges, and Future Trends. *Research Square (Research Square)*. https://doi.org/10.21203/rs.3.rs-4419881/v1

[112] Shwetali, K. H., Lahare Chaitali, Kherud Rohan, Ippar Ashwini, Kotame Vikrant, Pansare. (2025). AI-Driven Approaches to Enhance Pharmacovigilance. *Zenodo (CERN European Organization for Nuclear Research)*. https://doi.org/10.5281/zenodo.17558352

[113] Siddik, M., & Pandit, H. J. (2025). Datasheets for Healthcare AI: A Framework for Transparency and Bias Mitigation. *arXiv (Cornell University)*. https://doi.org/10.48550/arxiv.2501.05617

[114] Solanki, P., Grundy, J., & Hussain, W. (2022). Operationalising ethics in artificial intelligence for healthcare: a framework for AI developers. *AI and Ethics*, *3*(1), 223. https://doi.org/10.1007/s43681-022-00195-z

[115] Somara, S., Joshi, A. M., Mitra, K., Desai, S., Lundberg, M. S., Bhasin, S., & Hunsberger, J. (2025). Artificial Intelligence in Biotechnology and Pharmaceuticals: Evolution, Applications, and Regulatory Frontiers. *Current Stem Cell Reports*, *11*(1). https://doi.org/10.1007/s40778-025-00249-y

[116] Sonawane, S., & Baviskar, M. D. (2025). Risk- Based Validation of Software, Automation and Artificial intelligence in Pharmaceuticals. *Biosciences Biotechnology Research Asia*, *4*(22), 1279. https://doi.org/10.13005/bbra/3441

[117] Suksaeree, J. (2025). A Review of Artificial Intelligence (AI)-Driven Smart and Sustainable Drug Delivery Systems: A Dual-Framework Roadmap for the Next Pharmaceutical Paradigm. *Sci*, *7*(4), 179. https://doi.org/10.3390/sci7040179

[118] Tilala, M. H., Chenchala, P. K., Choppadandi, A., Kaur, J., Naguri, S., Saoji, R., & Devaguptapu, B. (2024). Ethical Considerations in the Use of Artificial Intelligence and Machine Learning in Health Care: A Comprehensive Review [Review of *Ethical Considerations in the Use of Artificial Intelligence and Machine Learning in Health Care: A Comprehensive Review*]. *Cureus*, *16*(6). Cureus, Inc. https://doi.org/10.7759/cureus.62443

[119] Tiwari, N. (2025). The Strategic Imperative for Healthcare Organizations to Build Proprietary Foundation Models. *arXiv (Cornell University)*. https://doi.org/10.48550/arxiv.2506.11412

[120] Tjondronegoro, D., Yuwono, E., Richards, B., Green, D., & Hatakka, S. (2022). Responsible AI Implementation: A Human-centered Framework for Accelerating the Innovation Process. *arXiv (Cornell University)*. https://doi.org/10.48550/arxiv.2209.07076

[121] Tsai, P. (2025). How to Build an AI Governance Framework with Data Governance. In *SpringerBriefs in philosophy* (p. 171). Springer Nature (Netherlands). https://doi.org/10.1007/978-3-031-85478-1_10

[122] Tuấn, Đ. A. (2024). Bridging the Gap Between Black Box AI and Clinical Practice: Advancing Explainable AI for Trust, Ethics, and Personalized Healthcare Diagnostics. *Preprints.Org*. https://doi.org/10.20944/preprints202409.1974.v2

Journal of Information Systems Engineering and Management
2025, 10(62s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

[123] Ueda, D., Kakinuma, T., Fujita, S., Kamagata, K., Fushimi, Y., Ito, R., Matsui, Y., Nozaki, T., Nakaura, T., Fujima, N., Tatsugami, F., Yanagawa, M., Hirata, K., Yamada, A., Tsuboyama, T., Kawamura, M., Fujioka, T., & Naganawa, S. (2023). Fairness of artificial intelligence in healthcare: review and recommendations [Review of *Fairness of artificial intelligence in healthcare: review and recommendations*]. *Japanese Journal of Radiology*, *42*(1), 3. Springer Science+Business Media. https://doi.org/10.1007/s11604-023-01474-3

[124] Weiner, E., Dankwa-Mullan, I., Nelson, W. A., & Hassanpour, S. (2024). Ethical Challenges and Evolving Strategies in the Integration of Artificial Intelligence into Clinical Practice. *arXiv (Cornell University)*. https://doi.org/10.48550/arxiv.2412.03576

[125] Weiner, E., Dankwa-Mullan, I., Nelson, W. A., & Hassanpour, S. (2025). Ethical challenges and evolving strategies in the integration of artificial intelligence into clinical practice. *PLOS Digital Health*, *4*(4). https://doi.org/10.1371/journal.pdig.0000810

[126] Wellnhofer, E. (2022). Real-World and Regulatory Perspectives of Artificial Intelligence in Cardiovascular Imaging. *Frontiers in Cardiovascular Medicine*, *9*. https://doi.org/10.3389/fcvm.2022.890809

[127] Wells, B. J., Nguyen, H., McWilliams, A., Pallini, M., Bovi, A., Kuzma, A. B., Kramer, J., Chou, S., Hetherington, T., Corn, P., Taylor, Y. J., Cuison, A., Gagen, M., Isreal, M., Akbilgiç, O., Barr, K., Bowers, A. F., Caffrey, R. G., Carroll, M. S., … Wright, G. (2025). A practical framework for appropriate implementation and review of artificial intelligence (FAIR-AI) in healthcare. *Npj Digital Medicine*, *8*(1), 514. https://doi.org/10.1038/s41746-025-01900-y

[128] Xia, B., Lu, Q., Zhu, L., Lee, S. U., Liu, Y., & Xing, Z. (2023). Towards a Responsible AI Metrics Catalogue: A Collection of Metrics for AI Accountability. *arXiv (Cornell University)*. https://doi.org/10.48550/arxiv.2311.13158

[129] . Youssef, E., Weddle, K., Zimmerman, L. H., & Palmer, D. (2025). Pharmacovigilance in Cell and Gene Therapy: Evolving Challenges in Risk Management and Long-Term Follow-Up [Review of *Pharmacovigilance in Cell and Gene Therapy: Evolving Challenges in Risk Management and Long-Term Follow-Up*]. *Drug Safety*. Adis, Springer Healthcare. https://doi.org/10.1007/s40264-025-01596-9

[130] . Zaki, N., Akor, A., Aburuz, S., & ZainAlAbdin, S. (2025). Beyond Accuracy: An Efficiency- and Safety-Aware Framework for Evaluating Clinical AI with Large Language Models. *bioRxiv (Cold Spring Harbor Laboratory)*. https://doi.org/10.1101/2025.10.14.25338039

[131] . Zhou, K., & Gattinger, G. (2024). The Evolving Regulatory Paradigm of AI in MedTech: A Review of Perspectives and Where We Are Today [Review of *The Evolving Regulatory Paradigm of AI in MedTech: A Review of Perspectives and Where We Are Today*]. *Therapeutic Innovation & Regulatory Science*, *58*(3), 456. Springer Science+Business Media. https://doi.org/10.1007/s43441-024-00628-3

[132] Zeb, S., FNU, N., Abbasi, N., & Fahad, M. (2024). AI in Healthcare: Revolutionizing Diagnosis and Therapy. *International Journal of Multidisciplinary Sciences and Arts*, *3*(3), 118. https://doi.org/10.47709/ijmdsa.v3i3.4546