TẠP CHÍ KHOA HỌC
TRƯỜNG ĐẠI HỌC TÂN TẠO

TAN TAO UNIVERSITY
From knowledge to the stars

BIOMEDICAL SCIENCE

AI in Healthcare: Promise, Pitfalls, and Future Directions

Nghia Phu Nguyen*,[1], Thy Dinh Van Le[1], Ngoc Q. N. Nguyen[1], Thuy Ngan Truong[1], Nha C. Chau[1], Jasmine Nguyen[2], Nhan Phu Nguyen[3], Thy Ha Anh Nguyen[4], Loc Vu[5] and Thach Nguyen[1,5,6]

[1]Cardiovascular Research, Methodist Hospital, Merrillville, Indiana, USA
[2]Department of Chemistry and Physics, Purdue University Northwest Hammond, Indiana, USA
[3]School of Computing, National University of Singapore, Singapore
[4]Faculty of Science, University of Montpellier, Montpellier, France
[5]Tan Tao University, School of Medicine, Tay Ninh Province, Vietnam
[6]Interventional Cardiology, St. Mary Medical Center, Hobart, Indiana, USA

*Corresponding author:
Nghia Phu Nguyen - Cardiovascular Research, Methodist Hospital, Merrillville, Indiana, USA.
Email: phunghiamd@gmail.com

Received: 01/04/2026. **Revised:** 24/04/2026. **Accepted:** 09/05/2026. **DOI:** 10.53901/tjs.2026.v01.issue02.art02

Abstract

Artificial intelligence (AI) has rapidly moved from a niche computational tool to an active participant in healthcare delivery and medical education. Its expanding role now extends beyond diagnosis to clinical trial screening platforms, decision support, physician-facing knowledge tools, and early pilots involving supervised prescription workflow. This review examines the evolution of healthcare AI, from traditional rule-based systems to machine learning, large language models, and multimodal AI, and explores its current applications across clinical practice, public health, and medical education. We discuss the major promises of AI, including improved diagnostic support, earlier disease detection, reduced administrative burden, scalable education, and augmented clinical reasoning. At the same time, we highlight important limitations and risks, including bias, hallucinations, limited generalizability, privacy concerns, automation bias, threats to humanistic care, and challenges to academic integrity. We further examine barriers to implementation, including technical integration, workflow disruption, regulatory uncertainty, and workforce preparedness. While AI offers substantial opportunities to improve healthcare systems and training environments, its impact will depend less on algorithmic capability alone and more on responsible integration, continuous oversight, and preservation of human judgment. Future efforts should prioritize trustworthy, equitable, and human-centered approaches that position AI as a tool to augment rather than replace clinician-patient relationship.

Keywords: Artificial intelligence; Healthcare AI; Medical education; Large language models; Clinical decision support; Human–AI collaboration; AI implementation; AI ethics.

Introduction

Artificial intelligence (AI) is increasingly reshaping healthcare and medical education.

Traditional AI in healthcare has evolved from early computer-based medical applications and rule-based decision-support systems to more advanced machine-learning tools used for image analysis, computer-aided detection, clinical prediction, and medication-related support [1]. Initially, most healthcare AI systems were designed for narrow, task-specific functions, such as detecting abnormalities, supporting diagnosis, improving workflow efficiency, or assisting with specific clinical and administrative decisions [2]. However, the field has rapidly shifted toward more complex and flexible systems, including generative AI, multimodal AI, workflow-based AI, and reasoning-support systems [2]. These newer models can generate text, summarize clinical information, integrate multiple data types, support communication, and assist with decision-making processes [3].

Several factors have contributed to the recent acceleration of AI adoption in healthcare. The widespread digitization of electronic health records, advances in cloud computing, the development of foundation models and large language models (LLMs), and the increasing availability of multimodal data have created an environment in which AI tools can be deployed more broadly [1]. This digital transformation was further accelerated by the COVID-19 pandemic [3]. As a result, the healthcare systems are forced to rapidly adopt digital technologies for clinical care, communication, education, and administrative workflows [3]. In addition, commercial deployment has increased the visibility and accessibility of AI tools, bringing them into hospitals, clinics, universities, and even direct patient-facing platforms [4].

As a result, AI roles are expanding into that of a clinical assistant, documentation tool, communication aid, educational resource, and administrative infrastructure [4]. In clinical practice, AI may assist physicians by summarizing patient records, drafting clinical notes, supporting triage, identifying high-risk patients, and improving workflow efficiency [3]. In medical education, AI may support personalized learning, automated feedback, case-based simulation, question generation, and student assessment [5].

Despite the rapid growth of AI, important gaps remain in the literature. Many studies focus on technical performance, diagnostic accuracy, and potential efficiency gains, while less attention is given to ethical, educational, and practical implementation challenges [6]. Key concerns include bias and inequity caused by underrepresented populations in training datasets, hallucinations and misinformation from generative AI, limited transparency in black-box systems, and unresolved issues related to privacy, patient confidentiality, secondary data use, and cybersecurity [3, 6, 7]. In addition, overreliance on AI may contribute to automation bias, deskilling, and reduced independent clinical reasoning [8]. AI may also affect humanistic care by weakening empathy, communication, and clinician–patient relationships [9]. Further concerns include ghost-writing, plagiarism, fabricated references, and AI-assisted misconduct [5]. These gaps highlight the need for a balanced review that examines both the promises and risks of AI and proposes responsible strategies for its integration into healthcare.

This review examines the current applications of AI in healthcare, its emerging roles in medical education, the major promises and pitfalls associated with its use, the challenges of implementation, and future directions for responsible integration. By addressing both opportunities and limitations, this review aims to provide a balanced perspective on how AI may transform clinical practice and medical education while maintaining patient safety, ethical standards, and the central role of human judgment.

Evolution of AI in Healthcare

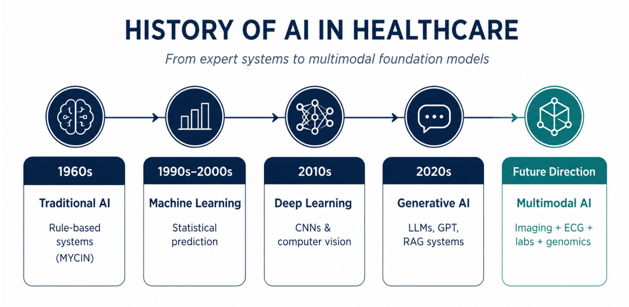

Figure 1: Evolution of AI in Healthcare.

Traditional AI

The first health information systems emerged in the 1960s, marking the beginning of automated processing of medical data [10]. During the 1970s, expert systems such as MYCIN demonstrated that computers could assist physicians in diagnostic reasoning and antimicrobial selection through predefined IF-THEN rules and logical inference [11]. MYCIN incorporated a consultation module for therapeutic decision support, an explanation module that justified its reasoning process, and a knowledge acquisition module that enabled experts to update clinical rules without direct programming. This architecture became a foundational model for later clinical decision support systems (CDSS) [12].

Machine learning era

Alongside the growing application of machine learning (ML) in fields such as physics and video games, healthcare also became an important area of interest for industry and research [13]. ML can generally be divided into supervised and unsupervised learning [14]. Supervised learning is commonly used for prediction and risk estimation based on labelled input data, whereas unsupervised learning is primarily used to identify hidden patterns or structures within medical datasets [13]. The emergence of deep learning marked a major transition in healthcare AI, shifting from manually engineered features for traditional ML models toward neural networks capable of learning directly from raw clinical data [15]. This advancement also accelerated the development of computer vision applications in healthcare, enabling automated analysis of medical images and videos for disease detection, image segmentation, and clinical decision support across multiple specialties [16, 17].

Generative AI and LLM era

LLMs have emerged as a major development in healthcare AI, enabling applications such as clinical documentation, medical text summarization, question answering, patient communication, medical education, and clinical decision support [18]. Despite the growing use in healthcare, LLMs remain vulnerable to hallucinations, inaccurate or fabricated medical information, and biased training data. Retrieval-augmented generation (RAG) has therefore been proposed as a potential solution, enabling models to retrieve relevant external information before generating responses [19]. This framework may improve factual accuracy, reliability, and personalization in healthcare applications while allowing more flexible integration of updated medical knowledge.

Multimodal AI

ML methods in healthcare have traditionally focused on using data from a single modality, limiting their ability to effectively replicate the clinical practice of integrating multiple sources of information for improved decision making [20]. Recently, multimodal AI has emerged as a future direction in healthcare by integrating heterogeneous clinical data, including medical imaging, laboratory results, electrocardiogram (ECG) signals, genomics, speech, and clinical text, into unified predictive models [21]. By combining multiple data modalities, these systems may improve diagnostic accuracy and support more personalized treatment [21].

Current Applications of AI in Healthcare

Diagnostic and Interpretive AI

Diagnostic and interpretive AI refers to systems that analyze clinical images or pathologic signals to detect disease, classify abnormalities, and support clinical decision-making. These tools are especially common in radiology, pathology, cardiology, and ophthalmology, where clinicians rely on Computed Tomography (CT), Magnetic Resonance Imaging (MRI), mammography, digital slides, ECGs, and retinal photographs [22–25]. To summarize the major clinical applications of diagnostic and interpretive AI, **Table** 1 compares its use across disciplines.

Table 1: Summary of Diagnostic and Interpretive Artificial Intelligence Applications, Benefits, and Limitations [22–27].

Clinical Area	Common Applications
Radiology / Medical Imaging	CT, MRI, and mammography interpretation; detection of lung nodules, intracranial hemorrhage, pulmonary embolism, tumors, and suspicious breast lesions
Pathology	Digital slide interpretation; cancer detection; identification of tumor cells, tissue patterns, margins, and biomarkers
Cardiology / ECG	Arrhythmia detection, atrial fibrillation prediction, and ECG-based risk prediction
Ophthalmology	Retinal image analysis; diabetic retinopathy and related eye disease detection

The main promise of diagnostic AI is improved speed, scalability, triage, and sensitivity. AI can process large volumes of images, slides, ECGs, and retinal photographs faster than clinicians alone, which may help identify urgent or subtle findings earlier [22–25]. However, safe implementation is limited by dataset bias, poor external validation, shortcut learning, and false confidence. Models may rely on scanner type, image labels, acquisition patterns, or institution-specific artifacts rather than true disease features, reducing generalizability and increasing the risk of biased predictions [22, 26, 27].

Clinical Decision Support Systems

CDSS provides patient-specific alerts, predictions, and recommendations to support medical decision-making. AI-based CDSS use electronic health record data, including vital signs, laboratory results, medications, diagnoses, and clinical notes, to predict risk, recommend treatments, and generate early warning alerts CDSS may help estimate risk of sepsis, mortality, deterioration, and readmission, allowing earlier monitoring, treatment, or escalation of care [28, 29]. The main promise of CDSS is earlier detection, standardized care, reduced cognitive burden, and personalized treatment. However, pitfalls include alert fatigue, automation bias, black-box decision making, and distribution shift [28, 29]. Excessive or low-value alerts can reduce clinician attention, while unclear AI outputs may encourage over-trust [30]. Model reliability may also decline outside the training environment, as shown by poor external validation of a sepsis prediction model [29]. Therefore, CDSS should support, not replace, clinician judgment [28–30].

Generative AI and LLMs in Clinical Practice

Table 2: Summary of Generative AI and LLMs Applications, Benefits, and Limitations in Clinical Practice [31–36].

Application	Main use and promise	Limitations
Clinical documentation: Ambient AI scribes, progress notes, discharge summaries, electronic health record (EHR) entries	Transcribes patient-clinician encounters and drafts structured clinical documentation. May reduce documentation burden, after-hours charting, and administrative burnout	May produce omissions, misinterpretations, hallucinations, or incomplete summaries; requires clinician review
Communication support: Patient education, portal replies, appointment messages, refill requests	Simplifies medical information and drafts patient-facing or administrative messages. May improve patient understanding, reduce cognitive task load, and support communication workflow	May generate inaccurate advice, omit safety details, misjudge health literacy, or create privacy risks
Knowledge synthesis: EHR summarization, literature review, guideline extraction	Summarizes clinical records, literature, lab trends, imaging histories, and guideline information. May reduce information overload and improve knowledge synthesis	May produce ungrounded reasoning, contextual omissions, hallucinations, or unreliable outputs without expert verification
Clinical reasoning augmentation: Differential diagnosis, evidence summaries, next-step suggestions	Organizes patient information and supports diagnostic reasoning. May reduce cognitive burden and help structure complex clinical cases	Evidence remains mixed; may hallucinate, produce incomplete reasoning, or encourage automation bias and false confidence

Generative AI and LLMs are progressively reshaping clinical documentation, communication support, knowledge synthesis, and clinical reasoning workflows [37–40]. These models reduce clinician cognitive fatigue and documentation workloads by drafting ambient visit notes, structuring portal message replies, translating patient-facing educational materials, and compressing disjointed electronic health record histories [31, 32, 41]. However, translating this raw linguistic capability into superior diagnostic outcomes remains a challenge. A randomized clinical trial demonstrated that providing physicians with an LLM did not significantly improve their final diagnostic reasoning scores over conventional resources, illustrating that human-AI collaboration is highly vulnerable to automation bias and flawed workflow design [33]. Because foundation models are built on structurally irregular electronic health record (EHR) architectures and remain susceptible to confident hallucinations, data omissions, and privacy risks, they must be restricted to supervised drafting aids that require mandatory human-in-the-loop expert review [34, 42]. **Table** 2 summarizes current uses of Generative AI and LLMs in healthcare.

AI in Public Health and Health Systems

AI-driven macro-informatics optimize healthcare infrastructure by supporting public health surveillance, outbreak forecasting, and hospital resource deployment [43–45]. In population health, ML models continuously ingest multi-stream datasets, including wastewater indicators, mobility matrices, climate trends, and digital social behavior, to create predictive "nowcasting" frameworks that flag infectious disease hotspots prior to traditional laboratory confirmation [46]. At the institutional level, neural networks analyze historical electronic health records to forecast emergency department surges, Intensive Care Unit bed constraints, and supply chain disruptions, enabling proactive instead of reactive clinical staffing [45].

While these tools offer immense potential to democratize specialist-level expertise in low- and middle-income countries (LMICs) via automated screening and telemedicine, their deployment remains constrained by algorithmic vulnerabilities [43]. Transferring models trained exclusively on high-income populations to resource-limited environments introduces systemic data inequity, causing performance degradation due to variations in baseline disease patterns, local hardware configurations, and cultural determinants of health [45, 47]. Consequently, to prevent AI from widening global health disparities, real-world deployment requires strict data governance, extensive local validation, and centralized public health oversight [43, 44].

AI in Medical Education and Workforce Preparedness

AI-Assisted Learning and Scalable Training

AI is increasingly being integrated into medical education as a tool to improve learning efficiency, accessibility, and personalization through adaptive learning systems, automated feedback, virtual simulations, and AI-assisted educational support (**Table** 3) [48, 49]. These applications are relevant to future workforce preparedness, as clinicians are likely to work in environments where AI-supported tools are embedded into clinical and educational workflows. AI training programs may improve learner satisfaction, knowledge acquisition, self-efficacy, and selected behaviour-related outcomes, although most remain introductory [50]. AI-assisted learning may therefore support scalable training by providing individualized pacing, rapid clarification, and access to educational resources in settings where faculty availability or mentorship is limited [47].

Table 3: AI-supported educational tools [48, 49, 51–53].

AI-supported educational tool	Educational function	Workforce competency supported
Adaptive learning systems	Personalize content, pacing, and remediation	Self-directed lifelong learning
Automated feedback and learning analytics	Identify gaps and guide targeted improvement	Reflective practice and performance monitoring
Virtual patients and simulation	Support case-based learning and clinical reasoning practice	Safe human-AI collaboration in decision-making
AI-assisted assessment	Generate items, feedback, and grading support	Assessment literacy and verification of AI outputs
Telemedicine and AI-mediated communication training	Practise digital communication and patient interaction	Digital empathy and patient-centred communication

However, the educational value of AI depends on how equitably and contextually it is implemented. In LMICs, AI-supported simulation, tele-mentoring, and automated feedback may help address constraints in educational infrastructure, but their effectiveness depends on stable internet access, locally relevant content, faculty development, and appropriate governance [47, 54]. Imported AI systems may embed linguistic, cultural, and clinical assumptions derived from dominant Western training data, which may not fully align with local educational contexts and could therefore reduce fairness, inclusivity, and educational relevance [54]. Thus, AI-assisted learning should be evaluated not only by efficiency or learner satisfaction, but also by its contribution to equitable and context-sensitive workforce training.

Clinical Reasoning in AI-Supported Environments

Clinical reasoning traditionally develops through uncertainty, reflection, feedback, and repeated exposure to patient complexity. Generative AI may support this process through virtual patient scenarios, differential diagnosis generation, and feedback [51]. However, early reliance on AI-generated interpretations may encourage cognitive outsourcing and reduce opportunities for learners to develop independent diagnostic judgment [55]. This concern has been described as extending beyond deskilling toward "never-skilling," where foundational reasoning habits may fail to consolidate because AI outputs are repeatedly encountered before learners construct their own understanding [56]. One possible solution is a Human–AI–Human approach [56], in which learners first form an initial clinical impression, then use AI to test assumptions or identify blind spots, and finally return to human judgment for contextual validation and accountability.

Assessment Validity and Educational Integrity

AI is also being used in assessment through multiple-choice question generation, clinical vignette development, automated feedback, grading support, and learning analytics [49]. These applications may reduce faculty workload and support formative assessment in resource-constrained settings. However, AI-generated assessments may contain weak distractors, implausible options, convergence errors, over-explained correct answers, and predictable answer patterns that reduce item discrimination and assessment validity [52]. AI is therefore better understood as an assessment-support tool rather than a replacement for expert oversight. Human review remains necessary to evaluate factual accuracy, cognitive depth, fairness, and alignment with learning objectives [57]. More broadly, AI may shift assessment toward formats that make reasoning visible, including oral justification, supervised case analysis, reflective comparison between learner and AI outputs, and structured clinical performance assessment.

AI Literacy and Future Workforce Readiness

AI literacy includes the ability to understand, evaluate, apply, and ethically interact with AI systems in professional practice [50]. Current AI education programs may improve confidence, knowledge, and self-efficacy, but many remain introductory and rarely evaluate higher-level outcomes such as workplace performance or sustained human–AI collaboration [50]. Future curricula may therefore need to expand toward competencies in bias recognition, hallucination detection, privacy protection, contextual validation, and responsible AI use in clinical workflows.

Workforce readiness also includes communication and relational competencies. As healthcare increasingly incorporates telemedicine, patient portals, AI-assisted documentation, and AI-mediated patient expectations, clinicians may require training in digital empathy and webside manner [53, 58]. Patients may also arrive with AI-generated interpretations of symptoms or treatment options, creating a clinical encounter in which physicians must evaluate, clarify, and contextualize algorithmic narratives alongside the patient's concerns [59]. AI literacy should therefore include epistemic literacy: the ability to judge how AI-generated information is produced, where it may be incomplete or misleading, and how it should be integrated into safe and accountable care.

AI literacy must also be developed across both learners and educators. Students may adopt AI informally for efficiency and cognitive support, whereas faculty may approach it more cautiously because of concerns regarding academic integrity, bias, privacy, and erosion of critical thinking [60]. This divide may fragment medical education unless institutions develop parallel AI literacy pathways, intergenerational mentoring, and governance structures that align students' digital fluency with faculty members' clinical and ethical expertise. Future medical education will therefore need to prepare clinicians to use AI-supported tools critically and ethically while preserving clinical reasoning, communication skills, assessment integrity, and professional accountability.

Promises of AI in Healthcare

AI promises benefits across clinical care, medical education, health-system operations, and public health, but these benefits remain highly dependent on context, validation, infrastructure, and human oversight. **Figure** 2 summarizes the major promise domains discussed below.

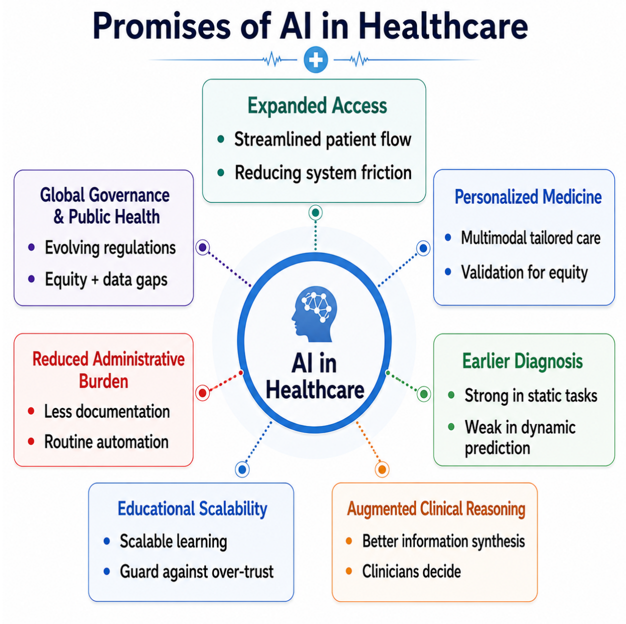

Figure 2: Overview of the evolving roles, opportunities, and challenges of artificial intelligence in healthcare.

Expanded Access

AI has shown potential to improve healthcare access through triage automation, workflow optimisation, and specialist support in resource-limited settings. An NHS-funded evaluation found a 73% reduction in GP waiting times and a 91% rate of automated appointment allocation at a single primary care site [61], and U.S. physician adoption of AI tools has reached 81% — more than double the rate recorded in 2023 [62]. In LMICs, AI-assisted TB-detection algorithms met WHO sensitivity standards for non-specialist community screening [63], and AI-supported colposcopy tools have been proposed to address critical specialist shortages in cervical cancer programmes [64]. However, sustained population-level access gains remain dependent on infrastructure investment, workflow integration, and locally validated models [65].

Personalized Medicine

Multimodal AI models are increasingly being used to support individualised risk prediction, treatment stratification, and tailored care. In oncology, multimodal models now outperform single-modality approaches in prognostic discrimination, demonstrating the potential for individualised risk stratification and treatment tailoring [66]. In cardiovascular medicine, deep learning ECG models have achieved cardiologist-level arrhythmia detection [67], and population-scale AI-based atrial fibrillation screening has been demonstrated via wearables [68]. Nevertheless, important evidence gaps remain, including limited outcome-level randomised trials demonstrating improved survival or quality of life, and persistent underdiagnosis in underrepresented populations when models are trained on non-representative datasets [69, 70].

Earlier Diagnosis

AI has demonstrated its most consistent diagnostic gains in standardised imaging and classification tasks. AI-based early warning systems can detect clinical deterioration sooner than traditional scoring methods, but real-world performance remains variable and false positives remain a challenge [29]. By contrast, radiology has demonstrated more consistent gains: AI-enabled worklist prioritisation reduces time-to-clinical-recognition of acute findings, including a 96% reduction in time-to-diagnosis for intracranial haemorrhage [71] and independent validation for acute neurological triage [72]. This distinction is important: radiology evidence often reflects faster recognition of existing pathology rather than earlier disease detection in the biological course. Current evidence therefore supports AI's role in prioritising high-risk cases identifiable from standardised inputs, whereas real-time bedside prediction remains dependent on dynamic data, workflow integration, and clinician interpretation.

Reduced Administrative Burden

Reduction of administrative burden is among the most consistently demonstrated benefits of healthcare AI. Clerical overload has been identified as a primary structural driver of physician burnout across specialities [73]. Ambient AI scribes reduced burnout by 13.9% across 263 clinicians in six health systems [74]; a 21.2% absolute reduction in burnout at Mass General Brigham and a 30.7% improvement in documentation-related wellbeing at Emory Healthcare were observed across 1,430 clinicians [75]; and EHR time fell by 13–16 minutes per eight patient hours across 8,581 clinicians at five academic centres [41]. Beyond scribing, AI-powered inbox management resolved 1.5 million of 4.7 million patient messages without physician involvement at Kaiser Permanente [42]. These improvements matter not only as efficiency gains but also as potential responses to clinician dissatisfaction and workforce attrition. However, efficiency gains may also reshape workforce distribution, underscoring the need for deliberate transition planning for administrative staff [76].

Educational Scalability

AI may expand the scalability of medical education through personalised feedback, documentation support, simulation, and differential-diagnosis scaffolding. Large language models have been shown to improve the completeness of differential diagnosis generation in clinician-oriented case evaluations [77, 78], and controlled studies demonstrate gains in documentation quality and clinical reasoning support when LLMs are used as structured learning aids [79]. These tools may be particularly valuable where faculty time, simulation resources, and specialist mentorship are limited. However, excessive reliance on AI-generated reasoning may weaken learners' tolerance for uncertainty and independent clinical judgment, especially if AI is used before learners attempt their own reasoning [56, 80].

Augmented Clinical Reasoning

AI augments clinical reasoning by improving information synthesis and completeness of differential diagnosis. LLMs perform at or above the USMLE passing threshold in structured scenarios [81, 82] and improve differential diagnosis completeness as an aid to clinicians in challenging real-world case evaluations [77–79]. However, benchmark performance in structured, text-based scenarios does not reliably transfer to real-world clinical contexts, where inputs are multimodal, temporally evolving, and embedded in socio-clinical complexity. Automation bias — the tendency to defer to AI outputs even when they conflict with clinical judgment — remains a recognised patient safety hazard across healthcare settings [83]. Current evidence therefore supports a collaborative augmentation model in which AI improves information synthesis while clinicians retain contextual interpretation and final decision-making authority.

Global Governance, Equity, and Public Health Preparedness

The global impact of healthcare AI will depend not only on technical performance but also on governance, equity, and data representation. Regulatory approaches differ substantially: the FDA has established adaptive pathways for AI/ML medical software [97], while the EU AI Act imposes mandatory conformity assessments for high-risk healthcare applications [114]. Equity considerations remain central because current AI development and evidence generation are concentrated in the USA and China, raising concerns about fairness and applicability in underrepresented populations [84]. At the public health level, AI-assisted genomic surveillance accelerated COVID-19 variant identification [115], and Natural Language Processing

Table 4: Major risk domains in healthcare AI: core mechanisms and potential consequences.

Risk domain	Core mechanism	Potential consequence
Bias and inequity [27, 69, 84–87]	Biased labels, proxy outcomes, subgroup underrepresentation, aggregate-only reporting, and latent demographic encoding	Unequal diagnosis, triage, risk stratification, resource allocation, or subgroup-level performance
Validity and post-deployment safety [29, 88, 89]	Dataset shift, limited external validation, local workflow differences, and performance drift after deployment	False reassurance, excessive alerts, missed cases, or unsafe local deployment
Hallucination and misinformation [90–92]	Fluent but unsupported, fabricated, incomplete, or clinically misleading generative outputs	Misleading recommendations, inaccurate summaries, documentation errors, or fabricated references
Opacity and accountability [93–96]	Limited interpretability, weak documentation, unclear intended use, and uncertain responsibility for AI-influenced decisions	Reduced contestability, impaired trust, and unresolved oversight or liability questions
Privacy and cybersecurity [97–100]	Secondary data use, imperfect de-identification, vendor access, online tracking, and cyber vulnerabilities	Confidentiality breach, loss of consent legitimacy, regulatory exposure, or erosion of public trust
Automation bias and deskilling [95, 101, 102]	Miscalibrated trust, cognitive offloading, and repeated dependence on AI assistance	Propagation of erroneous recommendations or reduced unaided clinical performance
Relational care [103–106]	Substitution of automated communication for contextual, patient-centred interaction	Reduced patient agency, inequitable communication quality, readability barriers, or weakened therapeutic relationships

Table 5: Major implementation challenges affecting the integration of artificial intelligence into routine clinical practice [4, 6, 9, 107–113].

Domain	Key implementation challenges	Potential consequences
Technical integration	Heterogeneous EHR architectures, inconsistent coding standards, fragmented datasets, workflow mismatch, and limited interoperability across institutions	Reduced model transferability, interoperability failure, workflow disruption, and unreliable cross-site deployment
Data security and privacy	Increased risks of unauthorized access, data leakage, and regulatory non-compliance during automated data exchange and multi-center integration	Privacy breaches, compromised patient confidentiality, and reduced institutional trust
Infrastructure and maintenance	Lack of sociotechnical infrastructure, feedback loops, post-deployment monitoring, data provenance systems, and continuous recalibration mechanisms	Model degradation, reduced scalability, poor sustainability, and unsafe long-term implementation
Clinical validation	Heavy reliance on retrospective studies, single-center datasets, limited demographic reporting, and insufficient external validation	Poor generalizability, subgroup uncertainty, and reduced reliability in diverse populations
Real-world implementation	Controlled experimental settings may not reflect routine clinical complexity, evolving workflows, or variable institutional resources	Performance decline after deployment, unintended consequences, and inconsistent clinical effectiveness
Human factors and usability	Limited clinician trust, insufficient explainability, poor usability, workflow disruption, and alert fatigue	Low adoption, increased cognitive burden, reduced efficiency, and clinician resistance
Regulatory uncertainty	Unclear liability, accountability, post-market surveillance, and inadequate approval pathways for adaptive AI systems	Legal ambiguity, weak oversight, and difficulty ensuring long-term safety
Workforce preparedness	Limited AI literacy, insufficient faculty training, lack of AI curricula, and inadequate continuing medical education	Reduced implementation readiness and unsafe or inconsistent AI use

(NLP)-based outbreak detection systems have demonstrated early-signal value from HealthMap's foundation work through more recent AI-driven advances [116, 117]. These issues highlight the importance of locally validated, equitable, and accountable AI implementation frameworks, particularly in LMIC settings where data gaps and infrastructure constraints may determine whether AI reduces or widens health inequities [54, 118, 119].

Pitfalls and Risks

Although AI may improve diagnostic accuracy, access to care, administrative efficiency, and scalable medical education, its risks emerge from the interaction between data, algorithms, clinicians, patients, and healthcare systems. **Table** 4 summarizes the major risk domains examined in the sections that follow.

Implementation Challenges

Despite the transformative potential of AI in healthcare, translating experimental success into routine clinical practice remains challenging. Beyond algorithmic performance, effective implementation requires integration into healthcare systems, rigorous validation, clinician acceptance, regulatory oversight, and workforce readiness, contributing to a persistent gap between AI promise and real-world clinical impact [6]. These challenges are summarized in **Table** 5.

Conclusion

AI is already reshaping how clinicians learn, make decisions, communicate, and deliver care. However, this review suggests that the main challenge is not whether AI can perform complex tasks, but how it can be integrated responsibly into healthcare systems built on trust, safety, and clinical judgment. Although AI may improve diagnostic accuracy, efficiency, access to care, and scalability of medical education, its benefits remain dependent on transparent governance, continuous validation, fairness, and appropriate human oversight.

Future progress will likely depend not only on stronger AI models, but also on workforce preparedness and human-centered implementation. Medical education may increasingly need to emphasize AI literacy, ethical reasoning, digital empathy, and human–AI collaboration so that clinicians can critically evaluate AI outputs while preserving independent clinical reasoning and accountability. At the same time, multimodal AI integrating imaging, laboratory data, ECG, genomics, speech, and clinical text may further support personalized and context-aware care. Ultimately, AI should augment rather than replace clinicians, while maintaining empathy, communication, patient trust, and equitable access to healthcare.

List of Abbreviations

AI: Artificial intelligence
CDSS: Clinical decision support system
COVID-19: Coronavirus disease 2019
CT: Computed tomography
ECG: Electrocardiogram
EHR: Electronic health record
EU: European Union
FDA: Food and Drug Administration
GP: General practitioner
HIPAA: Health Insurance Portability and Accountability Act
ICU: Intensive care unit
LLM: Large language model
LMIC: Low- and middle-income country
ML: Machine learning
MRI: Magnetic resonance imaging
NHS: National Health Service
NLP: Natural language processing
RAG: Retrieval-augmented generation
SARS-CoV-2: Severe acute respiratory syndrome coronavirus 2
TASS: Technical/Algorithm, Stakeholder, and Society
TB: Tuberculosis
USMLE: United States Medical Licensing Examination
WHO: World Health Organization

Acknowledgments

None

Ethics Approval and Consent to Participate

Not applicable

Funding

None to declare

Competing interests

None of the authors has conflicts of interest to declare

References

[1] Y. A. Fahim, I. W. Hasani, S. Kabba, and W. M. Ragab, "Artificial intelligence in healthcare and medicine: clinical applications, therapeutic advances, and future perspectives," *Eur. J. Med. Res.*, vol. 30, no. 1, p. 848, Sep. 2025, doi: 10.1186/s40001-025-03196-w.

[2] D. Schouten *et al.*, "Navigating the landscape of multimodal AI in medicine: A scoping review on technical challenges and clinical applications," *Med. Image Anal.*, vol. 105, p. 103621, Oct. 2025, doi: 10.1016/j.media.2025.103621.

[3] A. Al Kuwaiti *et al.*, "A review of the role of artificial intelligence in healthcare," *J. Pers. Med.*, vol. 13, no. 6, p. 951, Jun. 2023, doi: 10.3390/jpm13060951.

[4] P. Esmaeilzadeh, "Challenges and strategies for wide-scale artificial intelligence (AI) deployment in healthcare practices: A perspective for healthcare organizations," *Artif. Intell. Med.*, vol. 151, p. 102861, May 2024, doi: 10.1016/j.artmed.2024.102861.

[5] S. H. Abidi, O. Fabiyi, J. Almazan, M. Tariq, and F. Zehra, "The impact of AI-supported case-based learning in medical education: A scoping review protocol," OSF Registries, 2025, doi: 10.17605/OSF.IO/28E3G.

[6] R. A. El Arab *et al.*, "Bridging the gap: From AI success in clinical trials to real-world healthcare implementation—a narrative review," *Healthcare*, vol. 13, no. 7, p. 701, Mar. 2025, doi: 10.3390/healthcare13070701.

[7] J. L. Cross, M. A. Choma, and J. A. Onofrey, "Bias in medical AI: Implications for clinical decision-making," *PLOS Digit. Health*, vol. 3, no. 11, p. e0000651, Nov. 2024, doi: 10.1371/journal.pdig.0000651.

[8] P. E. Heudel, H. Crochet, Q. Filori, T. Bachelot, and J. Y. Blay, "Artificial intelligence in medicine: a scoping review of the risk of deskilling and loss of expertise among physicians," *ESMO Real World Data Digit. Oncol.*, vol. 12, p. 100693, Jun. 2026, doi: 10.1016/j.esmorw.2026.100693.

[9] L. T. Li, L. C. Haley, A. K. Boyd, and E. V. Bernstam, "Technical/algorithm, stakeholder, and society (TASS) barriers to the application of artificial intelligence in medicine: A systematic review," *J. Biomed. Inform.*, vol. 147, p. 104531, Nov. 2023, doi: 10.1016/j.jbi.2023.104531.

[10] D. V. Voshev, R. N. Shepel, N. A. Vosheva, and O. M. Drapkina, "Artificial intelligence in healthcare: historical trajectory, challenges and prospects (1960-2025)," *Prim. Health Care Russ. Fed.*, vol. 2, no. 3, pp. 35–47, Oct. 2025, doi: 10.15829/3034-4123-2025-72.

[11] E. H. Shortliffe, "MYCIN: A knowledge-based computer program applied to infectious diseases," in *Proc. Annu. Symp. Comput. Appl. Med.*

Care, 1977, pp. 66–69.

[12] E. H. Shortliffe, A. M. Randolph, G. B. Bruce, and J. C. William, "Computer-based medical decision making: From MYCIN to VM," in *Readings in Medical Artificial Intelligence: The First Decade*. Reading, MA, USA: Addison-Wesley Publishing Company, 1984.

[13] K. R. Foster, R. Koprowski, and J. D. Skufca, "Machine learning, medical diagnosis, and biomedical engineering research - commentary," *Biomed. Eng. OnLine*, vol. 13, no. 1, p. 94, 2014, doi: 10.1186/1475-925X-13-94.

[14] C. M. Bishop, *Pattern recognition and machine learning*, ser. Information science and statistics. New York, NY, USA: Springer, 2006.

[15] R. Miotto, F. Wang, S. Wang, X. Jiang, and J. T. Dudley, "Deep learning for healthcare: review, opportunities and challenges," *Brief. Bioinform.*, vol. 19, no. 6, pp. 1236–1246, Nov. 2018, doi: 10.1093/bib/bbx044.

[16] A. Esteva *et al.*, "Deep learning-enabled medical computer vision," *npj Digit. Med.*, vol. 4, no. 1, p. 5, Jan. 2021, doi: 10.1038/s41746-020-00376-2.

[17] J. Gao, Y. Yang, P. Lin, and D. S. Park, "Computer vision in healthcare applications," *J. Healthc. Eng.*, vol. 2018, pp. 1–4, 2018, doi: 10.1155/2018/5157020.

[18] D. Wang and S. Zhang, "Large language models in medical and healthcare fields: applications, advances, and challenges," *Artif. Intell. Rev.*, vol. 57, no. 11, p. 299, Sep. 2024, doi: 10.1007/s10462-024-10921-0.

[19] R. Yang *et al.*, "Retrieval-augmented generation for generative artificial intelligence in health care," *npj Health Syst.*, vol. 2, no. 1, p. 2, Jan. 2025, doi: 10.1038/s44401-024-00004-1.

[20] F. Krones, U. Marikkar, G. Parsons, A. Szmul, and A. Mahdi, "Review of multimodal machine learning approaches in healthcare," *Inf. Fusion*, vol. 114, p. 102690, Feb. 2025, doi: 10.1016/j.inffus.2024.102690.

[21] G. Azarfar *et al.*, "Responsible adoption of multimodal artificial intelligence in health care: promises and challenges," *Lancet Digit. Health*, vol. 7, no. 12, p. 100917, Dec. 2025, doi: 10.1016/j.landig.2025.100917.

[22] A. C. Yu, B. Mohajer, and J. Eng, "External validation of deep learning algorithms for radiologic diagnosis: A systematic review," *Radiol. Artif. Intell.*, vol. 4, no. 3, p. e210064, May 2022, doi: 11.1148/ryai.210064.

[23] Z. I. Attia *et al.*, "An artificial intelligence-enabled ECG algorithm for the identification of patients with atrial fibrillation during sinus rhythm: a retrospective analysis of outcome prediction," *The Lancet*, vol. 394, no. 10201, pp. 861–867, Sep. 2019, doi: 10.1016/S0140-6736(19)31721-0.

[24] V. Gulshan *et al.*, "Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs," *JAMA*, vol. 316, no. 22, pp. 2402–2410, Dec. 2016, doi: 10.1001/jama.2016.17216.

[25] D. S. W. Ting *et al.*, "Development and validation of a deep learning system for diabetic retinopathy and related eye diseases using retinal images from multiethnic populations with diabetes," *JAMA*, vol. 318, no. 22, pp. 2211–2223, Dec. 2017, doi: 10.1001/jama.2017.18152.

[26] S. Tripathi *et al.*, "Understanding biases and disparities in radiology AI datasets: A review," *J. Am. Coll. Radiol.*, vol. 20, no. 9, pp. 836–841, Sep. 2023, doi: 10.1016/j.jacr.2023.06.015.

[27] A. Brown, N. Tomasev, J. Freyberg, Y. Liu, A. Karthikesalingam, and J. Schrouff, "Detecting shortcut learning for fair medical AI using shortcut testing," *Nat. Commun.*, vol. 14, no. 1, p. 4314, Jul. 2023, doi: 10.1038/s41467-023-39902-7.

[28] J. Reis *et al.*, "Digital guardian angel supported by an artificial intelligence system to improve quality of life, well-being, and health outcomes of patients with cancer (ONCORELIEF): Protocol for a single arm prospective multicenter pilot study," *JMIR Res. Protoc.*, vol. 12, p. e45475, Apr. 2023, doi: 10.2196/45475.

[29] A. Wong *et al.*, "External validation of a widely implemented proprietary sepsis prediction model in hospitalized patients," *JAMA Intern. Med.*, vol. 181, no. 8, pp. 1065–1070, Aug. 2021, doi: 10.1001/jamainternmed.2021.2626.

[30] J. M. Baron, R. Huang, D. McEvoy, and A. S. Dighe, "Use of machine learning to predict clinical decision support compliance, reduce alert burden, and evaluate duplicate laboratory test ordering alerts," *JAMIA Open*, vol. 4, no. 1, p. ooab006, Mar. 2021, doi: 10.1093/jamiaopen/ooab006.

[31] P. Garcia *et al.*, "Artificial intelligence–generated draft replies to patient inbox messages," *JAMA Network Open*, vol. 7, no. 3, p. e243201, Mar. 2024, doi: 10.1001/jamanetworkopen.2024.3201.

[32] A. J. Thirunavukarasu, D. S. J. Ting, K. Elangovan, L. Gutierrez, T. F. Tan, and D. S. W. Ting, "Large language models in medicine," *Nat. Med.*, vol. 29, no. 8, pp. 1930–1940, Aug. 2023, doi: 10.1038/s41591-023-02448-8.

[33] E. Goh *et al.*, "Large language model influence on diagnostic reasoning: A randomized clinical trial," *JAMA Network Open*, vol. 7, no. 10, p. e2440969, Oct. 2024, doi: 10.1001/jamanetworkopen.2024.40969.

[34] M. Wornow *et al.*, "The shaky foundations of large language models and foundation models for electronic health records," *npj Digit. Med.*, vol. 6, no. 1, p. 135, Jul. 2023, doi: 10.1038/s41746-023-00879-8.

[35] H.-Y. Hsu, L.-W. Chen, W.-T. Hsu, Y.-W. Hsieh,

and S.-S. Chang, "Extracting clinical guideline information using two large language models: Evaluation study," *J. Med. Internet Res.*, vol. 27, p. e73486, Sep. 2025, doi: 10.2196/73486.

[36] S. Liu *et al.*, "Leveraging large language models for generating responses to patient messages—a subjective analysis," *J. Am. Med. Inform. Assoc.*, vol. 31, no. 6, pp. 1367–1379, May 2024, doi: 10.1093/jamia/ocae052.

[37] S. Aydin, M. Karabacak, V. Vlachos, and K. Margetis, "Large language models in patient education: a scoping review of applications in medicine," *Front. Med.*, vol. 11, p. 1477898, Oct. 2024, doi: 10.3389/fmed.2024.1477898.

[38] L. Bednarczyk *et al.*, "Scientific evidence for clinical text summarization using large language models: Scoping review," *J. Med. Internet Res.*, vol. 27, p. e68998, May 2025, doi: 10.2196/68998.

[39] P. Lee, S. Bubeck, and J. Petro, "Benefits, limits, and risks of GPT-4 as an AI chatbot for medicine," *N. Engl. J. Med.*, vol. 388, no. 13, pp. 1233–1239, Mar. 2023, doi: 10.1056/NEJMsr2214184.

[40] S. J. Shah *et al.*, "Physician perspectives on ambient AI scribes," *JAMA Network Open*, vol. 8, no. 3, p. e251904, Mar. 2025, doi: 10.1001/jamanetworkopen.2025.1904.

[41] L. S. Rotenstein *et al.*, "Changes in clinician time expenditure and visit quantity with adoption of artificial intelligence–powered scribes: A multisite study," *JAMA*, vol. 335, no. 16, pp. 1408–1417, Apr. 2026, doi: 10.1001/jama.2026.2253.

[42] V. X. Liu *et al.*, "Content of patient electronic messages to physicians in a large integrated system," *JAMA Network Open*, vol. 7, no. 4, p. e244867, Apr. 2024, doi: 10.1001/jamanetworkopen.2024.4867.

[43] D. B. Olawade, O. J. Wada, A. C. David-Olawade, E. Kunonga, O. Abaire, and J. Ling, "Using artificial intelligence to improve public health: a narrative review," *Front. Public Health*, vol. 11, p. 1196397, Oct. 2023, doi: 10.3389/fpubh.2023.1196397.

[44] P. Galange, R. Mather, B. Yaffe, M. Whelan, and M. Murti, "Commentary on the adoption of a test-based versus syndromic-based approach to outbreak declaration and management in hospital and institutional settings," *Can. Commun. Dis. Rep.*, vol. 50, no. 3/4, pp. 102–105, Apr. 2024, doi: 10.14745/ccdr.v50i34a03.

[45] H. Wu, X. Lu, and H. Wang, "The application of artificial intelligence in health care resource allocation before and during the COVID-19 pandemic: Scoping review," *JMIR AI*, vol. 2, p. e38397, Jan. 2023, doi: 10.2196/38397.

[46] H. Rilkoff, S. Struck, C. Ziegler, L. Faye, D. Paquette, and D. Buckeridge, "Innovations in public health surveillance: An overview of novel use of data and analytic methods," *Can. Commun. Dis. Rep.*, vol. 50, no. 3/4, pp. 93–101, Apr. 2024, doi: 10.14745/ccdr.v50i34a02.

[47] N. P. Nguyen and P. Tran, "Bridging the mentorship divide: how large language models could reshape medical workforce equity," *npj Digit. Med.*, vol. 9, no. 1, p. 29, Jan. 2026, doi: 10.1038/s41746-025-02167-z.

[48] Z. Ahsan, "Integrating artificial intelligence into medical education: a narrative systematic review of current applications, challenges, and future directions," *BMC Med. Educ.*, vol. 25, no. 1, p. 1187, Aug. 2025, doi: 10.1186/s12909-025-07744-0.

[49] J. S. Izquierdo-Condoy, M. Arias-Intriago, L. M. Corrales, and E. Ortiz-Prado, "Artificial intelligence in medical education: Transformative potential, current applications, and future implications," *JMIR Med. Educ.*, vol. 12, p. e77127, Feb. 2026, doi: 10.2196/77127.

[50] L. Woods *et al.*, "Assessing the effectiveness of artificial intelligence education and training for healthcare workers: a systematic review," *BMC Med. Educ.*, vol. 26, no. 1, p. 549, Mar. 2026, doi: 10.1186/s12909-026-08969-3.

[51] Y.-M. Line, C.-C. Chou, T.-H. Jaing, and C. T. C. Okoli, "Generative AI for clinical reasoning: A scoping review," *Teach. Learn. Nurs.*, vol. 21, no. 1, pp. e305–e312, Jan. 2026, doi: 10.1016/j.teln.2025.08.008.

[52] N. P. Nguyen, "Hidden pitfalls in AI-generated MCQs: A call for caution," *Asia Pac. Sch.*, vol. 11, no. 2, pp. 129–130, Apr. 2026, doi: 10.29060/TAPS.2026-11-2/LE3898.

[53] N. P. Nguyen, "Teaching webside manner and digital empathy in telemedicine education," *Med. Sci. Educ.*, vol. 36, pp. 493–494, Oct. 2025, doi: 10.1007/s40670-025-02560-z.

[54] N. P. Nguyen and P. Tran, "Teaching the machine to heal in Vietnam," *AI Soc.*, pp. s00 146–025–02 794–w, Dec. 2025, doi: 10.1007/s00146-025-02794-w.

[55] N. P. Nghia, "AI is making clinical reasoning optional—and that should worry us," *BMJ*, vol. 393, p. s871, May 2026, doi: 10.1136/bmj.s871.

[56] N. P. Nguyen, V.-P. Tran, A. C. M. Bui, P. Tran, and H. Nguyen, "How can junior doctors safeguard their development as clinicians in the age of AI?" *JME Pract. Bioeth.*, vol. 2, no. 1, p. e000053, Mar. 2026, doi: 10.1136/jmepb-2025-000053.

[57] J. M. Kowal, K. H. Bryant, D. Segall, and T. Kantrowitz, "Harnessing generative AI for assessment item development: Comparing AI-generated and human-authored items," *Int. J. Sel. Assess.*, vol. 33, no. 3, p. e70021, Aug. 2025, doi: 10.1111/ijsa.70021.

[58] A. Newnham, T. Tattersall, and J. Odendaal, "Do

medical schools need to adapt their curriculum in order to teach medical students 'webside' manner? A systematic review," *Med. Sci. Educ.*, vol. 35, no. 6, pp. 3173–3183, Sep. 2025, doi: 10.1007/s40670-025-02498-2.

[59] N. P. Nguyen and P. Tran, "The third presence in the clinic room: How artificial intelligence is reshaping the clinical encounter," *Mayo Clin. Proc. Digit. Health*, vol. 4, no. 2, p. 100354, Jun. 2026, doi: 10.1016/j.mcpdig.2026.100354.

[60] N. P. Nguyen and P. Tran, "The AI adaptation divide in medical education: A generational perspective," *PLOS Digit. Health*, vol. 5, no. 4, p. e0001338, Apr. 2026, doi: 10.1371/journal.pdig.0001338.

[61] Unity Insights, "Independent evaluation of smart triage at the groves medical centre," Health Innovation Kent Surrey Sussex, NHS-funded evaluation report, 2024.

[62] American Medical Association, "Physician survey on augmented intelligence: 2026 findings," AMA Center for Digital Health and AI, Tech. Rep., 2026, [Online]. Available: https://www.ama-assn.org/system/files/physician-ai-sentiment-report.pdf. [Accessed March 15, 2026].

[63] Z. Z. Qin *et al.*, "Tuberculosis detection from chest x-rays for triaging in a high tuberculosis-burden setting: an evaluation of five artificial intelligence algorithms," *Lancet Digit. Health*, vol. 3, no. 9, pp. e543–e554, Sep. 2021, doi: 10.1016/S2589-7500(21)00116-3.

[64] P. Xue, M. T. A. Ng, and Y. Qiao, "The challenges of colposcopy for cervical cancer screening in LMICs and solutions by artificial intelligence," *BMC Med.*, vol. 18, no. 1, p. 169, Dec. 2020, doi: 10.1186/s12916-020-01613-x.

[65] H. Alami *et al.*, "Organizational readiness for artificial intelligence in health care: insights for decision-making and practice," *J. Health Organ. Manag.*, vol. 35, no. 1, pp. 106–114, Dec. 2020, doi: 10.1108/JHOM-03-2020-0074.

[66] J. N. Acosta, G. J. Falcone, P. Rajpurkar, and E. J. Topol, "Multimodal biomedical AI," *Nat. Med.*, vol. 28, no. 9, pp. 1773–1784, Sep. 2022, doi: 10.1038/s41591-022-01981-2.

[67] A. Y. Hannun *et al.*, "Cardiologist-level arrhythmia detection and classification in ambulatory electrocardiograms using a deep neural network," *Nat. Med.*, vol. 25, no. 1, pp. 65–69, Jan. 2019, doi: 10.1038/s41591-018-0268-3.

[68] M. V. Perez *et al.*, "Large-scale assessment of a smartwatch to identify atrial fibrillation," *N. Engl. J. Med.*, vol. 381, no. 20, pp. 1909–1917, Nov. 2019, doi: 10.1056/NEJMoa1901183.

[69] L. Seyyed-Kalantari, H. Zhang, M. B. A. McDermott, I. Y. Chen, and M. Ghassemi, "Underdiagnosis bias of artificial intelligence algorithms applied to chest radiographs in under-served patient populations," *Nat. Med.*, vol. 27, no. 12, pp. 2176–2182, Dec. 2021, doi: 10.1038/s41591-021-01595-0.

[70] E. J. Topol, "High-performance medicine: the convergence of human and artificial intelligence," *Nat. Med.*, vol. 25, no. 1, pp. 44–56, Jan. 2019, doi: 10.1038/s41591-018-0300-7.

[71] M. R. Arbabshirani *et al.*, "Advanced machine learning in action: identification of intracranial hemorrhage on computed tomography scans of the head with clinical workflow integration," *npj Digit. Med.*, vol. 1, no. 1, p. 9, Apr. 2018, doi: 10.1038/s41746-017-0015-z.

[72] J. J. Titano *et al.*, "Automated deep-neural-network surveillance of cranial images for acute neurologic events," *Nat. Med.*, vol. 24, no. 9, pp. 1337–1341, Sep. 2018, doi: 10.1038/s41591-018-0147-y.

[73] T. D. Shanafelt *et al.*, "Relationship between clerical burden and characteristics of the electronic environment with physician burnout and professional satisfaction," *Mayo Clin. Proc.*, vol. 91, no. 7, pp. 836–848, Jul. 2016, doi: 10.1016/j.mayocp.2016.05.007.

[74] K. D. Olson *et al.*, "Use of ambient AI scribes to reduce administrative burden and professional burnout," *JAMA Network Open*, vol. 8, no. 10, p. e2534976, Oct. 2025, doi: 10.1001/jamanetworkopen.2025.34976.

[75] G. J. You *et al.*, "Ambient documentation technology in clinician experience of documentation burden and burnout," *JAMA Network Open*, vol. 8, no. 8, p. e2528056, Aug. 2025, doi: 10.1001/jamanetworkopen.2025.28056.

[76] Z. Obermeyer and J. J. Emanuel, "Predicting the future — big data, machine learning, and clinical medicine," *N. Engl. J. Med.*, vol. 375, no. 13, pp. 1216–1219, Sep. 2016, doi: 10.1056/NEJMp1606181.

[77] K. Singhal *et al.*, "Large language models encode clinical knowledge," *Nature*, vol. 620, no. 7972, pp. 172–180, Aug. 2023, doi: 10.1038/s41586-023-06291-2.

[78] D. McDuff *et al.*, "Towards accurate differential diagnosis with large language models," *Nature*, vol. 642, no. 8067, pp. 451–457, Jun. 2025, doi: 10.1038/s41586-025-08869-4.

[79] V. V. Eriksen, S. Möller, and J. Ryg, "Use of GPT-4 to diagnose complex clinical cases," *NEJM AI*, vol. 1, no. 1, Jan. 2024, doi: 10.1056/AIp2300031.

[80] T. w. Bickmore *et al.*, "Patient and consumer safety risks when using conversational assistants for medical information: An observational study of Siri, Alexa, and Google Assistant," *J. Med. Internet Res.*, vol. 20, no. 9, p. e11510, Sep. 2018, doi: 10.2196/11510.

[81] T. H. Kung *et al.*, "Performance of ChatGPT on USMLE: Potential for AI-assisted medical

education using large language models," *PLOS Digit. Health*, vol. 2, no. 2, p. e0000198, Feb. 2023, doi: 10.1371/journal.pdig.0000198.

[82] Z. Kanjee, B. Crowe, and A. Rodman, "Accuracy of a generative artificial intelligence model in a complex diagnostic challenge," *JAMA*, vol. 330, no. 1, pp. 78–80, Jul. 2023, doi: 10.1001/jama.2023.8288.

[83] D. Lyell and E. Coiera, "Automation bias and verification complexity: a systematic review," *J. Am. Med. Inform. Assoc.*, vol. 24, no. 2, pp. 423–431, Mar. 2017, doi: 10.1093/jamia/ocw105.

[84] L. A. Celi *et al.*, "Sources of bias in artificial intelligence that perpetuate healthcare disparities—a global review," *PLOS Digit. Health*, vol. 1, no. 3, p. e0000022, Mar. 2022, doi: 10.1371/journal.pdig.0000022.

[85] J. W. Gichoya *et al.*, "AI recognition of patient race in medical imaging: a modelling study," *Lancet Digit. Health*, vol. 4, no. 6, pp. e406–e414, Jun. 2022, doi: 10.1016/S2589-7500(22)00063-2.

[86] Z. Obermeyer, B. Powers, C. Vogeli, and S. Mullainathan, "Dissecting racial bias in an algorithm used to manage the health of populations," *Science*, vol. 366, no. 6464, pp. 447–453, Oct. 2019, doi: 10.1126/science.aax2342.

[87] D. A. Vyas, L. G. Eisenstein, and D. S. Jones, "Hidden in plain sight — reconsidering the use of race correction in clinical algorithms," *N. Engl. J. Med.*, vol. 383, no. 9, pp. 874–882, Aug. 2020, doi: 10.1056/NEJMms2004740.

[88] J. Feng *et al.*, "Clinical artificial intelligence quality improvement: towards continual monitoring and updating of AI algorithms in healthcare," *npj Digit. Med.*, vol. 5, no. 1, p. 66, May 2022, doi: 10.1038/s41746-022-00611-y.

[89] A. Y. Lee *et al.*, "Multicenter, head-to-head, real-world validation study of seven automated artificial intelligence diabetic retinopathy screening systems," *Diabetes Care*, vol. 44, no. 5, pp. 1168–1175, May 2021, doi: 10.2337/dc20-1877.

[90] E. Asgari *et al.*, "A framework to assess clinical safety and hallucination rates of LLMs for medical text summarisation," *npj Digit. Med.*, vol. 8, no. 1, p. 274, May 2025, doi: 10.1038/s41746-025-01670-7.

[91] S. Bedi *et al.*, "Testing and evaluation of health care applications of large language models: A systematic review," *JAMA*, vol. 333, no. 4, pp. 319–328, Jan. 2025, doi: 10.1001/jama.2024.21700.

[92] J. Gravel, M. D'Amours-Gravel, and E. Osmanlliu, "Learning to fake it: Limited responses and fabricated references provided by ChatGPT for medical questions," *Mayo Clin. Proc. Digit. Health*, vol. 1, no. 3, pp. 226–234, Sep. 2023, doi: 10.1016/j.mcpdig.2023.05.004.

[93] J. Amann, A. Blasimme, E. Vayena, D. Frey, and V. I. Madai, "Explainability for artificial intelligence in healthcare: a multidisciplinary perspective," *BMC Med. Inform. Decis. Mak.*, vol. 20, no. 1, p. 310, Dec. 2020, doi: 10.1186/s12911-020-01332-6.

[94] M. Ghassemi, L. Oakden-Rayner, and L. A. Beam, "The false hope of current approaches to explainable artificial intelligence in health care," *Lancet Digit. Health*, vol. 3, no. 11, pp. e745–e750, Nov. 2021, doi: 10.1016/S2589-7500(21)00208-9.

[95] S. Jabbour *et al.*, "Measuring the impact of AI in the diagnosis of hospitalized patients: A randomized clinical vignette survey study," *JAMA*, vol. 330, no. 23, pp. 2275–2284, Dec. 2023, doi: 10.1001/jama.2023.22295.

[96] Office of the National Coordinator for Health Information Technology, "Health data, technology, and interoperability: Certification program updates, algorithm transparency, and information sharing; correction," Federal Register, 2024, accessed: May 19, 2026. [Online]. Available: https://www.federalregister.gov/documents/2024/02/08/2024-02519/health-data-technology-and-interoperability-certification-program-updates-algorithm-transparency-and. [Accessed March 10, 2026].

[97] U.S. Department of Health and Human Services, "Use of online tracking technologies by HIPAA covered entities and business associates," HHS.gov, 2026, [Online]. Available: https://www.hhs.gov/hipaa/for-professionals/privacy/guidance/hipaa-online-tracking/index.html. [Accessed March 15, 2026].

[98] K. Moulaei, S. Akhlaghpour, and F. Fatehi, "Patient consent for the secondary use of health data in artificial intelligence (AI) models: A scoping review," *Int. J. Med. Inform.*, vol. 198, p. 105872, Jun. 2025, doi: 10.1016/j.ijmedinf.2025.105872.

[99] L. Rocher, J. M. Hendrickx, and Y.-A. De Montjoye, "Estimating the success of re-identifications in incomplete datasets using generative models," *Nat. Commun.*, vol. 10, no. 1, p. 3069, Jul. 2019, doi: 10.1038/s41467-019-10933-3.

[100] U.S. Food and Drug Administration, "Cybersecurity in medical devices: Quality management system considerations and content of premarket submissions," U.S. Department of Health and Human Services, Food and Drug Administration, Guidance for Industry and Food and Drug Administration Staff, Feb. 2026, [Online]. Available: https://www.fda.gov/regulatory-information/search-fda-guidance-documents/cybersecurity-medical-devices-quality-management-system-considerations-and-content-premarket. [Accessed March 1, 2026].

[101] K. Budzyń *et al.*, "Endoscopist deskilling risk after exposure to artificial intelligence in colonoscopy: a multicentre, observational study," *Lancet Gastroenterol. Hepatol.*,

vol. 10, no. 10, pp. 896–903, Oct. 2025, doi: 10.1016/S2468-1253(25)00133-5.

[102] K. Goddard, A. Roudsari, and J. C. Wyatt, "Automation bias: a systematic review of frequency, effect mediators, and mitigators," *J. Am. Med. Inform. Assoc.*, vol. 19, no. 1, pp. 121–127, Jan. 2012, doi: 10.1136/amiajnl-2011-000089.

[103] J. W. Ayers *et al.*, "Comparing physician and artificial intelligence chatbot responses to patient questions posted to a public social media forum," *JAMA Intern. Med.*, vol. 183, no. 6, pp. 589–596, Jun. 2023, doi: 10.1001/jamainternmed.2023.1838.

[104] A. Sauerbrei, A. Kerasidou, F. Lucivero, and N. Hallowell, "The impact of artificial intelligence on the person-centred, doctor-patient relationship: some problems and solutions," *BMC Med. Inform. Decis. Mak.*, vol. 23, no. 1, p. 73, Apr. 2023, doi: 10.1186/s12911-023-02162-y.

[105] W. R. Small *et al.*, "Large language model–based responses to patients' in-basket messages," *JAMA Network Open*, vol. 7, no. 7, p. e2422399, Jul. 2024, doi: 10.1001/jamanetworkopen.2024.22399.

[106] M. e. Tai-Seale *et al.*, "AI-generated draft replies integrated into health records and physicians' electronic communication," *JAMA Network Open*, vol. 7, no. 4, p. e246565, Apr. 2024, doi: 10.1001/jamanetworkopen.2024.6565.

[107] V. Curcin *et al.*, "Learning health systems provide a glide path to safe landing for AI in health," *Artif. Intell. Med.*, vol. 173, p. 103346, Mar. 2026, doi: 10.1016/j.artmed.2025.103346.

[108] O. Freyer, I. C. Wiest, J. N. Kather, and S. Gilbert, "A future role for health applications of large language models depends on regulators enforcing safety standards," *Lancet Digit. Health*, vol. 6, no. 9, pp. e662–e672, Sep. 2024, doi: 10.1016/S2589-7500(24)00124-9.

[109] B. Z. Hameed *et al.*, "Breaking barriers: Unveiling factors influencing the adoption of artificial intelligence by healthcare providers," *Big Data Cogn. Comput.*, vol. 7, no. 2, p. 105, May 2023, doi: 10.3390/bdcc7020105.

[110] M. Hassan, A. Kushniruk, and E. Borycki, "Barriers to and facilitators of artificial intelligence adoption in health care: Scoping review," *JMIR Hum. Factors*, vol. 11, p. e48633, Aug. 2024, doi: 10.2196/48633.

[111] J. Javed and S. Islam, "Artificial intelligence and electronic health records: a narrative review of current applications and challenges in pediatric surgery," *World J. Pediatr. Surg.*, vol. 8, no. 5, p. e001100, Oct. 2025, doi: 10.1136/wjps-2025-001100.

[112] F. Magrabi, D. Lyell, and E. Coiera, "Automation in contemporary clinical information systems: a survey of AI in healthcare settings," *Yearb. Med. Inform.*, vol. 32, no. 01, pp. 115–126, Aug. 2023, doi: 10.1055/s-0043-1768733.

[113] D. Windecker *et al.*, "Generalizability of FDA-approved AI-enabled medical devices for clinical use," *JAMA Network Open*, vol. 8, no. 4, p. e258052, Apr. 2025, doi: 10.1001/jamanetworkopen.2025.8052.

[114] European Parliament, "Regulation (EU) 2024/1689 laying down harmonised rules on artificial intelligence (Artificial Intelligence Act)," 2024, [Online]. Available: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:32024R1689. [Accessed March 17, 2026].

[115] E. B. Hodcroft *et al.*, "Spread of a SARS-CoV-2 variant through Europe in the summer of 2020," *Nature*, vol. 595, no. 7869, pp. 707–712, Jul. 2021, doi: 10.1038/s41586-021-03677-y.

[116] J. S. Brownstein, C. C. Freifeld, B. Y. Reis, and K. D. Mandl, "Surveillance sans frontières: Internet-based emerging infectious disease intelligence and the HealthMap project," *PLoS Med.*, vol. 5, no. 7, p. e151, Jul. 2008, doi: 10.1371/journal.pmed.0050151.

[117] J. S. Brownstein, B. Rader, C. M. Astley, and H. Tian, "Advances in artificial intelligence for infectious-disease surveillance," *N. Engl. J. Med.*, vol. 388, no. 17, pp. 1597–1607, Apr. 2023, doi: 10.1056/NEJMra2119215.

[118] N. Ndembi *et al.*, "Integrating artificial intelligence into African health systems and emergency response: Need for an ethical framework and guidelines," *J. Public Health Afr.*, vol. 16, no. 1, p. 876, Mar. 2025, doi: 10.4102/jphia.v16i1.876.

[119] N. Schwalbe and B. Wahl, "Artificial intelligence and the future of global health," *The Lancet*, vol. 395, no. 10236, pp. 1579–1586, May 2020, doi: 10.1016/S0140-6736(20)30226-9.