International Journal of Medical Informatics 213 (2026) 106362

Contents lists available at ScienceDirect

International Journal of Medical Informatics

journal homepage: www.elsevier.com/locate/ijmedinf

Review article

Human in the loop artificial intelligence in healthcare: applications, outcomes, and implementation challenges

David B. Olawade [a,b,c,*], Shamiul Bashir Plabon [d], Adeyinka Ojo [e], Muyiwa Ademola Ogunbona [f], Babajide David Makanjuola [g], Omobolaji Rosemary Olasilola [h]

[a] *Department of Allied and Public Health, School of Health, Sport and Bioscience, University of East London, London, United Kingdom*
[b] *Department of Research and Innovation, Medway NHS Foundation Trust, Gillingham ME7 5NY, United Kingdom*
[c] *Department of Public Health, York St John University, London E14 2BA, United Kingdom*
[d] *School of Nursing and Paramedic Science, Faculty of Life and Health Sciences, Ulster University, London Campus, EC1R 4TF, United Kingdom*
[e] *Department of Artificial Intelligence and Data Factory, Capgemini plc, London EC4V 4HN, United Kingdom*
[f] *Cell Biology and Genetics Unit, Department of Zoology, University of Ibadan, Ibadan, Nigeria*
[g] *Sheffield Business School, Department of Accounting, Banking and Finance, Sheffield Hallam University, Howard St, Sheffield City Centre, Sheffield S1 1WB, United Kingdom*
[h] *Department of Public Health, University of Hertfordshire, Hatfield AL10 9AB, United Kingdom*

A R T I C L E I N F O

Keywords:
Human in the loop
Artificial intelligence
Healthcare
Clinical decision support
Machine learning

A B S T R A C T

Background: The integration of artificial intelligence in healthcare has transformed clinical practice and research methodologies. However, concerns regarding algorithmic accountability, interpretability, and safety have necessitated human oversight in AI systems. Human in the loop artificial intelligence represents a collaborative paradigm where human expertise and machine intelligence converge to enhance decision making while maintaining ethical standards and clinical safety.
Aim: This review synthesizes current evidence on human in the loop AI in healthcare delivery and research, examining implementation frameworks, clinical outcomes, comparative advantages over fully automated and clinician-only approaches, and challenges.
Method: A comprehensive narrative review was conducted using PubMed, Scopus, Web of Science, and IEEE Xplore databases covering studies from 2018 to 2025. Data were thematically synthesized to identify patterns, frameworks, and outcomes. This narrative approach enables comprehensive conceptual synthesis across diverse HITL-AI applications and contexts.
Results: Human in the loop AI demonstrates significant applications across diagnostic imaging, clinical decision support, patient monitoring, drug discovery, and research data analysis. Evidence indicates improved diagnostic accuracy, reduced medical errors, enhanced patient safety, and increased clinician trust compared to both automated AI and traditional approaches. Implementation requires EHR interoperability, clear liability frameworks, adaptive training protocols, and quantum-safe cryptographic security. Challenges include workflow integration, regulatory gaps for adaptive systems, and sustainability concerns.
Conclusion: This review advances the field by synthesizing cross-domain implementation patterns, mapping collaboration models to risk-stratified contexts, identifying regulatory gaps for adaptive systems, and proposing future directions including post-quantum cryptographic integration, AI-driven adaptive architectures, and multi-center scalability frameworks for optimizing human–machine collaboration in healthcare.

1. Introduction

Artificial intelligence has emerged as a transformative force in modern healthcare, reshaping diagnostic processes, treatment planning, drug discovery, and healthcare research methodologies. The integration of machine learning algorithms, deep learning models, and natural language processing systems has enabled healthcare professionals to analyze vast datasets, identify complex patterns, and generate predictive

* Corresponding author at: Department of Allied and Public Health, School of Health, Sport and Bioscience, University of East London, London, United Kingdom.
E-mail address: d.olawade@uel.ac.uk (D.B. Olawade).

https://doi.org/10.1016/j.ijmedinf.2026.106362
Received 31 December 2025; Received in revised form 29 January 2026; Accepted 18 February 2026
Available online 19 February 2026

insights that were previously unattainable through conventional analytical approaches [1]. From computer assisted diagnostic systems in radiology to predictive models in intensive care units [2,3], AI technologies have demonstrated remarkable capabilities in augmenting clinical decision making and improving patient outcomes [4]. However, the autonomous deployment of AI systems in healthcare settings has raised critical concerns regarding algorithmic transparency, accountability, bias mitigation, patient safety, and ethical considerations [5,6]. These challenges have catalyzed the development and adoption of human in the loop artificial intelligence frameworks, which integrate human expertise and oversight into AI driven processes to ensure safe, ethical, and effective healthcare delivery [5,7].

Human in the loop AI represents a collaborative paradigm wherein human intelligence and machine learning algorithms work synergistically to accomplish tasks that could effectively perform independently [8]. Unlike fully automated AI systems that operate without human intervention, human in the loop frameworks incorporate human judgment at critical decision-making points, allowing clinicians, researchers, and other healthcare professionals to validate, refine, or override algorithmic recommendations based on contextual understanding, clinical experience, and patient specific factors [9,10]. This approach acknowledges that while AI systems excel at processing large datasets and identifying statistical patterns, human expertise remains essential for interpreting complex clinical scenarios, understanding nuanced patient needs, navigating ethical dilemmas, and ensuring that algorithmic outputs align with established medical standards and individual patient values. The comparative performance of HITL-AI relative to alternative approaches has emerged as a critical consideration for healthcare organizations. Evidence across multiple domains demonstrates that strategic integration of human expertise into AI decision-making processes yields superior outcomes compared to either fully automated AI or clinician-only workflows [11–13]. This synergistic advantage stems from combining AI's pattern recognition capabilities with human contextual understanding, ethical reasoning, and clinical judgment, creating a balanced ecosystem where automation efficiency complements human wisdom [9,10,14].

The application of human-in-the-loop AI in healthcare spans multiple domains including diagnostic imaging, where radiologists collaborate with AI algorithms to detect abnormalities in medical images [15–18]; clinical decision support systems, where AI generated recommendations are reviewed and validated by clinicians before implementation [19–21]; drug discovery and development, where researchers utilize AI models to identify potential therapeutic compounds while maintaining oversight of the discovery process [22–24]; patient monitoring systems, where automated alerts are triaged and validated by healthcare professionals [25–27]; and healthcare research, where AI assists in literature review, data analysis, and hypothesis generation under researcher supervision [1,28]. Each of these applications demonstrates the potential of human–machine collaboration to enhance healthcare quality, efficiency, and safety while maintaining the irreplaceable value of human judgment and expertise in clinical and research contexts.

Despite the growing adoption of human in the loop AI in healthcare, significant gaps persist in the literature regarding standardized implementation frameworks, optimal human AI interaction models, evaluation methodologies, and long-term clinical outcomes [29]. Critical operational considerations remain insufficiently addressed, including: (1) EHR integration pathways and interoperability standards for seamless HITL-AI deployment, (2) liability allocation frameworks when clinicians override AI recommendations, (3) regulatory compliance requirements for adaptive AI systems that evolve through human feedback, (4) scalability considerations for multi-center and cross-country implementations, and (5) emerging security requirements including quantum-safe cryptographic protection of healthcare data transmitted between AI systems and human operators [30–32]. Existing studies often focus on narrow applications or technical performance metrics without comprehensively examining the broader implications of human machine collaboration on clinical workflows, healthcare professional experiences, patient outcomes, and healthcare system sustainability. Furthermore, there is limited synthesis of evidence across different healthcare domains and settings, making it challenging for healthcare organizations, policymakers, and researchers to understand the full spectrum of opportunities [33], challenges, and best practices associated with human in the loop AI implementation [34]. This knowledge gap impedes the systematic and effective integration of human in the loop AI technologies into routine clinical practice and healthcare research. As summarized in Fig. 1, human-in-the-loop AI in healthcare can be conceptualized as a socio-technical framework in which AI systems and human expertise iteratively interact across multiple clinical and research domains to support safe, effective, and ethically aligned decision making.

This narrative review addresses these gaps by synthesizing current evidence on the application of human in the loop AI in healthcare delivery and healthcare research, with particular emphasis on operationalization strategies, comparative performance analysis, and future directions for next-generation HITL-AI systems. The rationale for this review stems from the urgent need to provide healthcare stakeholders with comprehensive, evidence-based guidance on implementing human in the loop AI systems that maximize benefits while mitigating risks. The novelty of this work lies in its: (1) holistic examination of HITL-AI across multiple healthcare domains to identify transferable implementation patterns, (2) explicit mapping of collaboration models to risk-stratified clinical contexts, (3) synthesis of operationalization strategies including EHR integration, regulatory compliance, and liability frameworks, (4) identification of critical regulatory gaps for adaptive HITL systems, (5) integration of emerging security considerations including quantum-safe cryptographic approaches for healthcare data protection, and (6) proposal of future directions encompassing post-quantum integration, AI-driven adaptive architectures, and multi-center scalability strategies. Unlike previous reviews examining isolated technical aspects or single-domain applications, this work synthesizes cross-domain evidence to generate actionable insights for clinical deployment across diverse healthcare settings.

The primary aim of this review is to synthesize existing literature on human in the loop AI applications in healthcare and healthcare research, examining implementation frameworks, clinical outcomes, comparative advantages, operationalization strategies, challenges, and future directions. Specific objectives include: (1) identifying key applications of human in the loop AI across healthcare domains with emphasis on comparative performance relative to automated AI and clinician-only approaches, (2) analyzing the impact of human in the loop AI on clinical decision making and patient outcomes, (3) evaluating implementation challenges and success factors including EHR integration, regulatory compliance, liability considerations, and emerging security requirements, (4) examining ethical and regulatory considerations with particular attention to gaps in oversight of adaptive AI systems, (5) mapping human-AI collaboration models to risk-stratified clinical contexts, and (6) proposing recommendations for future research and practice in human machine collaboration within healthcare contexts, including post-quantum cryptographic integration, AI-driven adaptive HITL architectures, and multi-center scalability strategies.

Fig. 1 highlights how data from diagnostic imaging, clinical decision support, patient monitoring, and healthcare research feed AI models whose outputs are interpreted, validated, and contextualized by clinicians, researchers, and patients, with human feedback continuously refining model performance and guiding ethically grounded, clinically sound decisions.

2. Methods

2.1. Review design and rationale

This narrative review was conducted following established

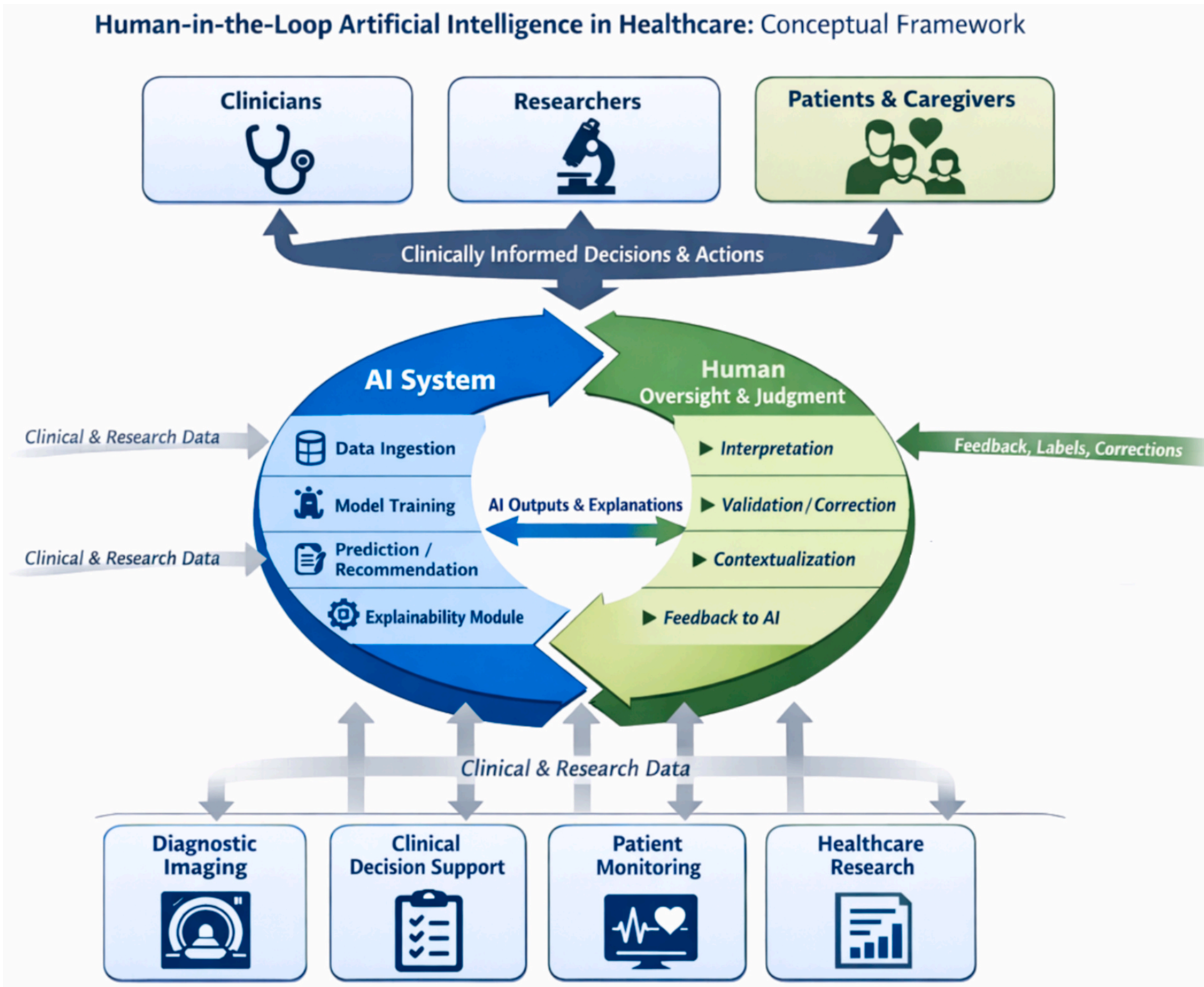

Fig. 1. Conceptual framework of human-in-the-loop artificial intelligence in healthcare, illustrating the bidirectional collaboration between AI systems and human expertise across clinical care and research domains.

methodological principles for qualitative synthesis of heterogeneous literature [35]. A narrative review approach was deliberately selected over systematic review or scoping review methodologies for several compelling reasons. First, the heterogeneity of HITL-AI implementations across diverse clinical domains (diagnostic imaging, clinical decision support, patient monitoring, drug discovery, healthcare research), varying AI architectures (deep learning, machine learning, natural language processing), different human-AI interaction models (verification, augmentation, active learning, human-in-command), and inconsistent evaluation metrics precludes meaningful application of PRISMA guidelines designed for homogeneous intervention studies. Second, the rapidly evolving nature of AI technology and implementation practices necessitates a flexible synthesis approach that can accommodate diverse study designs, theoretical frameworks, emerging applications, and real-world implementation reports rather than rigid inclusion criteria and standardized data extraction protocols typical of systematic reviews. Third, the primary objective of providing comprehensive conceptual understanding of HITL-AI across multiple healthcare domains, identifying cross-domain implementation patterns, and generating actionable insights for clinical practice and policy development aligns with narrative review's strength in synthesizing diverse evidence types to construct holistic understanding [35,36].

Importantly, this narrative review does not follow PRISMA guidelines, does not employ standardized data extraction tables, does not provide quantitative synthesis or *meta*-analysis of outcomes, does not include performance evaluation benchmarking, and does not present a PRISMA flow diagram. These methodological choices reflect the review's objectives of conceptual synthesis and actionable insight generation rather than quantitative effect size estimation. The thematic synthesis approach employed enables identification of implementation patterns, collaboration frameworks, operationalization strategies, and future directions that would be obscured by the reductionist requirements of quantitative meta-analysis.

2.2. Search strategy and study selection

A comprehensive search strategy was developed to identify relevant studies examining human in the loop AI applications in healthcare and healthcare research. The search was conducted across multiple electronic databases including PubMed, Scopus, Web of Science, and IEEE Xplore, ensuring broad coverage of biomedical, computer science, and health informatics literature. Search terms were carefully selected to capture relevant studies and included combinations of keywords such as "human in the loop," "human-in-the-loop AI," "interactive machine learning," "human machine collaboration," "artificial intelligence," "machine learning," "deep learning," "healthcare," "medicine," "clinical decision support," "diagnostic systems," "medical imaging," "patient care," and "healthcare research." Boolean operators were used to combine search terms effectively, and searches were limited to studies published between January 2018 and June 2025 to focus on recent developments in this rapidly evolving field.

Inclusion criteria for this review encompassed peer reviewed journal articles, conference proceedings, and technical reports that examined the application, implementation, or evaluation of human in the loop AI systems in healthcare settings or healthcare research contexts. Studies were required to explicitly address human oversight, interaction, or collaboration with AI systems, rather than focusing solely on fully automated AI technologies. Both empirical studies and theoretical frameworks were included to provide comprehensive insights into the current state of knowledge. Exclusion criteria included studies focusing exclusively on fully automated AI without human involvement, non-English language publications, abstracts without full text availability, and opinion pieces without substantial evidence or theoretical grounding. Studies focusing solely on technical algorithm development without healthcare application were also excluded.

Titles and abstracts were screened for relevance by the review team, followed by full-text review of selected articles to determine final inclusion. Data extraction focused on key themes including application domains, implementation frameworks, collaboration models, comparative performance where reported, outcomes reported, challenges identified, operationalization strategies, and recommendations provided. For studies reporting comparative outcomes, information on performance relative to automated AI or clinician-only approaches was extracted where available. However, given the narrative review design and absence of standardized data extraction protocols, formal quantitative pooling or meta-analysis was not conducted.

2.3. Data synthesis approach

A thematic synthesis approach was employed to organize and analyze the extracted data, allowing for identification of patterns, commonalities, and variations across studies. Themes were developed inductively through iterative reading and analysis of included studies, with regular team discussions to refine theme definitions and ensure consistency in interpretation. Major themes that emerged included: (1) application domains and comparative performance, (2) human-AI collaboration models mapped to clinical risk contexts, (3) implementation frameworks and operationalization strategies, (4) training and change management approaches, (5) ethical and regulatory challenges including gaps in adaptive system oversight, (6) data security considerations including quantum-safe cryptography, (7) unresolved tensions including liability allocation, and (8) future directions for adaptive HITL systems and scalability. This approach enabled comprehensive understanding of the multifaceted nature of human in the loop AI in healthcare while maintaining methodological rigor appropriate for narrative synthesis.

3. Applications of human in the loop AI in healthcare

3.1. Diagnostic imaging and radiology

Diagnostic imaging represents one of the most prominent application areas for human in the loop AI in healthcare. Radiological interpretation requires the analysis of complex visual information, pattern recognition, and integration of clinical context, making it an ideal domain for human machine collaboration [37]. AI algorithms, particularly deep learning models based on convolutional neural networks, have demonstrated impressive performance in detecting and classifying abnormalities in various imaging modalities including chest radiographs, computed tomography scans, magnetic resonance imaging, and mammography [16,18,38]. However, the integration of these algorithms into clinical workflows has largely followed a human in the loop approach wherein AI systems serve as decision support tools that augment rather than replace radiologist expertise [26]. Studies have shown that human in the loop implementations in radiology can improve diagnostic accuracy, reduce interpretation time, and decrease radiologist fatigue while maintaining the essential role of human judgment in final diagnostic decisions [12,15,17].

Table 1 summarizes key applications, AI functions, human roles, and reported outcomes across HITL-AI domains in healthcare, demonstrating consistent performance advantages when human expertise is strategically integrated into AI-driven processes. Evidence indicates that HITL-AI configurations outperform both fully automated AI systems and traditional clinician-only approaches across diagnostic accuracy, error reduction, and workflow efficiency metrics [12,15,17,39,40].

In breast cancer screening, human in the loop AI systems have been deployed to assist radiologists in mammography interpretation [13]. These systems typically provide preliminary assessments of mammographic images, flagging suspicious regions or lesions that warrant careful review. Radiologists then examine these flagged areas along with the entire image to make final diagnostic determinations [13]. Research has demonstrated that this collaborative approach can reduce false negative rates while maintaining low false positive rates, ultimately improving cancer detection rates compared to unassisted human interpretation [15,38]. This exemplifies the verification model of HITL-AI collaboration, where AI-generated outputs undergo systematic human validation before clinical action, an approach optimally suited for high-risk diagnostic contexts where both sensitivity and specificity are critical [9,13]. Similar applications have been developed for lung nodule detection in chest computed tomography scans, where AI algorithms identify potential nodules and radiologists validate these findings, assess nodule characteristics, and integrate imaging findings with patient history and clinical presentation [8]. The human in the loop approach in these contexts has proven valuable not only for improving diagnostic performance but also for building clinician trust in AI systems, as radiologists retain ultimate decision-making authority and can observe and learn from AI generated insights [9,10,13].

3.2. Clinical decision support systems

Clinical decision support systems represent another critical application domain for human in the loop AI in healthcare. These systems leverage AI algorithms to analyze patient data including electronic health records, laboratory results, vital signs, and imaging findings to generate diagnostic suggestions, treatment recommendations, risk predictions, or care pathway guidance [19–21]. Unlike autonomous decision systems, human in the loop clinical decision support systems present their outputs as recommendations that clinicians can accept, modify, or reject based on their clinical judgment and understanding of individual patient circumstances [41,42]. This approach acknowledges the complexity of clinical decision making, which involves not only data analysis but also consideration of patient preferences, clinical experience, ethical principles, and contextual factors that may not be fully captured in structured data [9].

Applications of human in the loop clinical decision support span multiple clinical specialties and care settings [43]. In emergency departments, AI powered triage systems analyze patient presentations and vital signs to suggest triage priorities, which emergency physicians then validate and adjust based on clinical assessment [44,45]. This represents the augmentation collaboration model, where AI risk stratification enhances rather than replaces clinical judgment, an approach well-suited for moderate-risk clinical contexts where efficiency and safety must be balanced [46]. In intensive care units, predictive models analyze continuous physiological monitoring data to identify patients at risk of clinical deterioration, generating alerts that intensivists review and act upon as clinically appropriate [47,48]. The human-in-command model employed in these critical care settings ensures physician authority over life-sustaining treatment decisions while leveraging AI vigilance for early warning detection [9,49]. In oncology, AI systems analyze tumor characteristics, genetic profiles, and treatment outcomes data to suggest personalized treatment options, which oncologists consider alongside tumor board discussions, patient preferences, and clinical guidelines [50–52]. Research evaluating these systems has generally shown improved clinical outcomes when human in the loop approaches are

Table 1
Clinical applications of human in the loop AI in healthcare with comparative performance indicators.

Application Domain	AI Function	Human Role	Collaboration Model	Reported Outcomes vs. Alternatives	References
Diagnostic Imaging	Detect abnormalities, flag suspicious regions	Validate findings, integrate clinical context, final diagnosis	Verification (high-risk contexts)	Improved diagnostic accuracy compared to AI-alone or clinician-alone; reduced interpretation time; enhanced cancer detection rates; lower false negative rates	[15–18]
Clinical Decision Support	Generate treatment recommendations, risk predictions	Review recommendations, consider patient preferences, final treatment decisions	Augmentation (moderate-risk) / Human-in-command (high-risk)	Reduced medical errors vs. unaided practice; improved guideline adherence; enhanced efficiency; higher clinician satisfaction than automated systems	[19–21,75]
Patient Monitoring	Analyze monitoring data, identify concerning patterns, generate alerts	Review alerts, assess patient status, determine interventions	Active learning (adaptive refinement)	Reduced alarm fatigue (up to 80% burden reduction); improved event detection vs. threshold-based systems; decreased readmissions; reduced provider burnout	[25–27]
Drug Discovery	Screen compound libraries, predict drug properties	Select candidates for testing, interpret results, guide research direction	Human-in-command (researcher-led exploration)	Accelerated discovery timelines vs. traditional methods; improved candidate selection; enhanced biological plausibility of identified compounds	[22–24]

Note: Performance advantages represent narrative synthesis of reported outcomes across reviewed studies. HITL-AI configurations consistently demonstrate superior performance compared to both fully automated AI systems and traditional clinician-only workflows across diagnostic accuracy, safety, efficiency, and user satisfaction metrics.

employed, including reduced medical errors, improved adherence to evidence based guidelines, enhanced efficiency, and increased clinician satisfaction compared to both unaided clinical judgment and fully automated decision systems [19,20,43].

3.3. Patient monitoring and care coordination

Patient monitoring systems increasingly incorporate human in the loop AI to manage the vast amounts of physiological and behavioral data generated by modern monitoring technologies. In hospital settings, continuous monitoring generates thousands of data points per patient per day, creating alert fatigue when traditional threshold-based alarm systems produce excessive false alarms. Human in the loop AI addresses this challenge by using machine learning algorithms to analyze monitoring data in context, identifying patterns that genuinely warrant clinical attention while suppressing false alarms [8,27,43]. Healthcare professionals receive intelligent alerts that have been filtered and prioritized by AI, allowing them to focus attention on patients who truly require intervention [49,53]. Studies have shown that this approach can reduce alarm burden by up to 80 percent while maintaining or improving detection of clinically significant events, ultimately enhancing patient safety and reducing healthcare professional burnout [54,55].

Remote patient monitoring represents another growing application of human in the loop AI, particularly for chronic disease management and post discharge care [27]. Patients use wearable devices or home monitoring equipment to collect data on vital signs, symptoms, activity levels, and medication adherence [25,56]. AI algorithms analyze these data streams to detect concerning trends or deviations from expected patterns, generating alerts to healthcare providers when intervention may be needed. Care coordinators or clinicians review these alerts, contact patients to assess their status, and determine appropriate interventions, which may range from medication adjustments to scheduling clinic visits or recommending emergency care [8]. This active learning model enables continuous refinement of alert algorithms based on clinician feedback, progressively improving positive predictive value while maintaining scalability [9,10]. Research has demonstrated that such programs can reduce hospital readmissions, improve chronic disease control, and enhance patient engagement in self-care while maintaining efficient use of healthcare professional time [27,56].

4. Applications of human in the loop AI in healthcare research

4.1. Literature review and knowledge synthesis

Healthcare research increasingly relies on human in the loop AI to manage the exponential growth of biomedical literature. With millions of research articles published annually, systematic identification and synthesis of relevant evidence has become extraordinarily time consuming using traditional manual methods. Human in the loop AI systems address this challenge by employing natural language processing and machine learning algorithms to screen literature, identify relevant studies, extract key data, and synthesize findings, while researchers maintain oversight and make final inclusion decisions [57]. These systems can rapidly screen thousands of abstracts based on predefined inclusion criteria, flagging potentially relevant studies for human review. Researchers then examine flagged articles to make final inclusion decisions, recognizing that human judgment remains essential for assessing study quality, relevance, and applicability to specific research questions [58].

Applications of human in the loop AI in systematic reviews and meta-analyses have demonstrated significant efficiency gains without compromising review quality. AI powered screening tools can reduce the time required for title and abstract screening while maintaining high sensitivity for relevant study identification. A recent study depicted a median screening load reduction of 47.1%, with studies reaching genuine recall at 95% efficiency [59]. Data extraction tasks, which traditionally require meticulous manual coding, can be accelerated through AI assisted extraction of study characteristics, outcomes, and effect sizes, with researchers verifying and correcting algorithmic outputs [60]. Some advanced systems employ active learning approaches wherein the AI model continuously learns from researcher decisions during the screening process, progressively improving its performance and reducing the burden of human review. This active learning collaboration model demonstrates iterative efficiency improvements, with human review requirements decreasing as the AI learns researcher preferences [57,58].

4.2. Clinical trial design and patient recruitment

Clinical trial design and patient recruitment represent challenging aspects of healthcare research that benefit substantially from human in the loop AI approaches. Designing optimal clinical trials requires consideration of numerous factors including patient eligibility criteria, sample size calculations, outcome measures, randomization strategies, and statistical analysis plans [61]. AI systems can analyze historical trial data, disease prevalence patterns, and patient characteristics to suggest optimal trial designs, which researchers then evaluate and refine based on scientific objectives, feasibility considerations, ethical requirements, and regulatory standards [62]. Similarly, patient recruitment, which is often the most significant bottleneck in clinical trial execution, can be enhanced through AI systems that screen electronic health records to identify potentially eligible patients. These systems apply complex eligibility criteria to large patient databases, generating lists of candidates that research coordinators then contact to assess interest, verify eligibility through detailed review, and obtain informed consent [63].

Research evaluating human in the loop AI for clinical trial applications has shown promising results. AI assisted patient recruitment systems have been reported to reduce recruitment timelines by 30 to 50 percent while identifying more eligible patients than traditional methods [64]. The human in the loop approach is critical in these contexts because final enrollment decisions must account for factors beyond electronic health record data, including patient preferences, social circumstances, comorbidities that may not be fully documented, and subjective clinical assessments that influence eligibility and trial suitability. Additionally, researchers must ensure that AI assisted recruitment does not inadvertently introduce selection biases or exclude underrepresented populations, requiring ongoing human oversight of recruitment patterns and proactive outreach to ensure diverse trial enrollment. Systematic monitoring of demographic representation in AI-assisted recruitment is essential to prevent algorithmic bias from perpetuating healthcare disparities [63,65].

5. Research data analysis and hypothesis generation

The analysis of complex healthcare datasets increasingly benefits from human in the loop AI approaches that combine algorithmic pattern recognition with researcher expertise. Modern healthcare research often involves analysis of high dimensional data including genomic sequences, proteomic profiles, medical imaging datasets, and longitudinal electronic health records containing thousands of variables [1,28]. Traditional statistical approaches may struggle to identify complex interactions and nonlinear relationships within such data, while fully automated machine learning approaches may generate spurious associations or identify patterns that lack biological plausibility [43,66]. Human in the loop AI addresses these limitations by enabling researchers to guide algorithmic exploration, interpret identified patterns in biological and clinical contexts, and generate testable hypotheses that advance scientific understanding [66].

Applications of human in the loop AI in research data analysis span multiple domains. In genomics research, AI algorithms can analyze whole genome sequencing data to identify genetic variants associated with disease risk, while researchers evaluate these associations for biological plausibility, validate findings through functional studies, and develop mechanistic hypotheses [67–69]. Recent advances include optimization-enabled deep learning approaches for disease detection from medical imaging data, where human-AI collaboration ensures biological validity of identified patterns while leveraging AI's pattern recognition capabilities [31]. In clinical outcomes research, machine learning models can identify patient subgroups with distinct treatment responses, which researchers then characterize clinically, investigate underlying mechanisms, and test through prospective studies [40,70,71]. In health services research, AI can detect patterns in healthcare utilization data that researchers examine to understand care delivery challenges and develop improvement interventions [1,20,72]. Throughout these applications, the human in the loop approach ensures that data driven discoveries are grounded in domain knowledge, subjected to appropriate validation, and translated into meaningful scientific advances rather than remaining as isolated algorithmic outputs.

6. Implementation frameworks and best practices

6.1. Human AI collaboration models

Successful implementation of human in the loop AI requires thoughtful design of human AI collaboration models that optimize the strengths of both humans and machines while mitigating their respective limitations. Several collaboration frameworks have emerged from research and practice, each with distinct characteristics, optimal use contexts, and risk-stratification considerations [29,46,73]. Table 2 explicitly maps these collaboration models to real-world clinical implementations with guidance for appropriate deployment across risk-stratified contexts.

The verification model involves AI generating outputs that humans verify before implementation, commonly seen in diagnostic imaging where AI flags abnormalities for radiologist confirmation [37]. This model is optimally suited for high-risk diagnostic contexts (e.g., cancer screening, stroke detection, critical pathology) where false negatives carry severe consequences and human expertise is essential for final diagnostic determination. The verification model requires clear protocols for human review, standardized verification workflows integrated into EHR systems, and mechanisms to track verification decisions for quality assurance and liability documentation [9,13].

The augmentation model features AI providing additional information or analysis to enhance human decision making, exemplified by clinical decision support systems that present risk scores alongside clinician assessments [46]. This model is appropriate for moderate-risk clinical decisions (e.g., emergency triage, medication dosing, discharge planning) where AI insights inform but do not determine clinical actions. Implementation requires user-centered interface design that presents AI outputs without cognitive overload, clear delineation of AI vs. human decision authority, and training protocols that emphasize critical evaluation of AI recommendations [44–46].

The active learning model creates iterative interaction where AI learns from human decisions to improve performance over time, used in literature screening where algorithms refine their criteria based on researcher selections [11,37]. This model excels in research contexts and low-to-moderate risk clinical applications (e.g., patient monitoring alert refinement, diagnostic screening prioritization) where continuous improvement is prioritized. Implementation requires robust feedback loop mechanisms, model retraining protocols, performance monitoring dashboards, and governance frameworks to ensure learning algorithms do not drift toward unintended behaviors [8,25].

The human in command model maintains humans as primary decision makers with AI serving purely supportive roles, prevalent in critical care settings where physicians make final treatment decisions after considering AI generated predictions [73]. This model is essential for very high-risk decisions (e.g., ICU treatment escalation/de-escalation, surgical planning, end-of-life care) where liability, ethical complexity, and patient values necessitate physician authority. Implementation requires AI outputs presented as suggestions rather than directives, explicit documentation that humans retain decision authority, and liability frameworks that clearly allocate responsibility to clinicians while acknowledging AI as a decision support tool [47–49].

Selection of appropriate collaboration models depends on multiple factors including clinical context, task complexity, risk level, regulatory requirements, and user preferences [74]. High risk clinical decisions typically require human in command or verification models to ensure adequate human oversight, while lower risk applications may effectively employ augmentation approaches. Task complexity influences optimal model selection, with highly complex tasks benefiting from augmentation approaches that leverage AI analytical capabilities while preserving human judgment for nuanced interpretation. User experience and trust levels also shape collaboration model effectiveness, as clinicians who are skeptical of AI may prefer human in command models initially, potentially transitioning to augmentation approaches as trust develops

Table 2
Human-AI collaboration models mapped to risk-stratified clinical contexts with implementation requirements.

Collaboration Model	Mechanism	Real-World Clinical Examples	Optimal Risk Context	Implementation Requirements	References
Verification Model	AI generates outputs; humans verify before action	Mammography: AI flags lesions → radiologist confirms; Pathology: AI detects abnormalities → pathologist validates	High-risk diagnostic (cancer screening, stroke detection, critical pathology)	EHR-integrated verification workflows; verification decision tracking; quality assurance protocols; liability documentation	[9,13,15,37]
Augmentation Model	AI provides analysis to enhance decisions; humans decide	Emergency triage: AI risk scores inform physician decisions; Medication dosing: AI suggests doses → clinician adjusts	Moderate-risk decisions (triage, medication management, discharge planning)	User-centered interfaces; clear authority delineation; critical evaluation training; cognitive load management	[44–46,73]
Active Learning Model	Iterative AI refinement from human feedback	Remote monitoring: alerts refined by clinician responses; Literature screening: AI learns from researcher selections	Low-moderate risk; Research contexts (adaptive monitoring, screening prioritization)	Feedback loop mechanisms; retraining protocols; performance monitoring; governance for algorithm drift prevention	[8,11,25]
Human-in-Command Model	Humans primary decision-makers; AI purely supportive	ICU: AI sepsis predictions → intensivist decides treatment; Surgery: AI risk assessment → surgeon plans approach	Very high-risk (ICU, surgical, life-sustaining decisions, end-of-life care)	AI as suggestions not directives; explicit human authority documentation; clear liability frameworks; decision tracking	[47–49,73]

Note: Collaboration model selection should align with clinical risk level, regulatory requirements, organizational liability frameworks, and user acceptance. Multi-center implementations require standardized model selection criteria and EHR interoperability protocols to ensure consistent human oversight across sites.

through positive experiences [10,29].

6.2. Training and change management

Effective implementation of human in the loop AI requires comprehensive training programs that prepare healthcare professionals to work effectively with AI systems. Training needs span multiple domains including technical understanding of AI capabilities and limitations, practical skills for interpreting AI outputs and integrating them into workflows, critical thinking skills for evaluating AI recommendations, and awareness of potential biases and failure modes. Many healthcare professionals have limited formal training in AI, machine learning, or data science, creating knowledge gaps that must be addressed to enable effective human AI collaboration. Training programs should be tailored to different user groups, with clinicians requiring different content than researchers, informaticians, or administrators. Effective training curricula should incorporate: (1) foundational AI literacy covering machine learning concepts and performance metrics, (2) domain-specific

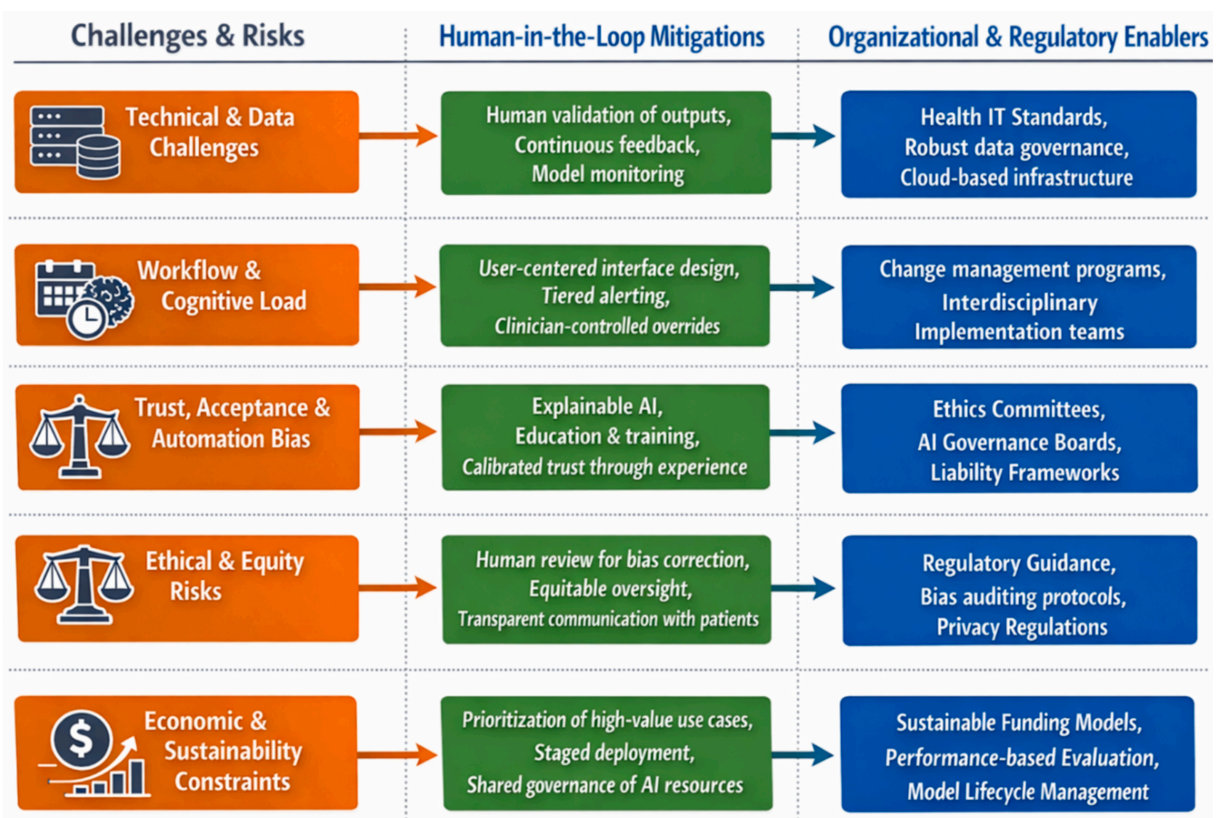

Fig. 2. Overview of key technical, workflow, trust-related, ethical, and economic challenges associated with implementing human-in-the-loop AI in healthcare, and corresponding mitigation strategies enabled by human oversight. The figure links these mitigations to organizational and regulatory enablers, emphasizing that effective governance, user-centered design, training, and ethical oversight are essential to realizing the benefits of human-in-the-loop AI while minimizing risks related to bias, safety, privacy, and sustainability.

application training focused on relevant clinical use cases, (3) hands-on simulation exercises in controlled environments, (4) competency assessment with proficiency requirements, and (5) ongoing continuing education as AI systems evolve [8,3676].

Change management processes are equally critical for successful human in the loop AI implementation. Introduction of AI systems often requires modifications to established workflows, role definitions, communication patterns, and decision-making processes, generating resistance if not managed thoughtfully [43]. Successful change management strategies include: early stakeholder engagement to understand concerns and preferences, transparent communication about AI system capabilities and limitations, gradual implementation with pilot testing before full deployment, ongoing support and feedback mechanisms to address problems as they arise, celebration of successes to build positive momentum, and leadership support demonstrating institutional commitment [36]. Organizations that treat human in the loop AI implementation as sociotechnical interventions requiring attention to both technology and human factors achieve better outcomes than those focusing narrowly on technical deployment [43]. As shown in Fig. 2, the successful deployment of human-in-the-loop AI in healthcare requires addressing interrelated technical, workflow, trust, ethical, and economic challenges through targeted mitigation strategies, supported by robust organizational and regulatory structures.

Table 3 outlines the major implementation challenges of human-in-the-loop AI across technical, workflow, trust, ethical, regulatory, and sustainability dimensions, alongside corresponding evidence-based mitigation strategies and practical operationalization considerations.

7. Impact on clinical outcomes and healthcare quality

The impact of human in the loop AI on clinical outcomes and healthcare quality has been evaluated across multiple studies, with generally positive findings when HITL-AI is compared to both fully automated AI systems and traditional clinician-only approaches [26]. In diagnostic applications, human in the loop AI has demonstrated improvements in diagnostic accuracy compared to both unassisted human interpretation and fully automated AI systems [26]. Evidence from radiology implementations indicates that radiologist-AI collaboration achieves higher sensitivity and specificity for abnormality detection than either radiologists or AI algorithms working independently [37]. This synergistic performance advantage stems from combining AI's superior pattern recognition in large-scale image analysis with radiologists' contextual integration of clinical history, prior imaging, and subtle findings that may lack clear patterns [9,13,15].

Patient safety represents another domain where human in the loop AI has demonstrated positive impact. Clinical decision support systems that alert clinicians to potential medication errors, drug interactions, or deviations from evidence-based guidelines have been shown to reduce preventable adverse events [43]. The human in the loop approach is critical for these applications because automated alerts without human review often generate excessive false positives, leading to alert fatigue and alert override behaviors that undermine safety benefits [54,55]. Systems that use AI to intelligently filter alerts and prioritize those requiring attention achieve better safety outcomes while minimizing workflow disruption compared to threshold-based automated alerting [54]. Similarly, patient monitoring systems employing human in the

Table 3
Implementation challenges and solutions for human in the loop AI in healthcare with operationalization strategies.

Challenge Category	Specific Challenges	Evidence-Based Solutions	Operationalization Requirements	References
Technical Integration	Interoperability with existing EHR systems; data quality issues; computational resource requirements; lack of standardized APIs	Adopt HL7 FHIR and SMART on FHIR standards; implement robust data governance frameworks; utilize cloud computing infrastructure; establish data quality monitoring protocols	EHR vendor collaboration; API development resources; cloud infrastructure investment; data quality dashboards; interoperability testing protocols	[86]
Workflow Integration	Disruption to established clinical workflows; increased cognitive load on clinicians; time constraints limiting AI interaction; inadequate workflow analysis pre-implementation	Conduct comprehensive workflow analysis pre-implementation; design user-centered interfaces minimizing cognitive burden; provide adequate training and ongoing support; implement phased rollout with workflow optimization	Workflow mapping tools; human factors engineering; time-motion studies; iterative interface refinement; clinician feedback mechanisms	[87]
Trust and Acceptance	Clinician skepticism about AI capabilities; lack of algorithmic transparency; concerns about professional autonomy; liability concerns; automation bias risks	Implement explainable AI (XAI) approaches providing interpretable outputs; demonstrate performance through rigorous pilot studies; clarify liability frameworks with legal counsel; establish trust calibration training	XAI tool integration; pilot study protocols with comparative performance data; institutional liability policies; trust assessment surveys; calibration training programs	[88]
Ethical and Regulatory	Algorithmic bias perpetuating health disparities; patient privacy and data security concerns; unclear regulatory pathways for adaptive AI; inadequate frameworks for continuously learning systems	Conduct systematic bias audits across demographic groups; implement privacy-preserving techniques (federated learning, differential privacy); engage regulatory bodies early in development; establish governance for adaptive systems	Bias audit protocols; diverse training data curation; privacy impact assessments; regulatory consultation processes; adaptive system governance committees	[30,32,53,77,82]
Sustainability and Scalability	High initial implementation costs; ongoing maintenance and model updating requirements; model performance degradation over time; scalability challenges for multi-center deployment	Develop business cases demonstrating value through reduced complications/ costs; establish model monitoring and retraining protocols; plan for continuous updates; create multi-center governance frameworks	ROI analysis tools; performance monitoring dashboards; retraining budgets; multi-center coordination mechanisms; scalability assessment frameworks	[85]
Data Security	Vulnerability to cyberattacks during EHR-AI data exchange; need for quantum-safe encryption; protection of sensitive health data in AI training/operation	Implement quantum cryptography with blockchain for EHR security; deploy attribute-based access control for PHR protection; utilize advanced encryption for data transmission; conduct regular security audits	Quantum-safe cryptographic infrastructure; blockchain integration for data integrity; access control systems; security monitoring protocols; incident response plans	[30,32]

Note: Operationalization of HITL-AI requires coordinated attention to technical, workflow, trust, ethical, regulatory, sustainability, and security dimensions. Multi-center implementations require additional coordination mechanisms, standardized protocols, and governance frameworks to ensure consistent implementation quality across sites.

loop AI have been associated with earlier detection of clinical deterioration, more timely interventions, and reduced mortality rates in hospital settings [8,27]. The combination of AI vigilance over continuous data streams and human clinical judgment about intervention appropriateness appears optimal for patient safety.

Healthcare efficiency and resource utilization also benefit from human in the loop AI implementation. Radiology departments using AI assistance have reported increased throughput and reduced interpretation times without compromising diagnostic quality [37]. Clinical decision support systems have been shown to improve adherence to cost effective care pathways and reduce unnecessary testing [19,75]. Remote patient monitoring programs leveraging human in the loop AI enable care teams to manage larger patient panels while focusing attention on those most likely to benefit from intervention [20,21]. These efficiency gains have potential to address healthcare workforce shortages and improve access to care. However, efficiency benefits must be balanced against implementation costs, which can be substantial including initial development/licensing, EHR integration, training, and ongoing maintenance. Successful implementations typically demonstrate clear value propositions through reduced complications, decreased length of stay, improved chronic disease management, or enhanced diagnostic accuracy that justify initial and ongoing investments in AI systems and human oversight infrastructure [20].

8. Ethical considerations and regulatory landscape

8.1. Ethical principles and considerations

Ethical considerations surrounding human in the loop AI in healthcare are multifaceted and require careful attention. Algorithmic bias represents a significant concern, as AI systems trained on historical data may perpetuate or amplify existing healthcare disparities if training data inadequately represent diverse populations or reflect biased clinical practices [65,76]. Human in the loop approaches partially mitigate this risk by enabling clinicians to identify and correct biased recommendations, but systematic bias monitoring and mitigation strategies remain essential [76]. Transparency and explainability constitute another ethical imperative, as patients and clinicians have legitimate interests in understanding how AI systems generate recommendations that influence healthcare decisions [53,77,78]. The black box nature of many machine learning algorithms, particularly deep neural networks, creates challenges for explainability, though recent advances in interpretable AI and explainable AI techniques offer promising approaches for making algorithmic reasoning more transparent [79].

Informed consent and patient autonomy require careful consideration in human in the loop AI contexts. Patients should be informed when AI systems contribute to their care and should have opportunities to understand how these systems work and to express preferences about their use [80]. However, the complexity of modern AI systems makes truly informed consent challenging to achieve, and healthcare organizations must balance detailed technical explanations against the need for understandable communication. Privacy and data security represent critical ethical concerns, particularly as HITL-AI systems require bidirectional data flow between AI algorithms and human operators. Beyond traditional encryption approaches, emerging quantum cryptographic techniques integrated with blockchain technology offer enhanced security for electronic health record exchange, achieving high data privacy rates and network security while enabling rapid encoding/decoding for clinical workflows [30,32]. Attribute-based access control mechanisms provide additional protection for personal health records in cloud environments, ensuring that only authorized clinicians can access patient data flagged by AI systems [32]. Human in the loop approaches involve human access to patient data flagged by algorithms, creating additional privacy considerations that must be managed through appropriate data governance, access controls, and privacy preserving technologies [13]. Liability and accountability questions arise when AI systems contribute to adverse outcomes, with ongoing debates about whether responsibility resides with clinicians who accept AI recommendations, healthcare organizations that deploy systems, AI developers, or some combination thereof. This liability allocation challenge is particularly acute when clinicians override AI recommendations, if the override prevents harm, liability is clear, but if overriding a correct AI recommendation leads to adverse outcomes, liability frameworks remain ambiguous [81].

Table 4 presents key ethical principles governing human-in-the-loop AI in healthcare, including beneficence, justice, autonomy, transparency, privacy, and liability, with associated considerations and implementation strategies addressing both established requirements and emerging challenges.

8.2. Regulatory landscape and gaps

The regulatory landscape for human in the loop AI in healthcare continues to evolve as regulatory agencies worldwide develop frameworks for AI oversight. In the United States, the Food and Drug Administration has established pathways for regulating AI-based medical devices, with different levels of scrutiny depending on risk classification and degree of human involvement in decision making [53,77,82]. Human in the loop systems that provide decision support rather than autonomous recommendations may face less stringent regulatory requirements than fully automated systems, though this distinction remains under active discussion. In Europe, the Medical Device Regulation provides the regulatory framework for AI systems used in healthcare, with recent proposals for specific AI regulation adding additional requirements around transparency, accountability, and human oversight [83]. These regulatory developments generally favor human in the loop approaches by requiring meaningful human control over high-risk AI applications, though specific requirements continue to evolve as regulators gain experience with diverse AI systems in healthcare contexts.

However, critical regulatory gaps persist, particularly for adaptive HITL systems that evolve through continuous learning from human feedback. Current regulatory frameworks generally assume static AI models that remain unchanged post-deployment, yet active learning HITL systems inherently modify their algorithms based on clinician decisions, patient outcomes, and operational feedback [84]. This creates regulatory uncertainty: (1) When does algorithm modification require re-submission for regulatory approval? (2) How should regulators assess safety and efficacy of systems that evolve over time? (3) What governance mechanisms ensure adaptive systems do not drift toward unsafe or biased behaviors? (4) How should post-market surveillance function for continuously learning systems where performance may change between assessments? These questions remain inadequately addressed in current regulatory frameworks, creating barriers to deployment of next-generation adaptive HITL-AI systems and requiring urgent regulatory modernization [82,84].

9. Challenges and limitations

Despite the promise of human in the loop AI in healthcare, numerous challenges limit its widespread adoption and effectiveness. Technical challenges include integration with existing health information technology systems, which often lack standardized application programming interfaces or data formats necessary for seamless AI integration. Data quality issues pose additional obstacles, as AI systems require high quality, complete, and accurately labeled data for optimal performance, yet healthcare data are often incomplete, inconsistent, or contain errors. Computational resource requirements for training and deploying sophisticated AI models can be substantial, requiring investments in computing infrastructure that may be prohibitive for smaller healthcare organizations. Model maintenance and updating present ongoing challenges, as AI systems may require retraining as patient populations change, medical knowledge evolves, or data distributions shift,

Table 4
Key ethical principles and considerations for human in the loop AI in healthcare.

Ethical Principle	Key Considerations	Implementation Strategies	Emerging Challenges	References
Beneficence and Nonmaleficence	Ensure AI systems improve patient outcomes without causing harm through errors or unintended consequences	Rigorous validation before deployment; continuous performance monitoring; clear escalation pathways for concerning outputs; adverse event reporting systems	Balancing rapid AI evolution with adequate safety testing; managing harm from AI-human interaction failures	[43,82]
Justice and Equity	Prevent algorithmic bias; ensure equitable access to AI benefits across diverse patient populations	Use diverse training data representing all demographic groups; conduct systematic bias audits; monitor performance across populations; provide human oversight to correct biased outputs; proactive outreach to underserved communities	Addressing intersectional bias across multiple demographic dimensions; ensuring equitable access to HITL-AI benefits in resource-limited settings	[65,76]
Autonomy and Informed Consent	Respect patient preferences; provide information about AI involvement in their care	Develop clear, accessible communication about AI use; allow patient choice in AI involvement when feasible; ensure human clinicians remain responsible for final decisions; document patient consent	Achieving truly informed consent given AI complexity; respecting patient autonomy when declining AI involvement may limit care quality	[80,81]
Transparency and Explainability	Make AI reasoning understandable to clinicians and patients to enable appropriate trust and oversight	Implement explainable AI (XAI) techniques providing interpretable outputs; provide clear documentation of system capabilities and limitations; train users in AI interpretation; maintain algorithmic transparency	Balancing model complexity/performance with interpretability; explaining adaptive algorithms that evolve over time	[53,77,78,82]
Privacy and Confidentiality	Protect patient data in AI training/operation; secure EHR-AI data exchange	Apply strong data governance; use privacy-preserving technologies (federated learning, differential privacy); implement role-based access controls; conduct privacy impact assessments; deploy quantum cryptography with blockchain for EHR security; implement attribute-based access control for PHR cloud storage	Securing data against quantum computing threats; protecting privacy in multi-center federated learning; managing privacy in continuously learning adaptive systems	[30,32,78,82]
Liability and Accountability	Clarify responsibility when AI contributes to clinical decisions; address liability for clinician overrides of AI recommendations	Establish institutional liability frameworks explicitly allocating responsibility; document AI role as decision support not autonomous decision-maker; maintain audit trails of AI recommendations and clinician actions; obtain malpractice insurance coverage addressing AI-assisted care; create review processes for adverse events involving AI	Ambiguous liability when clinicians override correct AI recommendations; responsibility allocation for harm from AI-human collaboration failures; liability gaps for continuously learning adaptive systems	[81,88]

Note: Ethical implementation of HITL-AI requires proactive attention to all principles, with particular emphasis on emerging challenges around liability allocation, quantum-safe privacy protection, and oversight of adaptive learning systems.

necessitating continuous investment in model monitoring and improvement.

Workflow integration challenges represent significant barriers to successful human in the loop AI implementation. Healthcare workflows are complex, context dependent, and often resistant to change, making introduction of AI systems disruptive if not carefully managed. The cognitive load imposed by AI systems can be substantial, as clinicians must not only perform their traditional tasks but also review AI outputs, assess their validity, and decide whether to accept, modify, or reject recommendations. Time constraints in busy clinical environments limit the attention healthcare professionals can devote to AI interaction, potentially leading to superficial engagement with AI outputs or reflexive acceptance without adequate critical evaluation. Alert fatigue remains problematic even in human in the loop systems, as poorly calibrated AI may generate excessive alerts that clinicians begin to ignore, undermining the intended safety and quality benefits.

Trust and acceptance challenges affect human in the loop AI adoption and effectiveness. Some clinicians express skepticism about AI capabilities, concerned about algorithmic errors, liability implications, or threats to professional autonomy. Lack of transparency in many AI systems exacerbates trust issues, as clinicians struggle to understand how systems generate recommendations and therefore have difficulty assessing when to trust AI outputs. Conversely, automation bias, wherein users over rely on automated recommendations without adequate critical evaluation, poses risks if clinicians develop excessive trust in AI systems and fail to identify erroneous outputs. Building appropriate trust, neither excessive skepticism nor blind acceptance, requires explainable AI designs, transparent communication about system performance, opportunities for users to develop experience with systems, and organizational cultures that support critical engagement with AI recommendations.

Economic and sustainability challenges also constrain human in the loop AI implementation. Initial costs for AI system development or licensing, validation studies, EHR integration, infrastructure upgrades, and comprehensive training can be substantial, often requiring significant capital investment before benefits materialize. Ongoing costs for system maintenance, model retraining, technical support, and human oversight add to the total cost of ownership. Demonstrating return on investment can be challenging, particularly when benefits accrue over long time horizons or manifest as avoided adverse events rather than measurable cost savings. Small healthcare organizations and resource limited settings may lack the financial resources, technical expertise, or supporting infrastructure necessary for successful AI implementation, potentially exacerbating healthcare disparities between well-resourced and under resourced settings [33]. Sustainability concerns arise as AI systems require ongoing maintenance and updating to remain effective, yet funding and attention may wane after initial implementation enthusiasm subsides [85].

10. Future directions and research priorities

Future research and development in human in the loop AI for healthcare should address several critical priorities. Standardized implementation frameworks are needed to guide healthcare organizations in deploying human in the loop AI systems effectively. Current implementations often reflect ad hoc approaches developed locally without benefit of evidence based best practices or lessons learned from other implementations. Development of standardized frameworks covering assessment of organizational readiness, selection of appropriate collaboration models matched to risk contexts, workflow analysis and redesign, EHR integration protocols, training program development, change management strategies, performance monitoring approaches, and continuous improvement processes would accelerate adoption and improve implementation success rates [29,34]. These frameworks should be flexible enough to accommodate diverse healthcare settings and use cases while providing sufficient structure to prevent common implementation pitfalls. Multi-center and cross-country comparative studies are particularly needed to identify which implementation strategies generalize across contexts and which require local adaptation based on healthcare system structure, regulatory environment, and organizational culture [33].

Evaluation methodologies for human in the loop AI require further development and refinement. Traditional approaches to evaluating medical technologies may not adequately capture the unique characteristics of human AI collaboration, which involves complex interactions between algorithmic performance, human decision making, and organizational contexts. Evaluation frameworks should assess not only technical performance metrics such as sensitivity and specificity but also measures of human AI collaboration quality, clinical workflow integration, user acceptance and satisfaction, impact on clinical outcomes, cost effectiveness, equity implications, and unintended consequences [21,43]. Longitudinal studies examining sustained performance and outcomes over extended time periods are particularly needed, as many current evaluations focus on short term pilot implementations that may not reflect long term effectiveness or sustainability.

Explainable and interpretable AI represents a critical research frontier for human in the loop applications. Current black box AI models, while often achieving high predictive accuracy, provide limited insight into their reasoning processes, making it difficult for clinicians to understand why systems generate particular recommendations [70]. Advances in explainable AI (XAI) techniques, including attention mechanisms, saliency maps, counterfactual explanations, and inherently interpretable models, show promise for making AI reasoning more transparent [79]. However, further research is needed to determine which explanation approaches are most useful for different clinical contexts, how explanations should be presented to optimize comprehension and appropriate trust, and whether enhanced explainability improves clinical decision making and patient outcomes. The optimal balance between model complexity and interpretability remains an open question requiring empirical investigation [74].

Bias detection and mitigation strategies require ongoing research attention to ensure human in the loop AI systems promote rather than undermine healthcare equity. While human oversight can identify and correct individual instances of biased AI outputs, systematic approaches to bias detection, monitoring, and mitigation are necessary to address bias at scale [76]. Research should examine how different sources of bias, including biased training data, biased algorithm design choices, and biased human oversight, interact to influence overall system fairness. Development and validation of bias auditing tools, fairness aware machine learning approaches, and diverse dataset curation strategies represent important research priorities [65]. Additionally, studies examining how human in the loop approaches compare to fully automated systems in terms of bias propagation and mitigation would inform optimal design choices for promoting equity.

Optimal human AI interaction designs require deeper investigation through human factors research and user centered design studies. Questions about how to present AI outputs to maximize appropriate utilization, how much explanation to provide, how to calibrate user trust, and how to design interfaces that support rather than impede clinical workflows remain incompletely answered. Research employing methods from human computer interaction, cognitive psychology, and clinical informatics can provide insights into effective interaction designs. Comparative studies evaluating different interface designs, interaction modalities, and collaboration frameworks in actual clinical settings would generate valuable evidence to guide system design [74]. Attention to diverse user needs and preferences is essential, as optimal designs may vary across clinical specialties, experience levels, and cultural contexts.

Post-quantum cryptographic integration represents an emerging priority for HITL-AI security. As quantum computing capabilities advance, current encryption methods protecting healthcare data exchange between AI systems and human operators face obsolescence [30]. Integration of quantum-resistant cryptographic algorithms with blockchain technology for EHR security, combined with federated learning approaches that enable model training without centralizing sensitive data, will be essential for maintaining patient privacy and data security in next-generation HITL-AI systems [30,32]. Research priorities include development and validation of quantum-safe cryptographic protocols optimized for clinical workflow requirements, evaluation of computational overhead introduced by advanced encryption, and establishment of standards for quantum-resistant healthcare AI security.

AI-driven adaptive HITL frameworks represent a transformative frontier where HITL systems not only leverage human feedback to improve algorithms but also adapt collaboration strategies based on context, user expertise, and task characteristics [10,29]. For instance, adaptive systems might dynamically adjust the collaboration model (verification vs. augmentation vs. human-in-command) based on clinician experience level, case complexity, or system confidence in its recommendations. Research priorities include: (1) development of meta-learning algorithms that optimize collaboration strategies across contexts, (2) establishment of governance frameworks ensuring adaptive systems evolve safely without regulatory oversight gaps, (3) creation of performance monitoring approaches for systems with dynamic collaboration models, and (4) investigation of how adaptive HITL systems impact clinician skill development and professional autonomy over time [29].

Scalability considerations for multi-center and international HITL-AI deployment require systematic investigation. While single-site implementations provide proof-of-concept, widespread healthcare transformation requires scalability across diverse settings with varying EHR systems, regulatory environments, clinical workflows, and organizational cultures [33]. Research priorities include: (1) development of interoperability standards enabling HITL-AI portability across EHR platforms, (2) establishment of federated learning frameworks enabling model training across institutions without compromising patient privacy, (3) investigation of how HITL-AI performance generalizes across geographic regions with different patient populations and clinical practice patterns, (4) creation of governance models for multi-institutional HITL-AI systems addressing questions of model ownership, liability allocation, and decision authority, and (5) evaluation of implementation strategies that successfully navigate diverse regulatory environments and cultural contexts [30,63]. Multi-center implementations offer opportunities to aggregate larger datasets for model training while requiring careful attention to data harmonization, privacy protection, and equitable benefit distribution across participating institutions.

11. Limitations of the review

This narrative review has several limitations that warrant acknowledgment. As a narrative rather than systematic review, this study does not follow PRISMA guidelines, does not employ standardized data extraction tables, does not provide quantitative synthesis or meta-analysis of outcomes, does not include performance evaluation benchmarking, and does not present a PRISMA flow diagram. These methodological choices reflect the review's objectives of conceptual synthesis and actionable insight generation across highly heterogeneous HITL-AI implementations rather than quantitative effect size estimation. The search strategy and study selection processes, while comprehensive, were not exhaustive and may have missed relevant studies. The lack of standardized quality assessment and risk of bias evaluation common in systematic reviews means that the included studies vary in methodological rigor, and findings should be interpreted accordingly. The rapidly evolving nature of AI technology means that some included studies may already be outdated, and emerging developments may not be fully captured. The review predominantly focuses on literature from high income countries with advanced health information technology infrastructure, potentially limiting generalizability to resource limited settings with different technological capabilities and healthcare delivery models.

Publication bias may affect the review findings, as studies demonstrating positive results are more likely to be published than those showing neutral or negative findings. Many implementations of human in the loop AI occur in proprietary commercial settings and may not be publicly documented in peer reviewed literature, creating a gap between research evidence and real-world practice. The heterogeneity of AI systems, clinical applications, implementation contexts, and evaluation methodologies across studies makes direct comparisons challenging and limits the ability to draw definitive conclusions about optimal approaches. Long term outcomes and sustainability of human in the loop AI implementations remain understudied, as most published evaluations examine relatively short time horizons. These limitations suggest that findings should be interpreted cautiously and that ongoing research is essential to build a more comprehensive and robust evidence base.

12. Conclusion

Human in the loop artificial intelligence represents a pragmatic and ethically sound approach to integrating AI into healthcare delivery and research. By combining the pattern recognition and data processing capabilities of machine learning with the contextual understanding, ethical reasoning, and clinical judgment of healthcare professionals, human in the loop systems can enhance healthcare quality, safety, and efficiency while mitigating risks associated with fully automated AI. Evidence from diverse applications including diagnostic imaging, clinical decision support, patient monitoring, and healthcare research demonstrates meaningful benefits including improved diagnostic accuracy, reduced medical errors, enhanced patient safety, and accelerated research processes. However, successful implementation requires careful attention to technical integration, workflow design, training and change management, ethical considerations, and regulatory compliance.

Significant challenges persist in realizing the full potential of human in the loop AI in healthcare. Technical hurdles including system integration, data quality, and computational requirements must be addressed through investments in health information technology infrastructure and data governance. Workflow integration challenges require user centered design approaches that minimize cognitive load and align with clinical processes. Building appropriate trust and acceptance necessitates transparent communication about AI capabilities and limitations, explainable AI designs, and organizational cultures that support critical engagement with AI recommendations. Economic constraints and sustainability concerns must be addressed through clear value demonstrations and sustainable funding models. Ethical and regulatory considerations around bias, privacy, accountability, and transparency require ongoing attention from healthcare organizations, technology developers, policymakers, and regulatory bodies.

Future progress in human in the loop AI for healthcare depends on coordinated efforts across multiple domains. Researchers must develop standardized implementation frameworks, refined evaluation methodologies, more explainable AI techniques, effective bias mitigation strategies, optimized human AI interaction designs, quantum-safe cryptographic security protocols, AI-driven adaptive HITL architectures, and multi-center scalability frameworks. Healthcare organizations must invest in supporting infrastructure, training programs, and change management processes necessary for successful implementation. Technology developers must prioritize explainability, fairness, user centered design, and quantum-resistant security in AI system development. Policymakers and regulators must establish clear guidelines that promote innovation while ensuring patient safety and ethical AI use, particularly addressing regulatory gaps for adaptive learning systems. Patients and the public must be engaged in conversations about AI in healthcare to ensure that implementations reflect societal values and priorities. Through these collective efforts, human in the loop AI can fulfill its promise of transforming healthcare for the benefit of patients, clinicians, researchers, and health systems.

CRediT authorship contribution statement

David B. Olawade: Writing – review & editing, Writing – original draft, Project administration, Methodology, Investigation, Formal analysis, Conceptualization. **Shamiul Bashir Plabon:** Writing – review & editing, Writing – original draft, Methodology, Investigation, Formal analysis. **Adeyinka Ojo:** Writing – review & editing, Writing – original draft, Methodology, Investigation. **Muyiwa Ademola Ogunbona:** Writing – review & editing, Writing – original draft, Methodology, Investigation. **Babajide David Makanjuola:** Writing – review & editing, Writing – original draft, Methodology, Investigation. **Omobolaji Rosemary Olasilola:** Writing – review & editing, Writing – original draft, Methodology, Investigation.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] K.P. Reddy, M. Satish, A. Prakash, S.M. Babu, P.P. Kumar, B.S. Devi, Machine learning revolution in early disease detection for healthcare: Advancements, challenges, and future prospects. In: 2023 IEEE 5th International Conference on Cybernetics, Cognition and Machine Learning Applications (ICCCMLA), Hamburg, Germany. IEEE, 2023, pp. 638-643. https://doi.org/10.1109/ICCCMLA58983.2023.10346963.

[2] D. van de Sande, M.E. van Genderen, J. Huiskens, D. Gommers, J. van Bommel, Moving from bytes to bedside: a systematic review on the use of artificial intelligence in the intensive care unit, Intensive Care Med. 47 (7) (2021) 750–760, https://doi.org/10.1007/s00134-021-06446-7.

[3] P. Hadweh, A. Niset, M. Salvagno, M. Al Barajraji, S. El Hadwe, F.S. Taccone, S. Barrit, Machine learning and artificial intelligence in intensive care medicine: critical recalibrations from rule-based systems to frontier models, J. Clin. Med. 14 (12) (2025) 4026, https://doi.org/10.3390/jcm14124026.

[4] M. Saqib, M. Iftikhar, F. Neha, F. Karishma, H. Mumtaz, Artificial intelligence in critical illness and its impact on patient care: a comprehensive review, Front. Med. 10 (2023) 1176192, https://doi.org/10.3389/fmed.2023.1176192.

[5] B. Aldosari, A. Alanazi, Pitfalls of artificial intelligence in medicine, Stud. Health Technol. Inform. 316 (2024) 554–555, https://doi.org/10.3233/SHTI240474.

[6] T. Pham, Ethical and legal considerations in healthcare AI: innovation and policy for safe and fair use, R. Soc. Open Sci. 12 (5) (2025) 241873, https://doi.org/10.1098/rsos.241873.

[7] E.B. Weiner, I. Dankwa-Mullan, W.A. Nelson, S. Hassanpour, Ethical challenges and evolving strategies in the integration of artificial intelligence into clinical practice, PLOS Digital Health 4 (4) (2025) e0000810, https://doi.org/10.1371/journal.pdig.0000810.

[8] E. Mosqueira-Rey, E. Hernández-Pereira, D. Alonso-Ríos, J. Bobes-Bascarán, Á. Fernández-Leal, Human-in-the-loop machine learning: a state of the art, Artif. Intell. Rev. 56 (4) (2023) 3005–3054, https://doi.org/10.1007/s10462-022-10246-w.

[9] I.G. Cohen, B. Babic, S. Gerke, Q. Xia, T. Evgeniou, K. Wertenbroch, How AI can learn from the law: putting humans in the loop only on appeal, npj Digital Med. 6 (1):160 (2023), https://doi.org/10.1038/s41746-023-00906-8.

[10] H. Yuan, L. Kang, Y. Li, Z. Fan, Human-in-the-loop machine learning for healthcare: current progress and future opportunities in electronic health records, Med. Adv. 2 (3) (2024) 318–322, https://doi.org/10.1002/med4.70.

[11] C. Reverberi, T. Rigon, A. Solari, C. Hassan, P. Cherubini, A. Cherubini, Experimental evidence of effective human–AI collaboration in medical decision-making, Sci. Rep. 12 (1) (2022) 14952, https://doi.org/10.1038/s41598-022-18751-2.

[12] M. Chen, Y. Wang, Q. Wang, J. Shi, H. Wang, Z. Ye, P. Xue, Y. Qiao, Impact of human and artificial intelligence collaboration on workload reduction in medical

image interpretation, npj Digital Med. 7 (1) (2024) 349, https://doi.org/10.1038/s41746-024-01328-w.
[13] A.C. Bodén, J. Molin, S. Garvin, R.A. West, C. Lundström, D. Treanor, The human-in-the-loop: an evaluation of pathologists' interaction with artificial intelligence in clinical practice, Histopathology 79 (2) (2021) 210–218, https://doi.org/10.1111/his.14356.
[14] S. Salloch, A. Eriksen, What are humans doing in the loop? Co-reasoning and practical judgment when using machine learning-driven decision aids, Am. J. Bioeth. 24 (9) (2024) 67–78, https://doi.org/10.1080/15265161.2024.2353800.
[15] B.N. Patel, L. Rosenberg, G. Willcox, D. Baltaxe, M. Lyons, J. Irvin, P. Rajpurkar, T. Amrhein, R. Gupta, S. Halabi, C. Langlotz, Human–machine partnership with artificial intelligence for chest radiograph diagnosis, npj Digital Med. 2 (1) (2019) 111, https://doi.org/10.1038/s41746-019-0189-7.
[16] B.O. Botwe, W.K. Antwi, S. Arkoh, T.N. Akudjedu, Radiographers' perspectives on the emerging integration of artificial intelligence into diagnostic imaging: the Ghana study, J. Med. Radiat. Sci. 68 (3) (2021) 260–268, https://doi.org/10.1002/jmrs.460.
[17] D.B. Larson, H. Harvey, D.L. Rubin, N. Irani, J.R. Tse, C.P. Langlotz, Regulatory frameworks for development and evaluation of artificial intelligence–based diagnostic imaging algorithms: summary and recommendations, J. Am. Coll. Radiol. 18 (3) (2021) 413–424, https://doi.org/10.1016/j.jacr.2020.09.060.
[18] T. Boeken, J. Feydy, A. Lecler, P. Soyer, A. Feydy, M. Barat, L. Duron, Artificial intelligence in diagnostic and interventional radiology: where are we now? Diagn. Interv. Imaging 104 (1) (2023) 1–5, https://doi.org/10.1016/j.diii.2022.11.004.
[19] C. Giordano, M. Brennan, B. Mohamed, P. Rashidi, F. Modave, P. Tighe, Accessing artificial intelligence for clinical decision-making, Front. Digital Health 3 (2021) 645232, https://doi.org/10.3389/fdgth.2021.645232.
[20] S. Secinaro, D. Calandra, A. Secinaro, V. Muthurangu, P. Biancone, The role of artificial intelligence in healthcare: a structured literature review, BMC Med. Inf. Decis. Making 21 (1) (2021) 125, https://doi.org/10.1186/s12911-021-01488-9.
[21] B. Vasey, M. Nagendran, B. Campbell, D.A. Clifton, G.S. Collins, S. Denaxas, A. K. Denniston, L. Faes, B. Geerts, M. Ibrahim, X. Liu, Reporting guideline for the early stage clinical evaluation of decision support systems driven by artificial intelligence, DECIDE-AI. Bmj. 377 (2022) e070904, https://doi.org/10.1136/bmj-2022-070904.
[22] M.A. Sellwood, M. Ahmed, M.H. Segler, N. Brown, Artificial intelligence in drug discovery, Future Med. Chem. 10 (17) (2018) 2025–2028, https://doi.org/10.4155/fmc-2018-0212.
[23] S.K. Saikin, C. Kreisbeck, D. Sheberla, J.S. Becker, A. Aspuru-Guzik, Closed-loop discovery platform integration is needed for artificial intelligence to make an impact in drug discovery, Expert Opin. Drug Discov. 14 (1) (2019) 1–4, https://doi.org/10.1080/17460441.2019.1546690.
[24] P.C. Tiwari, R. Pal, M.J. Chaudhary, R. Nath, Artificial intelligence revolutionizing drug development: Exploring opportunities and challenges, Drug Dev. Res. 84 (8) (2023) 1652–1663, https://doi.org/10.1002/ddr.22115.
[25] Z. Jeddi, A. Bohr, Remote patient monitoring using artificial intelligence. In: Artificial intelligence in healthcare, Academic Press, 2020, pp. 203-234. https://doi.org/10.1016/B978-0-12-818438-7.00009-5.
[26] S. Bakken, AI in health: keeping the human in the loop, J. Am. Med. Inform. Assoc. 30 (7) (2023) 1225–1226, https://doi.org/10.1093/jamia/ocad091.
[27] T. Shaik, X. Tao, N. Higgins, L. Li, R. Gururajan, X. Zhou, U.R. Acharya, Remote patient monitoring using artificial intelligence: current state, applications, and challenges, Wiley Interdiscip. Rev.: Data Min. Knowl. Discovery 13 (2) (2023) e1485.
[28] T. Davenport, R. Kalakota, The potential for artificial intelligence in healthcare, Future Healthcare Journal. 6 (2) (2019) 94–98, https://doi.org/10.7861/futurehosp.6-2-94.
[29] M. Puerta-Beldarrain, O. Gomez-Carmona, R. Sanchez-Corcuera, D. Casado-Mansilla, D. Lopez-de-Ipina, L. Chen, A multifaceted vision of the Human-AI collaboration: a comprehensive review, IEEE Access 13 (2025) 29375–29405, https://doi.org/10.1109/ACCESS.2025.3536095.
[30] P. Chinnasamy, L. Dhavamani, R.K. Ayyasamy, B.S. Kumar, A. Kiran, QuantumBlock health records: enhancing healthcare data security with quantum cryptography and blockchain technology, Clust. Comput. 28 (7) (2025) 474.
[31] D.K. Moorthy, P. Chinnasamy, P. Nagaraj, Optimization enabled ResNet features with transfer learning for Alzheimer's Disease Detection, Comput. Biol. Chem. 108613 (2025).
[32] Chinnasamy P, Siva Padmini P, Gogula S, Ramprathap K, Kumareshan N, Alagarsamy M. Securing sensitive health data: A creative solution for personal health record protection in cloud environment using attribute-based access control. In: AIP Conference Proceedings. Vol. 3300, No. 1. AIP Publishing LLC; 2025: 020100.
[33] B.Z. Wubineh, F.G. Deriba, M.M. Woldeyohannis, Exploring the opportunities and challenges of implementing artificial intelligence in healthcare: a systematic literature review, Urol. Oncol. 42 (3) (2024) 48–56, https://doi.org/10.1016/j.urolonc.2023.11.019.
[34] E. Coiera, S. Liu, Evidence synthesis, digital scribes, and translational challenges for artificial intelligence in healthcare, Cell Rep. Med. 3 (12) (2022) 100860, https://doi.org/10.1016/j.xcrm.2022.100860.
[35] B.N. Green, C.D. Johnson, A. Adams, Writing narrative literature reviews for peer-reviewed journals: secrets of the trade, J. Chiropr. Med. 5 (3) (2006) 101–117.
[36] S.I. Lemos, F.A. Ferreira, C. Zopounidis, E. Galariotis, N.C. Ferreira, Artificial intelligence and change management in small and medium-sized enterprises: an analysis of dynamics within adaptation initiatives, Ann. Oper. Res. 353 (1) (2025) 197–223, https://doi.org/10.1007/s10479-022-05159-4.
[37] A. Hosny, C. Parmar, J. Quackenbush, L.H. Schwartz, H.J. Aerts, Artificial intelligence in radiology, Nat. Rev. Cancer 18 (8) (2018) 500–510, https://doi.org/10.1038/s41568-018-0016-5.
[38] A. Pirrera, D. Giansanti, Artificial Intelligence and Breast Cancer Screening: Bibliometric Analysis, Diagnostics 13 (14) (2023) 2367.
[39] R.J. Chen, J.J. Wang, D.F. Williamson, T.Y. Chen, J. Lipkova, M.Y. Lu, S. Sahai, F. Mahmood, Algorithmic fairness in artificial intelligence for medicine and healthcare, Nat. Biomed. Eng. 7 (6) (2023) 719–742, https://doi.org/10.1038/s41551-023-01056-8.
[40] W.L. Bi, A. Hosny, M.B. Schabath, M.L/ Giger, N.J. Birkbak, A. Mehrtash, T. Allison, O. Arnaout, C. Abbosh, I.F. Dunn, R.H. Mak, Artificial intelligence in cancer imaging: clinical challenges and applications, CA: a cancer journal for clinicians. 69(2) (2019) 127-157. https://doi.org/10.3322/caac.21552.
[41] G. Improta, V. Mazzella, D. Vecchione, S. Santini, M. Triassi, Fuzzy logic–based clinical decision support system for the evaluation of renal function in post-Transplant patients, J. Eval. Clin. Pract. 26 (4) (2020) 1224–1234, https://doi.org/10.1111/jep.13302.
[42] M.A. Musen, B. Middleton, R.A. Greenes, Clinical decision-support systems, in: Biomedical Informatics: Computer Applications in Health Care and Biomedicine, Springer International Publishing, Cham, 2021, pp. 795–840, https://doi.org/10.1007/978-3-030-58721-5_24.
[43] S. Kumar, S. Datta, V. Singh, D. Datta, S.K. Singh, R. Sharma, Applications, challenges, and future directions of human-in-the-loop learning, IEEE Access 12 (2024) 75735–75760, https://doi.org/10.1109/ACCESS.2024.3401547.
[44] N. Villafuerte, S. Manzano, P. Ayala, M.V. García, Artificial intelligence in virtual telemedicine triage: a respiratory infection diagnosis tool with electronic measuring device, Future Internet. 15 (7) (2023) 227, https://doi.org/10.3390/fi15070227.
[45] R.A. El Arab, O.A. Al Moosa, The role of AI in emergency department triage: an integrative systematic review, Intensive and Critical Care Nursing. 89 (2025) 104058, https://doi.org/10.1016/j.iccn.2025.104058.
[46] M.J. McGrath, A. Duenser, J. Lacey, C. Paris, Collaborative human-AI trust (CHAI-T): a process framework for active management of trust in human-AI collaboration, Computers in Human Behavior: Artificial Humans. (2025) 100200, https://doi.org/10.1016/j.chbah.2025.100200.
[47] B. Gholami, W.M. Haddad, J.M. Bailey, AI in the ICU: In the intensive care unit, artificial intelligence can keep watch, IEEE Spectr. 55 (10) (2018) 31–35, https://doi.org/10.1109/MSPEC.2018.8482421.
[48] M. Mollura, L.W.H. Lehman, R.G. Mark, R. Barbieri, A novel artificial intelligence based intensive care unit monitoring system: using physiological waveforms to identify sepsis, Phil. Trans. R. Soc. A 379 (2212) (2021) 20200252, https://doi.org/10.1098/rsta.2020.0252.
[49] F. Li, S. Wang, Z. Gao, M. Qing, S. Pan, Y. Liu, C. Hu, Harnessing artificial intelligence in sepsis care: advances in early detection, personalized treatment, and real-time monitoring, Front. Med. 11 (2025) 1510792, https://doi.org/10.3389/fmed.2024.1510792.
[50] W.T. Tran, K. Jerzak, F.I. Lu, J. Klein, S. Tabbarah, A. Lagree, T. Wu, I. Rosado-Mendez, E. Law, K. Saednia, A. Sadeghi-Naini, Personalized breast cancer treatments using artificial intelligence in radiomics and pathomics, Journal of Medical Imaging and Radiation Sciences. 50 (4) (2019) S32–S41, https://doi.org/10.1016/j.jmir.2019.07.010.
[51] N. Biswas, S. Chakrabarti, Artificial intelligence (AI)-based systems biology approaches in multi-omics data analysis of cancer, Front. Oncol. 10 (2020) 588221, https://doi.org/10.3389/fonc.2020.588221.
[52] Z. Dlamini, F.Z. Francies, R. Hull, R. Marima, Artificial intelligence (AI) and big data in cancer and precision oncology, Comput. Struct. Biotechnol. J. 18 (2020) 2300–2311, https://doi.org/10.1016/j.csbj.2020.08.019.
[53] S.A. Alowais, S.S. Alghamdi, N. Alsuhebany, T. Alqahtani, A.I. Alshaya, S. N. Almohareb, A. Aldairem, M. Alrashed, K. Bin Saleh, H.A. Badreldin, M.S. Al Yami, Revolutionizing healthcare: the role of artificial intelligence in clinical practice, BMC Med. Educ. 23 (1) (2023) 689, https://doi.org/10.1186/s12909-023-04698-z.
[54] C.O. Fernandes, S. Miles, C.J.P. Lucena, D. Cowan, Artificial Intelligence Technologies for Coping with Alarm Fatigue in Hospital Environments because of Sensory Overload: Algorithm Development and Validation, J. Med. Internet Res. 21 (11) (2019) e15406, https://doi.org/10.2196/15406.
[55] J. Misurac, L.A. Knake, J.M. Blum, The effect of Ambient Artificial Intelligence Notes on Provider Burnout, Appl. Clin. Inf. 16 (2) (2025) 252–258, https://doi.org/10.1055/a-2461-4576.
[56] P.M. Patel, M. Green, J. Tram, E. Wang, M.Z. Murphy, A. Abd-Elsayed, K. Chakravarthy, Beyond the pain management clinic: the role of AI-integrated remote patient monitoring in chronic disease management–a narrative review, J. Pain Res. (2024) 4223–4237, https://doi.org/10.2147/JPR.S494238.
[57] S. Spillias, P. Tuohy, M. Andreotta, R. Annand-Jones, F. Boschetti, C. Cvitanovic, J. Duggan, E.A. Fulton, D.B. Karcher, C. Paris, R. Shellock, Human-AI collaboration to identify literature for evidence synthesis, Cell Reports Sustainability 1 (7) (2024), https://doi.org/10.1016/j.crsus.2024.100132.
[58] A. Blaizot, S.K. Veettil, P. Saidoung, C.F. Moreno-Garcia, N. Wiratunga, M. Aceves-Martins, N.M. Lai, N. Chaiyakunapruk, Using artificial intelligence methods for systematic review in health sciences: a systematic review, Res. Synth. Methods 13 (3) (2022) 353–362, https://doi.org/10.1002/jrsm.1553.
[59] C. Hamel, S.E. Kelly, K. Thavorn, D.B. Rice, G.A. Wells, B. Hutton, An evaluation of DistillerSR's machine learning-based prioritization tool for title/abstract screening–impact on reviewer-relevant outcomes, BMC Med. Res. Method. 20 (1) (2020) 256, https://doi.org/10.1186/s12874-020-01129-1.

[60] L. Schmidt, A.N.F. Mutlu, R. Elmore, B.K. Olorisade, J. Thomas, J.P. Higgins, Data extraction methods for systematic review (semi) automation: Update of a living systematic review, F1000Research 10 (2025) 401. https://doi.org/10.12688/f1000research.51117.3.
[61] J.L. Kwan, L. Lo, J. Ferguson, H. Goldberg, J.P. Diaz-Martinez, G. Tomlinson, J. M. Grimshaw, K.G. Shojania, Computerised clinical decision support systems and absolute improvements in care: meta-analysis of controlled clinical trials, BMJ 370 (2020), https://doi.org/10.1136/bmj.m3216.
[62] S. Askin, D. Burkhalter, G. Calado, S. El Dakrouni, Artificial intelligence applied to clinical trials: opportunities and challenges, Heal. Technol. 13 (2) (2023) 203–213, https://doi.org/10.1007/s12553-023-00738-2.
[63] R. Knevel, K.P. Liao, From real-world electronic health record data to real-world results using artificial intelligence, Ann. Rheum. Dis. 82 (3) (2023) 306–311, https://doi.org/10.1136/ard-2022-222626.
[64] D. Calaprice-Whitty, K. Galil, W. Salloum, A. Zariv, B. Jimenez, Improving Clinical Trial Participant Prescreening with Artificial Intelligence (AI): a Comparison of the results of AI-Assisted vs Standard Methods in 3 Oncology Trials, Ther. Innov. Regul. Sci. 54 (1) (2020) 69–74, https://doi.org/10.1007/s43441-019-00030-4.
[65] R.J. Chen, J.J. Wang, D.F. Williamson, T.Y. Chen, J. Lipkova, M.Y. Lu, S. Sahai, F. Mahmood, Algorithmic fairness in artificial intelligence for medicine and healthcare, Nat. Biomed. Eng. 7 (6) (2023) 719–742, https://doi.org/10.1038/s41551-023-01056-8.
[66] M. Alber, A. Buganza Tepole, W.R. Cannon, S. De, S. Dura-Bernal, K. Garikipati, G. Karniadakis, W.W. Lytton, P. Perdikaris, L. Petzold, E. Kuhl, Integrating machine learning and multiscale modeling—perspectives, challenges, and opportunities in the biological, biomedical, and behavioral sciences, npj Digital Med. 2 (1) (2019) 115, https://doi.org/10.1038/s41746-019-0193-y.
[67] Ó. Álvarez-Machancoses, E.J. DeAndres Galiana, A. Cernea, J. Fernández de la Viña, J.L. Fernández-Martínez, On the role of artificial intelligence in genomics to enhance precision medicine, Pharmacogenomics Pers. Med. (2020) 105–119, https://doi.org/10.2147/PGPM.S205082.
[68] S.A. Bagabir, N.K. Ibrahim, H.A. Bagabir, R.H. Ateeq, Covid-19 and Artificial Intelligence: Genome sequencing, drug development and vaccine discovery, J. Infect. Public Health 15 (2) (2022) 289–296, https://doi.org/10.1016/j.jiph.2022.01.011.
[69] Z. Dlamini, A. Skepu, N. Kim, M. Mkhabele, R. Khanyile, T. Molefi, S. Mbatha, B. Setlai, T. Mulaudzi, M. Mabongo, M. Bida, AI and precision oncology in clinical cancer genomics: from prevention to targeted cancer therapies-an outcomes based patient care, Inf. Med. Unlocked 31 (2022) 100965, https://doi.org/10.1016/j.imu.2022.100965.
[70] E. Zihni, V.I. Madai, M. Livne, I. Galinovic, A.A. Khalil, J.B. Fiebach, D. Frey, Opening the black box of artificial intelligence for clinical decision support: a study predicting stroke outcome, PLoS One 15 (4) (2020) e0231166, https://doi.org/10.1371/journal.pone.0231166.
[71] X. Yin, H. Liao, H. Yun, N. Lin, S. Li, Y. Xiang, X. Ma, Artificial intelligence-based prediction of clinical outcome in immunotherapy and targeted therapy of lung cancer. In: Seminars in cancer biology, Vol. 86, Academic Press, 2022, pp. 146-159. https://doi.org/10.1016/j.semcancer.2022.08.002.
[72] B. Meskó, G. Hetényi, Z. Győrffy, Will artificial intelligence solve the human resource crisis in healthcare? BMC Health Serv. Res. 18 (1) (2018) 545, https://doi.org/10.1186/s12913-018-3359-4.
[73] M. Vössing, N. Kühl, M. Lind, G. Satzger, Designing transparency for effective human-AI collaboration, Inf. Syst. Front. 24 (3) (2022) 877–895, https://doi.org/10.1007/s10796-022-10284-3.
[74] M. Nazar, M.M. Alam, E. Yafi, M.M. Su'ud, A systematic review of human–computer interaction and explainable artificial intelligence in healthcare with artificial intelligence techniques, IEEE Access 9 (2021) 153316–153348, https://doi.org/10.1109/ACCESS.2021.3127881.
[75] D. Zikos, N. DeLellis, CDSS-RM: a clinical decision support system reference model, BMC Med. Res. Method. 18 (1) (2018) 137, https://doi.org/10.1186/s12874-018-0587-6.
[76] M.G. Hanna, L. Pantanowitz, B. Jackson, O. Palmer, S. Visweswaran, J. Pantanowitz, M. Deebajah, H.H. Rashidi, Ethical and bias considerations in artificial intelligence/machine learning, Mod. Pathol. 38 (3) (2025) 100686, https://doi.org/10.1016/j.modpat.2024.100686.
[77] J. Zhang, Z.M. Zhang, Ethics and governance of trustworthy medical artificial intelligence, BMC Med. Inf. Decis. Making 23 (1) (2023) 7, https://doi.org/10.1186/s12911-023-02103-9.
[78] S. Reddy, S. Allan, S. Coghlan, P. Cooper, A governance model for the application of AI in health care, J. Am. Med. Inform. Assoc. 27 (3) (2020) 491–497.
[79] R. Amin, M.A. Al Ghamdi, S.H. Almotiri, M. Alruily, Healthcare techniques through deep learning: issues, challenges and opportunities, IEEE Access 9 (2021) 98523–98541, https://doi.org/10.1109/ACCESS.2021.3095312.
[80] T.P. Quinn, M. Senadeera, S. Jacobs, S. Coghlan, V. Le, Trust and medical AI: the challenges we face and the expertise needed to overcome them, J. Am. Med. Inform. Assoc. 28 (4) (2021) 890–894, https://doi.org/10.1093/jamia/ocaa268.
[81] G. Lorenzini, L. Arbelaez Ossa, D.M. Shaw, B.S. Elger, Artificial intelligence and the doctor–patient relationship expanding the paradigm of shared decision making, Bioethics 37 (5) (2023) 424–429, https://doi.org/10.1111/bioe.13158.
[82] S. Nasir, R.A. Khan, S. Bai, Ethical framework for harnessing the power of AI in healthcare and beyond, IEEE Access 12 (2024) 31014–31035, https://doi.org/10.1109/ACCESS.2024.3369912.
[83] S.M. Williamson, V. Prybutok, Balancing privacy and progress: a review of privacy challenges, systemic oversight, and patient perceptions in AI-driven healthcare, Appl. Sci. 14 (2) (2024) 675, https://doi.org/10.3390/app14020675.
[84] H.J. Park, Patient perspectives on informed consent for medical AI: a web-based experiment, DIGITAL HEALTH 10 (2024) 20552076241247938, https://doi.org/10.1177/20552076241247938.
[85] D. Vela, A. Sharp, R. Zhang, T. Nguyen, A. Hoang, O.S. Pianykh, Temporal quality degradation in AI models, Sci. Rep. 12 (1) (2022) 11654, https://doi.org/10.1038/s41598-022-15245-z.
[86] A. Jaleel, T. Mahmood, M.A. Hassan, G. Bano, S.K. Khurshid, Towards medical data interoperability through collaboration of healthcare devices, IEEE Access 8 (2020) 132302–132319, https://doi.org/10.1109/ACCESS.2020.3009783.
[87] T.K. Gandhi, D. Classen, C.A. Sinsky, D.C. Rhew, N. Vande Garde, A. Roberts, F. Federico, How can artificial intelligence decrease cognitive and work burden for front line practitioners? JAMIA Open. 6(3):ooad079 (2023) https://doi.org/10.1093/jamiaopen/ooad079.
[88] H. Smith, Clinical AI: opacity, accountability, responsibility and liability, AI & Soc. 36 (2) (2021) 535–545, https://doi.org/10.1007/s00146-020-01019-6.